TECHNICAL MEMORANDUM

535

Lessons learnt from the 1D+4D-Var assimilation of rain and cloud affected SSM/I observations at ECMWF

A. J. Geer, P. Bauer and P. Lopez

Research Department

Also released as EUMETSAT/ECMWF Fellowship Programme Research Report no. 17

September 2007

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

Cloud and rain affected observations from SSM/I have been assimilated operationally at ECMWF since June 2005, using a 1D+4D-Var method. A number of minor improvements have been made to the system over time; these are described. The rest of this paper examines the performance of the system, using departure statistics, forecast verification, observing system experiments (OSEs) and comparison to independent rainfall observations from the precipitation radar on the Tropical Rainfall Measuring Mission (TRMM). Cloud and rain affected SSM/I observations are shown to benefit forecast scores, principally in the tropics, and they benefit not just relative humidity but also vector wind. However, there are many areas where the system could still be improved. First, in the 1D+4D-Var method used here, only the TCWV amount retrieved by 1D-Var is passed to 4D-Var for assimilation. Cloud and rainfall amounts are not passed into the 4D-Var analyses. It appears that the forecast benefits come from the ability to assimilate total column water vapour in cloudy and rainy areas, which are not otherwise observed by satellites. The system is not yet making use of the observational information on cloud and rain. Second, the simplified moist physics operators used in the 1D-Var retrieval are shown to produce excessive amounts of rain compared to observations. The current bias correction scheme is partly able to mitigate the impact of this, but there are still residual biases, which mean that 1D-Var has a tendency to dry the atmospheric column in cloudy and rainy areas. A theme underlying all this work is the inability of the first guess to correctly predict the location and intensity of rainy and cloudy areas observed by SSM/I. Locations can be mismatched by 100 km to 200 km. This means that in 1D-Var, rain will often have to be completely removed in one retrieval, and completely added in another. Many implications of this are listed in this report, but the most important is that when observations are assimilated in separate observationally-determined 'clear-sky' and 'rainy' streams, a very large sampling bias can be created. Our recommendation is that in future all satellite observations with a rain and cloud information content are assimilated in a single stream, including cloudy and rainy radiative transfer when required. For SSM/I observations, this should be achieved with a future transition to direct 4D-Var of cloud and rain affected SSM/I observations, which will also help cure many of the problems listed above.

1 Introduction

In terms of both modelling and observations, it is perhaps in the area of clouds and precipitation that Numerical Weather Prediction (NWP) is least well-developed. Such phenomena occur on much smaller scales than the 25 to 60 km horizontal resolution of current forecast models. Satellite sensors in the infrared (IR) and visible are unable to see much below the cloud top. In contrast, sensors in the microwave are much less sensitive to cloud. Because they can provide information in the region of active weather systems, which are usually cloudy, microwave observations provide a large benefit to forecasts, particularly in the Southern Hemisphere (SH, e.g. English et al., 2000). Until recently, however, microwave observations remained unused in areas of heavier cloud and rain, because of the difficulty of simulating radiative transfer in such conditions, and the non-linearity of the moist physical processes involved. It is hoped that forecast benefits will come from assimilating more observations in such areas. One of the first operational uses of satellite data in such areas is the assimilation of cloud and rain-affected Special Sensor Microwave / Imager (SSM/I) radiances at the European Centre for Medium-range Weather Forecasts (ECMWF), which started in June 2005 (Bauer et al., 2006a,b). After two years of monitoring, experimentation, and further developments, this paper reviews the performance of the system. We identify a number of challenges that remain to be overcome before we can get the best out of the observations.

SSM/I observations are sensitive to surface properties, total column water vapour (TCWV), cloud ice and water, and rain and snow (Chevallier and Bauer, 2003). In order to assimilate all this information, an observation operator is required that links the analysis control variables (principally temperature, pressure, wind and water vapour, or transforms thereof) to the radiances seen by the satellite. For cloudy and rainy sky assimilation at ECMWF, the operator comes in two parts: a moist physics operator calculates the cloud and precipitation

profile associated with a particular profile of temperature and humidity, and from this, a radiative transfer operator calculates the observed radiances. The moist physics operator is composed of a convection scheme (Lopez and Moreau, 2005), which models the effect of clouds formed by sub-grid scale processes, and a large-scale condensation scheme (Tompkins and Janisková, 2004), which models clouds when they are formed by model-resolved processes. Given a profile of temperature, moisture, cloud and precipitation, and surface parameters, it is then possible to calculate the microwave radiances seen by the satellite, taking into account not just absorption and emission, but also the scattering effects of raindrops and ice particles (Bauer et al., 2006d).

The assimilation of rain-affected SSM/I observations fits into a broader intention to use both microwave and infrared observations in the presence of clouds and rain (Andersson et al., 2005) and builds on much previous work at ECMWF. SSM/I observations have been assimilated in clear skies for many years (Phalippou, 1996; Gérard and Saunders, 1999). A first step towards using observations in cloudy and rainy areas was to simulate brightness temperatures from the analyses and to compare them to their observed equivalents. Comparing to infrared observations Chevallier and Kelly (2002) concluded that cloud and rain were in general well modelled in the extratropics, but that the tropics were less well represented, particularly in the diurnal cycle. Comparing to microwave observations, Chevallier and Bauer (2003) found excessive rain and cloud amounts in the tropics. If the same information could be assimilated, it might be possible to start to correct such problems. To begin testing the assimilation of such data, moist physics operators were developed to diagnose cloud and rain amounts from the temperature and moisture profile. These were used to make 1D-Var retrievals from satellite retrieved rainfall rates (Marécal and Mahfouf, 2000) and directly from satellite radiances (Moreau et al., 2004). Such retrievals could then be assimilated into the 4D-Var system, leading to experiments with the 1D+4D-Var technique (Marécal and Mahfouf, 2002). Before these techniques could be made operational, it was clear that improvements would be needed in radiative transfer and the moist physics. With these improvements (Tompkins and Janisková, 2004; Lopez and Moreau, 2005; Bauer et al., 2006d), it became possible to assimilate rainy/cloudy SSM/I observations operationally (Bauer et al., 2006a,b) and also to experiment with assimilation of rainfall retrievals from the Precipitation Radar (PR) on the Tropical Rainfall Measuring Mission (TRMM) (Benedetti et al., 2005), cloud optical depths from MODIS (Benedetti and Janisková, 2007) and ground-based observations such as cloud and rain radars and rain gauges (Lopez et al., 2006; Lopez and Bauer, 2007).

Efforts to assimilate cloud and precipitation observations are of course not unique to ECMWF. Errico et al. (2007) provide references to many different experiments, and outline a number of issues that remain to be solved. In general, the choice has been made to assimilate derived products, such as precipitation rate, and this has been done both in the context of global NWP (e.g., Tsuyuki, 1997; Hou et al., 2004) and mesoscale modelling (e.g., Xiao et al., 2000). As yet, few of these experimental approaches have been made operational, and only at ECMWF are rain and cloud affected radiances (as opposed to products such as rain rates) assimilated. Of most relevance to this work are the experiments with 1D-Var retrievals from SSM/I performed at the Meteorological Service of Canada (MSC, Deblonde et al., 2007), though these retrievals are not yet assimilated in their global NWP system.

This paper first provides an overview of the 1D+4D method for assimilating SSM/I radiances in cloud and precipitation affected areas (Sec. 2.1), and then summarises a number of modifications that have been made to the system since its original operational implementation (Sec. 2.2). The rest of this paper describes the current performance of the system (Sec. 3), making use of departure statistics, forecast verification, Observing System Experiments (OSEs) and comparisons with independent data. Of these, the OSEs have been described in more detail by Kelly et al. (2007). We are currently expanding the 1D+4D system to assimilate SSM/I-like channels from TMI and AMSR-E, and SSMIS, giving additional coverage, and to extend the assimilation to land surfaces. These efforts will be described in separate notes. The ultimate intention in the area of rain and cloud affected microwave observations at ECMWF is to assimilate radiances directly in 4D-Var. Understanding

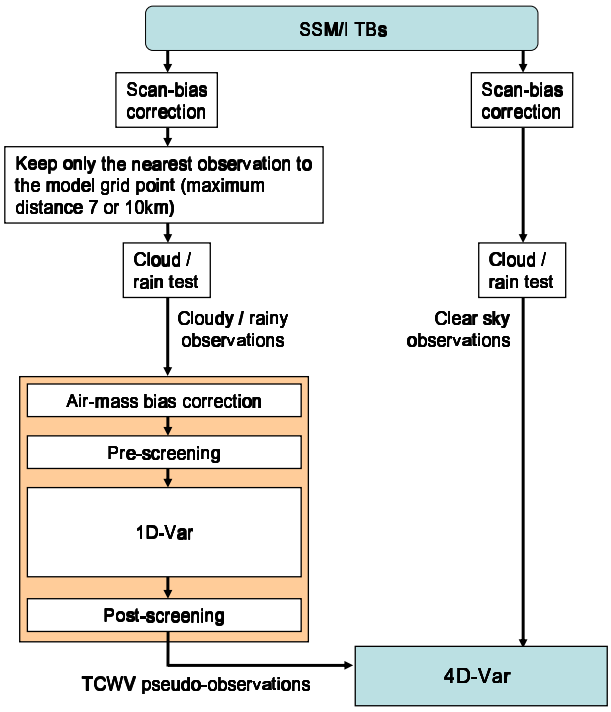

Figure 1: Flow of SSM/I observations in the ECMWF assimilation system.

the current performance of the 1D+4D-Var approach allows us to make recommendations for how best to do direct 4D-Var in the future.

2 1D+4D-Var method

2.1 Overview

ECMWF produces routine global analyses and 10-day forecasts from an assimilation system based on an atmospheric model with a semi-Lagrangian, spectral formulation. In the current operational cycle (32r2), the model has 91 levels in the vertical from the surface to an altitude of 80km, and a T799 horizontal resolution, corresponding to about 25km (see http://www.ecmwf.int/research/ifsdocs/ for documentation on earlier model versions; documentation on the latest version is not yet available). Global analyses of wind, temperature, surface pressure, humidity and ozone are produced twice-daily using four-dimensional multivariate variational assimilation (4D-Var; Rabier et al., 2000) using a staggered approach combining 6-hour and 12-hour time windows (Haseler, 2004). The analysis includes in-situ conventional data, satellite radiances from polar-orbiters, mainly from the AMSU-A and AIRS instruments, geostationary radiances, satellite-derived atmospheric motion vectors and surface winds from scatterometers. SSM/I radiances are assimilated, over oceans only, in two streams. Before assimilation, rain and cloud amounts are estimated directly from the observed brightness temperatures. Observations identified as clear sky are assimilated as radiances directly in 4D-Var. Cloud and rain-affected radiances are assimilated using a two-step 1D+4D-Var procedure (Bauer et al., 2006a,b). Figure 1 shows the flow of SSM/I observations. In this paper we will refer to observations selected for the clear-sky stream as 'clear' and those sent through the 1D+4D-Var path as 'rainy'.

The two-step procedure starts with a 1D-Var retrieval at each observation point. From this, a TCWV pseudo-

observation is derived and then assimilated in the main 4D-Var analysis. Although the mathematical basis of the 1D-Var is identical to that of the 4D-Var (see e.g., Rodgers, 2000), we will distinguish the two in this paper by describing the 1D-Var step as a 'retrieval', and the 4D-Var step as an 'analysis'.

Why did we start with a 1D+4D-Var approach, when experience shows that for most observations a direct 4D-Var approach is best? For example, at ECMWF between 1992 and 1996, a 1D-Var retrieval was used for assimilation of IR observations (Eyre et al., 1993), but the real forecast benefit from these observations only appeared after the introduction of 3D-Var and 4D-Var data assimilation systems (e.g. Bouttier and Kelly, 2001; Uppala et al., 2005), which in particular enabled the retrieval step to be eliminated and radiances to be assimilated directly. However, early experiments towards assimilation of rain and cloud observations at ECMWF (e.g. Marécal and Mahfouf, 2000, 2002; Moreau et al., 2004) have all been based around 1D-Var retrievals. One main reason is that the moist physics and radiative transfer of such observations tends to be highly non-linear. A 1D-Var retrieval is thought better able to handle such non-linearities than the full 4D-Var analysis (Marécal and Mahfouf, 2004). Another main reason is that separating the rainy retrievals from the rest of the system means it is much easier to monitor them, and it is easy to experiment with different approaches. The ability to apply quality controls after 1D-Var also allows us to eliminate bad retrievals. Such observations might cause convergence failures if used directly in the 4D-Var assimilation. The ultimate aim is to move to direct 4D-Var but we are building experience with the 1D+4D-Var approach first.

The SSM/I 1D-Var retrieval (Bauer et al., 2006a) makes a best estimate of the state of a single column of the atmosphere at the observation point, by combining the observed SSM/I radiances in channels 19v, 19h and 22v, with an a-priori (or background) taken from the first guess forecast of the 4D-Var assimilation, and including surface parameters and the temperature and moisture profile at the observation point. The weighting between observations and the background profile is determined by the observation and background error covariances. Background error covariances for temperature and moisture are taken from those used in the 4D-Var analysis. Observation errors are considered to be uncorrelated and for channels 19v, 19h, and 22v they are set to 3K, 6K, and 3K respectively. These errors account not just for the instrumental error but also that coming from the forward models, which is thought to be the largest part. Bauer et al. (2006a) used the Hollingsworth and Lönnberg (1986) method to derive these observation errors. The values of 3K, 6K and 3K come from rounding up to the nearest whole number. Deblonde et al. (2007) argue that such values provide a very strong weighting towards the observations in cloudy and rainy areas in 1D-Var retrievals, and recommend instead the use of much larger observation errors. However, this paper shows that the operational system is working acceptably with the current observation errors and Sec. 3.3.3 explains why we believe these error estimates are appropriate.

For the 1D-Var retrieval, the control vector, i.e. the state that is varied until the best estimate solution is found, contains only the humidity and temperature, and since cycle 31r1 was made operational in September 2006, the surface wind. Other surface parameters are held fixed. The vertical profile of cloud and precipitation is derived from the temperature and moisture profile using the moist physics operators. It is important to realise that the SSM/I radiances do not contain enough information to unambiguously reconstruct the full atmospheric state. Unlike a typical satellite retrieval, the 1D-Var retrieval depends strongly upon the model's first guess. In practice, the technique is able to produce a retrieval that matches the SSM/I observations very closely (see Fig. 7, Bauer et al., 2006a).

A TCWV amount is calculated from each 1D-Var retrieval, and around 20,000 of these pseudo-observations are assimilated into 4D-Var in each 12-hour window. The 1D-Var retrieval typically makes only very small changes to the first guess temperature profile; these temperature changes are not assimilated in 4D-Var. The TCWV observations cause changes in moisture, cloud, precipitation, and wind fields in the 4D-Var analysis (Bauer et al., 2006b). Observing System Experiments (Bauer et al., 2006b; Kelly et al., 2007, this paper) show that in general, the 1D+4D-Var assimilation improves forecasts of both humidity and dynamical variables, particularly in the tropics.

Experiment	SSM/I rain observations	Description
X	Off	Cycle 30r2 without SSM/I rain assimilation
A	On	Cycle 30r2
B	On	Cycle 30r2 + surface wind as a sink variable in the retrieval.
C	On	Expt. B + revised bias correction.
D	On	Expt. C + screening for excess falling snow in the model first guess.
E	On	Expt. D + changes to moist physics operators.
F	On	Cycle 31r1 (includes all of above improvements to 1D-Var, plus changes to other parts of the system.)
G	On	Cycle 31r1 + screening of poorly converged retrievals.
H	On	Expt. G + now includes retrievals previously rejected because of negative humidity.

Table 1: Experiments to test cumulative improvements to SSM/I rain assimilation, run for month of August 2005.

2.2 Recent improvements

This section summarises developments to the assimilation of rainy SSM/I radiances that have happened since it went operational in June 2005. OSE experiments (Bauer et al., 2006b, see their Fig. 8) showed that the original implementation of SSM/I rain assimilation was degrading temperature and geopotential forecast scores at around days two to three in the SH. Work over the past year has attempted to improve SH forecast scores, and to remedy a number of other problems that have become apparent. This section describes in turn each change to the system. To understand the impact of these changes, a series of experiments have been made in which each new modification has been added progressively to the full assimilation system. The experiments are listed in Table 1. The specific details of the experiments are given in later sections.

Because the ECMWF system is constantly evolving, these experiments were made with two different versions (known as cycles) of the ECMWF system. SSM/I rain assimilation was originally implemented in cycle 29r2. The experiments presented here were started within the experimental cycle 30r2 of the ECMWF system. This cycle had a bug that caused excessive upper tropospheric moisture, but we do not believe this significantly affected the results of our tests. Cycle 30r2 was not made operational, and instead the bug was fixed in a new cycle, 31r1, which included all improvements to the rainy SSM/I assimilation available at the time, along with many other changes to the ECMWF system. Cycle 31r1 went operational on 12th September 2006 and further modifications were tested within this new cycle. Modifications to the 1D+4D-Var assimilation subsequent to cycle 31r1 will be included in cycle 32r3 of the operational system.

Experiments were run for the month of August 2005, and were based on a slightly lower than operational resolution (T511 rather than T799) to save computer resources. For simplicity, only a 'delayed cutoff' (Haseler, 2004) assimilation stream is used, meaning that the assimilation time window is twelve hours in length, and no 'early delivery' analyses are made. Initial conditions came from the operational analyses. The control experiment X removes the rainy SSM/I observations from the cycle 30r2 system. The difference between experiments A and X shows us the impact of rain assimilation in cycle 30r2 of the ECMWF system.

The modifications to SSM/I rain assimilation were assessed by considering their impact on forecast scores. Forecasts from each experiment were verified against their own analyses. Figure 2 gives an illustrative example, based on the Root Mean Squared Errors (RMSE) of 3-day (72 hour) forecasts of geopotential height at 500hPa in the SH. The figure shows the difference in RMSE of experiments A to H compared to the control experiment X, normalised by the RMSE of experiment X. A positive difference indicates that RMSE is larger in

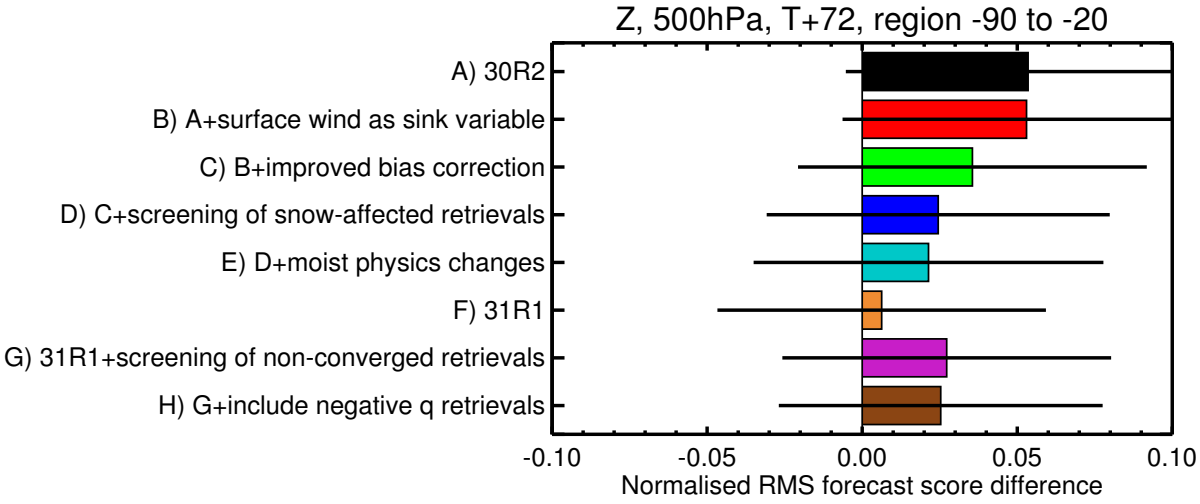

Figure 2: Normalised difference in RMSE forecast verification score for geopotential height at 500hPa, for the month of August 2005, taking the difference between experiments A to H (containing rainy SSM/I observations) and the control experiment X, from which rainy observations have been excluded. Forecast scores are based on the RMS difference between southern hemisphere T+72 forecasts and own analysis of geopotential height at 500hPa, for the month of August 2005. Positive differences in RMSE indicate a degradation in the experiment compared to the control; negative differences show an improvement. Error bars indicate the 90% confidence range for a significant difference between experiment and control.

the experiment than in the control. See appendix A for more details of the calculations. Figure 2 also functions as a key for Fig. 3, which summarises the scores over different forecast times and regions. From the black line on Fig. 3a we can see that, just as in the results of Bauer et al. (2006b) using cycle 29r2, at cycle 30r2 SSM/I rain assimilation was still degrading SH forecast scores in August, particularly in the 2 to 4 day range. Rain assimilation improves geopotential height scores at short ranges in the tropics in the baseline 30r2 system (Fig. 3b), and has no significant impact in the NH (Fig. 3c).

Figure 3 shows that cycle 31r1 provides forecast improvements compared to cycle 30r2 in a number of areas. The largest improvement is in the tropics, between forecasts days 1 and 6, and this appears to be statistically significant. The scores of 30r2 experiments (A, B, C and D) all cluster together, as do those from the 31r1 experiments (F, G and H). Clearly the modifications to the rain assimilation were irrelevant here, and the

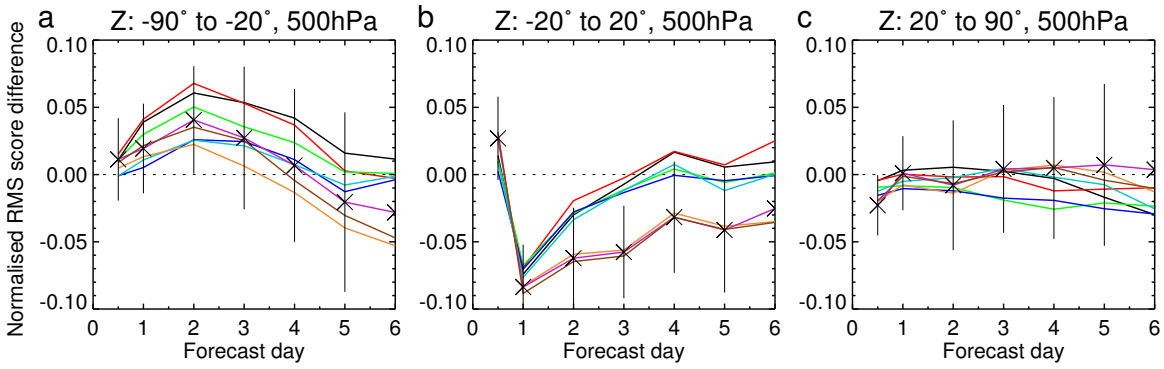

*Figure 3: As for Fig. 2 but showing RMSE differences by forecast time, and for **(a)** the SH; **(b)** the tropics and **(c)** the NH. Figure 2 functions as a key. The 90% confidence range is shown for experiment G only.*

forecast improvements were due to changes in another part of the system. In the NH, there seems to be little change in the forecast scores between any of the experiments. In the SH, cycle 31r1 (Expt. F) is substantially better than cycle 30r2 (Expt. A), but not to a degree that could be considered statistically significant. The changes to the rain assimilation introduced in experiments B, C and D have progressively improved forecast scores (i.e. RMSE errors have been reduced) in the SH between days 1 and 5. In experiment D, the forecast score degradation in the SH has been almost removed. These changes appear to have been responsible for a large part of the improvements made in 31r1 in the SH. Further changes to rain assimilation were tested on top of cycle 31r1 (experiments G and H). Experiment G causes a small degradation to appear in the SH again.

We describe the progressive changes to the system and the results of these experiments in more detail in the following sections. We examined forecast scores not just for geopotential height at 500hPa, but also for temperature, relative humidity, and vector wind, and at levels throughout the troposphere. They support the conclusions made here, but for space reasons, they are not shown.

2.2.1 Surface wind in the control vector

Surface winds have an influence on sea surface microwave emissivity (see e.g. Phalippou, 1996), but in the original version of SSM/I rain assimilation, sea surface wind speed was not allowed to vary between first guess and analysis. We hypothesised that if there were errors in first guess surface wind speed and these were being aliased into the temperature and humidity retrieval, this could cause problems. Surface wind speeds are particularly high in the Southern Ocean in winter. Hence, in experiment B, the surface wind speed (precisely, the 10m wind) was added to the 1D-Var control vector and allowed to vary during the retrieval. However, as Fig. 2 illustrates, this made little difference to forecast scores. Further, there was very little change in surface wind speed between 1D-Var first guess and retrievals (not shown). The most likely explanation is that in cloudy, rainy atmospheres, brightness temperatures are far more sensitive to changes in cloud properties than to changes in surface wind speed. In the retrieval, the specified background errors constrain the wind quite tightly, but allow moisture and cloud properties to vary strongly. The change was technically justified but it is clear that problems in the Southern Hemisphere were not linked to the surface wind.

2.2.2 Bias correction

1D-Var includes an air-mass dependent bias correction (see Fig. 1) which is derived offline, on a one-off basis. We make a linear regression between brightness temperature (TB) departures, and one or more 'predictors' drawn from the model first guess, such as TCWV. The resulting linear relationships between the predictors and TB departures are used to estimate the bias, and then remove it, prior to assimilation. In the original version, the only predictor used was TCWV, and the three SSM/I instruments (each on a different satellite) were all assumed to have the same biases.

A revised bias correction was tested in experiment C. The new scheme calculates separate coefficients for each SSM/I instrument, and there are now three predictors: TCWV, surface wind speed, and total column rain amount (taken from the 1D-Var first guess). Doing the SSM/I bias correction individually for each of the three satellites eliminated some small regional inter-satellite biases from the TCWV increment field (not shown). Figures 2 and 3 illustrate that this revised bias correction was responsible for substantial, though not in general statistically significant, improvements in forecast scores in the SH.

Further experience has, however, revealed a number of flaws in the new bias correction scheme. First, even after bias correction there are still large first guess (FG) departure biases in cloudy and rainy regions (see Sec. 3.3). Second, the choice of FG column rain as a bias predictor may be flawed. These problems are explained in more

detail in later sections.

2.2.3 *Elimination of retrievals with excessive snowfall*

Examination of the monthly mean 1D-Var TCWV increments (not shown) revealed that in southern high latitudes there were strong positive mean increases in TCWV. These were associated specifically with intrusions of cold, dry Antarctic air being advected northwards over progressively warmer seas. At the cold front associated with such intrusions, there was much convective activity, and this was being further enhanced between the first guess and the retrieval in 1D-Var.

Examination of a number of case studies in these areas (Bauer et al., 2006c) showed that while the 1D-Var retrievals were able to match observations well at 19v/h, 22v and 37v/h, there were still large departures in the 85v and 85h channels. In contrast to the other SSM/I channels, 85v and 85h are sensitive to snow. Channels 85v and 85h are not actively assimilated because microwave radiative transfer is thought to be less reliably simulated for snow than for rain, and the combined observation operator (moist physics and radiative transfer) is more non-linear at these wavelengths. Nevertheless, the departures were such that they suggested the 1D-Var retrieval was creating too much snow. The actively assimilated 19 GHz and 22 GHz channels are not very sensitive to snow. This means that snow is not well constrained by observations. For this reason, as well as the unreliability of radiative transfer in cases of heavy snow, it was decided to screen out observations containing too much falling snow in the first guess.

In experiment D, we rejected all retrievals with first guess integrated column snow amounts greater than 30% of the column rain amount. This removed most observations south of 45°S in August 2005. Observation numbers were also reduced in areas of tropical convection, and at high northerly latitudes. Section 3.1 examines the effect of this criteria in February 2007, which is similar to that in experiment D, except that it is the NH where most observations are removed. Figures 2 and 3 show that in experiment D this change had a positive (though again in general not significant) influence on forecast scores in the southern hemisphere. We appear to have eliminated the retrievals that were degrading the forecasts, though since we have eliminated most retrievals south of 45°S in order to do this, it may be that in future we can devise a more targeted screening.

2.2.4 *Changes to linearised moist physics*

The final changes tested at 30R2 (experiment E) were some minor improvements to the linearised moist physics used in the 1D-Var retrieval. These changes had a minor positive impact on forecast scores.

2.2.5 *Minimisation*

For each 1D-Var retrieval, the cost function is minimised using the M1QN3 software (Gilbert and Lemaréchal, 1989), which uses a quasi-Newtonian technique. The original 1D-Var implementation (Bauer et al., 2006a) was affected by a misunderstanding of the way the minimiser worked. As a convergence criterion, M1QN3 calculates the ratio of the norm of the cost function gradient compared to its value at the first iteration. This parameter is known as *EPSG* in the software. M1QN3 finishes its minimisation if either a target value of *EPSG* is reached, or if a maximum number of calls have been made to the observation operator and its adjoint, each pair of calls being known as a 'simulation' in the terminology of M1QN3. In the original implementation, the target value of *EPSG* was set so small that it was impossible to reach. Instead, M1QN3 always ran for the requested maximum 20 simulations before returning. The M1QN3 minimiser uses a 'line search' before each iteration, which means that it may make more than one simulation per iteration. The distinction between simulations

and iterations was not fully appreciated in the original implementation of the SSM/I rain assimilation. Hence, Fig. 10c in Bauer et al. (2006a) is of limited value, as it shows simply the histogram of the number of iterations corresponding to the fixed number of 20 simulations. It is not evidence of a well-behaved minimisation.

The original implementation produced two problems. The first is that many retrievals had converged long before the 20 simulations had been made, and that computer time was being wasted. The second is that some of the retrievals had failed to converge, but were being treated as good data, and the retrieved TCWV pseudo-observation was being assimilated in 4D-Var. 1D-Var does include a post-screening (Fig. 1 and Sec. 3.1) with a number of sanity checks to prevent bad observations reaching 4D-Var. These were not triggered, because when the minimisation had failed to converge, all that had happened was that the control vector remained close to the first guess. However, this did mean that some of the TCWV pseudo-observations, being fed into 4D-Var as new data, instead erroneously contained something very close to the first guess field. Up to 20% of SSM/I rain observations were affected.

To improve the situation, poorly converged retrievals are now rejected, by excluding 1D-Var minimisations that reduce the cost function gradient by less than a factor of 10 (e.g. $EPSG$ must be below 0.1). Additionally, a new target $EPSG$ of 0.01 is used, which means that many minimisations are considered to have converged after just a few iterations. The retrieved values of TCWV were found to be essentially no different compared to those from minimisations that had been forced to run for 20 simulations. This modification to SSM/I rain assimilation was tested in experiment G, on top of cycle 31r1. Figure 6 shows that forecast scores were degraded in the SH relative to cycle 31R1, though not by enough to be statistically significant. Forecast scores were barely affected elsewhere. A possible reason for the small degradation in the SH is that the analyses are now weighted more towards observational data than to the model first guess, which would likely increase the RMS differences between analyses and forecasts.

With this modification, 1D-Var is now working more correctly and efficiently. However, it is now clear that up to 20% of retrievals are failing to minimise. This could come for any number of reasons. Perhaps the rainy and cloudy atmosphere is simply too nonlinear. Perhaps the background and observation errors are too small in some cases. Perhaps first guess and observation are too far apart (see Sec. 3.5). Alternatively, 1D-Var makes use of a pre-conditioning based on the background error covariances (see Bauer et al., 2006a), and maybe this is sometimes insufficient. Future work will seek to understand whether such minimisation failures are inherent to rainy and cloudy atmospheres or whether there is a way to increase the number of successful minimisations.

2.2.6 Negative humidities in the retrieval

In the initial implementation of rainy SSM/I assimilation, approximately 10% of retrievals were being rejected because negative humidities were detected in the control vector during the 1D-Var minimisation. In almost all of these cases, the atmospheric profile consisted of a moist, cloudy, boundary layer (necessary to trigger the rainy, rather than the clear sky, SSM/I assimilation stream) overlaid by very dry layers in the mid and upper troposphere. Such cases were particularly prevalent in the subtropical oceanic high pressure areas. When the 1D-Var retrieval attempted to make the atmospheric column drier and less cloudy in such circumstances, it was possible for humidities in the dry upper layers to become negative. Since the forward model would fail if given negative humidities, such retrievals were aborted. This, however, meant there was a bias in such areas towards retrievals where the atmospheric column was being moistened instead. A simple solution was to reset any negative humidities to zero. This was tested in experiment H. It resulted in a small reduction in the biases in 1D-Var first guess departures in the affected areas (figures not shown) and a there was a minor, though not significant, improvement in SH forecast scores, particularly at the longer forecast ranges (Fig. 3).

2.2.7 *Modifications for passive monitoring and diagnostics*

A number of modifications have also been made that can have no effect on the results of the 1D-Var minimisation. First, the system has been expanded to allow the future use of other microwave sensors, such as AMSR-E and TMI. This will be described in a separate note. Second, a much larger range of diagnostics is now produced. These new diagnostics make possible many of the results shown in the rest of the paper.

For monitoring purposes, first guess TB departures are now calculated using the precipitation and cloud amounts of the full nonlinear first guess trajectory. These will be referred to as '4D-Var' first guess departures. These 4D-Var departures are helpful in understanding the future behaviour of direct 4D-Var assimilation. Note that the first guess departures used within 1D-Var still have to be calculated using the linearised 1D moist physics operators. We now also passively monitor the 'clear' observations, by calculating a first guess TB including the effects of cloud and rain.

For both the 1D-Var FG and retrieval and the 4D-Var FG and analysis, we record the TCWV, total column cloud liquid water and cloud ice, and total column rain and snow. The first three measures are easily computed from the model variables, which are stored as mass mixing ratios. The latter two diagnostics require a number of assumptions in their computation. In the model, rain is stored as a flux, in kg m^{-2} s^{-1}. To compute the total column integral in height or pressure, we would require the rain density or mass mixing ratio respectively. A conversion from rain and snow flux to density is already required for the radiative transfer calculations. Hence we use the same relationship for calculating the total column rain and snow. See Appendix B for the details. The relationship is based on a prescribed size distribution and fall speeds.

2.2.8 *Summary*

Early versions of SSM/I rain assimilation caused a small degradation in forecast scores in the SH in the 2 to 4 day forecast range during the southern winter. This degradation has been eliminated by screening out profiles with too much falling snow in the first guess, and by modifying the bias correction. Allowing surface wind speed to vary in the retrieval made little difference to forecast scores. These three modifications were included along with minor improvements to the moist physics and changes to other parts of the system in cycle 31r1 which became operational in September 2006.

A second set of modifications will be made operational in cycle 32r3. These are the improvement to the 1D-Var minimisation and a better treatment of negative humidities. These were tested here within cycle 31r1. There were also a number of changes for improved passive monitoring and the ability to assimilate sensors other than SSM/I. These do not affect the active retrievals and hence were not tested experimentally here. With the exception of the OSEs presented in Sec. 3.6, which are based on an unmodified cycle 31r1 configuration, the experiments presented in the rest of this paper include all these final modifications to 1D-Var.

3 Current status

This section explores the current status of the 1D+4D-Var assimilation of SSM/I. Results are based on experiments (see Table 2) made with cycle 32r1 of the ECMWF system, which already contains the modifications described in Secs. 2.2.1 to 2.2.4, and here we have also included the modifications described in Sec. 2.2.5, 2.2.6 and 2.2.7. To save computer resources, the horizontal resolution has been reduced to T511, rather than the full operational resolution of T799. The operational vertical resolution of 91 levels has been retained. For simplicity, only the delayed cutoff (12 hour time window) analyses are made. The experiments are intended

Experiment	SSM/I rain observations	Description
NORAIN	Off	Cycle 32r1 + modifications in Sec. 2.2.5, 2.2.6 and 2.2.7.
RAIN	On	Cycle 32r1 + modifications in Sec. 2.2.5, 2.2.6 and 2.2.7.

Table 2: Experiments for the month of February 2007.

to demonstrate how the operational forecasting system will work with the latest changes to rain assimilation. Experiment RAIN assimilates 'rainy' SSM/I observations. Experiment NORAIN is identical except that the 'rainy' SSM/I observations are removed. The experiments were run for the month of February 2007, taking initial conditions from the operational analyses for 12Z 31st January 2007.

3.1 Observation handling and quality control

This section summarises the observation flow in the most recent versions of the 1D+4D-Var system. Figure 4 shows the average number of observations passing through 1D-Var, based on the period 1st to 10th February 2007 in the RAIN experiment.

1D-Var operates on the model grid, using the SSM/I observation closest to that grid point. Observations are rejected if further than 10km from the centre of the grid point (7km for resolutions of T799 and higher). An initial screening eliminates observations over land, over ice, with model sea surface temperatures less than 275K (to avoid any chance of ice) and at latitudes higher than 60°. Fig. 4a shows the number of observations available for further processing after the initial screening. Currently we use only the SSM/I instruments onboard the Defence Meteorological Satellite Program (DMSP) F13 and F14 satellites, both of which have similar orbits. This results in quite patchy coverage; in future we hope to improve the situation by making use of microwave sensors onboard TMI and AMSR-E. Note that because of the grid-point selection, fewer observations pass through 1D-Var at the T511 resolution used here than in the operational T799 system.

Based on the observed brightness temperatures, roughly half of the observations are identified as clear sky (Fig. 4b). These observations are rejected as they will be separately used in the direct 4D-Var assimilation of clear sky SSM/I radiances (see Fig. 1). Observations are also rejected if there is excessive falling snow in the 1D-Var FG profile (Fig. 4c), as it is thought that radiative transfer simulations will be unreliable in such conditions (see Sec. 2.2.3).

The final first guess check eliminates observations where the FG departures are greater than 50K in any channel (Fig. 4d). The majority of these rejections occur in the inter tropical-convergence zone (ITCZ), the South Pacific convergence zone (SPCZ), the warm pool region, and at 20°S, 65°E, which was the location of typhoon Dora for much of the period shown. In other words, these rejections appear to be associated with areas of heavy cloud and rain. The 50K threshold was chosen to eliminate bad observations. However, we see in section 3.5 that rain systems are often poorly located in the first guess. It is quite possible that the observations may be clear yet the first guess has heavy rain, or vice-versa. We will see in the next section that the presence of particularly heavy cloud and rain can increase brightness temperatures by over 50K compared to a clear sky atmosphere even in the relatively insensitive 19v channel. Hence, to retain physically reasonable observations, this threshold should in the future be increased to 60K or 70K in channel 19v. Similar studies will be made for the other channels to determine more appropriate thresholds.

The remaining observations proceed to the 1D-Var minimisation. A variety of quality checks are made at each iteration of the minimisation, and a number are made after the minimisation completes. Roughly 20% of the remaining observations fail these checks (Fig. 4e). 98% of these failures come because the minimisation

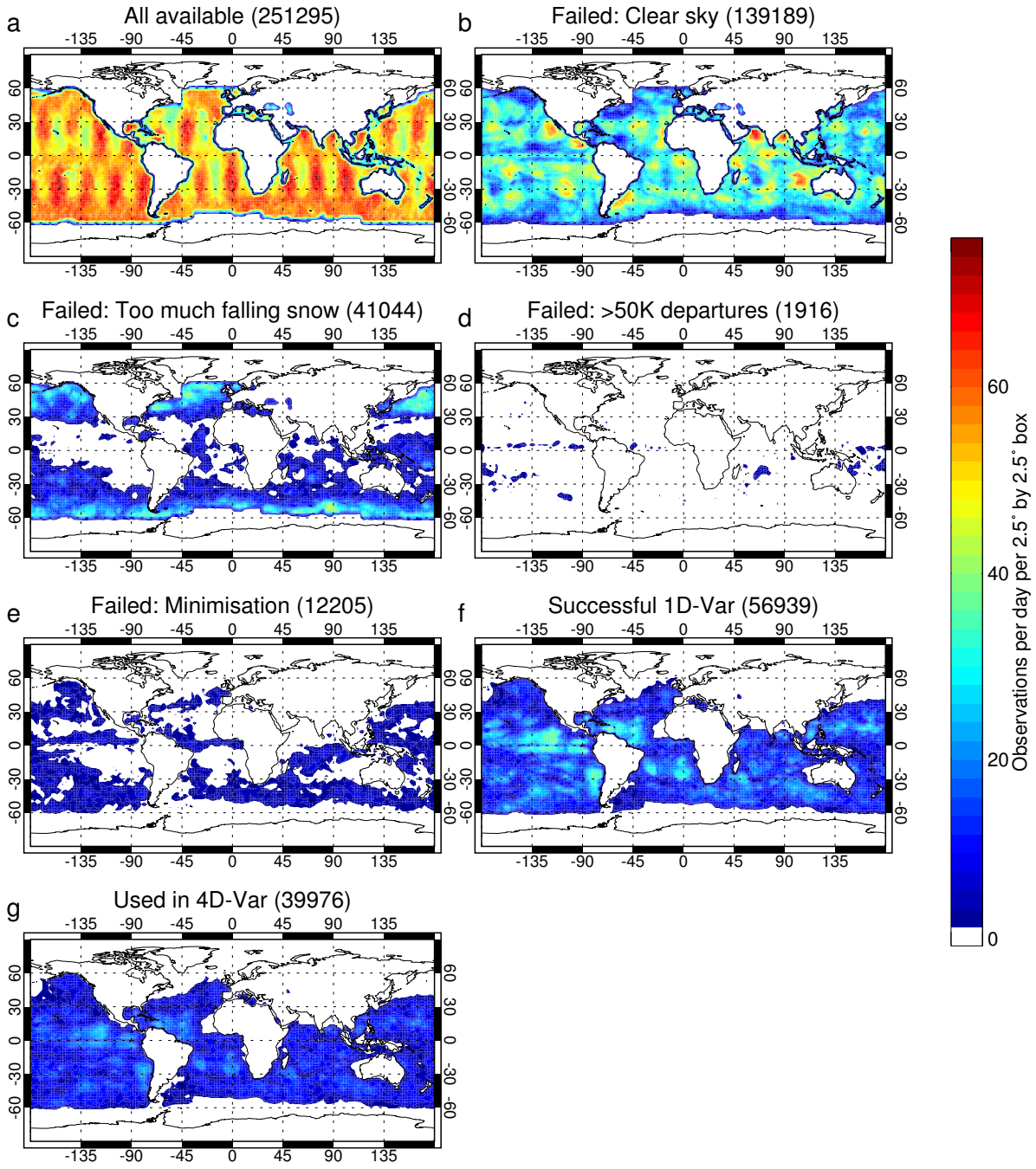

Figure 4: Average number of observations passing through 1D-Var, per day, based on the period 1st to 10th February 2007: (a) Within 10km of a grid point and passing initial screening; (b) Observation TBs indicate clear sky (not used in 1D-Var); (c) Model FG profile has too much falling snow; (d) Failed due to excessive FG departures; (e) Failure of 1D-Var minimisation; (f) Successful 1D-Var retrievals; (g) Used in 4D-Var

failed to converge. As discussed in Sec. 2.2.5, these minimisation failures are of some concern, and we do not yet understand whether such failures are inherent in the nonlinearity of rainy and cloudy situations, or whether results could be improved by using an alternative minimisation algorithm, a better pre-conditioning, or by some tuning of the existing algorithm.

Figure 4f shows the successful 1D-Var retrievals. TCWV pseudo-observations are calculated from the successful retrievals and passed to 4D-Var for assimilation. Within 4D-Var, observations are thinned in boxes of roughly 50 km by 50 km area, and only one observation is kept in each box. Considering that the spatial scales of rain-affected observations (see e.g. Sec. 3.5) can be much smaller than 50 km, it may be worth reconsidering this thinning strategy in the future. During 4D-Var, observations are also subject to variational quality control (Andersson and Järvinen, 1998b), which may reject observations if they are too far from the FG. Figure 4g shows the observations successfully used in 4D-Var. In this T511 experiment, about 40000 rain and cloud affected SSM/I observations are assimilated in 4D-Var each day. Because of the thinning, the number is also very similar in the T799 version (not shown), despite the larger number of 1D-Var retrievals available on the more closely space grid points.

3.2 Sensitivity of SSM/I to water vapour, cloud and precipitation

SSM/I observations are sensitive to surface properties, TCWV, cloud ice and water, and rain and snow (e.g., Chevallier and Bauer, 2003). In the absence of clouds or precipitation, however, observed brightness temperatures over the oceans are largely controlled by TCWV. Figure 5 illustrates a small sample of channel 19v brightness temperatures simulated from the full 4D-Var first guess trajectory fields of moisture, cloud and precipitation. As we see later, the 1D-Var first guesses, based on simplified moist physics, are less reliable, so we do not use them here. Brightness temperatures are shown against integrated total column amounts of the various atmospheric parameters. Brightness temperatures are strongly affected by TCWV, column rainfall, and column cloud water, with TCWV in particular showing quite a linear dependence. Thermal emission from cloud and rain causes brightness temperatures to increase: affected observations are those that deviate from this broadly linear dependence on TCWV. There is little influence from column snow or cloud ice.

Figure 6 illustrates TBs from channel 19h. Due to a lower ocean surface emissivity at the horizontal polarisation, channel 19h has more sensitivity to TCWV, cloud and precipitation, but this aside, its behaviour is very similar to 19v. For channel 22v, Fig. 7 shows that this channel is strongly dominated by TCWV variations. Cloud and rain do affect the channel, but only to the extent of increasing TB by a maximum of 20K over what could be expected through TCWV alone. Channels 37v/h and 85v/h are not assimilated so we do not show figures here. In broad terms, however, these channels show much more sensitivity to cloud and precipitation than do channels 19v and 19h. Also in channels 85v and 85h there is increased sensitivity to snow and cloud ice, and scattering by cloud and precipitation complicates matters further by acting to decrease brightness temperatures.

To investigate further the sensitivity of SSM/I brightness temperatures, we made a multiple linear regression (e.g., Wilks, 2006) between simulated brightness temperatures and four linear predictors: TCWV, column rain, column cloud water, and column snow. The sample was generated at roughly 1.6 million SSM/I observation locations between 1st and 10th February 2007, and is again based on the 4D-Var first guess atmospheric and surface fields. Table 3 shows the resulting gradients and correlation coefficients for channels 19v, 19h and 22v. The other four SSM/I channels are not assimilated operationally, and their dependence upon precipitation and cloud is much less linear. Hence there is little point in trying to understand them through linear regressions.

Table 3 shows that there is a strong correlation between TB and TCWV, and weaker correlations with the other predictors. For channel 22v, the correlation with TCWV is especially strong and the correlations with rain, snow and cloud water are especially weak. The negative dependence on rain is likely an artefact of the linear

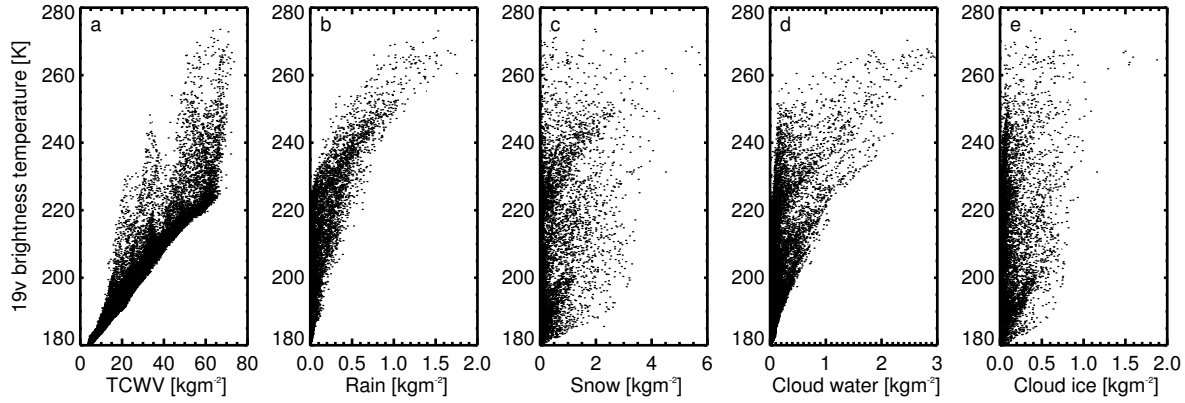

Figure 5: Modelled SSM/I channel 19v brightness temperatures, generated from profiles from the 4D-Var FG, plotted versus (a) TCWV, (b) Rain water column, (c) Snow column, (d) Column cloud liquid water, (e) Column cloud ice water. Sample is generated at observation points between 17Z and 18Z on 2nd February 2007.

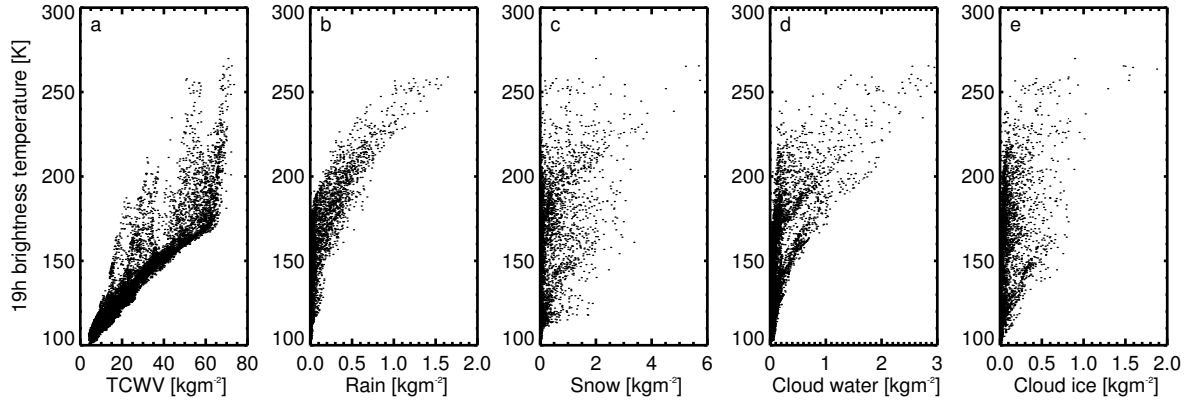

Figure 6: As Fig. 5 but for channel 19h.

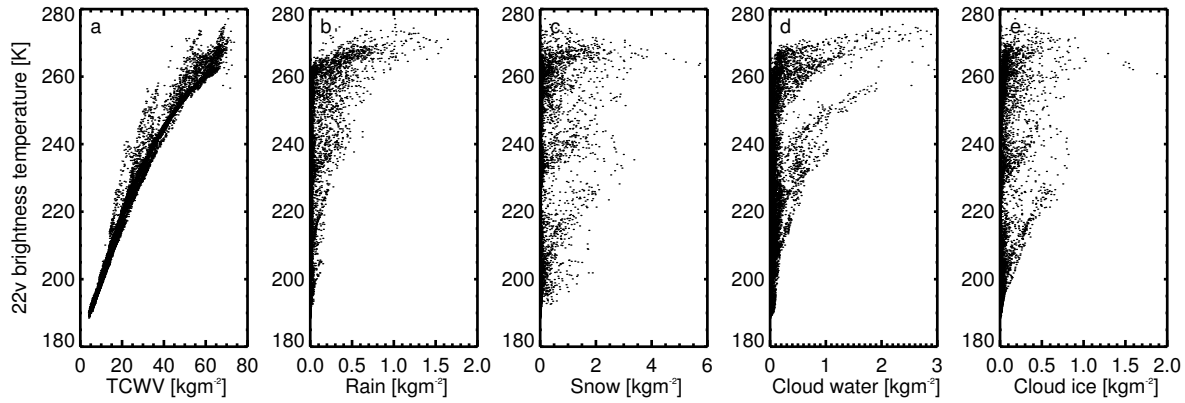

Figure 7: As Fig. 5 but for channel 22v.

	Channel	TCWV	Rain	Snow	Cloud water
Gradient	19v	0.77	9.80	1.62	9.59
	19h	1.21	7.77	8.17	25.8
	22v	1.25	-6.99	0.75	6.33
Correlation	19v	95	57	48	45
	19h	92	62	53	52
	22v	98	37	31	28

Table 3: Results from a multiple linear regression between brightness temperatures and TCWV, column rain, column snow and column cloud water. Gradient parameters are given in K (kg m^{-2})$^{-1}$ and linear correlation coefficients in %.

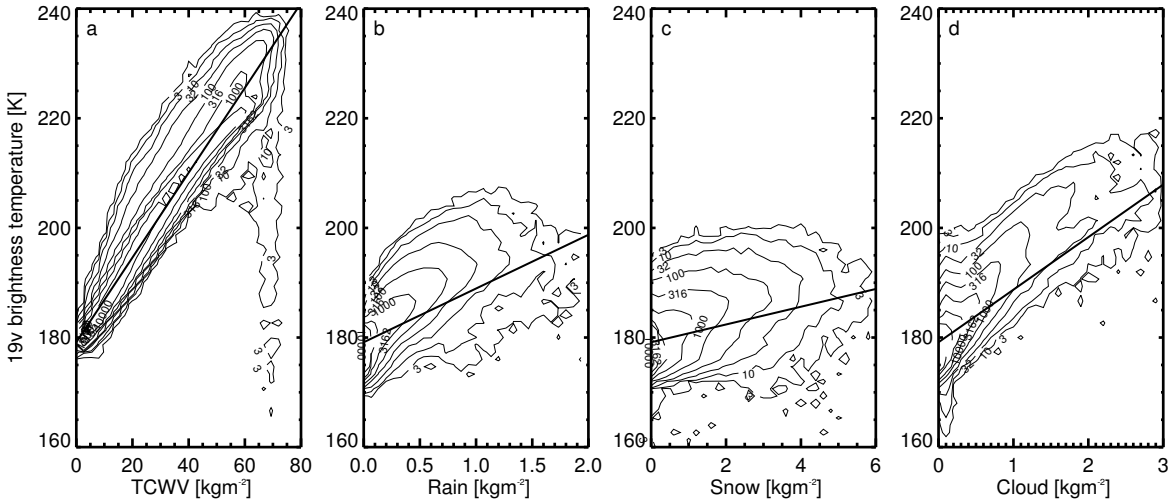

Figure 8: Two dimensional histograms showing modelled SSM/I channel 19v brightness temperatures against (a) TCWV, (b) Rain water column, (c) Snow column, (d) Column cloud water, after subtracting the estimated linear effect of the other three quantities. Isolines are logarithmically spaced at 3, 10, 32, 100, 316, 1000 and so on, and show the frequency of observations in an arbitrary-sized bin. The thick straight lines represent the linear fit derived from the multiple linear regression. The sample is from 1st to 10th February 2007.

regression and the fact that, as can be seen in Fig. 7, the dependence on TCWV deviates quite strongly from linearity compared to channels 19v and 19h.

Figure 8 shows, for channel 19v, the relationship between each predictor and brightness temperatures from which the estimated linear influence of the other three predictors has been subtracted. Of course, we would not wish to build a retrieval strategy on these linear regressions. Also, we are looking at correlations, which do not identify cause and effect. For example, TCWV is also strongly correlated with surface temperature, which also has an effect on brightness temperatures. However, this figure lets us to establish broad 'rules of thumb' linking radiances back to geophysical quantities, and these will be useful in the rest of the paper. The nearly linear dependence on TCWV is now much clearer (Fig. 8a). We can also see that our estimated linear dependencies on cloud and precipitation are surprisingly effective at predicting brightness temperatures. Also, we have ignored variations in the surface contribution to observed brightness temperature. However, compared to the effect of TCWV, rain and cloud, the effect of variations in the ocean surface emissivity appears to be minor. A final point to note is that cloud and precipitation column amounts follow, globally, a roughly log-normal distribution (see Sec. 3.3.2), and most of the sample is concentrated in very small column amounts, i.e. to the far left of the graph in Figs. 8b-d.

	19v	19h	22v	37v	37h	85v	85h
Successful 1D-Var	-1.08	-1.19	1.75	-3.05	-2.15	0.91	3.50
1D-Var rainy	-1.34	-1.71	1.62	-3.38	-3.03	1.03	3.22
1D-Var rainy and clear	-1.88	-2.43	1.06	-4.46	-4.57	-0.23	0.62
4D-Var rainy	-0.54	-0.30	2.02	-2.04	-0.40	0.68	3.07
4D-Var rainy and clear	-1.39	-1.55	1.30	-3.50	-2.66	-0.65	0.42

Table 4: Means of the FG departures (in K) for the samples used in Fig. 9, by SSM/I channel.

	19v	19h	22v	37v	37h	85v	85h
Successful 1D-Var	5.5	10.1	4.2	7.9	16.6	6.0	12.8
1D-Var rainy	7.1	13.2	4.8	9.1	19.4	7.9	13.8
1D-Var rainy and clear	5.6	10.4	4.1	7.5	16.1	6.7	11.4
4D-Var rainy	6.3	11.5	4.5	7.8	16.6	7.9	12.1
4D-Var rainy and clear	5.0	9.1	3.9	6.3	13.4	6.5	10.3

Table 5: Standard deviations of FG departures (in K) for the samples used in Fig. 9, by SSM/I channel.

We can make a number of final remarks. The SSM/I channels are, in a global oceanic context, primarily sensitive to TCWV. In channels 19v, 19h and 22v, assimilated in 1D+4D-Var at ECMWF, it is a minority of profiles, those with heavy cloud and precipitation, where brightness temperatures strongly deviate from what can be predicted by TCWV alone. For the majority of observations, where there are smaller or zero cloud or precipitation amounts (e.g. $<$0.2 kg m^{-2} rain, $<$0.4 kg m^{-2} cloud, or $<$0.5 kg m^{-2} snow), the effect on brightness temperatures is typically less than 5K.

3.3 First guess departures

The latest version of the 1D+4D-Var system calculates first guess brightness temperatures from both the 1D-Var first guess (in which the observation operator uses simplified moist physics to reconstruct the cloud and precipitation fields from profiles of first guess temperature and moisture, plus other parameters) and, for monitoring purposes, from a '4D-Var' first guess, taking the cloud and precipitation amounts generated by the full non-linear model during the first guess trajectory. These are calculated not only at the observationally-determined 'rainy' locations where 1D+4D-Var is applied, but also for the 'clear' observations which are currently assimilated directly in 4D-Var assuming clear-sky radiative transfer (see Fig. 1). Doing so helps us understand better the quality of the 1D-Var, and helps to plan for the direct 4D-Var of rainy radiances. FG departures are defined as observation minus FG. This section examines departures first from a spatial perspective, and then resolving by rain and cloud amount. All results in this section are from the RAIN experiment.

3.3.1 Spatial distribution

Figure 9 shows mean channel 19v departures over a ten day period, comparing 'rainy' skies with all skies ('rainy' + 'clear') and '4D' and '1D' first guesses. Tables 4 and 5 show respectively the mean departure and standard deviation, for each sample, along with those for the other SSM/I channels.

First we examine the first guess departures of those observations used actively in 1D-Var, with brightness temperatures simulated from cloud and precipitation fields generated using the simplified moist physics (Fig. 9a). These have been bias corrected, and only the observations that successfully produced a 1D-Var retrieval have been included. Despite this, the 1D first guess departures are typically negative in areas of cloud and rain.

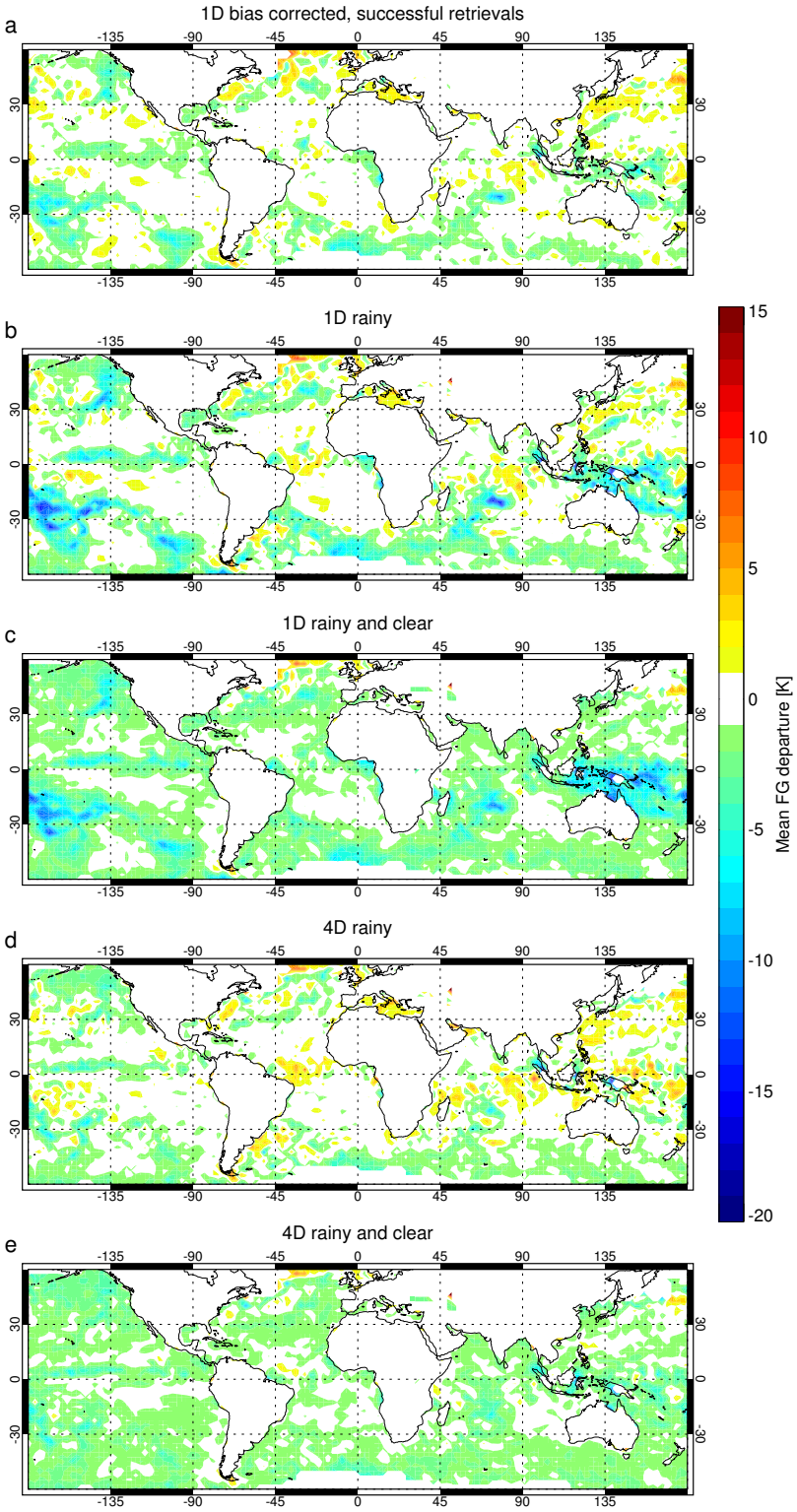

*Figure 9: Mean of channel 19v first guess departures (obs - FG), in 2.5° by 2.5° bins, for the period 1st to 10th February 2007: **a)** For all successful 1D-Var retrievals (by definition 'rainy'), based on the bias-corrected 1D-Var FG; **b)** For all 'rainy' observations, successful or not, and based on the 1D-Var FG before bias correction; **c)** As (b) but including all observations ('clear' and 'rainy'); **d), e)** as b), c) but from the 4D-Var FG.*

Negative departures are seen polewards of 30°, in the region of midlatitude frontal systems, and also in areas of deep convection: the warm pool area around the equator and 135E; the location of typhoon Dora during this period (20S, 65E); the ITCZ and SPCZ.

Figure 9b shows 1D departures before bias correction, and it includes all 'rainy' observations, whether they successfully produced a retrieval or not, and including those that would have been rejected by quality control because of FG departures greater than 50K (Fig. 4). Here, we see even larger negative biases. It appears that the 1D-Var first guess TB is excessively high in areas of heavy cloud and precipitation. We will see that this comes mainly from excessive cloud and rain in the simplified moist physics operators used by 1D-Var. The 1D-Var bias correction scheme includes the rain column amount as a predictor (see section 2.2.2), but looking at Fig. 9a it is clear that the scheme does not fully correct the bias in rainy and cloudy areas.

The negative bias becomes even more apparent (Fig. 9c) when all observations are included, not just those that are 'rainy'. We now include situations where the model's first guess is cloudy or rainy, but the observations themselves are 'clear'. In the current operational system, such observations pass through the clear sky, direct 4D-Var assimilation. In the clear sky stream, even if there is rain or cloud in the model first guess it is ignored in the forward calculation of radiances, and as a result the clear sky observations cannot directly correct the erroneous rain or cloud in the model. In parallel, the observations given to 1D-Var present an excessively 'rainy' view of the world. Hence, by previously considering statistics based only on the 'rainy' dataset, we have underestimated the size of the true first guess bias in the rainy stream. In the future, the rainy and clear streams should be merged. This would avoid such sampling biases, and also allow the model to be corrected in situations where the model is 'rainy' but the observation is 'clear'.

Figures 9d and e show that brightness temperatures from the 4D-Var first guess cloud and precipitation fields give much closer agreement with observations than do those from 1D-Var. In the all-sky comparisons, maximum biases are of order -7K from the 4D-Var first guess trajectory, and -20K from the 1D-Var estimate. Hence, we are currently in the process of modifying the simplified moist physics operators used in 1D-Var so that they produce estimates closer to those from the full, nonlinear model used in the 4D-Var FG trajectory. Nonetheless, Fig. 9e suggests that even the full model still has a bias towards excessive cloud or precipitation in rainy areas. In Sec. 3.5, comparisons with the Precipitation Radar (PR) on TRMM confirm that there is excess precipitation in both 4D-Var and 1D-Var first guesses.

We have so far concentrated on channel 19v, but similar conclusions can be drawn from examining most other channels. Figure 10 shows mean departures from the 4D-Var first guess for all seven SSM/I channels, for the combined 'clear' and 'rainy' sample. Channel 22v is an exception, because it is much less sensitive to clouds and precipitation than the others, and Fig. 10c reveals a broad positive bias, increasing towards the equator. This bias is easily removed for 1D-Var purposes using TCWV as a predictor (see Sec. 2.2.2).

With increasing sensitivity to cloud and rain, channels 19h, 37v and 37h show very similar patterns of negative departures to channel 19v, only with much greater magnitude (Fig. 10b, d and f). Again, this is evidence of excess cloud and rain in the 4D-Var FG. Departures are even more negative if the 1D instead of the 4D first guess is considered (see Table 4). Considering the standard deviations of combined 'clear' and 'rainy' observations (Table 5), we see that the 4D-Var first guess is always slightly closer to observations than is the 1D first guess, which again encourages work to match the simplified 1D moist operators with the results of the full nonlinear model. Channels 85v and 85h do not show the same patterns of negative departures as the 19 GHz and 37 GHz channels, which is likely explained by the fact that 85v and 85h are much more sensitive to snow and cloud ice, and that heavy cloud and precipitation also tend to cause decreases in brightness temperatures at this frequency, rather than to increase them.

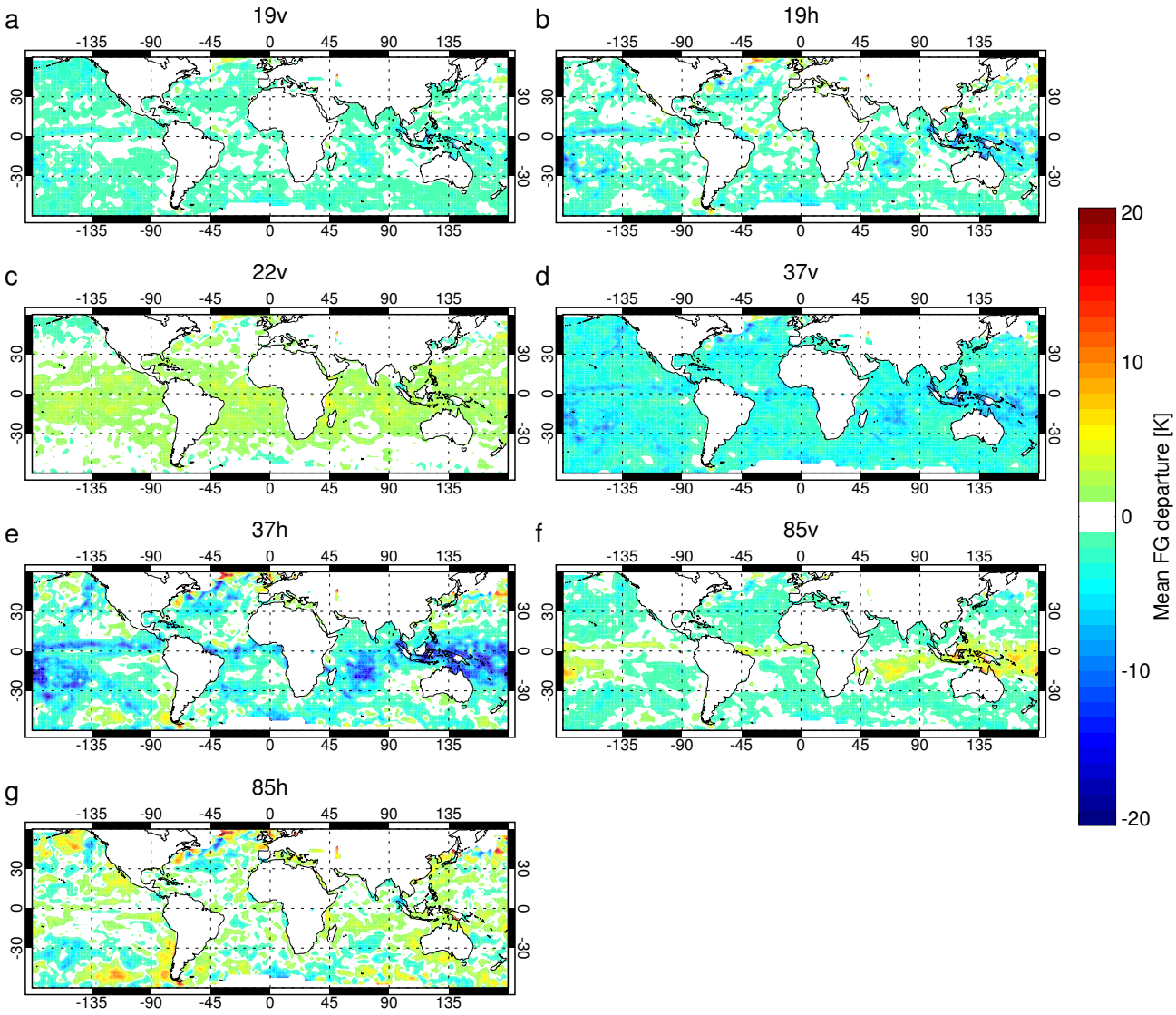

Figure 10: Mean of departures from 4D-Var first guess in both 'clear' and 'rainy' skies, for all seven SSM/I channels, for the period 1st to 10th February 2007.

3.3.2 *By rain and cloud amount*

Figures 11a and b and 12 a and b show, for channel 19v, the mean and standard deviations of the combined 'clear' and 'rainy' first guess departures. They have been binned respectively by first guess column rain amount and by first guess column cloud water. Both 1D-Var and 4D-Var FG departures are shown. For rain amounts of less than 10^{-3} kg m^{-2}, and for cloud water amounts less than roughly 10^{-2} kg m^{-2}, biases are between 0 and -2K, and standard deviations are, uniformly, 2K. FG standard deviations of 2K are very similar to those from the operational clear-sky 4D-Var path (see e.g. table 2 Bauer et al., 2006a, which shows a value of 1.76K). In contrast, towards the highest rain and cloud amounts on Figs. 11 and 12, standard deviations grow to over 15K and biases grow to -30K. However, it should be noted that these large biases are partly generated by the binning approach we have used, and are not necessarily indicative of a real bias. This is examined later in this section.

Figures 11c and 12c show the numbers of observations by FG rain and cloud amount. It is clear that the 1D moist physics produces more rain than does the full nonlinear model. In particular, there is a secondary peak in 1D rain at around 0.07 kg m^{-2} that is not seen in 4D rain. However, 4D rain does show a much smaller secondary peak at around 0.7 kg m^{-2}. In cloud, the picture is mixed. At the very largest cloud amounts (e.g. 1 kg m^{-2}), which have the strongest impact on TB (Sec. 3.2) the 1D moist physics is producing too much cloud. However, the 1D moist physics does not produce enough points with cloud in the middle range (around 0.1 kg m^{-2}), but does produce too many points with small cloud amounts, at around 0.01 kg m^{-2}.

In Figs. 11a and 12a, an apparently encouraging result is that means and standard deviations from the 4D-Var model fields and from the 1D-Var simplified moist physics operators are relatively similar. However, as we have just seen, a higher proportion of 1D-Var than 4D-Var estimates are found to the right of Figs 11a and 12a. Here, the departure means and standard deviations are larger, meaning that in global terms (e.g. Tables 4 and 5) 1D-Var does still show larger FG departure means and standard deviations than 4D-Var.

Figures 11c and 12c reveal another notable feature. The red part of these histograms shows the observations that have been identified as 'clear' by the SSM/I TBs; the blue part indicates 'rainy' observations. Very many observations that are identified as 'rainy' correspond to model first guesses with very little rain. Conversely, 'clear' observations often correspond to model first guesses showing quite heavy rain and cloud amounts. This would suggest that the model first guess is currently quite poor at generating rain and cloud at the locations observed by SSM/I. This also means that a significant fraction of the observations that are assimilated in the clear sky 4D-Var path correspond to rainy and cloudy first guesses. As noted already, the clear sky observation operator does not take cloud and rain into account, and hence the SSM/I observations are not currently being used, as we might hope, to suppress the erroneous cloud and rain in the model first guess.

A final notable feature of the histograms in Figs. 11c and 12c is that only 2% of 4D-Var model first guesses at the SSM/I observations points have column rain amounts or cloud water amounts that are too small to be included on the histogram. For the 1D-Var first guesses, the proportion is even smaller. In other words, cloud and rain is almost universally present in the model, though generally only in tiny amounts. This is a known feature of the moist physics parametrisations that we use, and the presence of very small amounts of cloud or rain should not to cause simulated TBs to diverge significantly from their clear-sky values. However, this issue may be worth revisiting in the future.

We have already mentioned that the large negative biases at high rain amounts do not necessarily indicate a biased first guess, and may be partly an artefact of our binning approach. To explain this, Fig. 13 shows a histogram of 4D-Var first guess departures against column rain amount, for both 'clear' and 'rainy' observations. For column rain amounts of 1 kg m^{-2}, it seems that most departures are negative. Thinking in binary terms, where the first guess is rainy, it is quite likely that some of the observations themselves will be clear. In these cases, the rainy model first guess will produce higher brightness temperatures than the clear observations (see

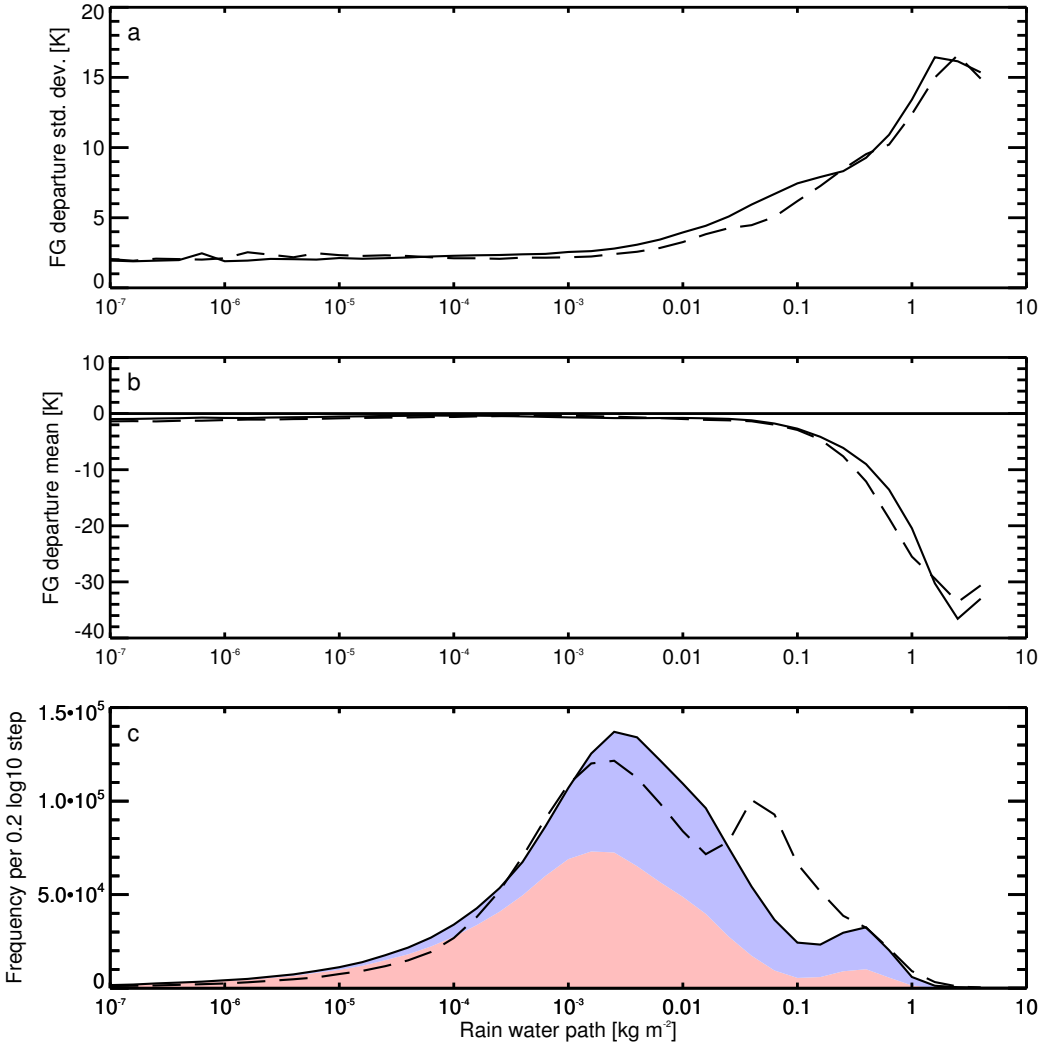

Figure 11: Channel 19v departure statistics for all 'clear' and 'rainy' observations, 1st - 10th February 2007, binned by first guess rain water path: ***a)*** *Standard deviation;* ***b)*** *Mean;* ***c)*** *Frequency per 0.2 log step in rain water path. Solid lines represent the 4D-Var FG; dashed lines represent the 1D-Var FG. The histogram of 4D-Var rain estimates has been divided into a red part, indicating observations identified as 'clear' and a blue part, for those identified as 'rainy'.*

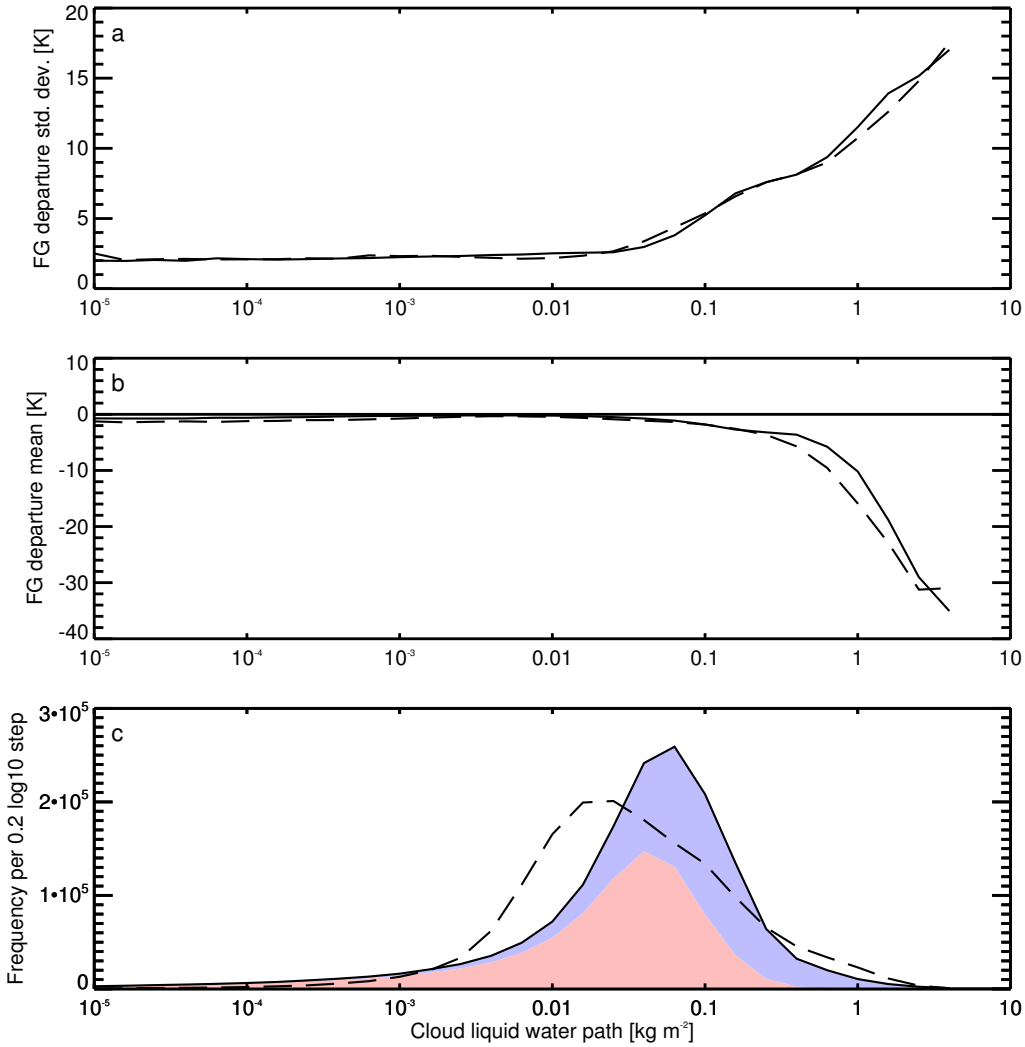

Figure 12: As Fig. 11 but binned by column cloud liquid water.

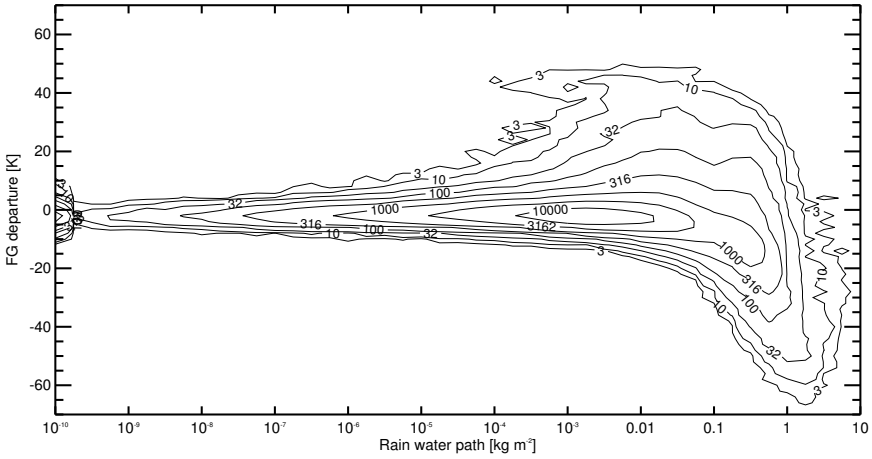

Figure 13: Histogram of channel 19v first guess departure statistics based on the 4D-Var first guess, binned by departure and by column rain amount. Contours of the number of observations per arbitrary-sized bin are drawn in logarithmic steps at 3, 10, 32, 100, 316, 1000 and so on. The lowest rain bin, at 10^{-10} kg m^{-2}, has been extended to include zero rain amounts, and it is interesting to note how few there are. Sample is the same as in Figs. 11 and 12

Sec. 3.2), resulting in negative departures. There are two likely reasons why we do not see positive departures at column rain amounts of 1 kg m^{-2}: First, we are near one extreme of the rain distribution, and it would be highly unlikely to find an observation with even more rain than in the first guess; Second, SSM/I TBs saturate at high rainfall amounts, i.e. after a certain point, TBs stop increasing. At column rain amounts of 0.01 kg m^{-2} there is a bulge of positive departures. Here, it is also natural that when the model first guess shows no little or no rain, some of the observations are rainy. Here, the departures will be positive, and a rainy observation can be up to 40K warmer than a less rainy or clear sky model first guess.

1D-Var currently uses first guess column rain amount as one of the predictors for bias correction (Sec. 2.2.2). It appears this is wrong in principle: we could simply be correcting for the artificial bias described above. However, Fig. 9 suggests that in our system, negative departures in rainy and cloudy areas are more than just a sampling issue. There should be no sampling bias in Fig. 9, since observations are divided into latitude-longitude bins and there is no other model or observation based selection. We could think up any number of modifications to the bias correction scheme, but a far better solution to the problem would be to improve the simplified moist physics operators so that they produce less rain and a better distribution of cloud.

3.3.3 Explaining our observation error estimates

Figures 11a and 12a help us to consider the study of Deblonde et al. (2007), who argue that the observation errors used in 1D-Var are too small, and that they cause an excessively strong weighting towards the observations in cloudy and rainy areas. In our system, observation errors are set to 3K, 6K, and 3K for channels 19v, 19h, and 22v respectively. These were derived by Bauer et al. (2006a) using the Hollingsworth and Lönnberg (1986) method. Deblonde et al. (2007) recommend the use of much larger observation errors (20K, 40K and 9K respectively) so as to be comparable with the size of the background errors when transformed to brightness temperature space. Bauer et al. (2006a) noted that background errors transformed to observation space were a factor 2 to 3 larger than the Hollingsworth and Lönnberg (1986) estimates of observation error. They hypothesised that background errors were overestimated in areas of cloud and rain. Here we propose an alternative explanation.

First, we note that this paper shows that the current operational system is working acceptably well with 'small' observation errors, notwithstanding the fact that many problems remain, particularly the excessive rain in the FG. As assumed by Hollingsworth and Lönnberg (1986), the standard deviation of first guess departures, σ_{FG}, should be related to the background error standard deviation in observation space, σ_B, and the observation error standard deviation, σ_R, by:

$$\sigma_{FG}^2 = \sigma_B^2 + \sigma_R^2. \quad (1)$$

Hence, σ_{FG} provides an upper bound to both σ_R and σ_B.

Figure 11a shows that for roughly half of the combined set of 'clear' and 'rainy' observations, i.e. those with column rain amounts less than 5×10^{-3} kg m^{-2}, σ_{FG} is less than 3K. For this sample, Eq. 1 suggests that channel 19v observation errors of 3K are actually at or above their upper bound. As the FG rain amount rises, however, σ_{FG} reaches a maximum of 15K. If we believe σ_R to be correct, Eq. 1 would imply $\sigma_B \simeq 14.7$K in such areas. We have already observed from Fig. 11c, and we will see in comparisons to PR observations (Sec. 3.5), that the location and intensity of areas of heavy rain and cloud is often poorly simulated in the first guess. The difference in channel 19v brightness temperature between a profile with heavy rain and light rain can easily be as large as 15K (see Fig. 8b). Our estimate of σ_B is compatible with this, suggesting that an observation error of 3K is perfectly reasonable. Finally, we note that such large FG errors are not just a result of the simplified moist physics used in 1D-Var. Figure 11a shows that 4D-Var FG error standard deviations are just as large.

We conclude that the observation errors and background errors used in 1D-Var are approximately correct. The apparent large discrepancy compared to the background errors in areas of heavy cloud and rain is explained by the high sensitivity of SSM/I TBs to the presence of cloud and rain, and the difficulty the first guess model has in simulating the exact locations and intensities of such phenomena.

3.3.4 Summary

This section makes a number of important points. First is that by dividing SSM/I observations into 'clear' and 'rainy' streams we become subject to a sampling bias. This sampling bias has historically caused us to underestimate the size of the FG biases in SSMI rainy 1D-Var. Additionally, when bias correction is derived separately for each stream, as it is at the moment, we are losing real information on the presence or absence of cloud and rain. A final problem with the current approach is that cloudy and rainy radiative transfer are ignored in the 'clear' stream, meaning that excess FG cloud and rain in these areas is not properly corrected. The ideal option for the future is to use a single stream including rainy and cloudy radiative transfer when required.

A second point is that both the 4D-Var and the 1D-Var model first guess appear to have too much rain and cloud compared to SSM/I. The simplified moist physics operators produce much more heavy rain than the full 3D model, leading to FG departures in regions such as the ITCZ and storm track regions of up to -20K. It should be possible to adjust the simplified moist physics for better performance to make their results closer to those from the full 3D model, and we are currently working on this. However, if we trust our radiative transfer, there is still a minor bias (of up to 7K in channel 19v TB) towards excessive rain or cloud even in the 4D-Var first guess trajectory.

The third point is that the model first guess appears quite poor at predicting rain or cloud exactly where SSM/I sees it. Data assimilation is thought to work best when making minor corrections to a good first guess. We have seen that it is quite normal in the current system for a completely clear first guess to be confronted with an observation of heavy rain, and vice-versa. This also resolves an apparent discrepancy between observation and background error estimates in areas of heavy rain noted by Bauer et al. (2006a) and Deblonde et al. (2007). We argue that 1D-Var observation errors are correctly estimated, and because the first guess model has difficulty

in simulating rain and cloud in the correct location, this can easily translate into large brightness temperature differences compared to observations.

3.4 Effect in 4D-Var

This section examines the effect of assimilating 'rainy' SSM/I observations on 4D-Var analyses and on the forecasts. 1D-Var makes a retrieval of the vertical profile of temperature, moisture, cloud and precipitation. However, only the TCWV amount is passed to 4D-Var, for assimilation as a pseudo-observation. TCWV pseudo-observations are not subject to bias correction. We are interested to know how well the information (especially on cloud and precipitation) is propagated into 4D-Var. Also we have identified a number of uncorrected biases in 1D-Var. Do these get passed into 4D-Var? The results in this section are based on the RAIN and NORAIN experiments for February 2007 (Table 2).

Figure 14 investigates the mean change in TCWV when rainy SSM/I observations are assimilated, and how this might be linked to the biases we have identified in the previous section. Panel a shows the February 2007 mean channel 19v departures in 1D-Var. Note that the colour scale is reversed compared to earlier figures. Even after bias correction, there are monthly mean patterns of negative departures in the ITCZ and in midlatitudes, coming from excessive rain and cloud in the 1D-Var first guess. These patterns are also reflected in the FG departures of the TCWV pseudo-observations presented to 4D-Var (Panel b). Areas with negative TB biases are associated with negative TCWV departures. However, many areas of the globe show relatively large positive TCWV departures, some corresponding to the subtropical dry regions, where the 19v departures are indeed positive, but too small to be seen on the colour scale used in Panel a. Because these regions are relatively dry, the TCWV departures look large when shown as a percentage in Panel b.

Figure 14c shows the difference between RAIN and NORAIN analyses of TCWV. RAIN analyses are generally drier than the NORAIN analyses. Drying has taken place in the regions where TCWV departures were negative, showing that this information has been correctly transferred into 4D-Var. However, the corresponding areas of positive TCWV departures appear to have been mostly ignored. Bauer et al. (2006c) show that there is a systematic drying effect in 4D-Var when 'rainy' SSM/I observations are assimilated. These observations are found in regions close to saturation, but because of the the way the relative humidity control variable is constructed (Hólm et al., 2002), positive moisture increments are penalised. Figure 14d shows that the patterns of drying (and a small amount of moistening) persist into the forecasts. The RAIN experiments are up to 4% drier in TCWV than in the NORAIN experiments in some areas. We can see that there are two reasons for this: (i) the 1D-Var forward model is biased towards producing excessive rain, resulting in mean reductions in TCWV in the 1D-Var retrievals and (ii) the penalisation of moistening increments by the relative humidity control variable in 4D-Var.

For the sample of all 1D-Var retrievals successfully assimilated in 4D-Var in February 2007, Fig. 15 shows histograms of the increment in 1D-Var versus the increment in 4D-Var at the observation location. Increments in 1D-Var come purely from the SSM/I observations. Increments in 4D-Var come not just from the TCWV pseudo-observations but from all other assimilated observations. Where the 1D-Var and 4D-Var increments agree, it cannot necessarily be attributed to the 1D-Var observations alone. However, where the 1D-Var and 4D-Var increments differ, this is a clear indication that the information retrieved in 1D-Var is being ignored in 4D-Var.

Figure 15a shows that 1D-Var TCWV increments are strongly correlated with increments in 4D-Var, which suggests that for TCWV, 1D+4D-Var is working well. In contrast, 1D-Var increments in rain appear to show absolutely no correlation with those in 4D-Var (Fig. 15b). This may be because of the difficulty of simulating rain at small scales, or because the TCWV pseudo-observation simply does not have much control over rain in

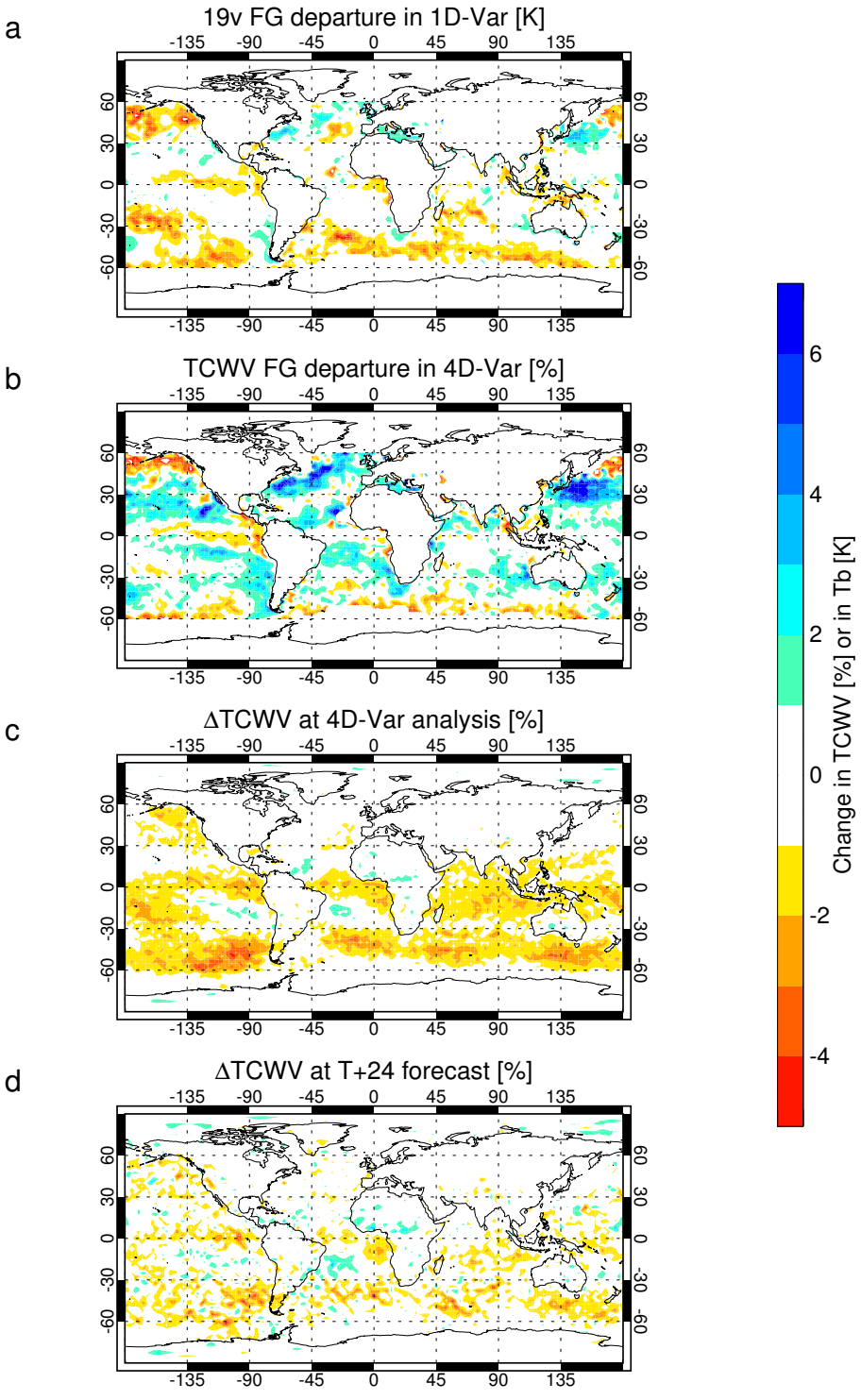

Figure 14: February 2007 mean in 2.5° by 2.5° boxes: ***(a)*** *Channel 19v TB departure (obs - 1D-Var first guess) for successful 1D-Var retrievals;* ***(b)*** *Departure of 1D-Var TCWV pseudo-observations in 4D-Var [100×(obs - first guess)/first guess];* ***(c)*** *TCWV differences between RAIN and NORAIN analyses [100×(RAIN - NORAIN)/RAIN];* ***(d)*** *As (c) but for T+24 forecasts.*

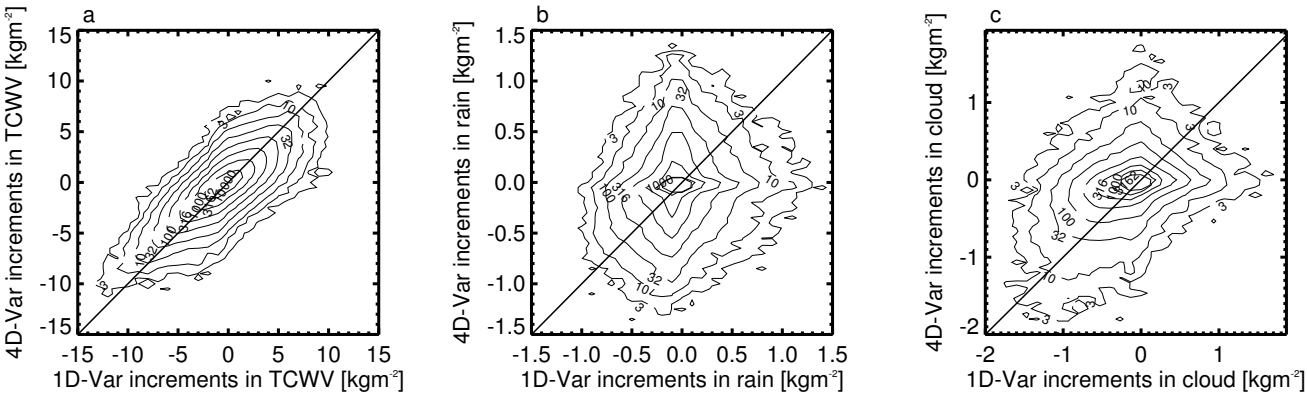

*Figure 15: 2D histograms of increments (analysis or retrieval - first guess) at SSM/I observation points in 4D-Var against those in 1D-Var: **a)** TCWV, **b)** Column rain amount, **c)** Column cloud amount. Sample is all successful 1D-Var observations for February 2007. The 1:1 line is overplotted. Contours are in logarithmic steps, i.e. 1, 3, 10, 32, 100, 316 etc.*

the 4D-Var model. It is clear that we cannot yet claim to have a rain assimilation system. Instead we have a system that makes best use of TCWV information in cloudy and rainy areas. However, there are encouraging hints of a correlation in cloud increments (Fig. 15c). Where 1D-Var significantly reduces cloud, in many cases 4D-Var does too. This should perhaps be expected given the direct link between relative humidity and cloud formation. The fact that increases in cloud are not transferred to 4D-Var may be linked with the fact that, due to the current moisture control variable, moistening increments are suppressed in areas near saturation.

Figure 16 makes a similar analysis in brightness temperature space. Remember that only channels 19v, 19h and 22v are assimilated. The other four channels (of which only three are shown on the figure, for space reasons) are passively monitored. Channel 22v has the strongest dependence on TCWV, and the least on cloud and rain. As we might expect, it is in this channel that 4D-Var increments best match those in 1D-Var. However, the distribution does not follow the 1:1 line. The 4D-Var increments are in general smaller than the 1D-Var increments. In transferring information from 1D-Var to 4D-Var, we have lost the vertical structure of the humidity changes, as well as all information on cloud and precipitation. It is likely that both of these factors contribute to the reduced increments in 4D-Var. Channels 19v and 19h show some sign of 4D-Var increments being similar to those in 1D-Var, but mainly when those increments are negative, i.e. they come from a drying of the atmosphere or a reduction in cloud or rain. This is consistent with what we saw in column cloud in Fig. 15c. The hint of agreement between 1D-Var and 4D-Var increments in channels 37v and 37h is encouraging, considering they are not assimilated. In contrast, there is no agreement in channel 85v, in agreement with our belief that snowfall and ice cloud are not well handled in the radiative transfer.

3.5 Comparisons to independent rain observations

This section compares surface rain rates from the 4D-Var first guess and analysis, and from the 1D-Var first guess and retrievals, against observations from the Precipitation Radar (PR) on TRMM (Kummerow et al., 1998). TRMM follows an inclined orbit designed to provide coverage from roughly 37°S to 37°N, with an equator crossing time that varies from day to day, enabling the whole diurnal cycle to be sampled over the course of a month. Following a change in the operating altitude of TRMM in 2001, the horizontal resolution of PR is 5 km. The radar is scanned across-track, making 49 separate measurements or 'rays', giving a swath width of 250 km. PR is sensitive only to rainfall amounts greater 0.7 mm h^{-1}.

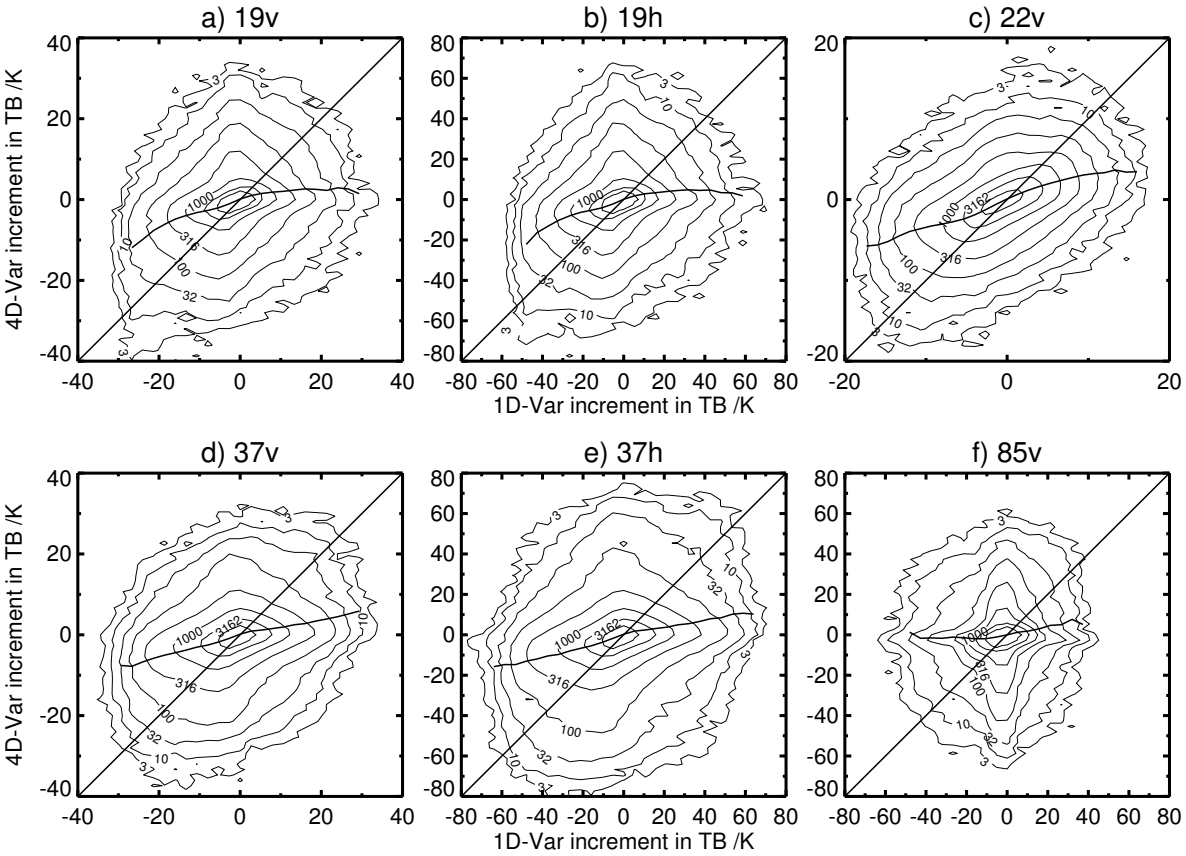

Figure 16: 2D histograms of increments (analysis or retrieval - first guess) at SSM/I observation points in 4D-Var against those in 1D-Var, as for Fig. 15, only in brightness temperature space. The straight bold line is the 1:1 line. The wiggly bold line shows the mean 4D-Var increment in bins delineated by their 1D-Var increment. Channel 85h is omitted for space reasons.

PR measures the amount of radar reflected from raindrops. By assuming a certain distribution of drop sizes and fall speeds, this can be converted into a vertical profile of rain rates. In contrast, our 1D-Var retrievals from SSM/I come from passively sensing the top of the atmosphere microwave radiation, which comes from the emission and scattering from many things including rain drops. Many space-based rain retrievals are available, but most use passive-microwave techniques. PR is particularly suitable for these comparisons because it uses a different measuring principle, and also as result it is not affected by surface emission, gas absorption, cloud ice or water, or (for rain retrievals below the freezing level) snow.

In this section we use version 6 of the '2A25' rain retrievals (Iguchi et al., 2000), which have been downloaded from http://daac.gsfc.nasa.gov/. Because PR measures a vertical profile of rain, but cannot always see to the surface due to radar ground clutter, some investigators have chosen the rain rate at 2km altitude as representative of the surface. The PR data files provide both a 'near surface rain rate' and 'estimated surface rain', both essentially based on radar reflectivities in the lowest clutter-free bin, but with some small differences in the algorithm (JAXA/NASA, 2005). The figures we show are based on the 'estimated surface rain'; we tried all three alternative estimates, but the choice makes essentially no difference to our ocean-only comparisons.

Assumptions on drop size distributions and fall speeds are needed to convert PR reflectivity measurements into a rain rate. Unfortunately we cannot use the approach of Masunaga et al. (2002), who found that a better way of comparing PR to microwave-based TMI rain observations was by looking at rain density, a quantity more closely related to the radiative transfer. This eliminates the assumptions on drop size and fall speed. However, our moist physics models represent rain in flux terms, and as noted in Sec. 2.2.7, similar assumptions need to be made in 1D-Var to convert these fluxes into densities for the simulation of SSMI TBs. For simplicity we have chosen to compare the PR rain rates to the model rain rates. We should be aware of the possibility that our drop size and fall speed assumptions are incompatible with those used in the PR rain rate retrievals, but for the moment we have not investigated this further. Masunaga et al. (2002) also show that PR is sensitive, as well as to rain, to the larger frozen hydrometeors. However, this is only an issue above the freezing level and should not affect our comparisons at the surface.

3.5.1 Monthly mean

Figure. 17 shows monthly mean surface rain rates from the 1D-Var and 4D-Var first guess at SSM/I observation locations, as well as the PR monthly mean, for February 2007. Observations have been binned on a 2.5° by 2.5° grid. Fig. 18 shows the zonal mean rain rates, again for oceans only. The comparison confirms what we already saw in the SSM/I departure statistics (Sec. 3.3): 1D-Var first guess produces an excessive amount of rain. Rain amounts in the first guess trajectory from 4D-Var are also higher than PR, by between 20% and 140% depending on latitude. However, we can get a rough feeling for the quality of PR by noting that zonal mean rainfall from TMI (also on TRMM) has been shown to be roughly 30% to 40% higher than PR (Kummerow et al., 1998). It is still not clear which of those two estimates is best. Hence we should be cautious in saying that 4D-Var rain is overestimated on the basis of comparisons to PR alone. However, the SSM/I FG departure statistics (Sec. 3.3) do also suggest that 4D-Var overestimates rain amounts.

Ignoring the mean differences, the distribution of rain over the globe is similar in the model and in PR observations, but the smaller details on the 2.5° by 2.5° grid do not agree. These small-scale differences are unlikely to be significant, since monthly mean rainfall can be strongly affected by just one or two extreme events. Shin and North (1988) estimated that monthly means from TRMM sensors would have a sampling error of about 8% to 12% even on a 5° by 5° grid.

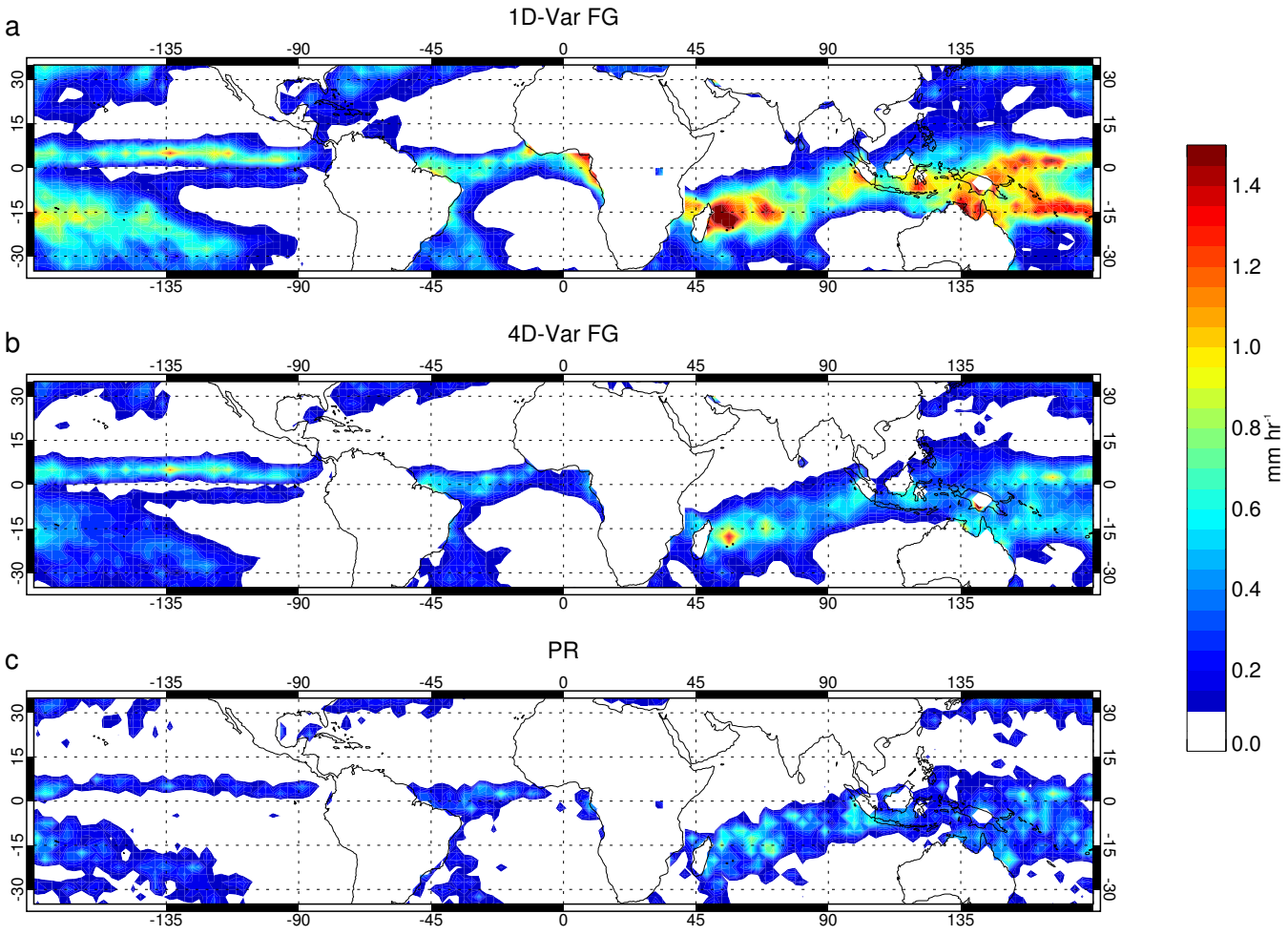

*Figure 17: February 2007 mean surface rain rate, in mm h^{-1} in 2.5° by 2.5° bins, from the first guess at all SSM/I locations (i.e., both clear and rainy) simulated by **a)** the 1D-Var first guess, **b)** the 4D-Var first guess, and **c)** from TRMM PR 2A25 surface rain estimates, shown over ocean points only.*

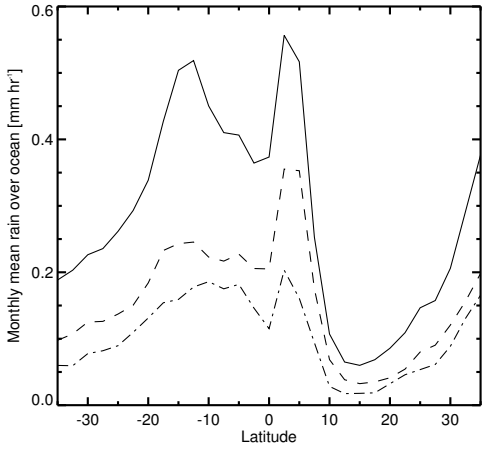

Figure 18: Zonal mean February 2007 mean surface rain rate, over oceans only, in mm h^{-1} from the 1D-Var first guess (solid), the 4D-Var first guess (dashed) and TRMM PR 2A25 (dot-dashed).

3.5.2 Colocations

One of the most powerful aspects of PR is the ability to reveal the structure of rain at high spatial resolution. In this section we make near-exact colocations between PR and the 1D-Var retrievals at SSM/I locations. Because the equator crossing time of TRMM's orbit varies from day to day, it is possible to make colocations between SSM/I and PR within a few minutes in time, and with an exact match in space. Each day, such colocations are found at only a small range of latitudes, but these latitudes change day by day, and over the month of February 2007, we build up colocations at all longitudes, and all latitudes between 36.4°S and 36.4°N.

Figure 19 gives an example of the quality of the colocations in an area of frontal rainfall in the North Atlantic. All observations shown on the figure come from within a 6 minute period. In this case TRMM is at its most northerly point and is flying from west to east over the comparison area. The DMSP satellite (with SSM/I onboard) is travelling nearly north to south, on a descending node. TRMM intersects just a few SSM/I swaths.

Figure 20 shows the wider meteorological context for this example, with forecast surface pressure at 0900 UTC overlaid by the SSM/I brightness temperature difference between channels 19h and 19v. This quantity becomes less negative the more TCWV, cloud, and rain is present, thus highlighting the frontal systems. The least negative (smallest) differences indicate heavy rain. There is a deep low at 39°N, 68°W, with heavy frontal rain. There are a number of other frontal features in the mid Atlantic associated with various troughs and a much less intense low. At first glance, the first guess and SSM/I observations look very similar. However, if we look closely at the region of Fig. 19, which is shown by a rectangular box spanning 10° of longitude and 2° of latitude, we start to see differences. The locations of heavy rain (e.g. areas with 19h-19v differences smaller than an arbitrarily chosen threshold of -20K) differ between the first guess and observations by from 1° to 2°, i.e. roughly 100 to 200 km. The extents and intensities of these rain areas do not match between first guess and observations either. This reinforces what we saw in Sec. 3.3, that first guess rain and cloud is often very different from observations.

Figure 19a shows the PR observations at their full 5km resolution. A perhaps surprising feature of PR is that the vast majority of points show surface rainfall rates of zero. In one sample day, only 4% of measurements show non-zero rainfall, though it should be noted that the instrument cannot see rain amounts lower than 0.7 mm h^{-1}. The colocation shown in Fig. 19 has been specially chosen because of the presence of rain. Even here, rainfall is restricted to very localised areas. Note that a usual assumption is that the insensitivity of PR to small rain

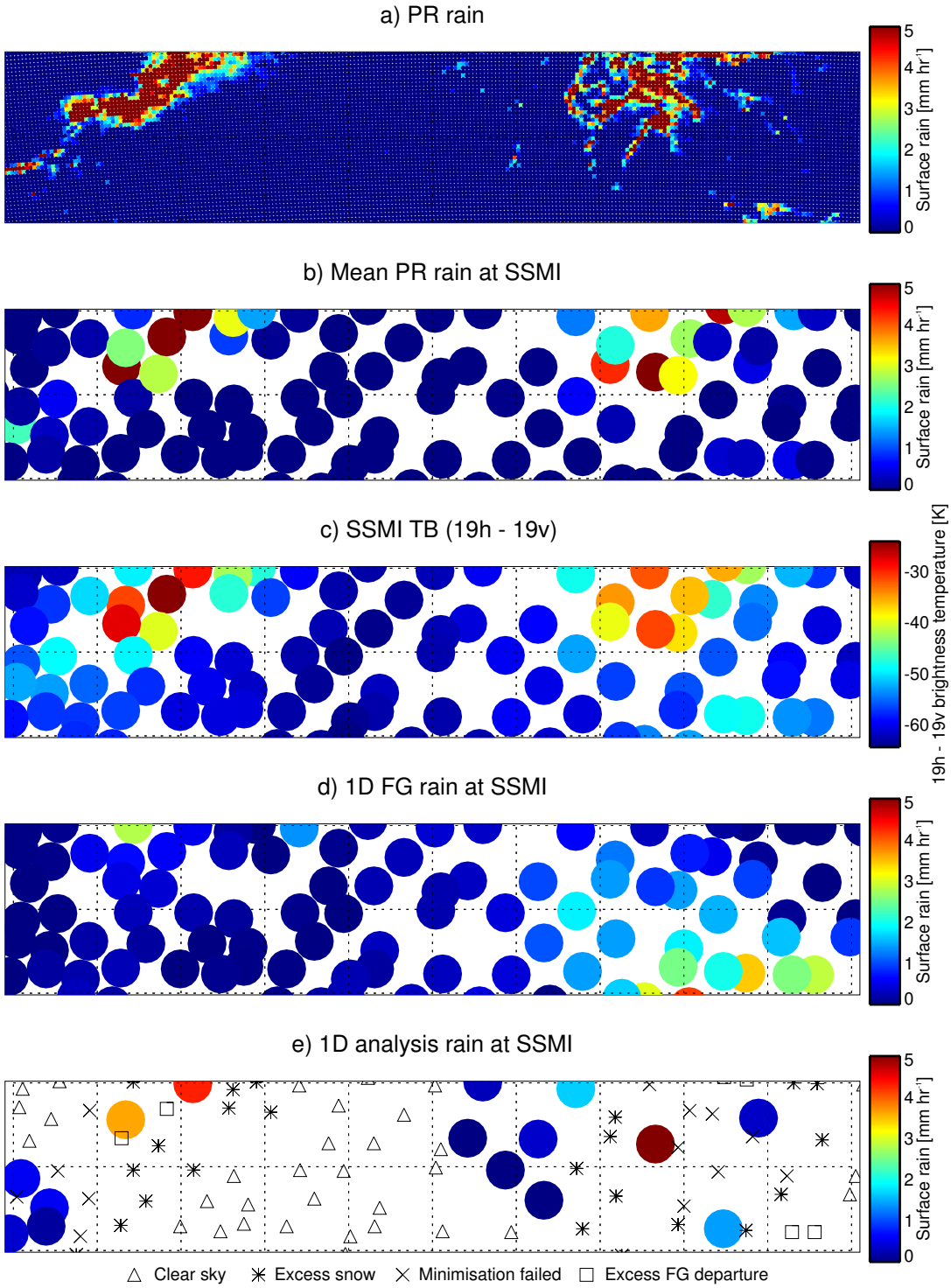

Figure 19: Example of colocations between PR and SSM/I, in the North Atlantic, showing an area from 34°N to 36°N and 48°W to 38°W. Observations were made within the period 09:34 to 09:40 UTC on 23rd February 2007: ***a)*** *PR surface rain estimates at their original 5km resolution, with most rainy observations exceeding the 5 mm h^{-1} colour scale maximum;* ***b)*** *PR rain rates averaged within a 25km radius of the SSM/I observations used by 1D-Var;* ***c)*** *Difference between SSM/I observed brightness temperatures in channels 19h and 19v;* ***d)*** *1D-Var first guess surface rain at SSM/I locations;* ***e)*** *1D-Var retrieved surface rain, and in cases where an retrieval was not produced, symbols indicate the reason. Observations are represented by circles of roughly 50km diameter (SSM/I and 1D-Var) and 5km diameter (PR).*

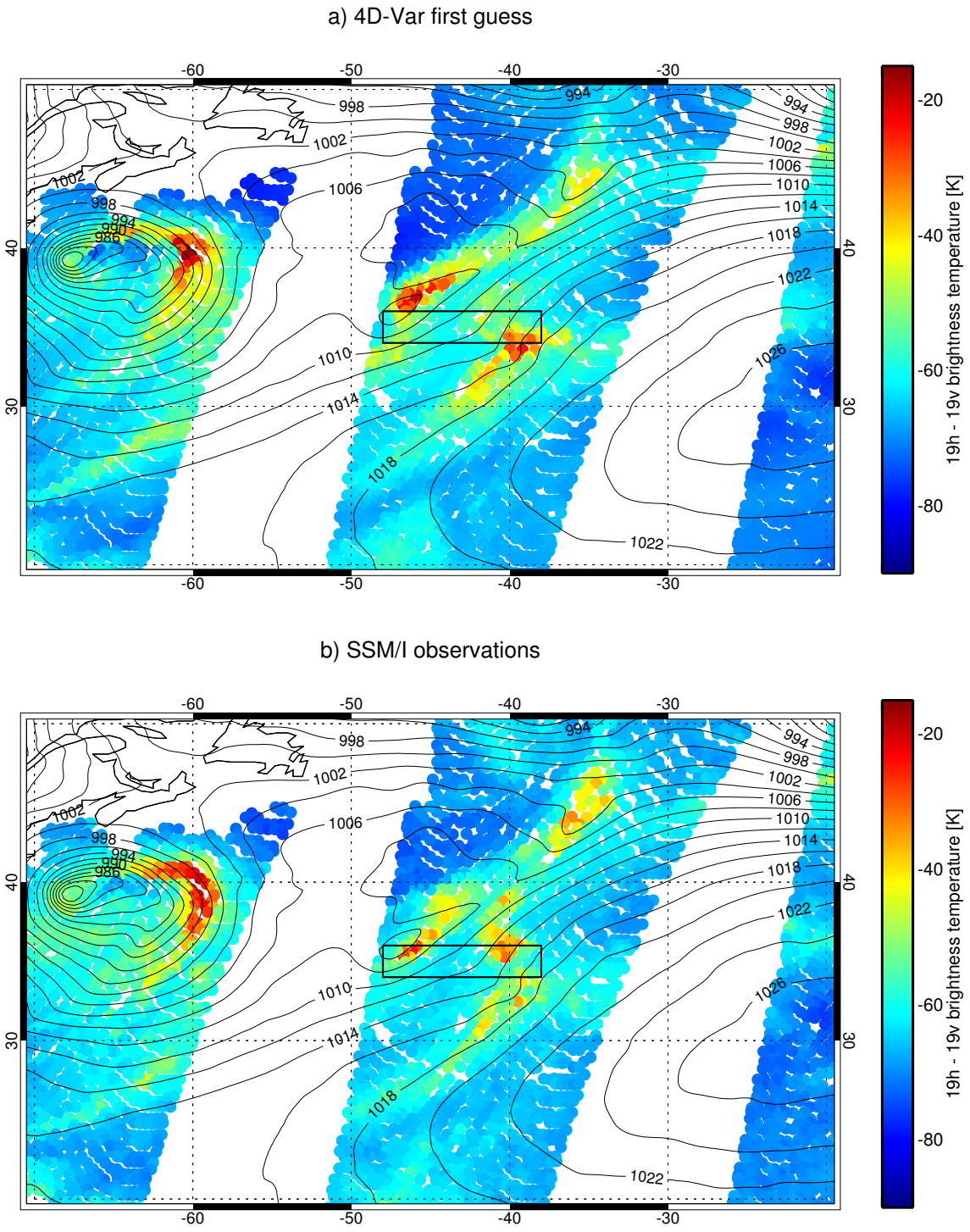

Figure 20: First guess pressure at mean sea level in the North Atlantic for 09:00 UTC on 23rd February 2007, overlaid by the channel 19v - 19h brightness temperature difference at SSM/I 1D-Var locations, for the DMSP-F13 satellite only: ***a)*** *Simulated from the 4D-Var first guess;* ***b)*** *Observed by SSM/I. The region of Fig. 19 is shown by a rectangular box.*

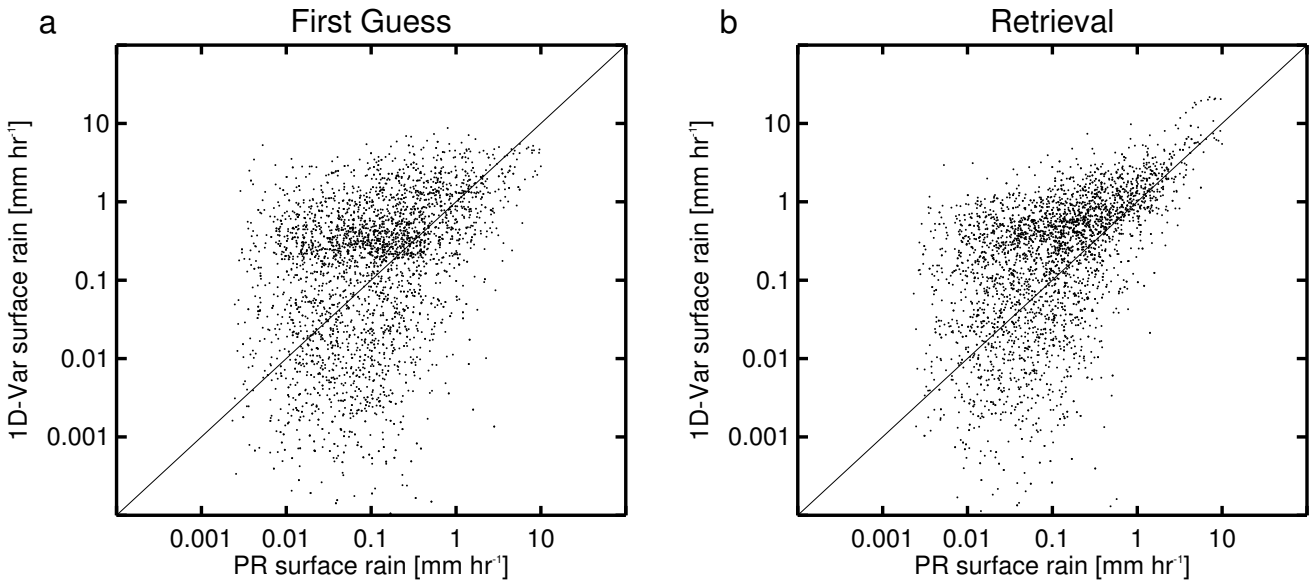

*Figure 21: Scatter plots of PR surface rain rate (averaged over the SSM/I footprint) versus **a)** first guess and **b)** retrieval from 1D-Var, for February 2007, restricted to points where a successful 1D-Var retrieval was made. Colocations are made within 7.5 minutes of the SSM/I observation time, and mean PR rain rates are calculated from all PR observations within a 25km radius of the SSM/I centre point. By definition, the figure can only show colocations where both PR and 1D-Var rain rates are greater than* 10^{-4} *mm* h^{-1}.

amounts does not in most cases affect the determination of mean rainfall (e.g. Ikai and Nakamura, 2003).

To compare with SSM/I, we average all PR observations within a 25km radius of the SSM/I location. This is shown in Figure 19b. Recall that 1D-Var is done on grid-points, using the closest SSM/I observation with a maximum mismatch of 10 or 7 km depending on the grid resolution. SSM/I observations that do not match these criteria have been discarded, so this figure does not represent the full spatial sampling of SSM/I. Nevertheless, SSM/I is oversampled and the actual footprint is 70 by 45 km for channels 19v and 19h, and 60 by 40 km for channel 22v (Spencer et al., 1989). The resolution of the model grid points themselves is roughly 25km. The figure represents SSM/I observations and 1D-Var retrievals as circles of approximately 25km radius centred on the SSM/I observation location.

Figure 19c shows the observed SSM/I channel 19h - 19v difference at 1D-Var locations. The SSM/I observations with the least negative 19v-19h differences in Fig. 19c correspond, visually at least, with the locations where PR shows rain (Fig. 19b). This gives confidence that we are comparing like with like. SSM/I 19v-19h differences of order -50K indicate the wider area of cloud and high TCWV surrounding the rainy areas.

Figure 19d shows surface rainfall in the 1D-Var first guess. The model first guess captures neither of the active rain systems seen by PR and SSM/I, but instead it shows rain in the lower right corner of the domain. The 4D-Var first guess rainfall (Fig. 20a) is qualitatively very similar to the 1D-Var first guess.

Figure 19e shows the performance of the 1D-Var retrieval. About half of points are identified as 'clear' so are not used in 1D-Var. In the area of the rain systems (both FG and observed), almost all retrievals are abandoned due to one of three factors: excessive snowfall in the first guess, excessive first guess departures, or failed minimisations. This indicates that we need to re-consider both the snowfall and the FG departures check, and also to improve the success of the minimisation. However, of the successful retrievals, there are three in the areas of heavy rain shown by PR, and in all three of these, rain has been retrieved by 1D-Var where there was none in the FG.

		PR	
		Clear	Rain
1D-Var first guess	Clear	238	66
	Rain	5275	3472
1D-Var retrieval	Clear	252	67
	Rain	5261	3471

Table 6: Rain contingency table, with rain defined as surface fluxes greater than 10^{-4} mm h^{-1}, for February 2007 colocations between PR and a successful 1D-Var retrieval.

Figure 21 shows a scatter plot of all colocations in February 2007 between PR and successful 1D-Var retrievals, but only for the $\simeq$3500 cases where both the PR and 1D-Var rain rate is greater than 10^{-4} mm h^{-1}. As before, all PR observations within 25km of the SSM/I observation have been averaged; typically about 80 PR observations are averaged together. The maximum allowed time difference between PR and SSM/I is $\pm$7.5 minutes. As we might expect from the previous example, the agreement between PR and FG rain is not good. 1D-Var retrieved rain does appear visually closer to that measured by PR, especially for rain rates above 0.1 mm h^{-1}. However, the improved agreement is not really at the level where it is worth trying to examine quantitative measures of fit: 1D-Var retrievals are simply not particularly close to PR.

Table 6 examines the full set of colocations, including those where PR sees no rain, in terms of a contingency table. Remember that we are looking at the 'rainy' set of SSM/I observations: they have already been selected on the basis of cloud or rain being detected in the SSM/I brightness temperatures. The majority of 1D-Var first guesses and retrievals see rain above the 10^{-4} mm h^{-1} threshold, but in only 40% of cases does PR confirm this. Here, this indicates that, as well as producing too much rain, 1D-Var first guess produces rain in too many places. Roughly half of the 1D-Var first guesses show rain fluxes between 1×10^{-4} and 5×10^{-2} mm h^{-1}. Hence a lot of this excessively widespread rain is associated with very low rainfall amounts. This excessively widespread rain is thought to come from the simplified large scale convection scheme. Table 6 shows that the problem is not corrected in the 1D-Var retrieval either.

3.6 Observing System Experiments

In this section we make use of Observing System Experiments (OSEs) to investigate the impact of SSM/I rain data assimilation on forecasts and analyses. An OSE is simply an assimilation experiment in which certain observations are added or removed, and the results are compared to a control experiment. Hence, for example experiments X and A created an OSE in Sec. 2.2, as did experiments RAIN and NORAIN in Sec 3.4. In this section we examine the results from a more comprehensive set of OSEs, covering all the satellite observing systems used by ECMWF, of which the SSM/I clear-sky and rain assimilation results have been summarised by Kelly et al. (2007). These experiments were made with cycle 31r1 of the ECMWF system, using T511 horizontal resolution instead of the operational T799 (the 91 vertical levels of the operational system were, however, retained in these experiments). The experiments are summarised in Table 7. In this section, we summarise and expand on the results of Kelly et al. (2007) as regards rainy SSM/I assimilation.

The control experiment, identical to the operational system apart from the change of resolution, is similar to experiment F in Sec. 2.2. However, the experimental design of these OSEs has many improvements compared to the experiments in Secs. 2.2 and 3.4. Experiments were run for a long period, from 1st June 2006 until 15th August 2006. For the first 14 days, variational bias control (VarBC) was allowed to adapt to the changed observational system, before being fixed for the remainder of the experimental period. Statistics are based on the period 15th June until 15th August 2006, excluding the first two weeks of spin-up, and hopefully avoiding

Experiment	SSM/I rain observations	Description
BL	Off	Cycle 31r1 but with all satellite observations removed except one AMSU-A.
BL + RAIN	On	As BL but with SSM/I rainy observations added
CTRL - RAIN	Off	As CTRL but with SSM/I rainy observations removed
CTRL	On	Cycle 31r1 control

Table 7: Observing system experiments run for 1st June to 15th August 2006.

any bad effects from drifts in the bias correction during the experiment. The longer period and the better control of VarBC mean that we can be more confident in the results of these experiments compared to those presented earlier.

As well as the control (CTRL), which includes the full set of operational observations, there is a "low baseline" experiment (BL), which includes only conventional observations and data from a single AMSU-A instrument, that on NOAA-16. The AMSU-A observations help constrain dynamical fields in the Southern Hemisphere, where there are few conventional observations. In this section we examine experiments where SSM/I rain data assimilation was added to the low baseline (BL + RAIN), and where it was removed from the full system (CTRL - RAIN). Figure 22 helps summarise the results presented by Kelly et al. (2007). It shows the forecast skill, calculated as the RMS difference between forecasts and analyses valid at the same time (see Appendix A). Here, we have used the operational analyses to verify against. The reduction in forecast error between the baseline (BL) and the control (CTRL) experiments represents the influence of satellite observations, excluding the one AMSU-A already included in BL. This influence is largest in relative terms in the tropics, and smallest in the Northern Hemisphere (NH), where analyses from the baseline experiment are already well constrained by conventional observations and the single AMSU-A. In the NH, the addition of all the other satellite systems results in only 12 h of additional forecast skill. For example in Fig. 22c, 48 h forecasts of 1000hPa vector wind from CTRL have the same skill as 36 h forecasts from BL.

The main impact of adding rain assimilation to the baseline comes in the tropics (e.g. Figs. 22b and e). Rain assimilation is beneficial: in almost all regions and levels, its addition to the baseline improves forecasts. The beneficial impact is found not just in the directly-observed moisture fields (e.g. relative humidity, Figs. 22d-f) but also in the dynamical fields (e.g. vector wind, Figs. 22a-c). Kelly et al. (2007) show that SSM/I observations (both clear-sky and rainy) are responsible for large mean changes in moisture, particularly in the tropical lower and mid troposphere. They also show that most of the mean differences in TCWV between the baseline and the control can be replicated simply by adding both clear and rainy sky SSM/I assimilation to the BL experiment, without the assimilation of any more satellite data types.

The impact of removing rain assimilation from the control is extremely small in comparison to the effect of adding it to the baseline, but there does seem to be a degradation in forecast scores when it is removed. The small size of the impact indicates that the complete observing system has much redundancy. This redundancy is not limited to the SSM/I rain observations. OSEs were performed for other satellite observation types. The removal of any single one of these has only a tiny impact on the quality of forecasts. It is only by removing most of the other satellite observations that the maximum possible impact of any particular satellite observation type can be seen. These conclusions are highly consistent with those of Lopez and Bauer (2007) who experimentally assimilated rain observations from radar and rain gauges over the United States.

In the rest of this report, we are concerned with the incremental impact of rain observations within the full observing system. We see that such incremental impacts are typically small, and as shown in this paper, such impacts can be dominated by biases rather than by the true information content of the observations. It is useful

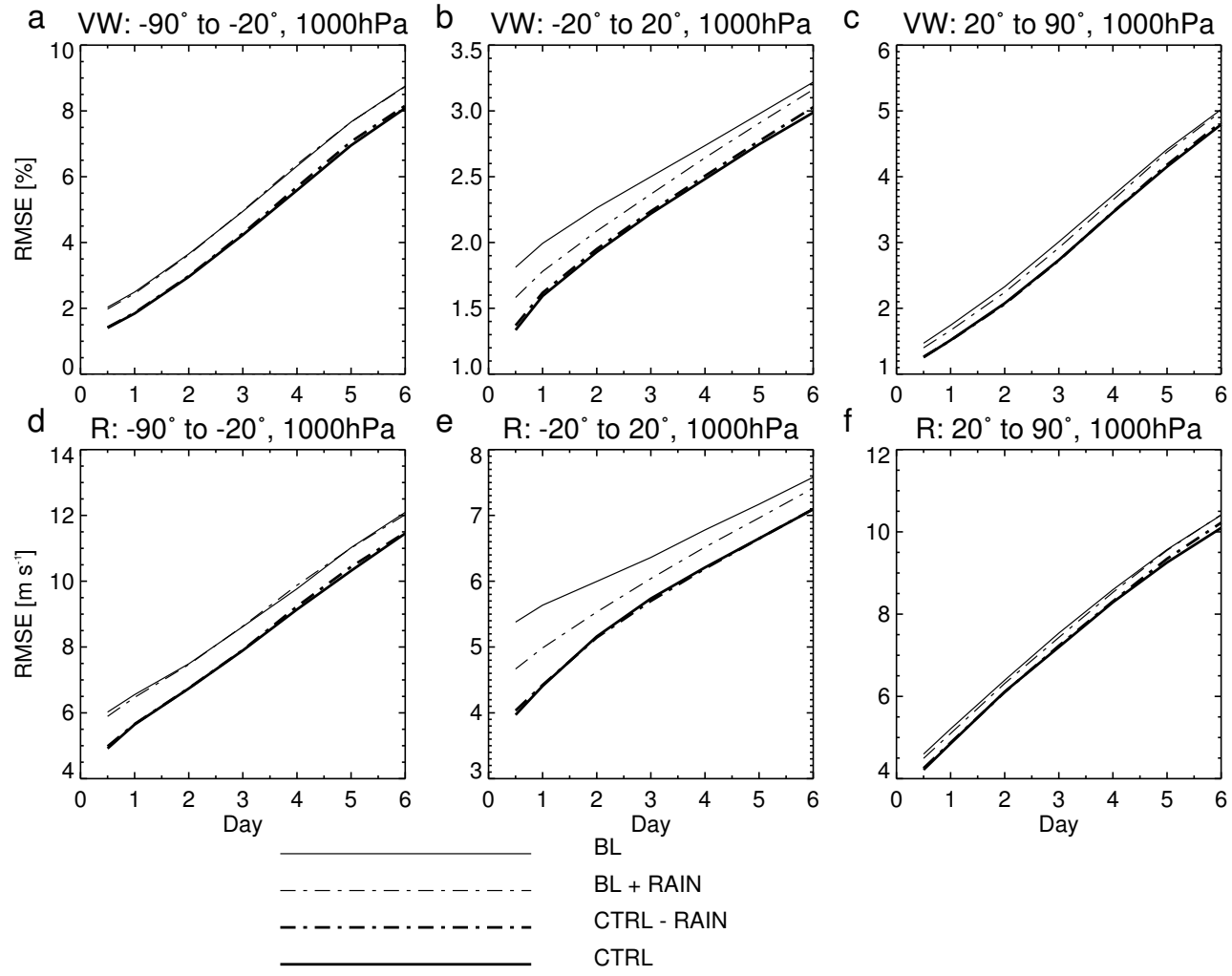

Figure 22: Root mean square errors (RMSE) against forecast time, in vector wind (VW, in m s^{-1}, top row) and relative humidity (R, in %, bottom row) at 1000hPa, in the Southern Hemisphere (a,d), tropics (b,e), and Northern Hemisphere (c,f). Errors are calculated as the RMS difference between experiment forecasts and operational analyses.

to remind ourselves, with Fig. 22, that in the absence of competing satellite observation types, the 1D-Var assimilation of rainy SSM/I observations brings large forecast benefits, particularly in the tropics.

In Fig. 22, we have calculated forecast errors by comparison to a common reference, the analyses produced operationally by ECMWF. This is because the operational analyses are undoubtedly a better representation of the atmosphere than those from the low baseline experiments. However, we should be cautious in our interpretation of the difference between the CTRL and CTRL-RAIN experiments in Fig. 22. In general, in a pair of experiments like this (a situation known as 'data-denial'), if analyses from the control are used for verification of both control and experiment, the short range forecasts (up to 24h) from the control gain a spurious advantage over forecasts from the experiment. The presence of additional observations in the control causes perturbations in the analysed fields which are carried into the short range forecasts. It is irrelevant whether these perturbations are beneficial or detrimental to the analyses; the short range forecasts from the control are almost guaranteed to be closer to the verifying analyses than the forecasts from the data-denial experiment. Another way to look at this is that if we are trying to test a particular observation type, we should avoid making the a-priori assumption that the control analyses (containing this observation type) are better than those from the experiment with the observation type removed. In Fig. 22, we have verified against operational analyses (made with a different cycle, 30r1) in an attempt to avoid this effect. However the operational analyses still share with the CTRL experiment the presence of rainy SSM/I observations. In the rest of this report, we have compared forecasts to the analyses from which they were generated. This approach is more appropriate to data-denial experiments where a single observation type is removed.

Figures 23, 24 and 25 show the differences in forecast errors between the CTRL and CTRL-RAIN experiments in relative humidity, temperature, and vector wind. Here, the forecasts are verified against their own analyses.

Figure 23 shows that in the data-denial experiments the principal beneficial impact of SSM/I rain assimilation on relative humidity is at 700hPa in the tropics and subtropics, between 30°S and 30°N. This has been seen consistently in all experiments we have made, no matter what version of the ECMWF system or time of year, including those at cycle 29r2 described by Bauer et al. (2006b) and at cycle 30r2 by Bauer et al. (2006c). Improvements persist as far as the T+72 forecast, though by this stage they are in general not statistically significant. A small area of degradation at T+24 at 50°N suggests continuing problems in the winter hemisphere in the mid-latitudes.

Figure 24 shows that the impact on temperature is more mixed. Near the surface, there is positive impact to the south of the equator and there are degradations to the north. In the tropical upper troposphere, rain causes better forecasts at 200hPa and worse forecasts at 300hPa. This dipole structure was not seen in the original cycle 29r2 experiments of Bauer et al. (2006b). However, it has appeared in similar OSEs (not shown) from cycle 30r2 onwards. It is linked to the presence of small mean temperature changes over the forecast range, i.e. spin-up. It is hard to avoid temperature spin-up effects in the tropics and the ECMWF system is never entirely steady over the forecast range. However, the tropical temperature spin-up changed in character from cycle 30r2 onwards, likely as a result of model changes allowing supersaturation with respect to ice. This spin-up is present in the system regardless of whether the SSM/I rainy observations are included or not. However, the drying effect of SSM/I rain assimilation (see Sec. 3.4) causes decreased CAPE in the tropics, decreased high cloud, and changes to the amount of latent heating in the tropical upper troposphere, causing a slight cooling during the forecast. Because of the nature of the underlying spin-up, the effect on forecast scores is different at different levels. At 200hPa, where there are increasing temperatures over the forecast range, the extra cooling reduces the amount of spin-up and gives better RMS forecast scores. At 300hPa, where there are decreasing temperatures over the forecast time, the cooling causes an increased spin-up and worse scores. These effects are small (order 0.1K to 0.2K over the forecast range) but they do indicate that efforts should be made to reduce the drying effect of SSM/I observations, i.e. to remove the biases in the system.

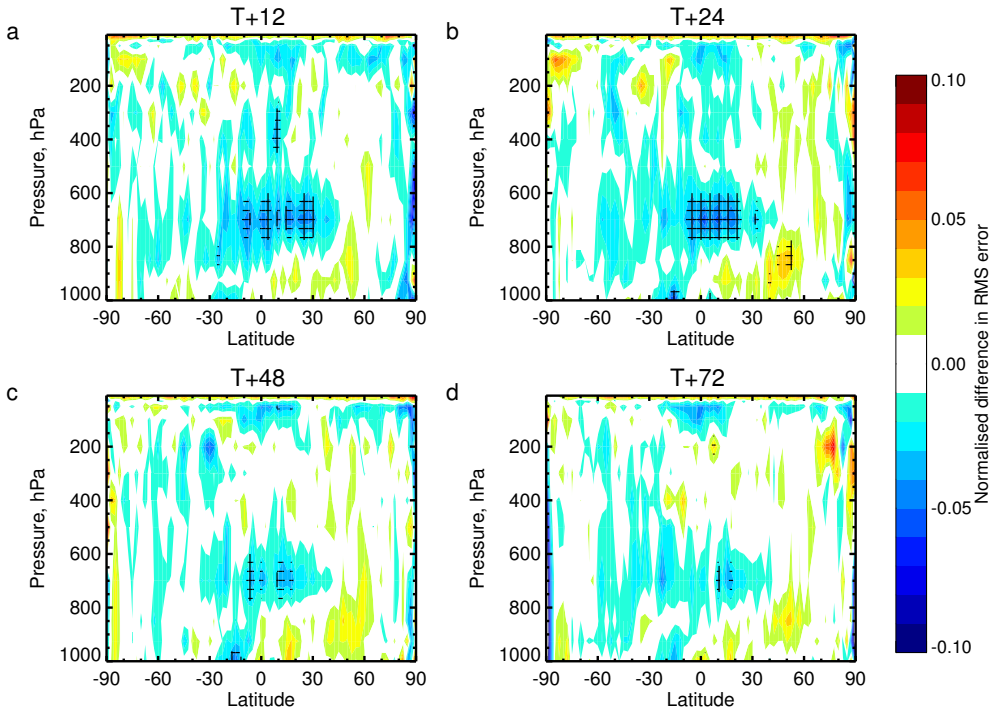

Figure 23: Normalised difference in RMSE relative humidity errors between CTRL and CTRL-RAIN experiments, calculated as $(RMSE_{CTRL} - RMSE_{CTRL-RAIN})/RMSE_{CTRL-RAIN}$*, and resolved by latitude and by pressure level, and shown for forecast times of 12, 24, 48 and 72 hours. Cross-hatching indicates differences that are statistically significant. Negative (blue) contours represent areas where the CTRL experiment has a lower RMSE than CTRL-RAIN, and thus produces better forecasts. Statistics are based on the period 15th June to 15th August 2006, and in contrast to Fig. 22, errors are calculated as the difference between forecast and own analysis. See Appendix A for details of the calculations.*

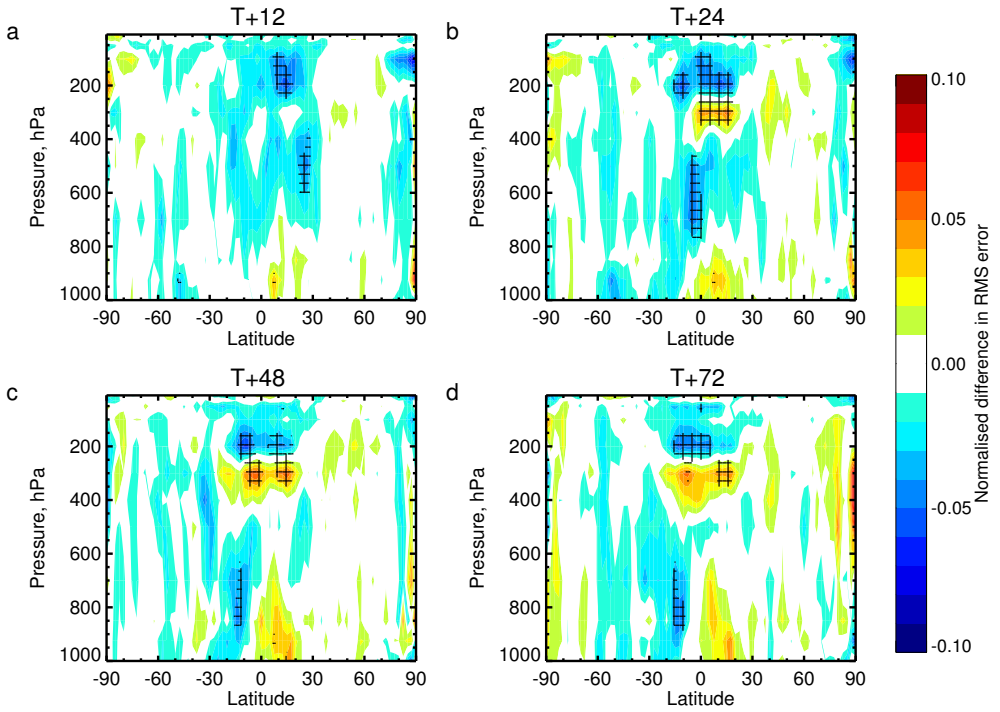

Figure 24: As for Fig. 23 but for temperature.

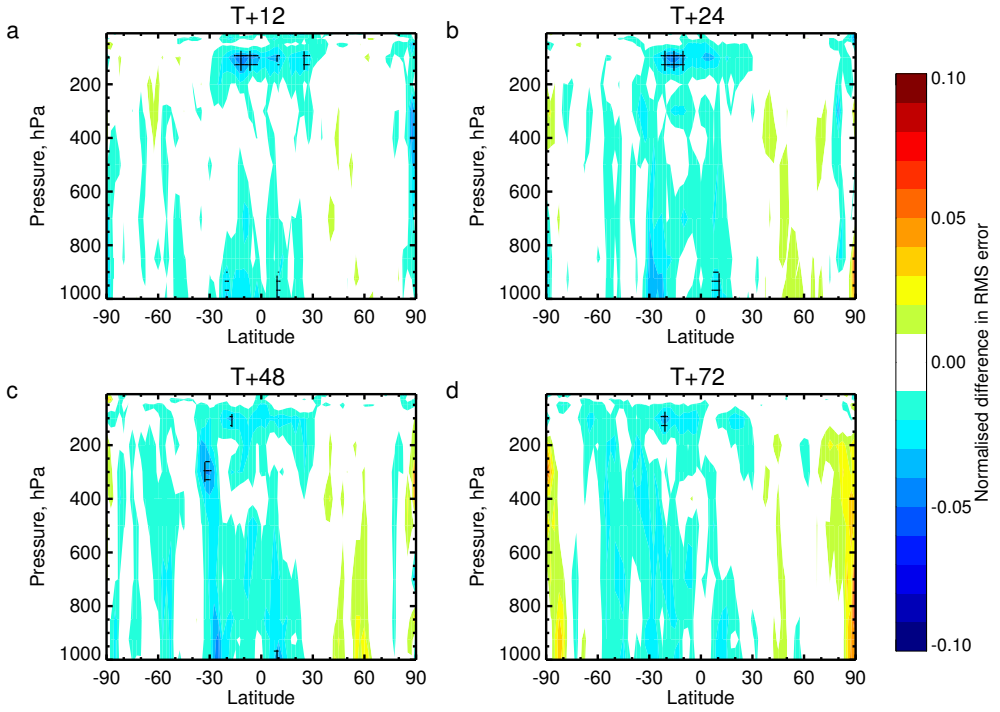

Figure 25: As for Fig. 23 but for vector wind.

SSM/I rain assimilation improves forecasts of vector wind throughout the tropics and subtropics and through the full height of the troposphere, though in no cases can these changes be considered statistically significant (Fig. 25). However, we have already seen that when SSM/I rain observations are assimilated into the low baseline experiment, there are substantial improvements in vector wind (Fig. 22b). This would suggest that the minor improvements in vector wind forecasts shown in Fig. 25, while not statistically significant, are linked to a real improvement in forecasts.

Equivalent figures can be created from the RAIN and NORAIN experiments examined in earlier sections. The RAIN and NORAIN experiments are for February 2007 and cycle 32r1, as opposed to summer 2006 and cycle 31r1 for the CTRL and CTRL-RAIN experiments examined in this section. Figures are not shown here, but the important features, such as the positive impact in RH at 700hPa in the tropics, and the degradation in temperature at 200hPa, are also apparent in the forecast error differences between RAIN and NORAIN. Differences are in general less statistically significant, because of the shorter period.

4 Conclusion

Cloud and rain affected observations from SSM/I have been assimilated operationally at ECMWF since June 2005, using a 1D+4D-Var method. A number of minor improvements have been made to the system over time; these have been detailed in section 2.2. The rest of this work examines the performance of the system after two years of operations and summarises a number of problems that still need to be dealt with before rain-affected observations can be used to their full potential. Rain assimilation has previously been tried in many case-study experiments, and many different issues have been raised from a largely theoretical perspective (e.g. Errico et al., 2007). In this study we examine a more practical side to the problem. While issues such as for example the nonlinearity of the moist physics operators are important, it appears that our currently available operators (Tompkins and Janisková, 2004; Lopez and Moreau, 2005) do work reliably in 1D-Var and 4D-Var. The issues we face on a practical level are those of biases in models and operators, of making sure we correctly sample cloud, clear-sky and rain in our observations, and the continuing poor quality of first guess rain and cloud distributions.

To understand how the system works, we have made use of a number of short (1 to 2 months) experiments run offline. The results from them are representative of many other similar experiments that have been run whilst developing the system. Four main tests have been presented. First, we have examined departure statistics of first guess brightness temperatures. An important improvement has been to start monitoring 'clear' observations as well as 'rainy', and to look at first guesses calculated from the nonlinear model's rain and cloud, instead of just those from 1D-Var, which uses simplified moist physics operators to diagnose rain and cloud amount from the moisture and temperature profile. Second, we have examined data-denial experiments to test the impact of the 'rainy' observations on 4D-Var. For the first time, we have also examined how well the moisture, rain, and cloud information is transmitted from 1D-Var to 4D-Var. Third, we have compared 1D-Var retrievals to independent observations from the precipitation radar on TRMM. Fourth, we have examined OSE experiments, which have been separately performed by Kelly et al. (2007). An important development has been the use of 'low baseline' experiments, from which the majority of satellite observations have been removed, to more clearly see the impact of a particular observation type.

Rain assimilation improves RMS forecast scores. In experiments where the rain observations are removed from the full observing system, we see a statistically significant improvement in tropical humidity at around 700 hPa, and minor but not significant improvements in vector wind. When adding rain assimilation to a low baseline system, there are very significant improvements in tropical moisture forecasts, as well as in wind and (not shown here) temperature forecasts. These results justify the operational use of the current rain assimilation

system. We see, however, that it is the additional TCWV information provided by these measurements in rainy and cloudy areas that improves analyses and presumably forecasts. Observational information on rain and cloud is not properly used in the 1D+4D-Var technique. The 1D+4D-Var technique is only a first step and we want to learn where we can improve. A principle conclusion from this study is that further developments are best made within a direct 4D-Var framework. The more detailed conclusions are subdivided between general issues, i.e. those likely to affect any rain assimilation system, specific recommendations for short-term improvements to the 1D+4D-Var system, and an outlook for the direct 4D-Var approach.

4.1 General issues

- **Mislocation of first guess rain and cloud**. The locations of areas of heavy rain and cloud can be 100-200km out of place in the model first guess compared to observations from PR and SSM/I. Because rain systems are often of similar size to the mislocation error, many rainy areas in the first guess correspond to 'clear' SSM/I observations, and vice-versa. This is true not just of 1D-Var, which uses simplified moist physics operators to generate rain and cloud, but also of the forecast model used in 4D-Var. Rain assimilation has a difficult job: to remove rain in one place and create it in another. It is often said that data assimilation works best when making small corrections to a reasonably accurate first guess. In areas of rain and cloud we have not yet reached that stage. The problem is likely to become even more significant as model resolution increases and it is hard to imagine that first guess forecasts over six or twelve hours could ever produce rain and clouds in exactly the right place at the right time. It might be possible to force rain and clouds in the model first guess to be closer to observations, either by using very short time window data assimilation, or by using the Kalman Filter. Another approach might be to assimilate area-averages of the cloud and rain information. Of course this problem is already well known, but it is still worth emphasising that it must always be taken into account when designing our rain and cloud assimilation strategies. In the implementation of the 1D+4D-Var system, its impact was perhaps not fully appreciated.

- **Sampling biases**. The 1D+4D-Var technique has been implemented only for observations which are themselves identified as 'rainy'. 'Clear' observations are instead assimilated directly in 4D-Var without simulating the radiative effect of any rain or cloud present in the first guess. This is wrong because first guess and observed cloud and rain are often completely mislocated. There are two impacts: (a) the first guess is not properly corrected in areas where the first guess is rainy but the observations are clear; (b) there is a very large sampling bias. If bias correction is derived separately for each stream, as it is at the moment, we are losing real information on the presence or absence of cloud and rain. Since the match between first guess and observed rain is not very good, it does not make sense to draw the distinction between 'clear' and 'rainy' scenes from the observations alone. The ideal option for the future is not to make this distinction at all, and to use a single stream including rainy and cloudy radiative transfer when required. This will likely be true for all satellite observations with a rain or cloud information content.

- **Model biases in rain and cloud**. We have seen that the 1D-Var first guess monthly mean surface rainfall is roughly 100% higher than that in the 4D-Var first guess. However, even the 4D-Var first guess rain appears to be overestimated, based on both the SSM/I TB departures, and comparisons to independent data from TRMM PR. The presence of the bias in the 1D-Var first guess has historically been masked in the rainy SSM/I FG departure statistics due to the sampling bias described above: the statistics were formerly based on only rainy and cloudy observations. Observations are of course bias corrected before being assimilated in 1D-Var, using a number of predictors and a multiple linear regression. This works well enough to enable the rain assimilation to make a positive impact on forecast scores. However, a lot of regional bias remains, such that there is typically a drying tendency in rainy and cloudy areas. The

problem becomes even worse in the higher frequency SSM/I channels, which are much more sensitive to cloud and rain. These cannot be assimilated while such large FG biases remain. Ongoing work is attempting to improve the simplified moist physics operators. More generally, it is worth asking if a standard linear-regression type bias correction is appropriate in the presence of cloud and rain. Such techniques were found to work well for temperature and water vapour assimilation (e.g. Eyre et al., 1993) where the biases came from instrumental effects, or from inadequacies in the clear-sky radiative transfer. In contrast it may be very difficult to correct the patterns of bias produced by a complicated and non-linear parametrisation of cloud and rain using just simple linear correlations. Biases can also be related to the presence or absence of cloud and rain. The problem is that as long as the rain and cloud models remain biased, or until we can devise a bias correction model sufficiently accurate to remove these biases, we will continue to assimilate biases as if they were actual observational information.

- **Validity of the 1D+4D-Var method** Early experiments with the 1D+4D-Var suggested that the assimilation of just TCWV into 4D-Var was able to improve the 4D-Var rain rates (e.g. Marécal and Mahfouf, 2002). Experience with the current operational system suggests this is not the case. Comparing the increments in 1D-Var and 4D-Var, TCWV information is transferred correctly into the 4D-Var analysis, cloud information is partly transferred (and only when it corresponds to the removal of cloud), but rain information is not transferred at all. Hence, we cannot yet claim to have a complete rain assimilation system. Instead we have a system that makes best use of TCWV information in cloudy and rainy areas.

A final important point has not been examined in detail here, but as explained by Bauer et al. (2006c), it will also be necessary to improve the 4D-Var control variable so that rain and cloud are included in some way. The current RH control variable (Hólm et al., 2002) penalises moistening increments in regions close to saturation. In rain and cloud data assimilation, we would actually want this extra moisture to produce more rain and cloud in the analyses.

4.2 Improvements for 1D+4D-Var at ECMWF

This paper has identified a number of areas where the 1D+4D-Var system needs to be improved:

- **Information transfer to 4D-Var**. Because of the choice to assimilate only the TCWV retrieval into 4D-Var, we have seen that much information is lost. If time permits, we can try assimilating vertical profiles of retrieved temperature, humidity, cloud and rain. However, since direct 4D-Var will eliminate this problem, it may not be worth much effort trying to improve 1D+4D-Var.
- **Excess rain in the simplified moist physics operators**. We are working on tuning the simplified moist physics operators to be more like the full forecast model, and closer to observations. The alternatives would be much less desirable: for example, we could try to find an improved, and possibly non-linear bias correction scheme, but initial attempts have not been successful. Using first guess rain as a predictor is not a good idea: there will always be a spurious bias identified at high rain amounts.
- **FG departure threshold**. Because of the possibility of complete mislocation of cloud and rain between first guess and observations, the current first guess departure threshold of 50K is too tight for most channels. The presence or absence of cloud and rain can make up to 70K difference to brightness temperatures, so good observations are being needlessly discarded. However, the threshold could be reduced in channel 22v, where a 20K threshold would be more appropriate.
- **Snow threshold**. Observations are currently discarded where the fraction of snow in the column is too high. The problem is that the rain and cloud amounts may be so small as to be irrelevant, yet a fractional

threshold may still be triggered. We need to move to a threshold that takes into account the absolute snow amount.

- **Minimisation problems**. Roughly 20% of 1D-Var retrievals fail to minimise. Much work is still required to decide what may be causing this. Possible candidates include badly specified error covariances (or similarly, that first guess and observation are too far apart), excessive non-linearity in the forward models, or poor pre-conditioning.

- **Bias correction**. The bias correction scheme does not fully eliminate the bias caused by excessive rain in the simplified moist physics. We have also seen that using first guess rain as a predictor is not a good idea: there will always be a spurious bias identified at high rain amounts. The first priority is to tune the simplified moist physics operators to reduce bias, and that work is already under way. The 1D-Var bias correction scheme will be reconsidered once we have done this.

- **Observation errors**. There seems no obvious need to use larger observation errors as recommended by Deblonde et al. (2007). Their argument is based on the projection of background errors in temperature and moisture into radiance space. In our system this results in a standard deviation of order 20K in channel 19v (Bauer et al., 2006a). In contrast, we use an observation error error of only 3K. This means the 1D-Var retrievals will draw close to the observations. We argue here that these numbers are compatible with observed first guess departure standard deviations in rainy and cloudy areas. The large background errors appear to be representative of the fact that the mislocation of cloud and rain in the first guess can easily result in first guess TB errors of greater than 20K. Finally, another reason to weight observations strongly in the 1D-Var retrieval is the fact that the first guess is used twice in the 1D+4D-Var method.

4.3 Recommendations for 4D-Var

Our aim for the future is to move to direct 4D-Var of SSM/I radiances. Compared to 1D+4D-Var, we expect a much better transfer of rain and cloud information into the 4D-Var analyses. However, this will still be limited by the known problems with the normalised relative humidity control variable, which suppresses moistening increments near saturation (Bauer et al., 2006c). It is hoped that in the future a total moisture control variable may allow better control over cloud and rain. Direct 4D-Var will also benefit from a less biased first guess rain and cloud amount compared to 1D-Var, which must rely on the simplified moist physics operators. However, even the full forecast model produces too much rain, and this will need to be dealt with either by bias correction or by improving the forecast model. However, we have seen from 1D-Var that bias correction is likely to be difficult.

The principal recommendation for direct 4D-Var of SSM/I radiances is that the distinction between 'clear' and 'rainy' observation streams should be removed. The observation operator must take into account rain and cloud whenever it appears in the model. Direct 4D-Var will also use first guess screening criteria similar to 1D-Var, and hence will also benefit from the improvements listed above.

Acknowledgements

Sabatino DiMichele, Chris O'Dell, Hans Hersbach, Deborah Salmond, Graeme Kelly and Sami Saarinen have all been very helpful in this work. Alan Geer was funded by the EUMETSAT fellowship programme.

The TRMM PR data used in this study were acquired as part of the TRMM mission. The algorithms were developed by the TRMM Science Team. The data were processed by the TRMM Science Data and Information

System (TSDIS) and the TRMM Office; they are archived and distributed by the Goddard Earth Sciences Data and Information Services Center (GES DISC). TRMM is an international project jointly sponsored by the Japan National Space Development Agency (NASDA) and the U.S. National Aeronautics and Space Administration (NASA) Office of Earth Sciences.

A Forecast scores

For a general introduction to forecast scores see Wilks (2006). Scores presented in this paper are based on the root mean squared error, $RMSE_j$, of a single forecast, j, valid at one date and time:

$$RMSE_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - a_i)^2}, \tag{2}$$

where f_i is the value of a parameter at one grid point of the forecast being verified and a_i is its equivalent from an analysis valid at that date. For the statistics presented in this paper, forecasts and analyses have been interpolated onto pressure levels and onto a regular 2.5° by 2.5° latitude-longitude grid. Index i runs over all n grid points of one forecast at one pressure level, either for one latitude, or for a range of latitudes, such as 90°S to 20°S. When such statistics span more than one latitude, an area weighting factor is included in the summation, but for clarity this is omitted from the equations presented here.

In order to calculate the RMSE of vector wind, Eq. 2 is replaced by:

$$RMSE_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(u_i^f - u_i^a)^2 + (v_i^f - v_i^a)^2}, \tag{3}$$

where u_i^f and v_i^f are the first guess zonal and meridional wind components and u_i^a and v_i^a are the equivalents from the verifying analysis.

In this paper we examine the time-means of the daily forecast scores:

$$\overline{RMSE} = \frac{1}{m}\sum_{j=1}^{m} RMSE_j, \tag{4}$$

where j might for example run over all m T+24 forecasts in one month. To compare two experiments, A and B, we can calculate the normalised difference in forecast scores between them:

$$\Delta\overline{RMSE} = \frac{\overline{RMSE}_A - \overline{RMSE}_B}{\overline{RMSE}_B}. \tag{5}$$

The construction of these measures from a set of daily values means that it is possible to calculate the statistical significance of the difference in forecast scores between two experiments. The set of daily forecast scores from each experiment can be treated as forming a population. We then apply the standard statistical test for the significance of differences in the mean of two populations (see e.g. Wilks, 2006). Here, we have assumed that daily forecast scores are independent of each other, an assumption which is not valid in atmospheric situations and which will result in inflated confidence estimates. However, a more problematic assumption has also been made: that the variance of daily estimates over a period of a month or two is representative of the true variance of such estimates over all of time. This is of course not the case and practical experience shows that forecast verification scores can change markedly if the tests are done in different years or in different seasons (e.g., Andersson et al., 1998a). Hence, we believe that the forecast verification statistics and their confidence estimates should be treated with appropriate caution.

	a [mm h^{-1} (g m^{-3})$^{-1}$]	b
Rain	20.89	1.15
Snow	29.51	1.10

Table 8: Coefficients for Eq. 6

B Computations of total column rain and snow

This appendix describes how we calculate total column rain and snow amounts from the model rain and snow fields, which are computed as fluxes, in kg m^{-2} s^{-1}.

By assuming exponential size distributions for both rain and snow, constant density with size, and a certain distribution of fall speeds, the following relation between precipitation rate, PR, in mm h^{-1} and water content, WC, in g m^{-3}, has been derived:

$$PR = a \times WC^b. \tag{6}$$

Coefficients a and b are given in Table 8. By using different coefficients, the relation can be applied either to link rain rate and liquid water content, or snow rate and frozen water content. In our calculations, the conversion of snow flux to a rate in mm h^{-1} assumes the density of snow to be 100 kg m^{-3} and the density of water to be 1000 kg m^{-3}.

Total column rain water or snow water C is calculated from the integral of rain or snow density ρ_P with altitude, z, from the surface to the top of the atmosphere (TOA), in SI units:

$$C = \int_{surface}^{TOA} \rho_P dz \tag{7}$$

For our purposes, this equation is discretized on model levels and z is calculated from the model's pressure and temperature fields using the hydrostatic relation.

References

Andersson, E., P. Bauer, A. Beljaars, F. Chevallier, E. Hólm, M. Janisková, P. Kållberg, G. Kelly, P. Lopez, A. Mcnally, E. Moreau, A. J. Simmons, J.-N. Thépaut, and A. M. Tompkins (2005). Assimilation and modeling of the atmospheric hydrological cycle in the ECMWF forecasting system. *Bull. Am. Met. Soc. 86*, 387–402.

Andersson, E., J. Haseler, P. Undén, P. Courtier, G. Kelly, D. Vasiljevič, C. Brankovi, C. Cardinali, C. Gaffard, A. Hollingsworth, C. Jakob, P. Janssen, E. Klinker, A. Lanzinger, M. Miller, F. Rabier, A. Simmons, B. Strauss, J.-N. Thépaut, and P. Viterbo (1998a). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III: Experimental results. *Quart. J. Roy. Meteorol. Soc. 124*, 1831–1860.

Andersson, E. and H. Järvinen (1998b). Variational quality control. *ECMWF Tech. Memo., 250, available from http://www.ecmwf.int*.

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

Bauer, P., P. Lopez, D. Salmond, A. Benedetti, S. Saarinen, and E. Moreau (2006b). Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var. *Quart. J. Roy. Meteorol. Soc. 132*, 2307–2332.

Bauer, P., P. Lopez, D. Salmond, and A. J. Geer (2006c). Assimilation of cloud and precipitation affected microwave radiances. *ECMWF Tech. Memo., 502, available from http://www.ecmwf.int*.

Bauer, P., E. Moreau, F. Chevallier, and U. O'Keeffe (2006d). Multiple-scattering microwave radiative transfer for data assimilation applications. *Quart. J. Roy. Meteorol. Soc. 132*, 1259–1281.

Benedetti, A. and M. Janisková (2007). Assimilation of MODIS cloud optical depths in the ECMWF model. *ECMWF Tech. Memo., 515, available from http://www.ecmwf.int*.

Benedetti, A., P. Lopez, P. Bauer, and E. Moreau (2005). Experimental use of TRMM precipitation radar observations in 1D+4D-Var assimilation. *Quart. J. Roy. Meteorol. Soc. 131*, 2473–2495.

Bouttier, F. and G. Kelly (2001). Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Quart. J. Roy. Meteorol. Soc. 127*, 1469–1488.

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

Chevallier, F. and G. Kelly (2002). Model clouds as seen from space: Comparison with geostationary imagery in the 11-μm window channel. *Mon. Weath. Rev. 130*, 712–722.

Deblonde, G., J.-F. Mahfouf, B. Bilodeau, and D. Anselmo (2007). One-dimensional variational data assimilation of SSM/I observations in rainy atmospheres at MSC. *Mon. Weath. Rev. 135*, 152–172.

English, S. J., R. J. Renshaw, P. C. Dibben, A. J. Smith, P. J. Rayer, C. Poulsen, F. W. Saunders, and J. R. Eyre (2000). A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Quart. J. Roy. Meteorol. Soc. 126*, 2911–2931.

Errico, R. M., P. Bauer, and J.-F. Mahfouf (2007). Issues regarding the 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. *Quart. J. Roy. Meteorol. Soc. 119*, 1427–1463.

Gérard, E. and R. Saunders (1999). Four-dimensional variational assimilation of Special Sensor Microwave / Imager total column water vapour in the ECMWF model. *Quart. J. Roy. Meteorol. Soc. 125*, 3077–3102.

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

Haseler, J. (2004). Early-delivery suite. *ECMWF Tech. Memo., 454, available from http://www.ecmwf.int*.

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 A 38*, 111–136.

Hólm, E., E. Andersson, A. Beljaars, P. Lopez, J.-F. Mahfouf, A. Simmons, and J.-N. Thepaut (2002). Assimilation and modelling of the hydrological cycle: ECMWF's status and plans. *ECMWF Tech. Memo., 383, available from http://www.ecmwf.int*.

Hou, A. Y., S. Q. Zhang, and O. Reale (2004). Variational continuous assimilation of TMI and SSM/I rain rates: Impact on GEOS-3 hurricane analyses and forecasts. *Mon. Weath. Rev. 132*, 2094–2109.

Iguchi, T., T. Kozu, R. Meneghini, J. Akawa, and K. Okamoto (2000). Rain-profiling algorithm for the TRMM precipitation radar. *J. App. Met. 39*, 2038–2052.

Ikai, J. and K. Nakamura (2003). Comparison of rain rates over the ocean derived from TRMM Microwave Imager and Precipitation Radar. *J. Atmos. Ocean Tech. 20*, 1709–1726.

JAXA/NASA (2005). Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar algorithm: Instruction manual for version 6. *TRMM project document, 180pp., issued by TRMM Precipitation Radar team, Japan Aerospace Exploration Agency (JAXA) and National Aeronautics and Space Administration (NASA), available from http://www.eorc.jaxa.jp/TRMM*.

Kelly, G. A., P. Bauer, A. J. Geer, P. Lopez, and J.-N. Thépaut (2007). Impact of SSM/I observations related to moisture, clouds and precipitation on global NWP forecast skill. *Mon. Weath. Rev., submitted*.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson (1998). The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Ocean. Tech. 15*, 809–817.

Kummerow, C., J. Simpson, O. Thiele, W. Barnes, A. Chang, E. Stocker, R. Adler, A. Hou, R. Kakar, F. Wentz, P. Ashcroft, T. Kozu, Y. Hong, K. Okamoto, T. Iguchi, H. Kuroiwa, E. Im, Z. Haddad, G. Huffman, B. Ferrier, W. S. Olson, E. Zipser, E. A. Smith, T. T. Wilheit, G. North, T. Krishnamurti, and K. Nakamura (1998). The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. App. Met. 39*, 1965–1982.

Lopez, P. and P. Bauer (2007). "1D+4D-Var" assimilation of NCEP stage IV radar and gauge hourly precipitation data at ecmwf. *Mon. Weath. Rev. 135*, 2506–2524.

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

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

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

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

Marécal, V. and J.-F. Mahfouf (2004). Experiments on 4D-Var assimilation of rainfall data using an incremental formulation. *Quart. J. Roy. Meteorol. Soc. 129*, 3137–3160.

Masunaga, H., T. Iguchi, R. Oki, and M. Kachi (2002). Comparison of rainfall products derived from TRMM Microwave Imager and Precipitation Radar. *J. App. Met. 41*, 849–862.

Moreau, E., P. Lopez, P. Bauer, A. M. Tompkins, M. Janisková, and F. Chevallier (2004). Variational retrieval of temperature and humidity profiles using rain rates versus microwave brightness temperatures. *Quart. J. Roy. Meteorol. 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. *Quart. J. Roy. Meteorol. Soc. 122*, 327–355.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons (2000). The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteorol. Soc. 126*, 1148–1170.

Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: Theory and Practice*. World Scientific.

Shin, K.-S. and G. North (1988). Sampling error study from rainfall estimate by a satellite using a stochastic model. *J. App. Met. 27*, 1218–1231.

Spencer, R., H. Goodman, and R. Hood (1989). Precipitation retrieval over land and ocean with the SSM/I: Identification and characteristics of the signal. *J. Atmos. Ocean Tech. 6*, 254–273.

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

Tsuyuki, T. (1997). Variational data assimilation in the tropics using precipitation data. Part III: Assimilation of SSM/I precipitation rates. *Mon. Weath. Rev. 125*, 1447–1464.

Uppala, S. M., P. W. Kållberg, A. J. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. V. D. Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. McNally, J.-F., Mahfouf, J.-J., Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevič, P. Viterbo, and J. Woollen (2005). The ERA-40 re-analysis. *Quart. J. Roy. Meteorol. Soc. 131*, 2961–3012.

Wilks, D. (2006). *Statistical methods in the atmospheric sciences* (2 ed.). Academic Press.

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