TECHNICAL MEMORANDUM

488

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

Peter Bauer, Philippe Lopez, Deborah Salmond, Angela Benedetti, Sami Saarinen and Marine Bonazzola

Research Department

European Centre for Medium-Range Weather Forecasts, UK

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

February 2006

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

Series: ECMWF Technical Memoranda

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

Contact: library@ecmwf.int

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

Abstract

This paper presents the operational implementation of a 1D+4D-Var assimilation system of rain affected satellite observations at ECMWF. The first part describes the methodology and performance analysis of the 1D-Var retrieval scheme in clouds and precipitation that uses SSM/I microwave radiance observations for the estimation of total column water vapour (TCWV). This part describes the technical implementation of the TCWV observations in 4D-Var as well as the impact analysis.

The effect of the TCWV observations implied by precipitation on the 4D-Var analyses is significant and the total information content is comparable to that of SSM/I, HIRS and AMSU-B radiances. Regions with systematic drying in the analysis persist throughout the forecast while moistening is removed by precipitation after 1-2 days. The corresponding divergence increments reflect the feedback between moisture and dynamics. Forecast evaluation using model analyses exhibits mostly positive relative humidity forecast scores, in particular at 700 hPa and in the Tropics. Some short-term negative forecast scores are observed for geopotential near 1000 hPa and in the Southern hemisphere between days 2-4. Wind scores vary greatly between regions and different forecast lengths. Tropical cyclone tracking forecasts are only slightly affected by a reduced location error spread through the rain assimilation. Comparison to dropsonde observations of wind and temperature shows improvement as does TCWV analysis validation against independent observations from Jason radiometer data. The system has been implemented operationally in June 2005 and will be further developed towards a direct 4D-Var assimilation of radiances in clouds and precipitation.

1 Introduction

The assimilation of data relating to precipitation imposes several new challenges on a numerical weather prediction (NWP) system. The solution approach depends on several factors, particularly on whether the NWP system is used for process studies or for operational purposes. The main challenge for a variational assimilation system is the fact that moist physical processes must be accounted for in the minimization and that the corresponding model parameterizations are likely to be non-linear and potentially discontinuous. Most global and operational systems rely on incremental formulations of variational data assimilation because it produces the best trade-off between computational efficiency and an objective, three or four-dimensional, physically consistent system (Courtier et al. 1994). The incremental approach assumes linear model behaviour in the vicinity of the short-range forecast state that represents the background constraint in the analysis. The non-linear physics therefore represents an apparent contradiction to this assumption.

This contradiction has motivated several studies in the past that investigate the minimization performance employing more or less regularized versions of moist physics parameterization schemes (Vukićević and Errico 1993, Errico and Raeder 1999). Convection parameterizations as employed in global models were easily identified as the major source of non-linearities and discontinuities. Zupanski and Mesinger (1995) demonstrated the importance of regular physics parameterizations in the minimization but also the importance of accurate non-linear models in the forward integrations (model trajectory in four-dimensional variational assimilation, 4D-Var). The accuracy of non-linear models used in the integration and their consistency with linearized model versions used in the minimization may therefore produce a conflict.

Non-incremental variational assimilation systems can only be operated on individual cases for experimentation purposes or, in a global operational framework, with less spatial resolution than currently achievable with incremental systems. Only few studies exist that permit general conclusions on system design that are not neutralized by over-simplifications of other fundamental problems that had to be made for successful experimentation. In many cases, the results are strongly driven by the definition of case studies, observation errors, assimilation window length and cost function (Zou and Kuo 1996, Tsuyuki 1997, Xiao and Zou 2000) and none of these frameworks generalize to a potential operational set-up.

Other issues are related to formal requirements of the variational formulation such as unbiased observations with well defined and Gaussian-shaped error statistics as well as accurate and regularized observation operators that translate between model and observation space (Errico et al. 2000). Some of these issues are less apparent in physical adjustment and nudging schemes (e.g. Krishnamurti and Bedi 1996), however, these will not be considered here because they are not relevant to the data assimilation system that is implemented at ECMWF.

Operational systems have to be numerically stable and computationally efficient. This requirement as well as the lack of accurate process understanding seriously limit the choice of the employed observations and models (observation operators). The potential excitation of gravity waves has also been recognized by several authors (Tsuyuki 1997, Fillion 2002). However, it has been demonstrated that conservative use of rainfall observations in regional and global modelling systems is possible (Marécal and Mahfouf 2000, Marécal and Mahfouf 2002, Tsuyuki et al. 2003, Treadon et al. 2003). In these systems, data screening is an efficient way of ensuring linearity and avoiding excessive small-scale model perturbations that may cause gravity waves.

Another important development was the choice of microwave radiances as observations instead of rain rates. Differences between observed and modelled radiances are more Gaussian-shaped and less biased than differences between observed and modelled rain rates. The gradient of the combined moist physics - radiative transfer observation operator is always non-zero for radiance perturbations even in the absence of rain because radiances are also sensitive to modifications of temperature, moisture and surface state. The gradient of the moist physics parameterizations, however, is always zero if the first-guess rain is zero (Moreau et al. 2002).

Bauer et al. (2006a) showed that the largest contribution to the operator's non-linearity originates from the moist physics parameterizations, that errors may be characterized realistically and that biases are rather small. Multiple-scattering radiative transfer codes are computationally efficient and accurate, and tangent-linear and adjoint versions are available (Bauer et al. 2006b). Here, observations from the Special Sensor Microwave / Imager (SSM/I) were chosen because they are operationally available to the NWP community since 1987 and have been proven to be very stable and accurate.

A two-stage approach, i.e. a 1D-Var retrieval of an intermediate 'pseudo-observation' from radiance observations before assimilating it in 4D-Var, was operationally implemented at ECMWF. This method has several advantages. Firstly, a potentially non-linear observation operator can only be applied in an incremental assimilation to a very limited extent via additional outer loop updates of the model state. The non-linear 1D-Var, however, updates the model state during minimization and may therefore better cope with non-linearities. These mainly occur for more intense and convective precipitation (Bauer et al. 2006a). Secondly, the two-stage approach facilitates quality control because the output and performance of the 1D-Var retrieval can be used to screen out data before they affect the 4D-Var analysis. Marécal and Mahfouf (2000, 2002) have successfully demonstrated the benefits of such method for rain rate observations. It has also proven to be successful in the framework of early implementations of clear-sky radiance assimilation in the past (e.g. Eyre et al. 1993, Phallippou 1996).

The second part of this paper builds upon the 1D-Var retrieval product that is the total column water vapour (TCWV) retrieval in the presence of clouds and precipitation. The technical implementation related to observation operators, assumptions on errors and quality control are described in the first part (Bauer et al. 2006a). In Section 2, the paper describes the characterization of TCWV observation statistics in the global 4D-Var analysis and the definition of background and observation errors. Both are linked to the available choices of observation quantities (TCWV) and control variables (normalized relative humidity) in the ECMWF system on 28 June 2005. The global impact of TCWV observations on the analysis is illustrated in Section 3 by an illustration of the systematic modification of moisture, precipitation and divergence and by a calculation example of the relative contribution of TCWV observations to the total observation information content. The impact on model forecasts is presented in Section 4 as well as the forecast evaluation against model analyses and independent

observations over several months of experimentation. The paper is concluded by a discussion in Section 5.

2 Implementation

2.1 Computing

Currently, all calculations are carried out on an IBM p690 Power4+ cluster at ECMWF. Each of the processors has a 1.9 GHz clock rate. The model fields required in the data assimilation system are currently interpolated to observation locations and the observation operators are applied to the interpolated fields. Since the observation operator that is applied in the rain assimilation is very complex and requires many more parameters than available through the existing route, the 1D-Var rain assimilation calculations are carried out in grid-point space. This ensures that all fields that are output from the model physics can be used directly. This development also paves the way for the direct assimilation of cloud and rain affected radiances in 4D-Var.

For each model timestep (currently 12 minutes) the grid-point nearest to each SSM/I observation is matched to the observation. This is done within a fixed search radius of the observation and, if no gridpoint is found in this radius, the observation is not used. Grid-points are matched with only one observation and are removed from the search once a match has been found. This search has been coded such that the matching of observations to grid-points is identical if the number of processors is changed.

Since the distribution of the SSM/I observations is not homogeneous over the grid-point space (caused by the application over oceans only) there is a severe load imbalance of the rain assimilation work between the processors, whose workload distribution uses a geographical decomposition. This has been resolved by determining how much work each processor will have to do at each timestep from the number of observation/grid-point matches. Those processors which have much work to do pass work to processors that otherwise would have little or no work to do using message passing. This load-balancing has reduced the time consumed for the rain assimilation by a factor of 3-4 for a T_L799 L91 run (spectral model truncation at wavenumber 799 and 91 model layers). A full run of 4D-Var at T_L799 L91 takes about 80 minutes on 796 processors of the IBM p690+ cluster at ECMWF. The 1D-Var rain assimilation costs about 12% of the time for the 4D-Var when a 10 km search radius is used. The number of observations for 10, 7.5, and 5 km search radii is 62,000, 47,000, 32,000 and the computational cost is 9, 6, and 4 minutes, respectively.

2.2 Error definition

2.2.1 (i) Observation errors

As in the case of radiances (Bauer et al. 2006a), the determination of TCWV observation errors is not straightforward. The pseudo-observation values of TCWV originate from the 1D-Var retrievals and therefore the observation errors correspond to the 1D-Var retrieval errors. This depends on the background error assumed for the 1D-Var control vector (temperature and specific humidity), the observation error defined for the radiances and the accuracy of the observation operator.

An indirect method already applied by Bauer et al. (2006a) for the estimation of the radiance observation errors is the one proposed by Hollingsworth and Lönnberg (1986). This approach is based on the assumption that the observation error can be deduced from the spatial covariance of first-guess departures, σ_{FG}. This method assumes that the first-guess departures can be calculated with data from an observational network that is densely populated and whose neighboring observations are spatially uncorrelated. If the covariances are composed

of observation, σ_R, and background error, σ_B, contributions, the covariances with separation distance $d > 0$ only contain a background term while at $d = 0$ the covariances are $\sigma_{FG}^2 = \sigma_B^2 + \sigma_R^2$. Assuming that σ_B^2 is the same at $d = 0$ and for very small spatial offsets, σ_R^2 can then be obtained from the extrapolation of the spatial covariance curve from $d > 0$ to $d = 0$. However, in the case of TCWV error estimation, the Hollingsworth-Lönnberg method is less appropriate because the TCWV 'observations' are not true observations but originate from 1D-Var retrievals which were constrained already with the first-guess TCWV. Therefore both background and 'observations' contain TCWV information that is also spatially correlated.

For the linear case and for an error-free operator, the 1D-Var analysis error covariance matrix, **A**, can be directly derived as the covariance of the maximum *aposteriori* solution to the inverse problem:

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \tag{1}$$

with background error covariance matrix **B**, observation plus modelling error covariance matrix **R**, and Jacobian matrix of the observation operator **H** (e.g. Rodgers 2000). The problem of non-linearity may enter the inversion through the *apriori* information by a non-Gaussian probability distribution function and/or the observation operator. The former issue was shown to be less critical by Bauer et al. (2006a) in that for the lower three SSM/I channels, the first-guess radiance departure distributions have Gaussian shapes and exhibit very small and correctable biases. The linearity of the observation operator - here, the moist physics parameterizations and the multiple-scattering radiative transfer model - is ensured for the majority of the cases with larger deviations from linearity occurring for stronger convection and deeper cloud systems. This suggests that it is worth evaluating Eq. (1) due to the lack of other error estimation options. The Jacobian matrix is derived using the adjoint of the observation operator because the dimension of the observation vector is 7 (number of SSM/I channels) compared to the dimension 120 (60 model levels for temperature and specific humidity) of the state vector. The error standard deviation of TCWV is then calculated from:

$$\sigma_{A,TCWV} = \frac{1}{g} \left(\Delta \mathbf{p}^T \mathbf{A} \Delta \mathbf{p} \right)^{1/2} \tag{2}$$

with the vector of pressure differences between adjacent layer interfaces, $\Delta \mathbf{p}$. Figure 1 shows the resulting scatter plot from 198,114 cases produced with four 12-hour 4D-Var analyses on September 1-2, 2004. Figure 1a shows the realization of Eq. (1-2) as dots. For comparison, the average Hollingsworth-Lönnberg values per TCWV interval are overplotted as solid lines. The first conclusion is that Eq. (1) produces larger error estimates than those from the Hollingsworth-Lönnberg method even though the error gradient with TCWV is similar. Secondly, there are two regimes of errors which may roughly be approximated by linear fits with an offset of $1 kg m^{-2}$ and slopes that differ by a factor of two. Each regime is defined by the magnitude of **H** because **R** is constant. If the ratio between σ_A and σ_B is displayed (Fig. 1b) the same two regimes are identified. The reduction of 1D-Var analysis errors with respect to background errors is found to be between a factor of 1/1.8 and a factor of 1/3. The bulk of situations, however, belongs to the sample with larger improvements.

It was investigated whether the error regime could be classified according to the amount of TCWV, the rain amount, the activation of convection, the radiance first-guess departures, the TCWV or rain increments, the 1D-Var convergence or the linearity of **H**. Only a limited dependence on rain amount and active convection could be identified, i.e., with more rain and more active convection in a larger number of cases exhibited errors belonging to the class with larger errors. However, no unique classification criteria could be defined and therefore no justification for the simplification of the error calculation by curve-fitting (Marécal and Mahfouf 2002) in the context of the employed observation operators was found. This requires an explicit error calculation with Eq. (1)-(2).

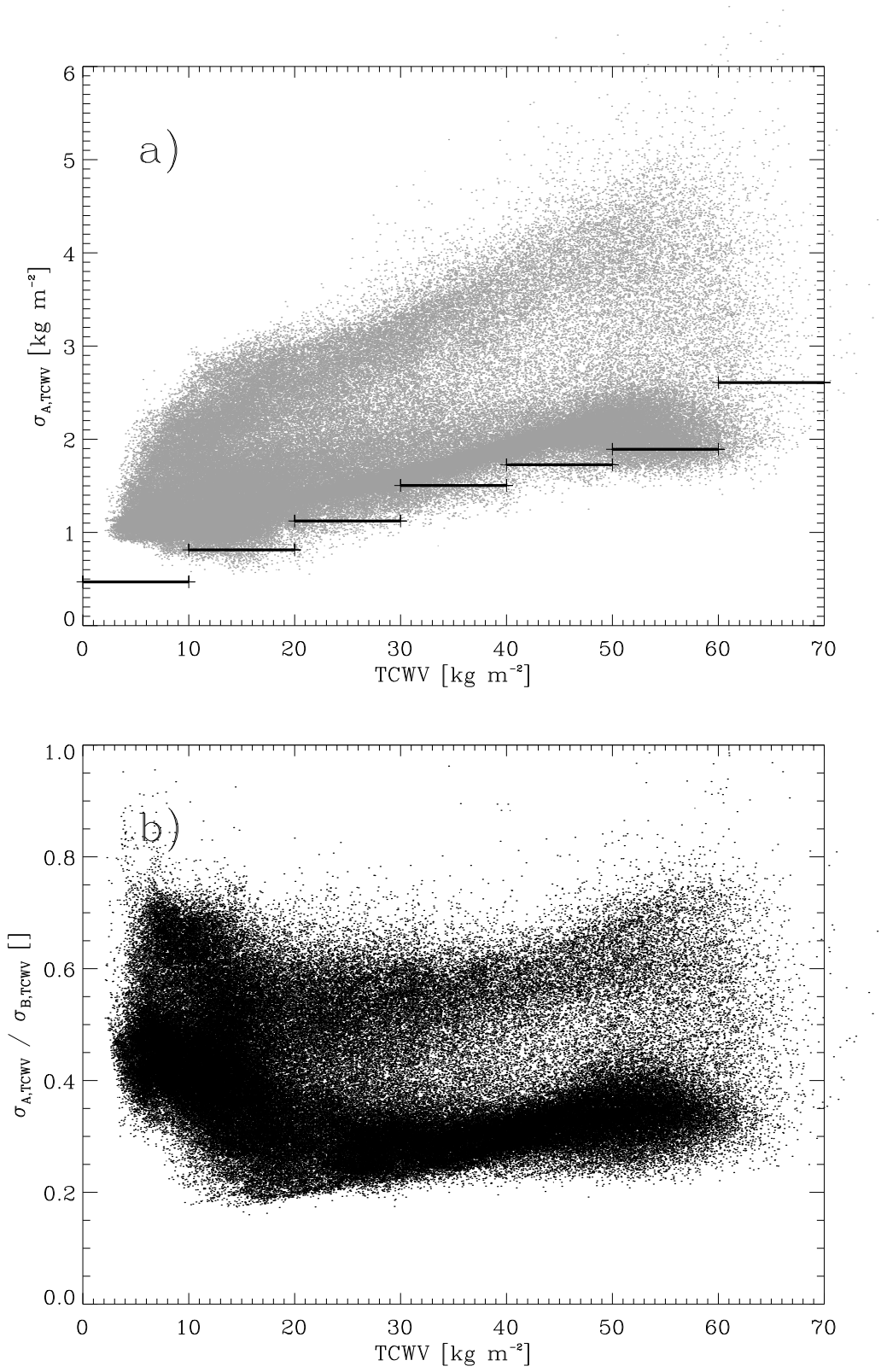

Figure 1: Analysis error standard deviation obtained from Eq. (1)-(2) (dots) as a function of TCWV and mean errors obtained from Hollingsworth-Lönnberg method (a; horizontal bars). Ratio of analysis and background error standard deviations (b).

2.2.2 (ii) Background errors

The specification of the background error covariance matrix (**B**-matrix) is an important aspect of the variational analysis since it is primarily through this matrix that the information from observations is distributed in the vertical. This applies to both 1D-Var and 4D-Var because the TCWV increments from the 1D-Var retrieval are vertically distributed according to the local background error covariance matrix. The definition of **B** is based on the method of the perturbed analysis ensemble described in Zagar et al. (2005). According to this method, the analysis state is statistically perturbed according to their random error characteristics. The resulting short-range forecasts are then also perturbed and their spread is assumed to represent the random error associated with the short-range forecast fields. In the most recent version of the ECMWF model, this background error correlation calculation has been modified by adding a wavelet decomposition for introducing spatial inhomogeneity (Fisher 2004). Although it would be desirable to use an identical definition for the 1D-Var, this would be computationally very expensive due to the technical limits imposed by the fact that the 1D-Var is run at the forward trajectory level and not at the analysis level. Therefore, the original correlation calculation method is applied in the 1D-Var while the 4D-Var analysis employs the wavelet decomposition.

Another important issue is whether the existing background error formulation requires a separation into rain affected and clear-sky error covariances. For this purpose, the so-called NMC-method (after National Meteorological Center, now National Centers for Environmental Prediction, NCEP; Parrish and Derber 1992) was applied using the operational ECMWF model. The method assumes that the mean differences between forecasts of different lengths valid for the same target dates, i.e., initialized with a delay, are representative of forecast errors and can therefore be used to quantify forecast error covariance statistics. The main advantage of the NMC-method is that the forecasts for calculating the statistics are available in the operational archives and no separate and computationally expensive run of the forecast or the analysis system is required. On the other hand, the computational cost is the main drawback for ensemble-type methods. The disadvantage of the NMC-method is that it requires forecast lengths (24-48 hours) that are larger than those of the short-range forecast that is used as a background in the analyses (3-15 hours). For our purpose, this disadvantage is not so relevant because the method is only employed for the investigation of forecast error statistics inside vs. outside precipitation.

The NMC-method was applied using three months (December 2004 - February 2005) of 24-hour and 48-hour forecasts of specific humidity, q, and temperature, T, with the same target dates, respectively. The rain - no rain distinction was obtained from a 3-hour forecast of large-scale and convective precipitation and by applying a minimum threshold of 0.05 mm/h. The statistics were derived for specific humidity, q, and temperature, T, at seven pressure levels. The resulting mean vertical correlations between q and T at 1000, 500 and 200 hPa with the other levels are presented in Fig. 2 for the entire globe (Figs. 2a, c, e) and the Tropics (30S-30N, Figs. 2b, d, f). The solid lines represent all gridpoints while the dashed lines show rainy gridpoints. For both q and T, no substantial difference can be identified over either tropical areas or the entire globe. Different vertical correlation patterns could be expected inside precipitation due to the temperature and moisture structure that is modified by latent heat release and evapouration. However, the existing formulation does not produce a significant precipitation signature. Therefore the operational formulation of **B** was applied for both 1D-Var and 4D-Var.

In Fig. 3. error standard deviations are displayed in terms of relative humidity, σ_R, because this parameter, scaled by its variance, is used as the moisture control variable at ECMWF. Mean profiles from the NMC-method statistics are compared with the error standard deviation formulation from Rabier et al. (1998) that was employed in operations. This formulation was obtained from radiosonde-model comparisons and contains model background temperature and relative humidity as predictors. A separate correction for lower layers was introduced, that also uses the pressure difference between the actual pressure and 800 hPa. Figure 3 compares

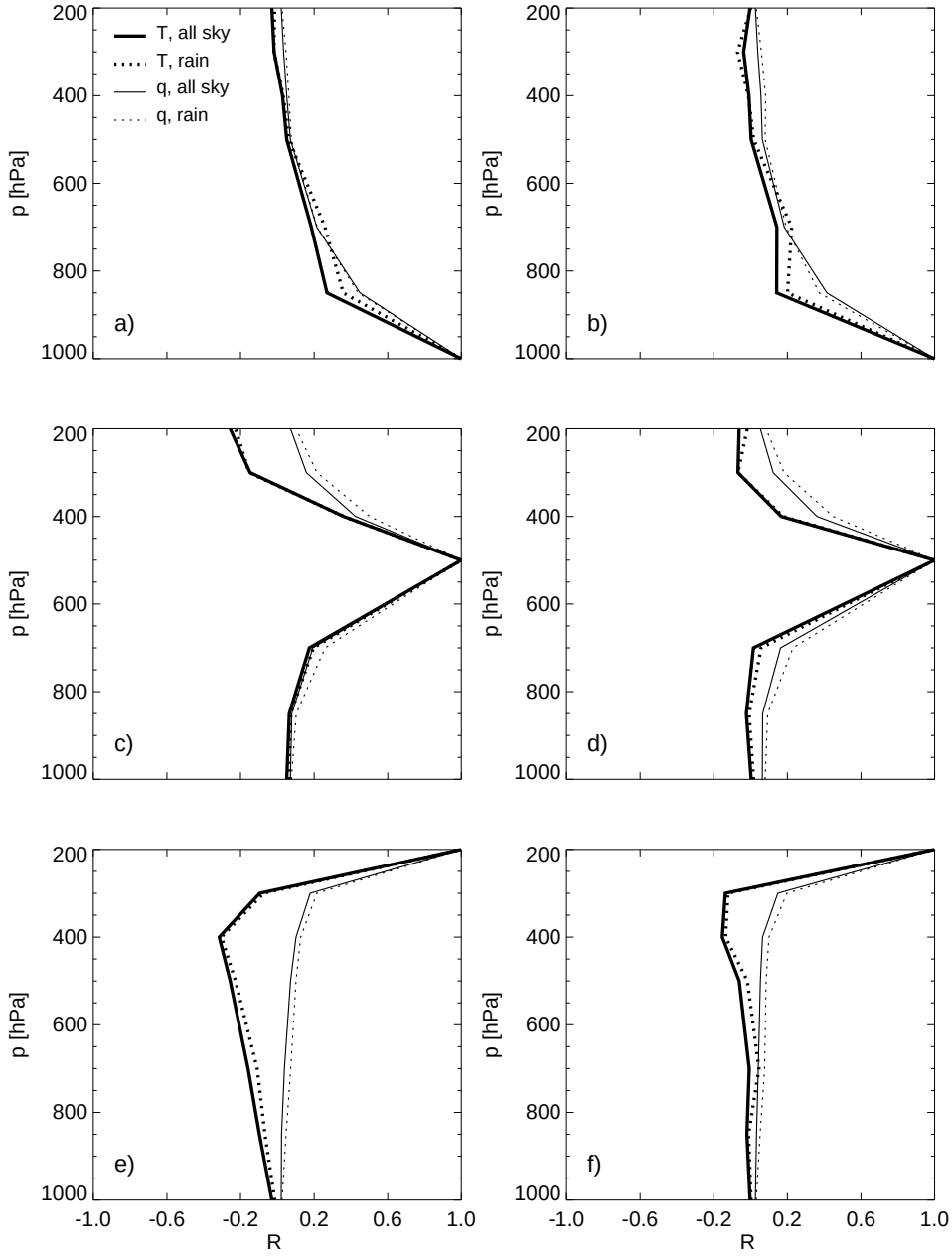

Figure 2: Temperature (thick solid: all data, thick dotted: rain areas) and specific humidity (thin solid: all data, thin dotted: rain areas) short-range forecast error correlations obtained from NMC-method with respect to levels 1000 (a, b), 500 (c, d), 200 hPa (e, f) for the whole globe (a, c, e) and the Tropics (b, d, f).

the globally averaged error standard deviations (Fig. 3a) and those for the Tropics (Fig. 3b). The NMC-profiles are included as well. As is the case for the error correlations, the difference between rain covered areas and clear skies is not significant regardless of the employed method and independent of latitude band. Again, this result does not advocate a distinction of background errors inside vs. outside precipitation given the operational

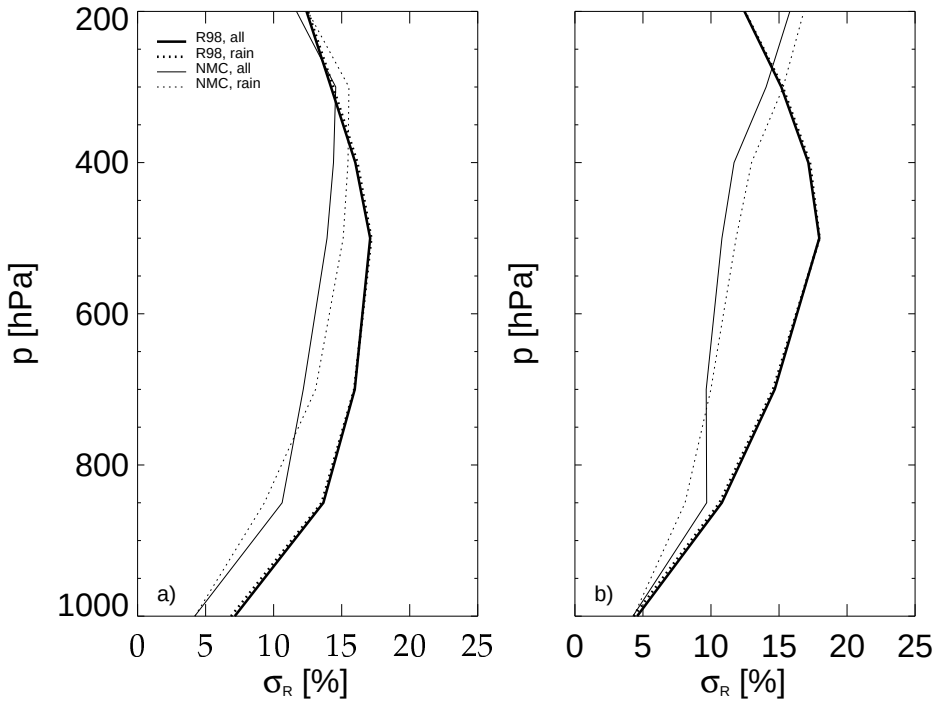

Figure 3: Relative humidity background error standard deviation (Rabier et al. 1998, RA98), thick solid: all data, thick dotted: rain areas; NMC-method, thin solid: all data, thin dotted: rain areas) short-range forecast error standard deviations for the whole globe (a) and the Tropics (b).

formulation.

3 4D-Var analysis evaluation

Several three-months experiments were run to evaluate the impact of the TCWV pseudo-observations in clouds and precipitation on analyses and forecasts. The experiments were carried out with the ECMWF forecasting system cycle CY29R2. CY29R2 is the first operational model version in which the rain assimilation is active. Table 1 summarizes the experiment names and characteristics. The *NEWOPS* and *NEWOPS − NORAIN* experiments follow the operational configuration in that there are 4 analyses per day, namely two shifted 12-hour assimilation window (between 9-21 and 21-9 UTC; delayed cut-off data assimilation, DCDA) analyses, as well as two 6-hour assimilation window (early delivery data assimilation, DA) analyses valid at 00 and 12 UTC. The former produce the first-guess fields for the latter and only the DA analyses initialize the medium-range forecasts (Haseler 2004). To limit the computational cost, the *RAIN* and *NORAIN* experiments were configured such that only two DCDA analyses were performed per day to initialize the medium-range forecasts.

3.1 Observation statistics

Figure 4 shows a time series from *RAIN* between 08-10/2004 of 4D-Var first-guess (FG) and analysis (AN) departure biases and standard deviations for the Northern hemisphere (30-60N, Fig. 4a), the Tropics (30S-30N, Fig. 4b) and the Southern hemisphere (30-60S, Fig. 4c), respectively. The departures are very stable and show

Table 1: Summary of rain assimilation and control experiments.

Experiment identifier	Type	Period
RAIN	DCDA analysis	08-10/2004
NORAIN	DCDA analysis without rain assimilation	08-10/2004
OLDOPS	DA+DCDA analysis (operations)	04/2005-06/2005
NEWOPS	DA+DCDA analysis (operations)	07/2005-present
NEWOPS-NORAIN	DA+DCDA analysis (operations) without rain assimilation	23-31/08/2005

that only in the Tropics a small bias of 0.4 kgm^{-2} remains, which corresponds to about 18% of the standard deviation and about 6-8% of the total TCWV amounts. In mid-latitudes, the biases are nearly zero. In the winter hemisphere, the standard deviations are smaller by 50% than in the summer hemisphere due to the lower abundance of moisture in the atmosphere. This result confirms that the observation operator, the bias correction (that is only applied to radiances), and the 1D-Var retrievals perform well and stabley so that no second bias correction prior to the assimilation of TCWV in the 4D-Var analysis is required.

For the month of September 2004 (*RAIN*), the global TCWV statistics show a FG (AN) mean departure of 0.29 (0.09) kgm^{-2} and FG (AN) departure standard deviation of 1.59 (1.09) kgm^{-2}. The mean observation error is 1.94 kgm^{-2}. The sample size is 1,472,320.

3.2 Analysis increments

At this stage, the assimilation of SSM/I radiances in rain affected areas only 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-correlations exist between moisture and dynamics. 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.

Figure 5 shows a three-month average of normalized TCWV differences between *RAIN* and *NORAIN* from the analyses (Fig. 5a), the 24-hour (Fig. 5b), 48-hour (Fig. 5c), and 72-hour (Fig. 5c) forecasts. The graphs therefore show the relative impact of the rain assimilation on top of the operational analysis. The TCWV-differences were normalized with the control TCWV analyses/forecasts to show the relative impact and to avoid that the effect in mid-latitude areas is underestimated due to the lower moisture abundance. The hatched areas in Fig. 5 indicate where the differences are statistically significant. The significance was calculated with a t-test applied to the TCWV-differences against the 0-difference and with a 5% significance level.

The mean analysis increment 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 increment differences are fairly similar to those produced by the 1D-Var (Bauer et al. 2006a, their Fig. 14a) even though the latter show more spatial detail because 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

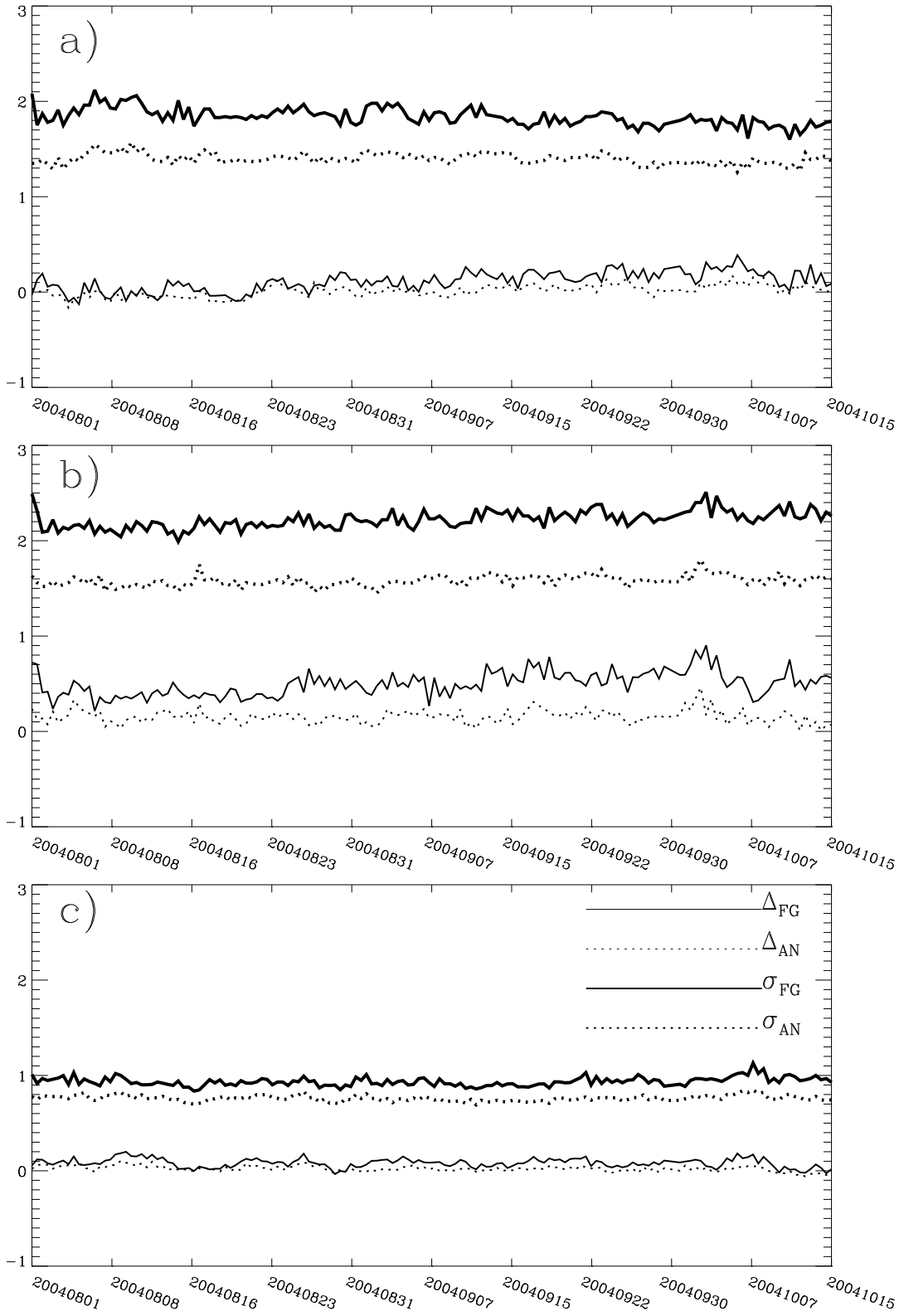

Figure 4: Time series between August 1 and October 15 of TCWV first-guess (FG) and analysis (AN) mean departures (Δ) and departure standard deviations (σ) for Northern hemisphere (a), Tropics (b), and Southern hemisphere (c) from experiment RAIN. Units are $kg\,m-2$.

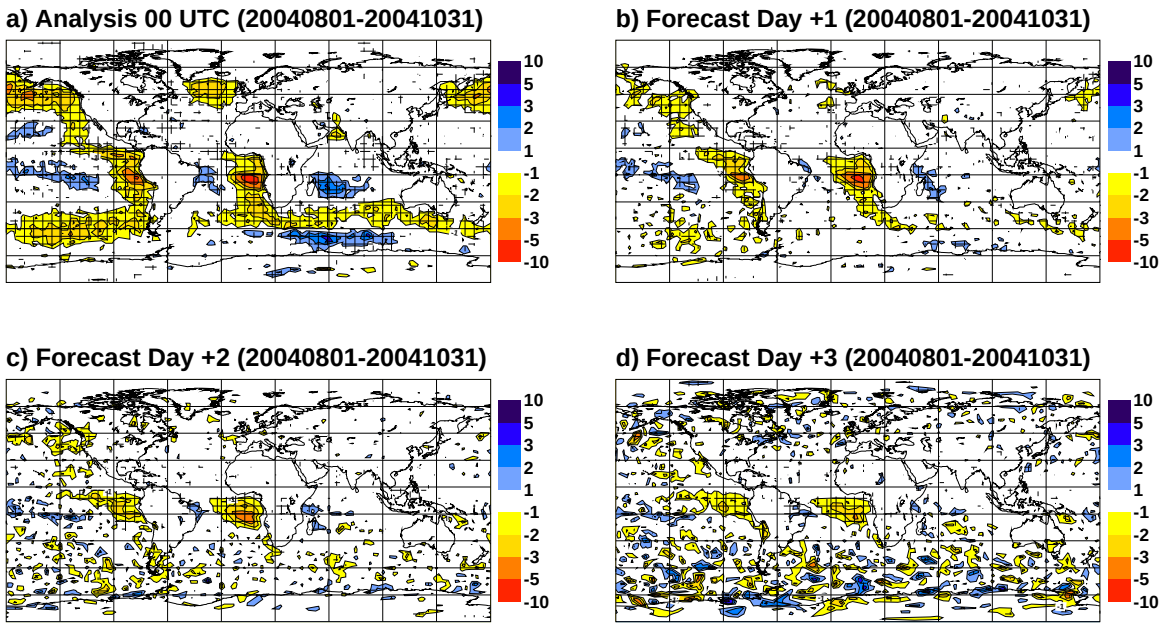

Figure 5: Mean normalized analysis (a) and 1-day (b), 2-day (c), 3-day (d) forecast difference of TCWV between RAIN and NORAIN experiments for period August-October 2004. Units are in % and crosses indicate statistical significance (95% significance interval with t-test on analysis differences against 0-difference).

and temperature served as a control variable, while they are combined to a scaled relative humidity control variable in 4D-Var (Andersson et al. 2005). 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 5b-d show how the analysis increments are propagated into the forecast. Again, mean increment 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. This can be seen in Fig. 6 that shows the mean differences of the forecasted precipitation (in mm) between 24 and 96 hours. The patterns are rather noisy but the smaller impact on precipitation in areas of TCWV-reduction than in areas of moistening can be identified. 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

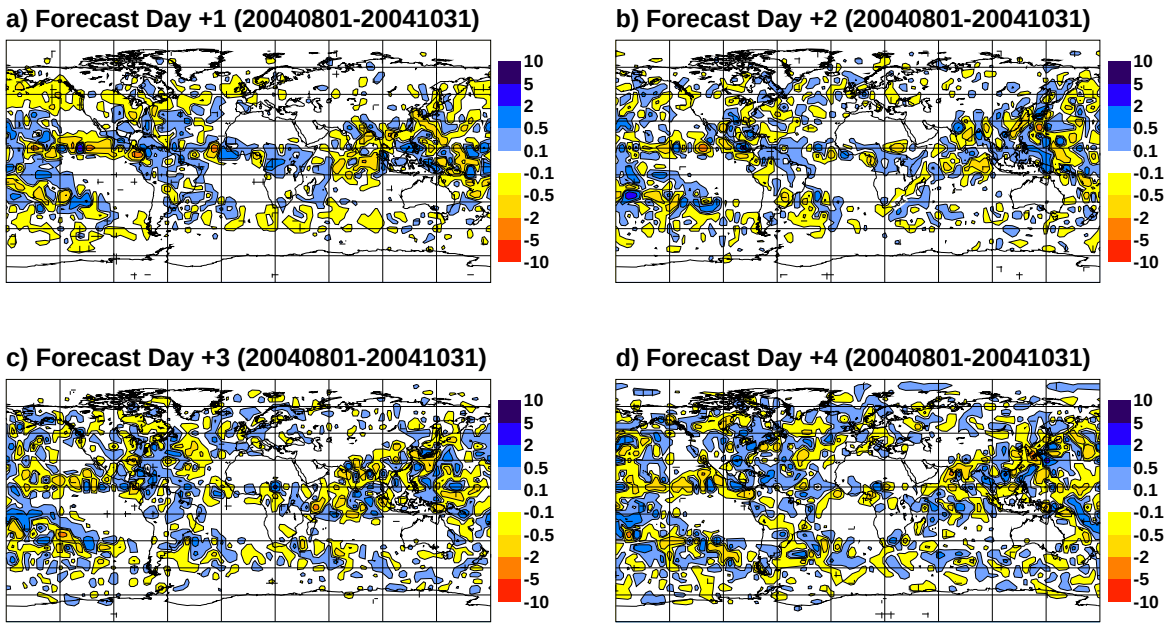

Figure 6: Mean 24-hour accumulated precipitation forecast difference for 1-day (a), 2-day (b), 3-day (c), and 4-day (d) forecasts between RAIN and NORAIN experiments for period August-October 2004. Units are in mm and crosses indicate statistical significance (95% significance interval with t-test on analysis differences against zero-difference).

significantly changed.

Figure 7 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 not be reproduced within, e.g., a 3D-Var system. 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.

3.3 Sensitivity to observations

It is convenient to estimate the individual impact of observations inside a data assimilation system that is constrained by a large number of diverse observations by evaluating the information content that is associated with an individual observation. Recently, this concept was formalized based on the ECMWF analysis system in

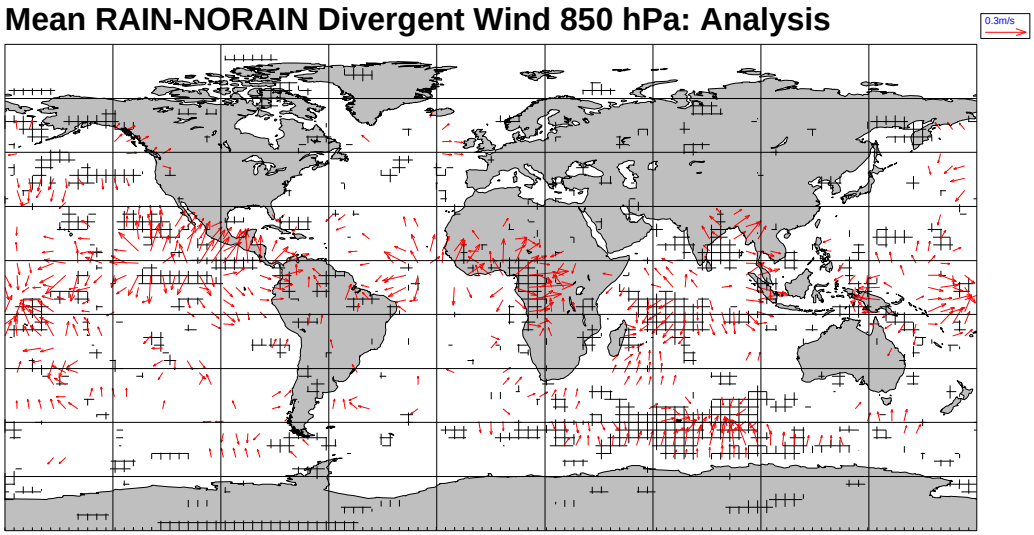

Figure 7: Mean 850 hPa divergent wind component difference from RAIN and NORAIN analyses. Calculated over period August-October 2004. Crosses indicate statistical significance (95% significance interval with t-test on analysis differences against 0-difference).

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.

Table 2 summarizes the *DFS* of each observation type from a single analysis (August 27, 2005, 00 UTC) in the *RAIN* experiment. Note that the *DFS* is the convolution of the information content per observation with the number of observations per observation type. 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. Moreover, the pattern of the geographical distribution of high influence (not shown here) resembles the precipitation distribution and also shows larger spatial variability than the influence maps of clear-sky satellite observations.

Table 2: Information content (DFS in %) per observation type in 12 UTC analysis on August 27, 2005, from NEWOPS (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).

Observation type	Number	DFS [%]
AIRS	1,280,872	30.94
AMSU-A	627,019	17.91
Aircraft	151,745	6.66
HIRS	94,886	5.84
Temp	66,029	5.03
Satob	110,704	4.95
TCWV	45,742	4.76
Meteosat	101,585	4.25
AMSU-B	66,038	4.22
Quikscat	113,702	3.64
Pilot	53,004	3.31
Synop	62,186	2.87
GOES	59,162	2.84
SSM/I	84,164	2.06
Drift buoys	4,564	0.60
Ozone	11,010	0.11

4 Forecast evaluation

Forecast evaluation is generally carried out with model analyses as a reference. However, this section will also include an evaluation with respect to observations that can be considered accurate and representative. Figure 8 shows the 3-month average relative humidity root mean square error (RMSE) difference between *RAIN* and *NORAIN*. The RMSE's were calculated from the forecasted fields, *FC*, and the respective analyzed fields, *AN*, with:

$$RMSE = \left[\frac{1}{N} \sum_{N} (FC - AN)^2 \right]^{1/2} \tag{3}$$

RMSEs are here displayed as zonal cross sections of relative humidity for the 24, 48, 72, and 120-hour forecasts. Negative RMSE-differences identify a better performance by *RAIN* and positive differences show better *NORAIN* forecasts. *RAIN* performs better throughout the Tropics between 40S-40N and all model levels. The RMSE-differences are about 5% and less with respect to the relative humidity RMSE's of the *NORAIN* experiment that are overplotted as isolines. A small negative impact of the rain assimilation can be observed near the surface between 40S-60S which is related to the area of strong moistening in the Southern Indian Ocean (see Fig. 5). This area is affected by strong surface winds. A potential model bias at high wind speeds may influence the 1D-Var retrieval and therefore degrade the analysis. This possibility will be further investigated in the future. In any case, the memory of the rain assimilation vanishes with increasing forecast length and becomes very small at day 5. Figure 9 shows similar statistics for temperature and reveals that RMSE-differences are very small but again a slightly negative impact near 500 hPa between 40S-60S is noted. For temperature, it is not obvious that this deterioration is initiated near the surface and it quickly dissipates after 24-48 hours (Figs. 9c-d).

Figures 10-12 show the *RAIN* and *NORAIN* RMSE forecast scores for relative humidity, geopotential height and wind vector, stratified into Northern hemisphere (N.Hem., 30N-90N), Southern hemisphere (S.Hem., 30S-

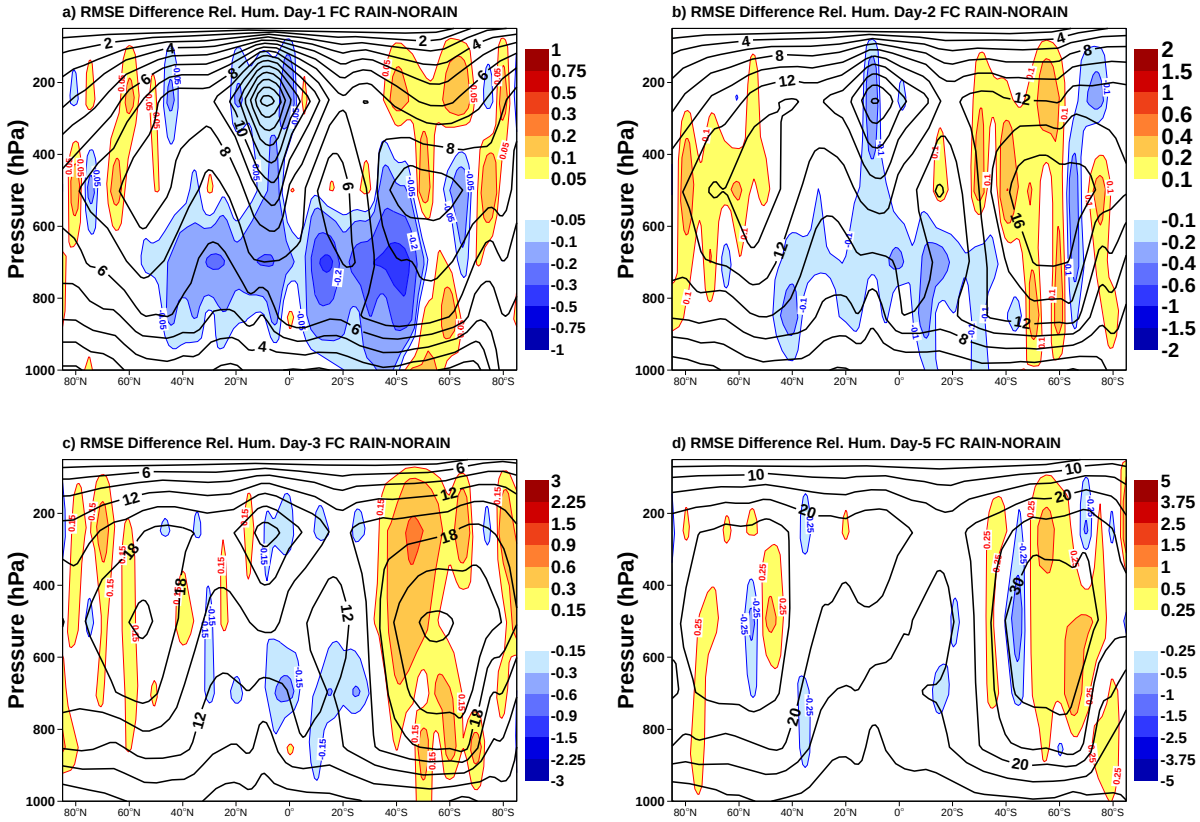

Figure 8: Zonal mean relative humidity forecast skill (as normalized difference of root mean square errors between RAIN and NORAIN) for 24-hour (a), 48-hour (b), 72-hour (c) and 120-hour (d) forecasts. Calculated over period August-October 2004 against own analysis, respectively. Isolines how NORAIN relative humidity root mean square errors.

90S), Tropics (30S-30N) and Northern Pacific (N.Pac., 20-75N, 140E-120W). Only levels at 1000, 700, 500, and 200 hPa are shown. The forecasts were again verified against own analyses and are displayed as:

$$score = 100 \left(\frac{RMSE_{RAIN} - RMSE_{NORAIN}}{RMSE_{NORAIN}} \right) \tag{4}$$

The error bars refer to the 95% significance level that was calculated from a t-test of the RMSE-difference against zero. The figures show that depending on area and level, a large variety of results are obtained with both positive and negative differences. In most cases, the statistical significance disappears after 3-4 days for geopotential and wind vector. For relative humidity (Fig. 10) and at 1000 hPa, the 12-hour forecast performance is between 1-3% worse for all areas. This negative score dissipates after 24 hours. At higher altitudes the performance is more neutral. As was seen in Fig. 8, the relative humidity scores are positive at 700 hPa through out the 10-day forecast in the Tropics. Also the other areas exhibit better performance by *RAIN* than *NORAIN*, however the scores tend to become neutral after 3-4 days. This signal is not seen in the geopotential height forecasts (Fig. 11). Despite the fact that these are more noisy beyond 72 hours, the short-range forecast performance varies from slightly better (S.Hem., 24 hours; Tropics, 48 hours) to worse (N.Pac., 24-72 hours; S.Hem., 72-120 hours). A similar distribution of results is obtained for the wind vectors (Fig. 12) with better scores in the Tropics (24-96 hours, 700-1000 hPa) and rather neutral scores elsewhere.

The forecast skill of tropical cyclone tracks was investigated separately for the rather active Atlantic systems

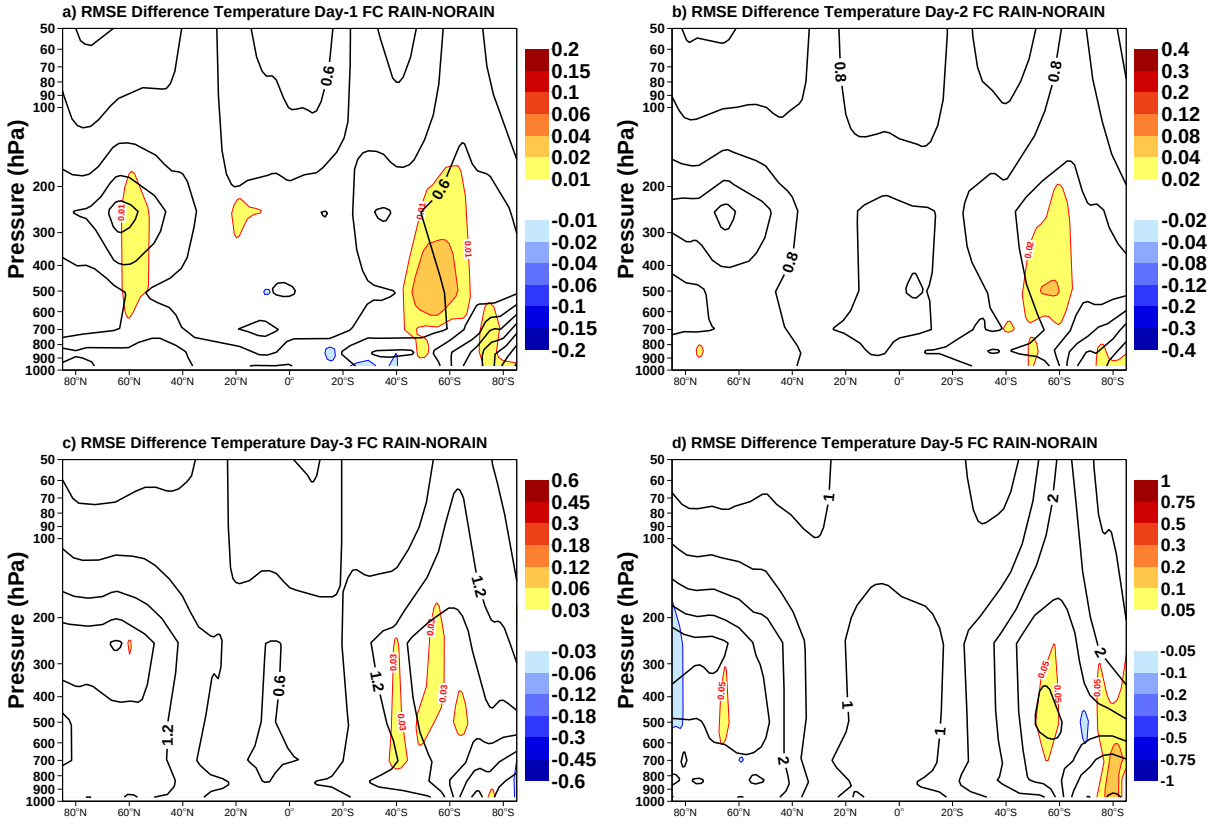

Figure 9: As Fig. 8 for temperature (in K).

over the period of August-September 2004. 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). The tracking algorithm currently in use at ECMWF is based on the identification of the cyclone's centre. The search for the centre is usually performed in a box or radius around a first-guess location. The first-guess location is obtained by extrapolating past track positions or by advection of the cyclone with the steering flow. 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, in particular as employed in the minimization (80-120 km). This overestimation of centre pressure reduces during the forecast because the forecast is carried out at the highest model resolution (here 40 km).

Figure 13 summarizes Δd (Fig. 13a), Δp (Fig. 13b) for the available number of cases (Fig. 13c). Both mean and standard deviations for experiments *RAIN* (solid) and *NORAIN* (dotted) are shown. The sample size reduces with increasing forecast length because cyclones may dissipate and the tracking routine fails in identifying the weakening centre of the cyclones. Mean location errors increase from 100 km to 550 km. This error mainly consists of an along-track error because, apart from central pressures 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 withing 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.

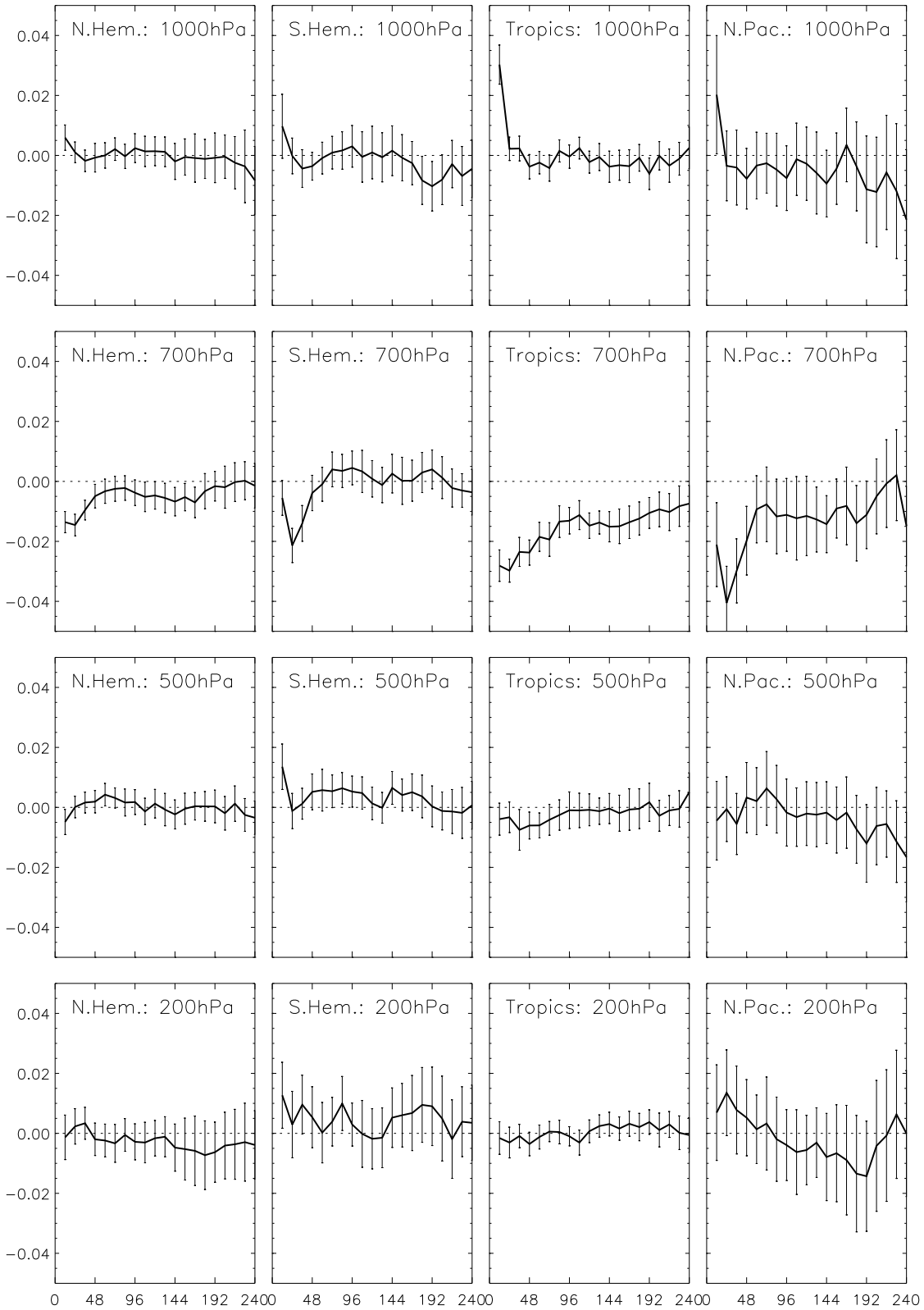

Figure 10: Normalized root mean square relative humidity error difference between experiments RAIN and NORAIN for Northern hemisphere (N.Hem.), Southern hemisphere (S.Hem.), Tropics and Northern Pacific (N.Pac.) and for 1000, 850, 500, 200 hPa levels. Error bars indicate 95% (t-test) confidence intervals. Calculated over period August-October 2004 against own analysis, respectively.

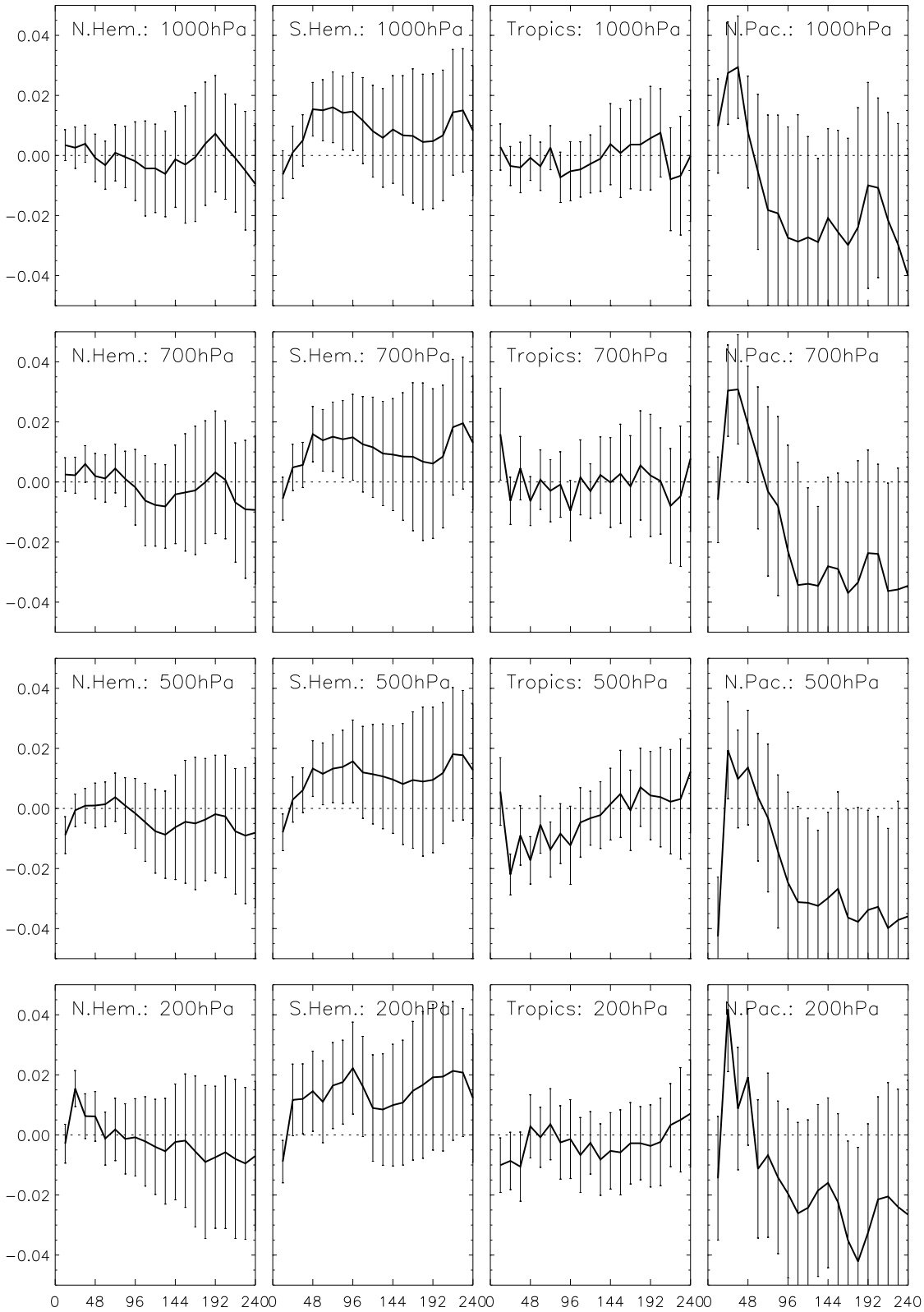

Figure 11: As Fig. 10 for geopotential height.

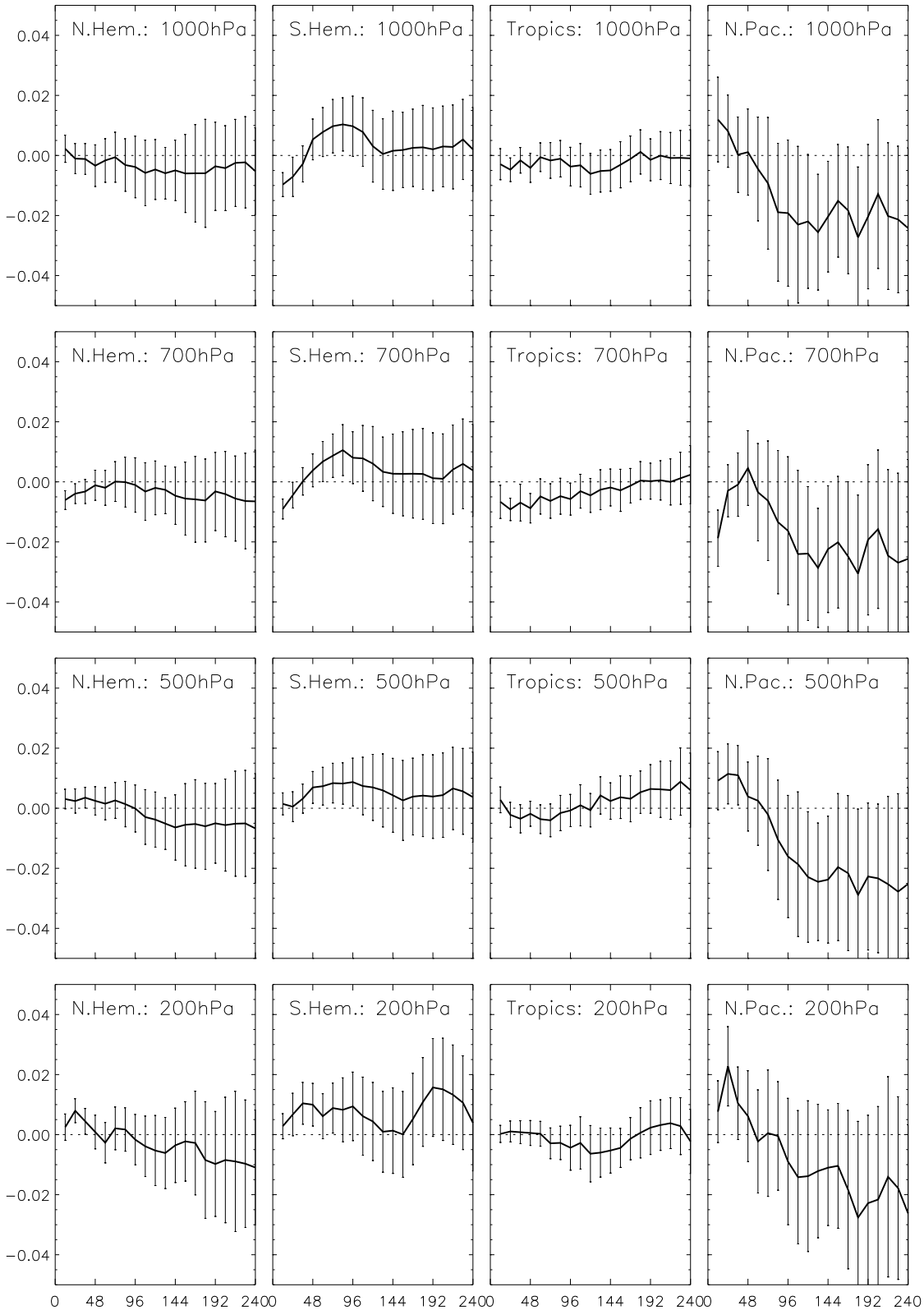

Figure 12: As Fig. 10 for vector wind.

The 4D-Var tropical cyclone analysis performance in comparison to dropsonde observations in the region 15N-40N and 50W-100W is displayed in Fig. 14. Figure 14a-b shows the first-guess (thick) and analysis departure means and standard deviations of wind (Δ_w, σ_w) and temperature (Δ_T, σ_T; Figs. 14c-d) between *RAIN* (dotted) and *NORAIN* (solid) for 08-10, 2004. 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*, biases, Δ_w, are larger throughout the troposphere. Secondly, the analysis standard deviations are very similar while the biases of *NORAIN* remain larger. For temperature, a similar but much weaker behaviour is observed. Again, the biases in *NORAIN* remain larger after the 4D-Var analysis.

Similar statistics have been calculated for hurricane Katrina that caused immense destruction in the south-east of the US in late August 2005. In Fig. 15, experiment *NEWOPS* refers to the operational model configuration that includes the rain assimilation while experiment $NEWOPS-NORAIN$ represents a copy of *NEWOPS* with a deactivated rain assimilation (see Tab. 1). Here, *NEWOPS* produces significantly smaller first-guess standard deviations in wind and slightly smaller ones in temperature. Analysis biases are slightly worse for *NEWOPS*. In summary, *RAIN* produces better biases and, at least in the case of hurricane Katrina, *NEWOPS* produces much better standard deviations.

Table 3 summarizes the statistics of independent TCWV retrievals from Jason Microwave Radiometer (JMR) that was launched on the Jason-1 satellite in December 2001 for providing observations related to air-sea interactions through combined radiometer-altimeter measurements. The radiometer carries three channels at 18.7, 23.8, and 34.0 GHz and points in nadir direction. The radiance measurements are used to correct for the excess path delay through the atmosphere experienced by the radar altimeter signal due to water vapour and suspended cloud liquid water. The path delay is equivalent to an estimation of TCWV and cloud liquid water path and can therefore be used for an independent evaluation of model TCWV. The statistics in Tab. 3 represent the September 2005 TCWV means of JMR retrievals, ECMWF model analyses from cycles C29R1 (*OLDOPS*, prior to rain assimilation) and CY29R2 (*NEWOPS*), the biases, standard deviations and correlations of each model version with respect to JMR-derived TCWV, respectively. For all areas, biases and standard deviations are reduced by about 10% and correlations are increased when the rain assimilation is present. Similar figures were obtained for August 2005 (not shown here). Generally, the model tends to overestimate TCWV by $\approx$0.5-1 kgm^{-2} in the winter hemisphere and the Tropics and by $\approx$1-1.5 kgm^{-2} in the summer hemisphere. These patterns correspond to the extensive negative TCWV increments produced by the rain assimilation that are evident in Fig. 5a in these areas for the period 08-10 2004.

5 Discussion

This paper presents the results of the implementation of the 1D+4D-Var assimilation of rain affected SSM/I radiances at ECMWF. The first part of the paper described the implementation of the 1D-Var algorithm that retrieves TCWV in the presence of clouds and precipitation over oceans using microwave radiances. The second part focused on the global model analysis and forecast performance when these new observations were included.

An important result is that a global rain assimilation is feasible in terms of computing time and stability. With 50-80,000 active rain observations in our configuration, the total increase in computing time is of the order of 5-10% within the entire 4D-Var analysis. No 4D-Var minimization problems have been noted. In this case, the implementation of a two-stage 1D+4D-Var assimilation is clearly advantageous because the non-linear 1D-Var and multiple screening stages before and after the 1D-Var permit better protection from instability. Future developments towards the direct assimilation of rain affected radiances will show how stringent the data screening has to become.

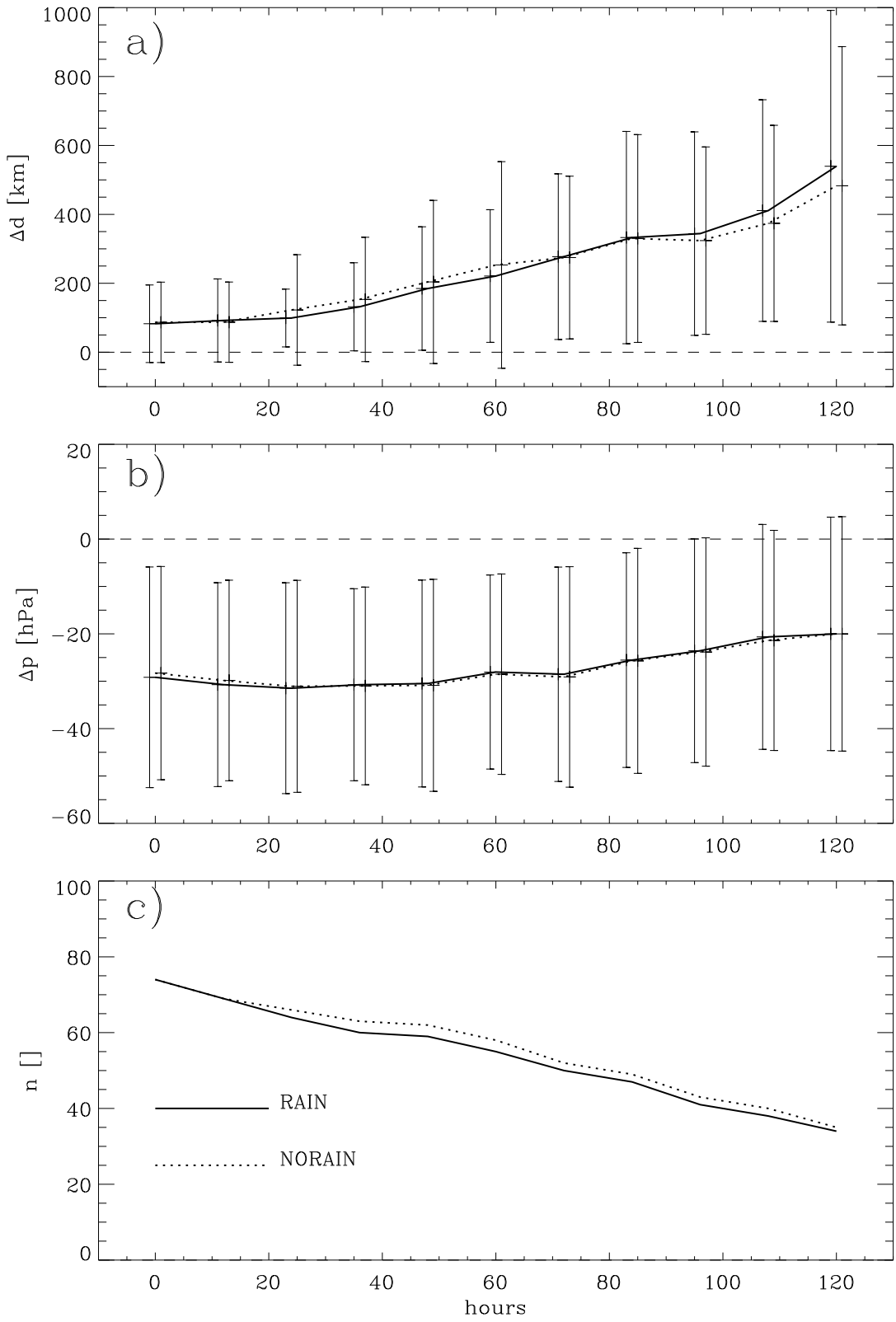

Figure 13: Mean and standard deviations of Tropical cyclone location distance error Δd (a; in km) and mean sea-level pressure error Δp (b; in hPa) and number of forecasts, respectively, for experiment RAIN and NORAIN. Error bars were offset by ± 1 hour with regard to valid forecast time to highlight differences.

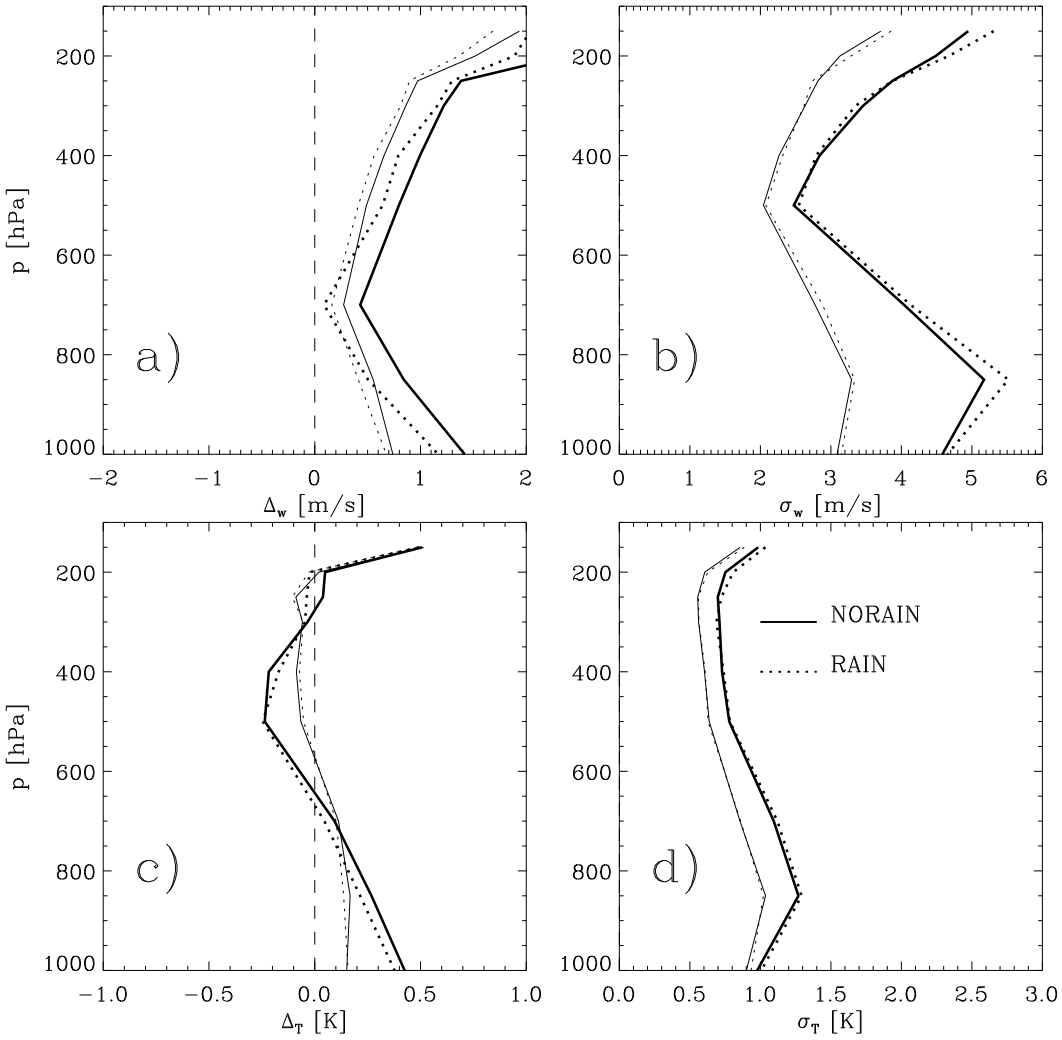

Figure 14: Biases, Δ and standard derviations, σ, between dropsonde observations and model first-guess (thick solid, thick dashed) or model analysis (thin solid, thin dashed) for experiment RAIN and NORAIN. Wind speed bias (a) and standard deviation (b), temperature bias (c) and standard deviation (d) for period August 1 - October 30, 2004, in the Caribbean.

The TCWV observation statistics showed that they are almost unbiased so that the only bias correction to be applied is the correction to radiances before the 1D-Var. Most of the systematic moisture increment structures that were produced by the 1D-Var remain after the 4D-Var analysis. The definition of moisture background errors near saturation causes positive increments to be smaller, less widespread and to dissipate faster in the forecast through precipitation. Globally, the model spin-down, that is the excessive generation of precipitation from systematic moistening of the lower atmosphere in the analysis, could be reduced. Areas with systematic drying persisted longer throughout the forecasts. In all cases however, the mean 4D-Var increments amounted to less than 5% of the total TCWV abundance. The forecasted precipitation patterns resemble the TCWV increment structures but show increasing noise with increasing forecast length when the rain assimilation is compared to a control experiment. In areas with large spatial extension of systematic TCWV increment structures the divergence increments followed the moisture redistribution indicating an active feedback between moisture and dynamics.

Forecast evaluation showed positive relative humidity scores in the Tropics near 700 hPa where the primary

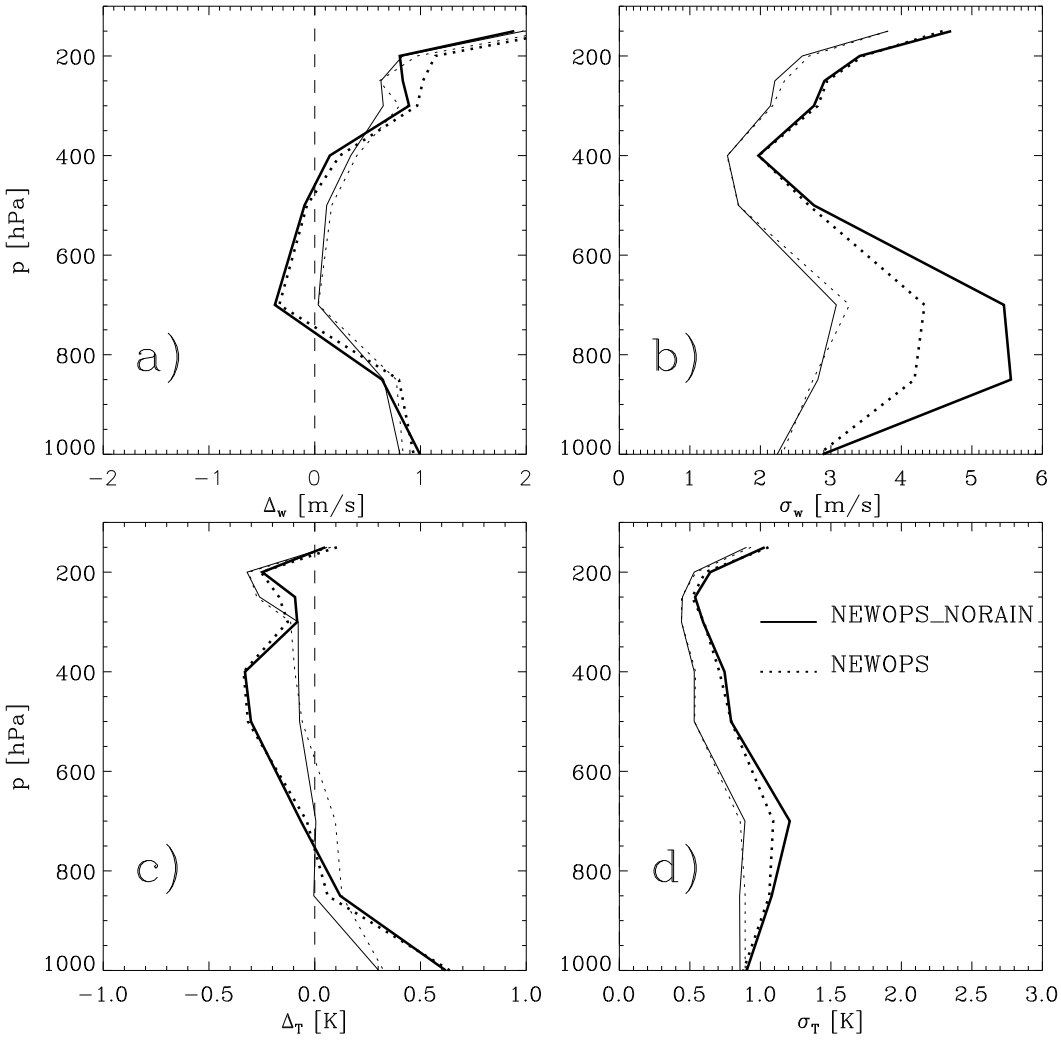

Figure 15: As Fig. 14 for model versions NEWOPS and NEWOPS – NORAIN in period August 23-30, 2005, tropical cyclone Katrina in the Caribbean.

analysis signal is produced. Near the surface, some local negative performances were observed but they only lasted for 12 hours. The geopotential height forecasts in the Southern hemisphere were slightly negative at all selected levels past day 2 (neutral to positive before). However, the noise in score differences between two experiments strongly increased past day 3-4 thus the statistical significance of mean differences was reduced.

The skill in tropical cyclone forecasts was rather similar between the rain assimilation experiment and its control given the inherent forecast variability. However, mean cyclone location was slightly improved in the rain assimilation experiment while forecast variability (location standard deviation) was much smaller than in the control. The evaluation of wind and temperature analyses using dropsonde observations in the vicinity of cyclones showed neutral-to-better performance of the rain assimilation. Comparison of model analyses with independent satellite-derived TCWV exhibited a 10% improvement with the model cycle that contains rain assimilation. The regions of greatest improvement correspond to the areas of largest systematic TCWV increments introduced by the rain affected observations.

We have not presented a dedicated validation of 1D-Var precipitation analyses or 4D-Var precipitation forecasts at this stage. This is for several reasons:

Table 3: JMR TCWV retrieval and model analysis comparison for September 2005. Units are $kg\,m^{-2}$.

Parameter	Global	North. hemisph.	Tropics	South. hemisph.	North. Atlant.	North. Pac.	South East. Pac.	Mediterr.
CY29R1:								
Number	81,955	22,233	25,294	35,036	10,656	11,263	5,932	627
Mean Model	25.2515	27.1074	40.5710	13.1665	25.0540	28.6561	29.2813	26.6023
Mean Jason	24.2384	25.8825	39.4168	12.3890	23.6507	27.6068	28.4461	24.5860
Bias (Jason-Model)	-1.0131	-1.2249	-1.1542	-0.7774	-1.4033	-1.0493	-0.8352	-2.0163
Standard deviation	1.7697	1.9388	2.0485	1.3703	1.9792	1.8681	1.6390	2.0368
Correlation	0.9938	0.9892	0.9865	0.9838	0.9883	0.9895	0.9902	0.9411
CY29R2:								
Number	81,955	22,233	25,294	35,036	10,656	11,263	5,932	627
Mean Model	25.1993	27.0159	40.4480	13.1649	24.9734	28.5603	29.2128	26.3803
Mean Jason	24.2384	25.8825	39.4168	12.3890	23.6507	27.6068	28.4461	24.5860
Bias (Jason-Model)	-0.9499	-1.1334	-1.0312	-0.7759	-1.3227	-0.9535	-0.7668	-1.7942
Standard deviation	1.6928	1.8338	1.9454	1.3561	1.8767	1.7609	1.5713	1.8157
Correlation	0.9943	0.9903	0.9876	0.9849	0.9894	0.9907	0.9909	0.9537

The radiance departures that provide the observational constraint in the 1D-Var assimilation are mainly related to the integrated liquid water path of precipitation since only the first three SSM/I channels are used. The radiance departure standard deviations are reduced by a factor of four after the analysis and the biases are insignificant. This suggests a much better fit of total water path after the 1D-Var on average. The freezing level height is linked to the temperature profile that is rather well predicted by the global model. This means that the retrieved surface rain rate is better than the first-guess rain rate since the rain profile does not exhibit a large vertical variability (see also radiance vs. rain rate assimilation by Moreau et al. 2003). This is further supported by the fact that also in the upper four channels the 1D-Var analysis departure biases and standard deviations are better than the first-guess ones. While this is only an indirect validation that will be affected by modeling errors, the comparison to rain retrievals from other sources will suffer from different sampling and additional retrieval errors contained in the derived reference products.

The mean rain increments (Bauer et al. 2006a) are rather small so that, given the uncertainties of oceanic rain products, the relative difference between 1D-Var first-guesses and analyses will not produce statistically significant offsets to the available reference data. The ECMWF model produces forecasts with accumulated rain amounts over 12-24 hour periods. These accumulations are produced on a 40 km grid and with 15 minute time steps (to be 25 km and 12 minute time steps with the next model cycle). We performed various intercomparisons with continental accumulated rain products on a daily basis and, say, 1-degree grids (radar network data over the US, daily rain-gauge product over Australia). These intercomparisons produced little evidence of a mean improvement or deterioration of land surface rainfall by the rain assimilation. Over oceans, only TRMM PR data could be considered a reference data set because all other products suffer from greater intrinsic uncertainties than forecast differences to be identified. However, even a TRMM PR product is generated from a very different sampling strategy (narrow swath, about 0-2 overpasses per day, instantaneous estimate at 4 km resolution) so that differences in representing area-averaged and accumulated precipitation overwhelm relative differences between two model experiments.

Given the complexity of the system and the potential destabilization of the 4D-Var analysis by the assimilation of observations where only few other observations are present and where strong non-linear effects may be produced by the moisture analysis, the current implementation must be considered a great success. As to date,

this is the first global operational 4D-Var system that employs rain affected observations from satellites. The impact on the analysis is significant and similar or larger than those from HIRS, SSM/I and AMSU-B sensors.

Apart from technical details, future areas of improvements are linked to the direct assimilation of radiances because the entire 4D-Var observation operator (including dynamics) would be involved in the minimization. This may require a list of other substantial changes to the system such as the change of the control variable (from relative humidity to total water), a higher model resolution and the activation of moist physics parameterizations earlier in the minimization. These modifications will impose more computational cost and potentially a more restrictive or more optimal data screening. However, these difficulties will be compensated by the assimilation of rain affected observations that is consistent with other radiance observations and that will produce a more balanced interaction between moisture information and model dynamics in the four-dimensional analysis.

Acknowledgements

The authors are grateful to Thomas Jung for the provision of the diagnostics and help with their interpretation, to Carla Cardinali for the calculation of the information content, to Saleh Abdalla for the evaluation of Jason data with model cycles CY29R1 and CY29R2. We thank John Hague for his idea and implementation of the load balancing and Jean-Noël Thépaut for many valuable discussions on the subject as well as Mats Hamrud and Jan Haseler for helping with the code implementation. The work was partly funded by the European Space Agency (ESA) under contract No. 17193/03/NL/GS. A. Benedetti was partly funded by the Cloudsat project (contract NAS5-99237).

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, 2005: Assimilation and Modelling of the Hydrological Cycle. *Bull. Amer. Meteorol. Soc.*, **86**, 387-402.

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

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

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

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.

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

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

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

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

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

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

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.

Haseler, J., 2004: The early-delivery suite. *ECMWF Technical Memorandum*,**No. 454**, 35 pp.

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

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.

Parrish, D.F., and J.C. Derber, 1992: The National Meteorological Center's Spectral Statistical-interpolation Analysis System. *Mon. Wea. Rev.*,**120**, 1747–1763.

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

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

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

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. *Proc. ECMWF/GEWEX Workshop on Humidity Analysis, 8-11 July 2002*, 33–47.

Tsuyuki, T., 1997: Variational data assimilation ibn 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. *Proc. ECMWF/GEWEX Workshop on Humidity Analysis, 8-11 July 2002*, 59–67.

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

Vukićević, T., and R.M. Errico, 1993: Linearization and adjoint of parameterized moist diabatic processes. *Tellus*,**45A**, 493–510.

Xiao, Q., X. Zou, and Y.-H. Kuo, 2000: Incorporating the SSM/I-derived precipitable water and rainfall rate inot a numerical model: A case study for the ERICA IOP-4 cyclone. *Mon. Wea. Rev.*,**128**, 87–108.

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.

Zagar, N., E. Andersson, and M. Fisher, 2005: Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *Q. J. Roy. Meteor. Soc.*,**131**, 987–1011.

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