Towards the use of satellite data for cloud and rain assimilation

by

J.-F. Mahfouf, V. Marécal , M. Janiskova, F. Chevallier and J. Barkmeijer

European Centre for Medium-Range Weather Forecasts

1. Importance of cloud and rain observations for data assimilation

In principle, data assimilation systems should include all available observations of the atmosphere to provide the best initial state for numerical weather prediction (NWP) models. However, current operational data assimilation systems mostly document the atmosphere in terms of wind, temperature and surface pressure. Far fewer observations are used regarding atmospheric water vapour (mostly from the surface network and from radiosondes) and cloudy/rainy areas are poorly sampled since satellite data are systematically rejected in such conditions. As an example only 25% of satellite radiances from TOVS/ATOVS are actually used in the ECMWF data assimilation system. There are several reasons to explain this. First, there is no obvious relation between the initial conditions of an NWP model and cloud/rain products. This link is in fact complex through the description of sub-grid scale processes that are not yet adequately modelled. Cloudy and rainy systems are characterized by high spatial and temporal variabilities. This raises issues about the ability of global models with rather coarse resolutions to describe these variabilities and about their predictability (for medium range forecasts). On the other hand, it seems that more observations are required to improve the numerical prediction of rapidly developing mesoscale systems (e.g. flash floods) and of tropical cyclones. The density of the synoptic network is not adequate to provide all the ingredients in the initial conditions to accurately forecast these severe storms. The tropical circulation being strongly coupled to the diabatic heating from cumulus convection and the sparsity of the conventional network in these regions make cloud and rain products from satellites very attractive.

2. Satellite products

A large number of satellite data on clouds and precipitation is currently available. When measuring radiances at the top-of-the-atmosphere, satellites do not distinguish between cloudy and clear-sky situations. Therefore radiances from atmospheric sounders and imagers on operational polar orbiting and geostationary satellites in the infra-red and microwave regions of the electromagnetic spectrum contain information about cloud and precipitation. Geophysical products on clouds and rain can then be derived from radiances using retrieval algorithms. Rainfall rates can be obtained from microwave imagers such as SSM/I or TMI and from VIS/IR imagers from geostationary satellites (see Simmer (1996) for a review). Cloud properties such as cloud cover, cloud top pressure, cloud optical thickness or cloud liquid water content can also be derived using a-priori information (ISCCP products are a well known example described in Rossow (1996)). New satellite products related to clouds and precipitation will be available in the near future (less than 10 years). High resolution radiances from infra-red spectrometer (AIRS) and interferometer (IASI) will allow temperature soundings of the troposphere/ stratosphere with improved vertical resolution and accuracy. They will also provide more accurate radiances in cloudy situations that could be exploited. Microwave imagers such as AMSR on board ADOES-II and EOS-Aqua having additional low frequency channels with respect to SSM/I (at 10 GHz) will allow better rainfall retrievals. The TRMM follow-on mission (GPM) expected in 2007 will focus on increasing the temporal sampling of microwave rain rate estimates using a constellation of eight satellites. As initiated by TRMM, several explorer missions will be launched with active sensors (radar, lidar) on board. They should provide less ambiguous information on hydrometeors (cloud and ice water, rain water) than passive retrievals. The vertical structure will also be sampled more precisely than with current satellite sounders. The drawback of active instruments is to have a very narrow field of view (less than 1 km) that can lead to representativeness problems. Therefore they need to be associated with imagers over a broader swath (few tens of km).

When examining theoretical extinction coefficients in Figure 1 over the infra-red and microwave spectrum, some considerations can be given on the adequacy of these two spectral regions for cloud and rain retrievals. In the infra-red region (4 to 100 µm), clouds have a much higher absorption than rain. Therefore rain is hidden by clouds and for thick clouds the signal mostly comes from their top. Any rain rate retrieval in the infra-red region is based on an indirect information on cloud top temperature (such as the GPI from Arkin and Meiner (1987)). When increasing wavelength (between 1 and 50 mm), cloud absorption becomes much weaker than rain absorption. Therefore rain can be directly detected in this spectral region. Clouds being less opaque in the microwave region they can be sensed over much deeper layers than in the infra-red. The dominance of the rain signal means that cloud information can be contaminated when a small amount of rain droplets is present. Careful checks are thus required for cloud retrievals. From this simple examination, it can be concluded that retrievals are easier for clouds in the infra-red region and easier for rain in the microwave region. This contrasted behaviour of clouds and rain in these two spectral regions should also be exploited for satellite retrievals.

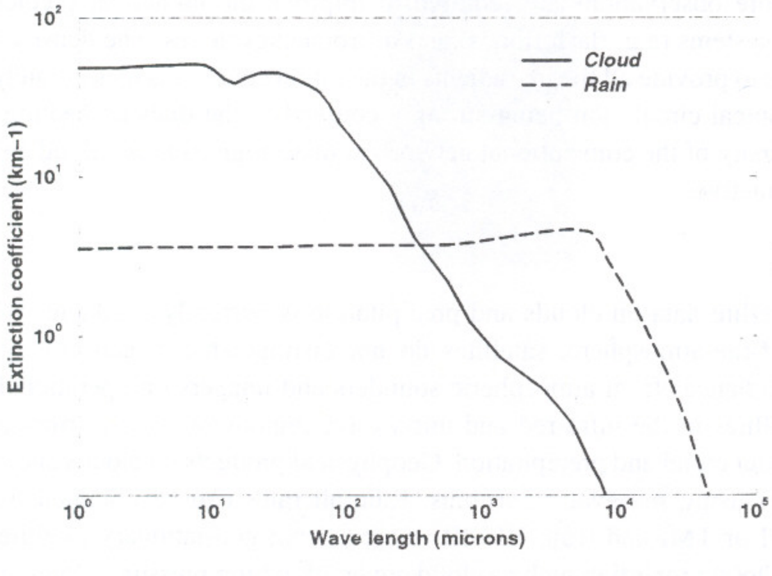


Figure 1: Theoretical cloud and rain extinction coefficients for a nimbostratus cloud (mode radius of 6 mm) and a raining cloud having a 50 mm/hour rainfall rate (mode radius of 6 mm) based on data from Deirmendjian (1975

3. Strategies for cloud and rain assimilation

One important requirement for the assimilation of cloud and precipitation in a NWP model is to have an accurate forward modelling of the observation counterparts. The difficulty of such forward modelling explains why cloud and precipitation products are not yet extensively used in data assimilation. This requirement is likely to be fulfilled in the near future due to on-going developments in NWP research. More realistic cloud and convection schemes are being developed. For example, prognostic cloud schemes provide a more physical description of cloud production and dissipation through an explicit link with other physical processes and the dynamics (Tiedtke, 1993). The horizontal resolution of global NWP models is increasing (the current ECMWF model has an equivalent resolution of about 40 km). The scales described by NWP models are thus becoming more comparable to the pixel size of radiometers. Data assimilation has also made important progress during recent years. The variational approach (3D/4D-Var) allows the use of observations that are non-linearly linked to model variables. Clear-sky radiances can then be assimilated directly without having to produce first temperature and humidity retrievals (Andersson *et al.*, 1994). New types of observations can be introduced in variational data assimilation as long as an observation operator (to produce the model equivalent of observations) is available and observation errors are known. Finally, as stated in the previous paragraph, new satellites missions will provide more accurate measurements of clouds and precipitation (additional frequencies, synergy between active and passive instruments).

The most straightforward strategy for the assimilation would be to use satellite radiances in cloudy and rainy areas (particularly in the variational context). There the quantification of observation errors is easy (mostly instrumental and representativeness) and a preprocessing of the data before entering the analysis system can be applied (quality control, bias corrections). On the other hand, the forward modelling is difficult. In addition to temperature and water vapour, vertical profiles of hydrometeors (liquid and ice water contents for rain and cloud particles) need to be specified to solve the radiative transfer equation. Radiative properties depending upon the shape, the size and the density of particles must also be known (and they can be very complex for the ice phase), together with overlap assumptions of cloud cover on the vertical.

The alternative to satellite radiances is the assimilation of satellite derived geophysical quantities (surface rain rate, cloud liquid water content,...). For derived products the quantification of associated errors is generally difficult since retrievals involve various steps where prior information is introduced. However, the observation operator is simpler since it only involves moist physical processes (cloud and precipitation parameterizations) instead of a radiative transfer modelling.

Given the advantages and difficulties associated with each approach, it is possible to favour a direct assimilation of cloudy radiances and an assimilation of satellite derived rain rates. Indeed, in the microwave and infra-red regions scattering can be neglected for cloud particles which makes the radiative transfer equation easier to solve. Information on cloud fraction, liquid and ice water contents is available in most NWP models. Radiances in rainy conditions are more difficult to model. The radiative transfer equation should include scattering processes which are expensive and difficult to solve. Indeed, there are many uncertainties associated to the description of the microphysics of ice (and melting) particles. Most global NWP models do not have prognostic equations for rain water (liquid and ice). Model time steps are large enough to assume that any local rain production in the atmosphere falls instantaneously to the ground. Using satellite rain rate retrievals allows an easier comparison between model and observations but it does not mean that the above difficulties are solved. These uncertainties are present in the retrieval process and should be reflected in the associated observation errors.

4. Assimilation techniques

During the past twenty years, there has been a large number of feasibility studies performed on the assimilation of rainfall rates both at mesoscale and global scale (see Falkovich et al. (2000) for a recent study referencing previous works). Most of these studies were based on empirical methods (nudging or physical initialization). Empirical methods are generally easy to implement but they implicitly assume "perfect" observations and therefore tend to nudge too strongly the initial model state to fit the data (at the expense of other observations and model balances). More recently, variational assimilation of clouds and precipitation has been developed both in one-dimensional and four-dimensional frameworks. Variational retrievals are consistent with statistics of observations and background (a-priori) errors and with other data included in the assimilation. The drawback of this method that it is more expensive than nudging techniques, involves important technical developments (linearized versions of observation operators) and relies on the validity of the tangent-linear approximation (which may be a problem when strongly non-linear physical processes are considered).

4.1 Nudging assimilation of rainfall data

Physical initialization was developed by Krishnamurti *et al.* (1984, 1993) for improving the initial state of a global NWP model for tropical forecasts. The procedure can be divided in three steps. First an accumulated precipitation analysis Ra is performed at the resolution of the model using various sources of data (raingauges, infra-red and microwave retrievals). From this analysis rainfall increments are constructed ΔR=Rf-Ra, where Rf is the accumulated precipitation from the forecast model. Then, an inversion procedure is defined in order to convert rainfall increments ΔR into modifications of temperature and humidity profiles (ΔT and Δq). Krishnamurti *et al.* (1984) have shown that it is possible to invert a Kuo type convection scheme when assuming that the moisture convergence is entirely converted into rain and that corrections to the humidity profile conserve the total column water vapour. Other inversions have been defined such as a rescaling of the heating profile by convection according to the ratio between observed and simulated rainfall (Puri and Miller, 1990; Manobianco *et al.*, 1994). The last step is to link the thermodynamical profile changes to modifications of the dynamics. Krishnamurti *et al.* (1984) integrate their model over a period between 24 to 48 hours prior the initial forecast time, where dynamic fields are nudged toward an atmospheric analysis and where humidity fields are modified by the inversion of the convection scheme. The divergent wind is significantly changed by imposing a smaller nudging coefficient. Diabatic normal mode initialisation has also been used by various authors (Heckley *et al.*, 1990; Kasahara *et al.*, 1996). They impose a modified diabatic heating profile (from observed precipitation) in the model equations to produce a dynamically balanced initial state.

4.2 Variational assimilation of cloud and rain observations

In variational data assimilation the initial state x of a NWP model is defined by the minimum of an objective function measuring the distance of a model trajectory to a background state x^b (short-range forecast) and to all available observations y over an assimilation window (typically 6 hours). The chosen scalar product is an Euclidian distance weighted by the inverse of the covariance matrices of background and observation errors **B** and **R**:

$$J(x) = 1/2(x - x^b)^T\mathbf{B}^{-1}(x - x^b) + 1/2\,(H(x) - y)^T\mathbf{R}^{-1}(H(x) - y)$$

For a given observation y, the observation operator $H(x)$, includes the time propagation of the initial model state to the observation time and horizontal/vertical interpolations of the control vector x at the observation location. If the observation is a satellite radiance, the observation operator also includes a radiative transfer model to convert temperature and humidity profiles. For an assimilation of surface rainfall rates, the observation operator includes the moist physical processes generating precipitation: deep cumulus convection and large-scale condensation.

The minimization of J requires its gradient to be known at each iteration of a numerical algorithm (e.g. conjugate gradient, Quasi-Newton):

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

The operator $\mathbf{H}^T$ is the transpose of the tangent-linear observation operator. For a variational assimilation of precipitation it corresponds to the linearized physical parameterizations. This operator can be obtained by computing explicitly the elements of its Jacobian matrix in finite differences. This technique is only suitable when x is a low dimension vector (as in a 1D-Var context). For a realistic 4D-Var assimilation this operator has to be obtained by the "adjoint technique" as explained in Le Dimet and Talagrand (1986). The adjoint technique is much cheaper than explicit Jacobians but requires an explicit coding of the adjoint model.

5. Usefulness of linearized physics

As explained previously, a variational assimilation of cloud and precipitation requires the linearization of physical processes to solve the minimization problem. Physical processes are characterized by strong non-linearities and thresholds than can make the tangent-linear approximation very poor. Therefore, it is worth asking about the usefulness of linearized physical processes in a tangent-linear NWP model. Two aspects of this question are illustrated here: the impact of physical processes on the tangent-linear approximation and the capacity of linearized physics to identify unstable modes not described by an adiabatic model (they are important to quantify forecast errors for various applications).

5.1 Tangent-linear approximation

The evolution of analysis increments over a period of 24 hours produced by a tangent-linear model (TLM) including physical processes is compared with pairs of non-linear model (NLM) integrations with physics (and considered in this study as the truth). The model is a low resolution version of the ECMWF forecast model (Δx=200 km with 31 vertical levels).

The evolution of zonal wind increments near the surface is illustrated in Figure 2 for three versions of the TLM: an adiabatic version [a], a version with a simplified vertical diffusion scheme from Buizza (1994) [b], a version with a comprehensive package of physical processes described in Mahfouf (1999) [c]. The lack of surface friction in the adiabatic TLM leads to evolved increments which are much too large with respect to the NLM. The inclusion of a simple diffusion scheme reduces significantly the size of the increments, but they are still overestimated. The inclusion of a stability-dependent diffusion scheme in the TLM produces an evolution of increments in better agreement with the NLM evolution. It is important to indicate that Mahfouf (1999) has modified the linearized vertical diffusion scheme to prevent excessively large derivatives of the exchange coefficients around neutral stability.

The role of physical processes is further examined for global mean absolute errors associated with specific humidity. The inclusion of moist physical processes reduces the mean absolute error of the TLM with respect to the NLM by 14%. Improvements are significant in mid-latitudes disturbances (large scale condensation) as well as in the tropics (deep moist convection). The complete physical package in the TLM improves by 26% the time evolution of specific humidity increments compared to an adiabatic TLM. The effect of the dry physics is dominant in the planetary boundary layer through the vertical diffusion parametrization scheme.

5.2 Tropical singular vectors

The ECMWF linearized physical package has been used to compute the most unstable growing perturbations over a two-day period (known as singular vectors). Singular vectors are used to define the initial conditions of the Ensemble Prediction System (EPS) at ECMWF (Molteni *et al.*, 1996). They are computed operationally over the extra-tropical hemispheres and are usually associated with the development of baroclinic instabilities. The inclusion of linearized physical processes for extra-tropical singular vectors does not produce radical changes of their structure. The impact in the tropics is more significant since latent heat release by convection is an essential ingredient for the maintenance of tropical disturbances. Tropical singular vectors (TSVs) targeted over the Tropical Pacific Ocean and estimated using an almost adiabatic linearized model are compared with TSVs obtained from a linearized model with more physical processes including deep cumulus convection (Figure 3). Wind

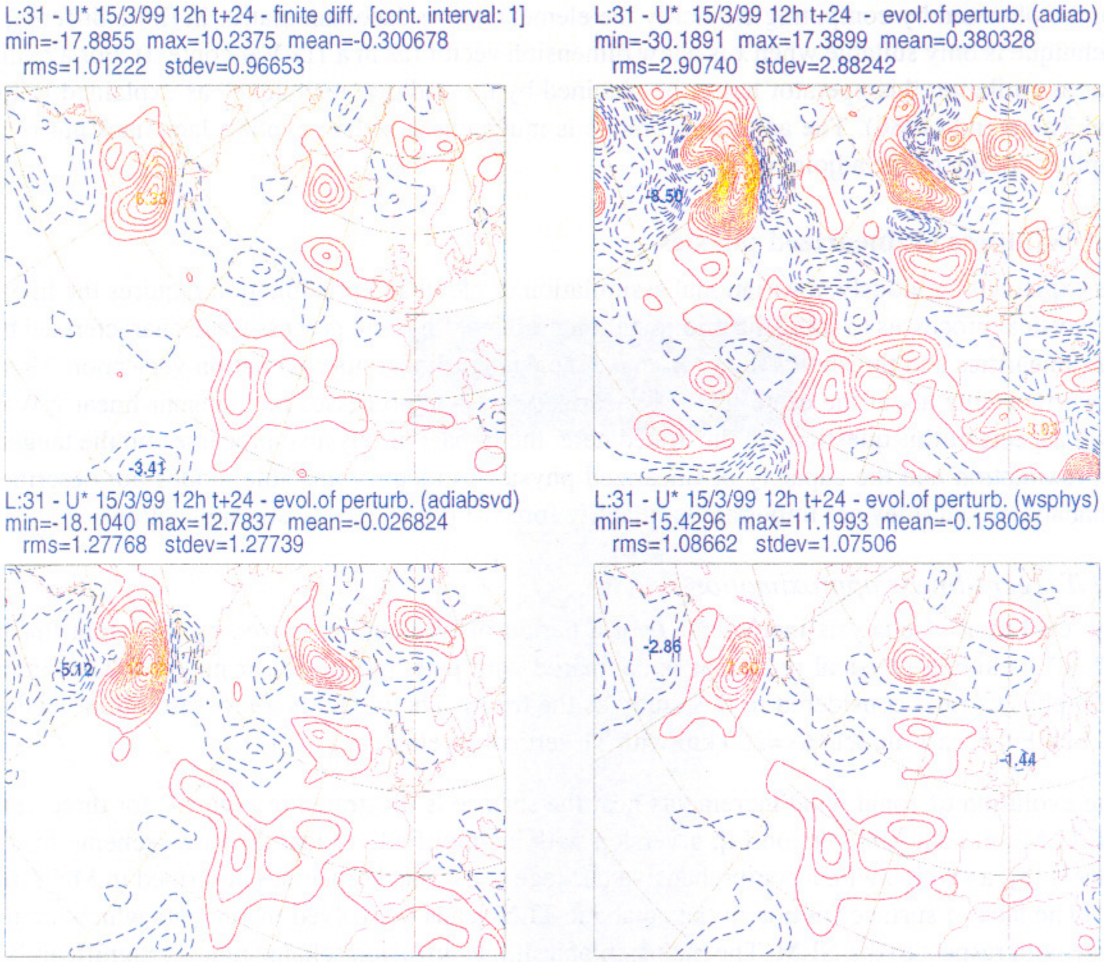


Figure 2 : Evolution of analysis increments for the zonal wind at the lowest model level (~ 30 metres after a 24-hour evolution obtained from the difference between two non linear integrations with full physics (upper left panel) and from an adiabatic tangent-linear model (upper right panel), a tangent-linear model with a simplified vertical diffusion scheme (lower left panel) and a tangent-linear model with a comprehensive package of physical processes (lower right panel).

perturbations at 500 hPa are more intense when the linearized physics is included and the stream function presents a more defined dipole structure on each side of the centre of the cyclone. The next step is to examine if such differences in the initial conditions reflect in terms of spread of the EPS. The various forecast trajectories of the cyclone are shown in Figure 4 together with the observed trajectory, the ensemble mean trajectory and the trajectory of the deterministic model. There are clearly more members in the ensemble that span the actual observed trajectory using "diabatic" TSVs. More results can be found in Puri *et al.* (2001). These conclusions are in agreement with simple linearized models of tropical cyclone developments where diabatic heating has to specified.

6. One-dimensional variational assimilation of rainfall rates

The variational assimilation of rainfall rates has been first studied with a one-dimensional column model by Fillion and Errico (1997). Given an observed surface precipitation rate R_0 with an associated error σ_0 and an instantaneous rainfall rate $R(x)$ produced by applying model atmospheric profiles of temperature and specific humidity to moist physical parameterizations, an optimal profile x minimizes the following cost-function:

$$J(x) = 1/2(x - x^b)^{\mathrm{T}}\mathbf{B}^{-1}(x - x^b) + 1/2\ [(R(x) - R_0)/\sigma_0]^2$$

Experiments have been performed by Marécal and Mahfouf (2000a) using the moist physical processes of the ECMWF operational model for moist convection (Tiedtke, 1989) and with a simple moist adjustment scheme to generate large-scale precipitation from supersaturations. Initial profiles are taken from 6-hour forecasts and tendencies from the dynamics and other physical processes are added to update the profiles (but they are not control variables as such). Background statistics defined by the matrix **B** are taken from the ECMWF operational system (the unbalanced temperature is used which is a reasonable assumption in tropical areas and implies no correlations with the wind field). The observation error is assumed to be 25% of the surface rainfall rate with a lower bound at 0.01 mm/hour (to reflect the uncertainty in rain occurrence by retrieval algorithms). Rainfall rates are obtained from the NASA retrieval algorithm applied to the microwave imager TMI on board TRMM (2A12 product). This product being provided at a much higher resolution (6 km) than the ECMWF grid size (60 km), observed rain rates are averaged at the model resolution using a simple binning technique. Figure 5ab shows the behaviour of the minimization for a particular point. The convergence is very fast (about 5 iterations). After 16 iterations, the norm of the gradient has decreased by five orders of magnitude. In order to understand the vertical structure of the increments, the rainfall Jacobians **R** and the standard deviation of background errors for temperature and humidity are depicted in Figure 5cd. A high sensitivity of the convection scheme is found near the surface where the cloud parcel ascent starts. There is a significant sensitivity from cloud base up to the melting level and Jacobians are much smaller above

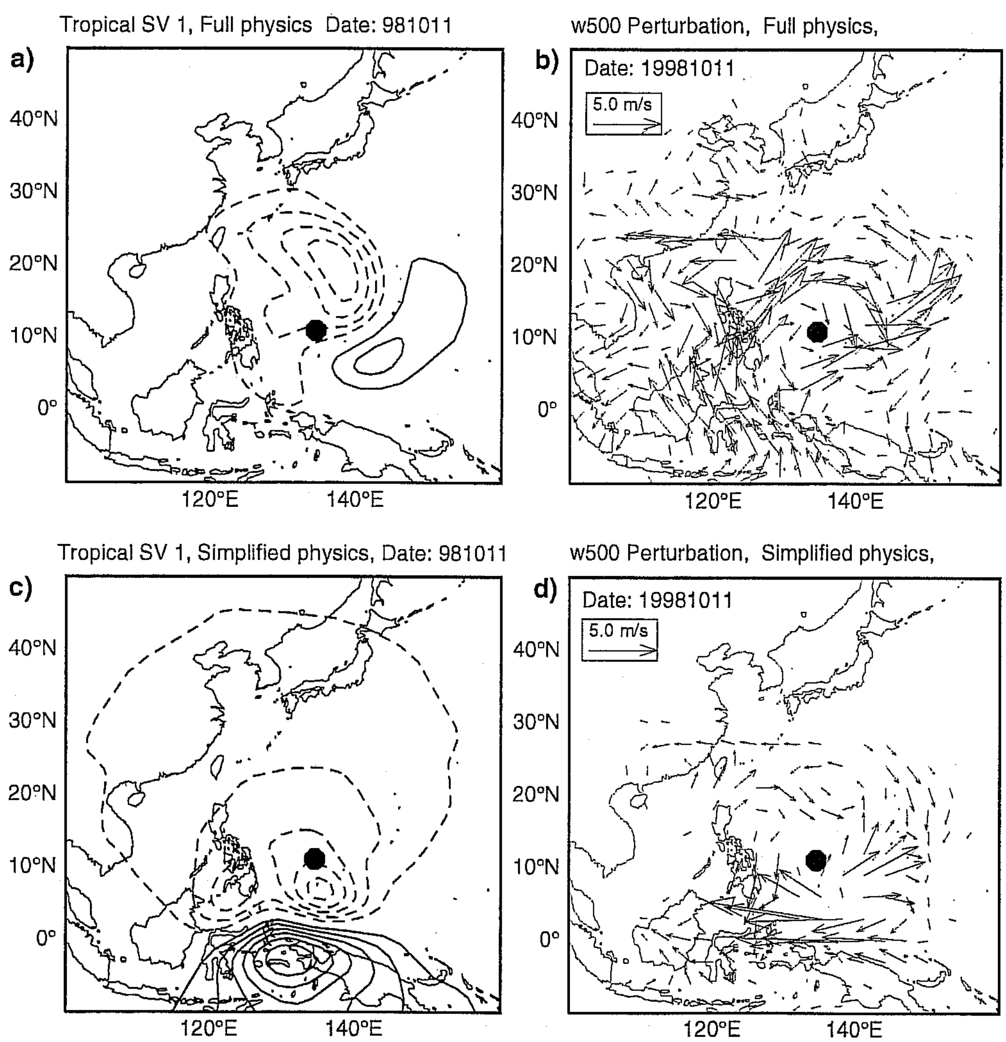


Figure 3: Stream function (in m2/s2) and wind vector perturbations (in m/s) at 500 hPa for the first singular vector based on full linearized physics (top panels) and simplified physics (bottom panel). The cyclone location is indicated by a large dot.

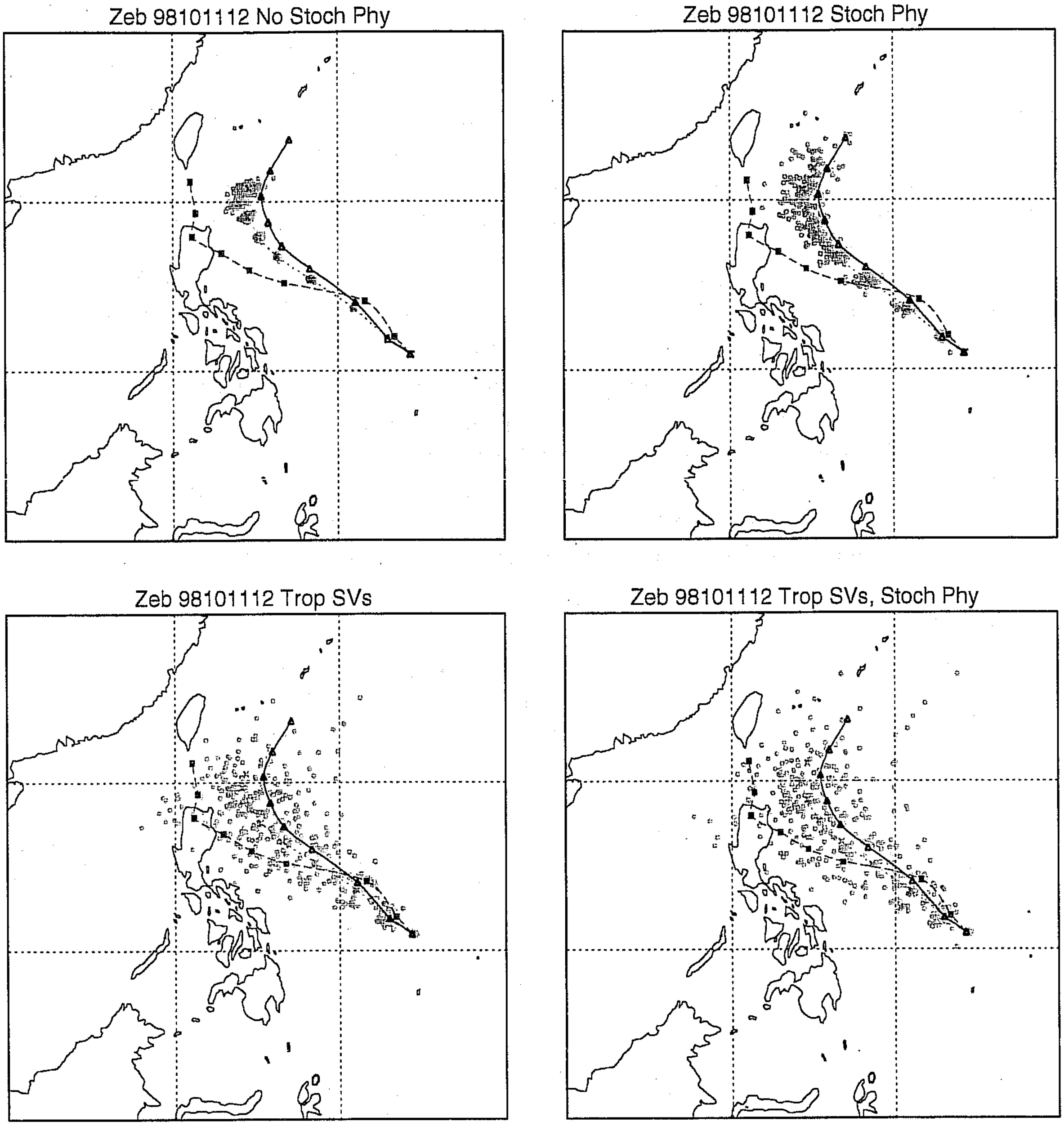


Figure 4 :Tracks for tropical cyclone Zeb based on operational analyses (squares), operational forecasts (triangles), ensemble means (grey line) and individual ensemble members (dots). The tracks are over four days, positions are plotted every 12 hours, and are for EPS runs with no stochastic physics (top left), with stochastic physics (top right), tropical SVs (bottom left) and tropical SVs+stochastic physics (bottom right).

that level. Background errors are fairly uniform on the vertical for temperature, whereas for humidity there is a peak of variance around 850 hPa (level 50) with an exponential decrease aloft and smaller values in the boundary layer. These values do not correspond specifically to rainy systems but are assumed to represent averaged forecast errors. Flow dependent forecast errors would require the development of a Kalman filter (even though 4D-Var is able to implicitly built dynamic structure functions).

Temperature increments shown in Figure 5e have a maximum (positive) value at the melting level associated with the large (negative) value of the rainfall Jacobian (since 1D-Var has to decrease model rainfall to better fit the observation). Increments have a rather broad vertical structure produced by the cross-correlations in the matrix **B**. Humidity increments are maximum around level 50 where the

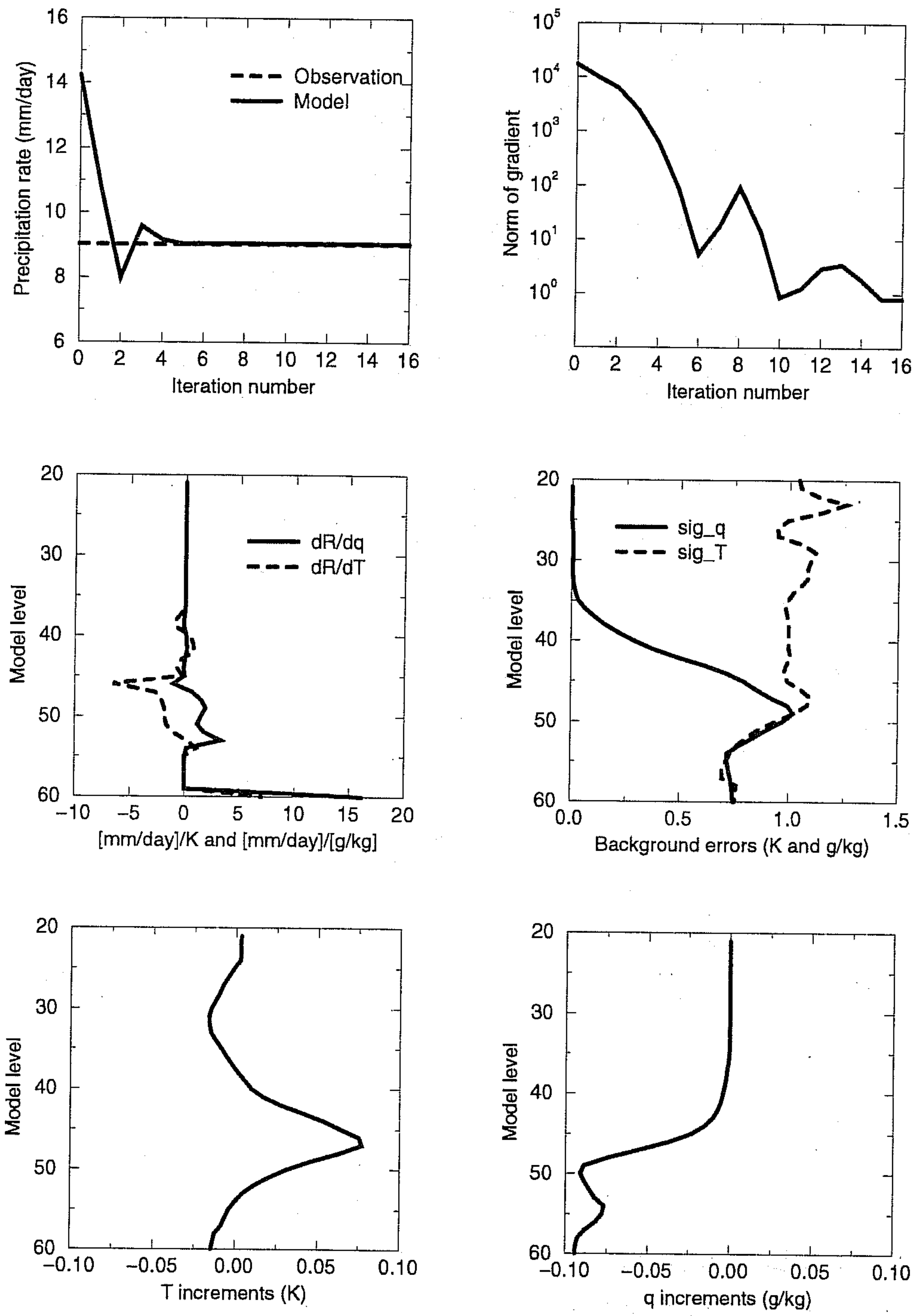


Figure 5: 1D-Var assimilation of rainfall rates. The following quantities are plotted: precipitation rate at each iteration of the minimization with the observed value (upper left), norm of the gradient of the cost-function (upper right), Jacobian vector of the sensitivity of surface rainfall rate to modifications in temperature and specific humidity (middle left), standard deviation of forecast errors for temperature and specific humidity (middle right), analysis increments for temperature (lower left), analysis increments for specific humidity (lower right).

background error is the largest (Figure 5f). Large increments are also found in the boundary layer where, despite smaller background errors, the rainfall Jacobian **R** with respect to humidity is maximum.

Despite rather small increments (with respect to forecast errors) the model fits the observation very closely at the end of the minimization. This is a consequence of the highly non-linear behaviour of *R*. The high sensitivity of the convection scheme leads to large Jacobians: a small change in temperature and humidity modifies significantly surface precipitation produced by the convection scheme. To examine more globally the performance of the 1D-Var, Table 1 summarizes results obtained for two 6-hour periods. There are about 2000 points where the minimization is successful. Around the same number of points are rejected a-priori because rain is present in the observations and moist processes do not produce any precipitation at the surface. For such profiles, the 1D-Var is unable to trigger precipitation. Given the tropical coverage of TRMM, the number of profiles where precipitation includes a stratiform component is small (about 10%). Finally a small number of profiles are rejected after quality control checks (too large first guess departure, minimization failure). In terms of bias and standard deviations for the first-guess and analyzed precipitation, Table 2 reveals that the standard deviation is reduced by a factor two from first-guess departures to analyses departures. The bias remains small both for the background and for the analysis.

Table 1: Results of 1D-Var analysis for two 6-hour periods

Date	Number of analysed profiles	Number of profiles where Rb=0 and R0>0	Number of profiles with stratiform rainfall	Number of profiles rejected
11 February 1998 at 00 UTC	2685	2680	383	80
26 August 1998 at 12 UTC	2667	2196	151	10

Table 2: Statistics of 1D-Var results for two 6-hour periods (expressed in mm/day)

Date	Background departure (O-B) Mean	Background departure (O-B) Std	Analysis departure (O-A) Mean	Analysis departure (O-A) Std
11 February 1998 at 00 UTC	-0.12	2.01	0.11	0.80
11 February 1998 at 00 UTC	-0.22	2.18	0.09	1.09

7. Four-dimensional variational assimilation of rainfall rates

The 1D-Var assimilation has been coupled to the ECMWF 4D-Var system as shown in Figure 6. The adjusted humidity profiles from the 1D-Var are vertically integrated to produce "pseudo-observations" of total column water vapour (TCWV) for assimilation in the 4D-Var system. This approach has been chosen for various reasons. First, as shown in detail by Marécal and Mahfouf (2000a) the 1D-Var mostly modifies humidity profiles to provide surface rainfall rates closer to observations. Therefore by discarding temperature profiles, not too much information from the retrieval is lost. From a technical point of view, as TCWV is already assimilated from SSM/I retrievals, this type of observation is easier to introduce in 4D-Var than an observed surface rainfall rate. An observation error is assigned to each TCWV retrieval using a fit of the analysis error produced by the 1D-Var (see Marécal and Mahfouf, 2000b for details). The observations are introduced at the model grid resolution (60 km) and are assumed to be uncorrelated. This may not be entirely exact but it is likely that in rainy systems the horizontal correlation between observations is smaller than in clear sky areas. However, given the low

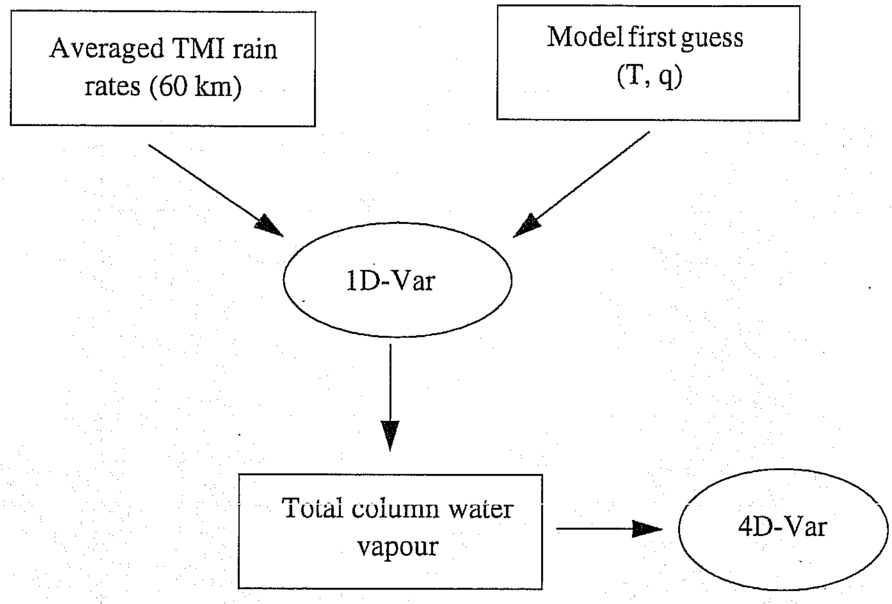


Figure 6: Methodology for a four dimensional variational assimilation of total column water vapour in rainy areas

resolution (200 km) of the 4D-Var minimization it is likely that some redundancy exists for these observations. The use of water vapour information from the 1D-Var is complementary to the one given by SSM/I TCWV retrievals since the first type of observations samples rainy areas and the second type samples clear sky areas. Several two-week 4D-Var assimilation periods have been run using TCWV retrieval from the 1D-Var using TMI derived rain rates. The number of observations assimilated per analysis cycle (i.e. every 6 hour) is about 2000 (which is roughly the same number as for SSM/I TCWV retrievals). However, the background departures are much smaller in rainy areas (due to the large sensitivity of the response of convection) than in clear-sky situations: 0.5 kg/m^2 for the standard deviation compared to 3.5 kg/m^2. The global impact on the humidity analysis is shown in Figure 7 where the averaged RMS (root mean square) values of TCWV increments are compared for the experiment and for a control. There is a systematic reduction of the increments in the tropical belt which indicates that the model first-guess is closer to observations and that smaller corrections are required to the background state at each analysis cycle when assimilating TCWV in rainy areas. For a period including the development of the hurricane Bonnie (August 1998), Figure 8a shows that the analyzed trajectory of the cyclone is closer to the observed "best track" when TMI rain rates are (indirectly) assimilated in 4D-Var. This effect is more pronounced at the early stages of the development of the cyclone. When the cyclone has reached a mature stage and the trajectory is already well described in the control, TMI rain rates act to deepen its core. There is a maximum difference of 5 hPa on the 25-08-1998 at 12 UTC (Figure 8b). This shows that rather small modifications on the water vapour field in active regions can have a non negligible impact on the resulting dynamics in 4D-Var assimilation. The impact on the quality of medium-range forecasts is also globally positive, even though the improvement in the trajectory of hurricane Bonnie is not systematic. By drying the atmosphere where precipitation is not observed, the assimilation of TMI rain rates has a positive impact on the model spin-down for the hydrological cycle (Figure 9). Lower rain rates are produced at the beginning of the forecasts in better agreement with the model equilibrium (reached after two days). The modification of tropical humidity

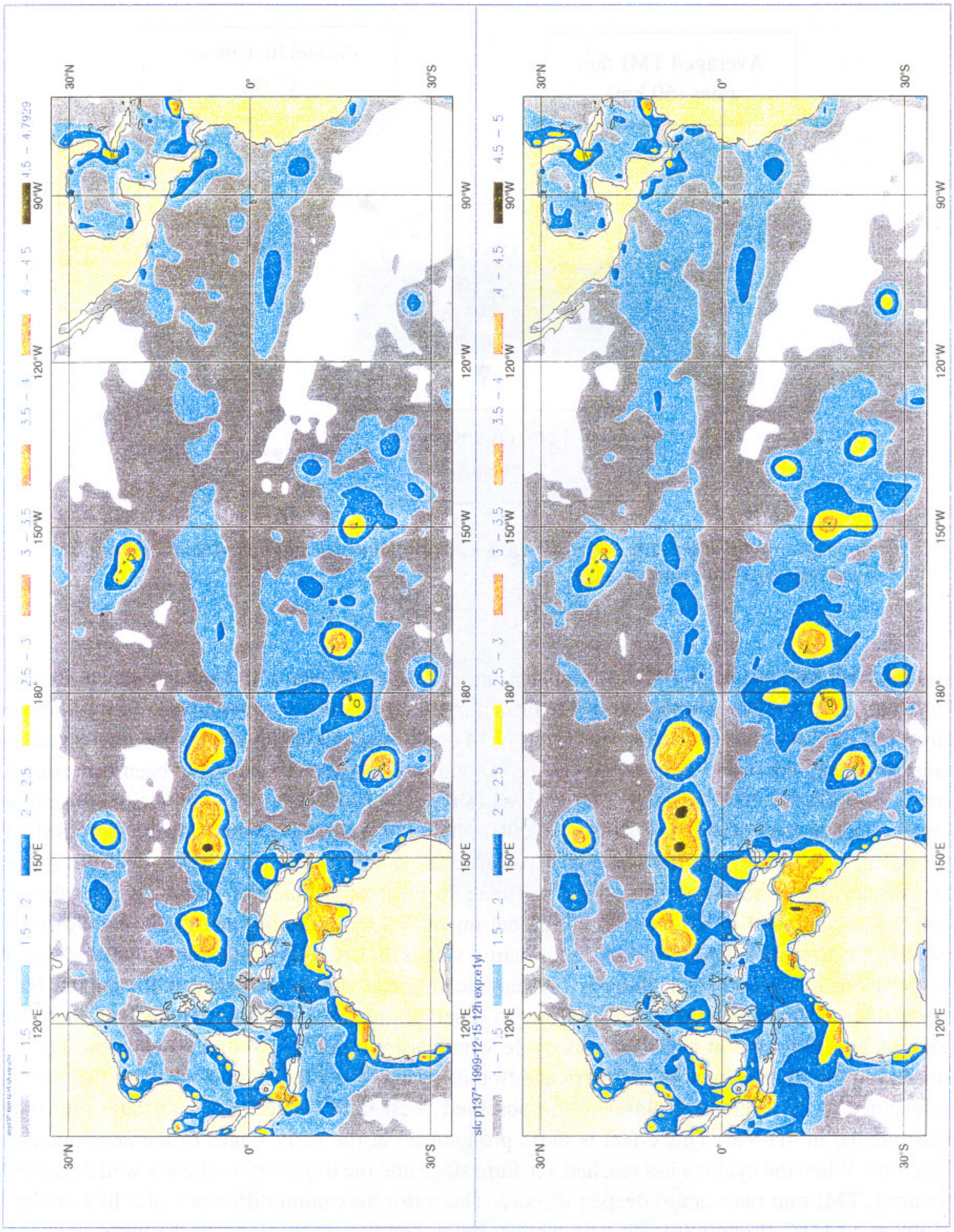
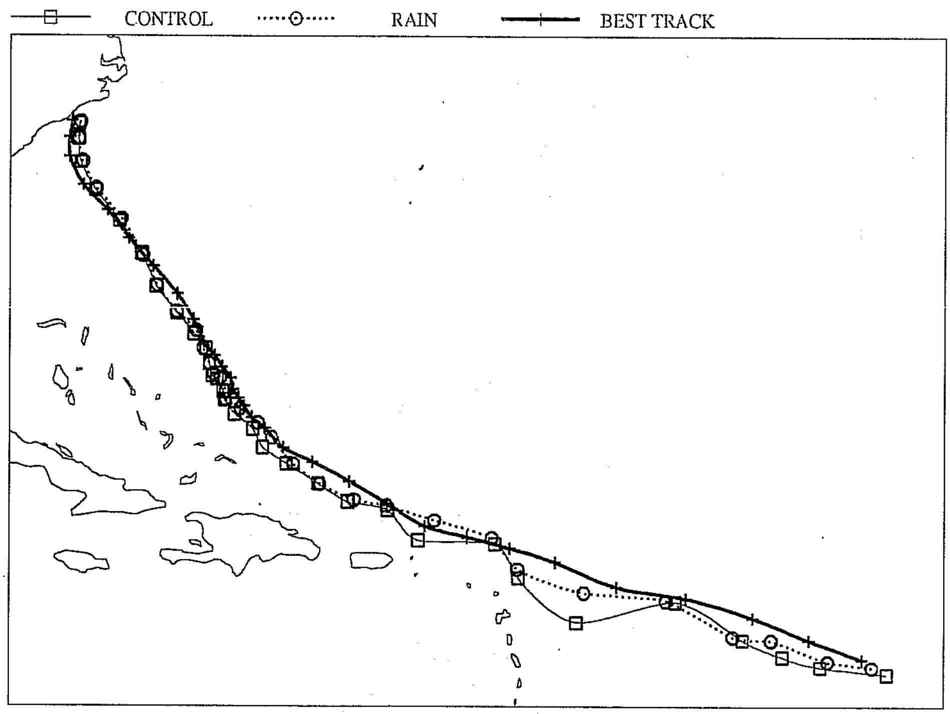


Figure 7: Root-mean square increments of TCWV in kgm-2 averaged over a 3-week period (15-12-1999 to 05-12/2000) for a control assimilation (right panel) and an assimilation with TMI 2A12 rain rates (left panel).



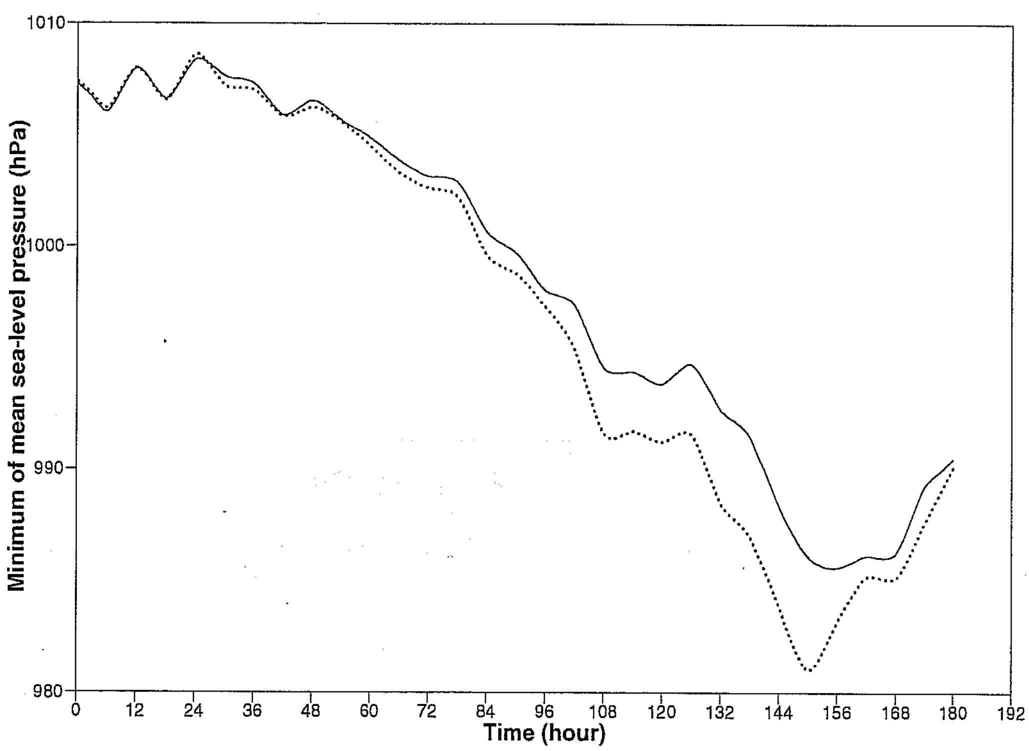


Figure 8: Analysed track of hurricane Bonnie . (a) location from 19 August 1998 at 12000 UTC to 27 August at 0000 UTC. (b) temporal evolution of the minimum of mean sea level pressure in the hurricane (initial time corresponds to 19 August 1998 at 1200 UTC). Symbols are every 6 hours.

has also a direct impact on wind scores. Figure 10 shows that the root-mean-square error for the forecast wind vector at 850 hPa is significantly reduced up to day 4 when TMI rain rates are assimilated.

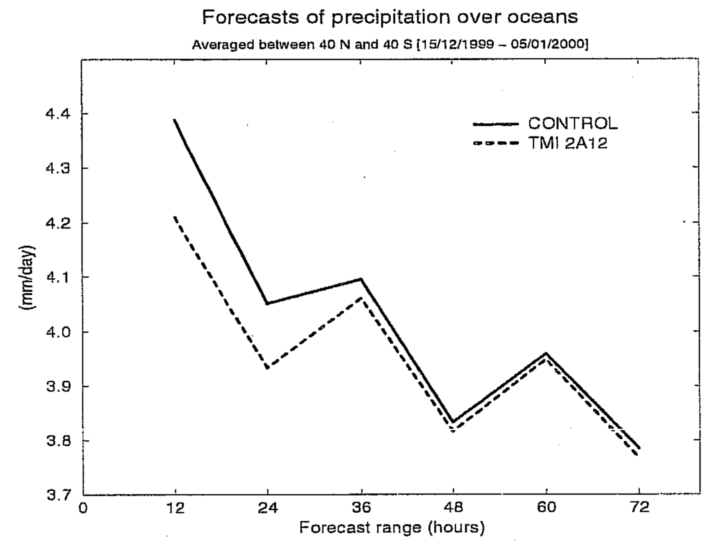


Figure 9: Time evolution of global precipitation in the tropical belt (40 S to 40 N) averaged over 22 forecasts for a control experiment and an experiment with TMI rain rate assimilation.

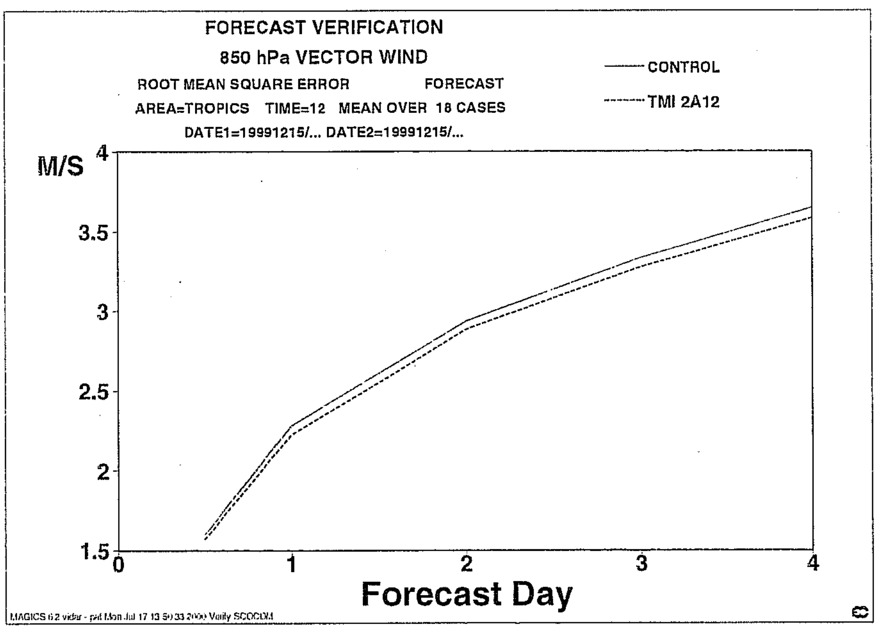


Figure 10: Tropical wind scores at 850 hPa averaged over 18 forecasts, each verified against its own analysis for a control experiment and an experiment with TMI rain rate assimilation

8. Toward cloud assimilation at ECMWF

8.1 Validation of cloud products

Before assimilating cloud properties in an NWP model a direct validation of cloud fields is first required. Such validations were initially done by comparing monthly means with climatological satellite products such as ISCCP for cloud cover and ERBE for top-of-the-atmosphere radiative fluxes. These comparisons have recently focused at ECMWF on the validation of short-range forecasts for TOA and surface fluxes (Chevallier and Morcrette, 2000) and on the evaluation of the vertical structure of clouds using lidar data from the space shuttle (Miller *et al.*, 1999) as well as ground based radar (Hogan *et al.*, 2001). Comparisons using weather regime and composite classifications have also started with ISCCP data (Klein and Jakob, 1999).

8.2 Simulation of cloudy radiances

Another type of validation prior to the assimilation of cloud properties can be performed by a simulation of cloudy radiances. The current fast radiative transfer code RTTOV (Saunders *et al.*, 1998) used at ECMWF for the inversion of clear-sky radiances has been adapted by F. Chevallier (personal communication) to account for cloudy situations. In the infra-red, cloud optical properties are parametrized according to Ebert and Curry (1992) for ice particles and Smith and Shi (1992) for water droplets. The parameterization of Liebe (1989) is used for cloud absorption in the microwave. Using the cloud overlap assumption proposed by Raisanen (1998) full sky radiances can be simulated using

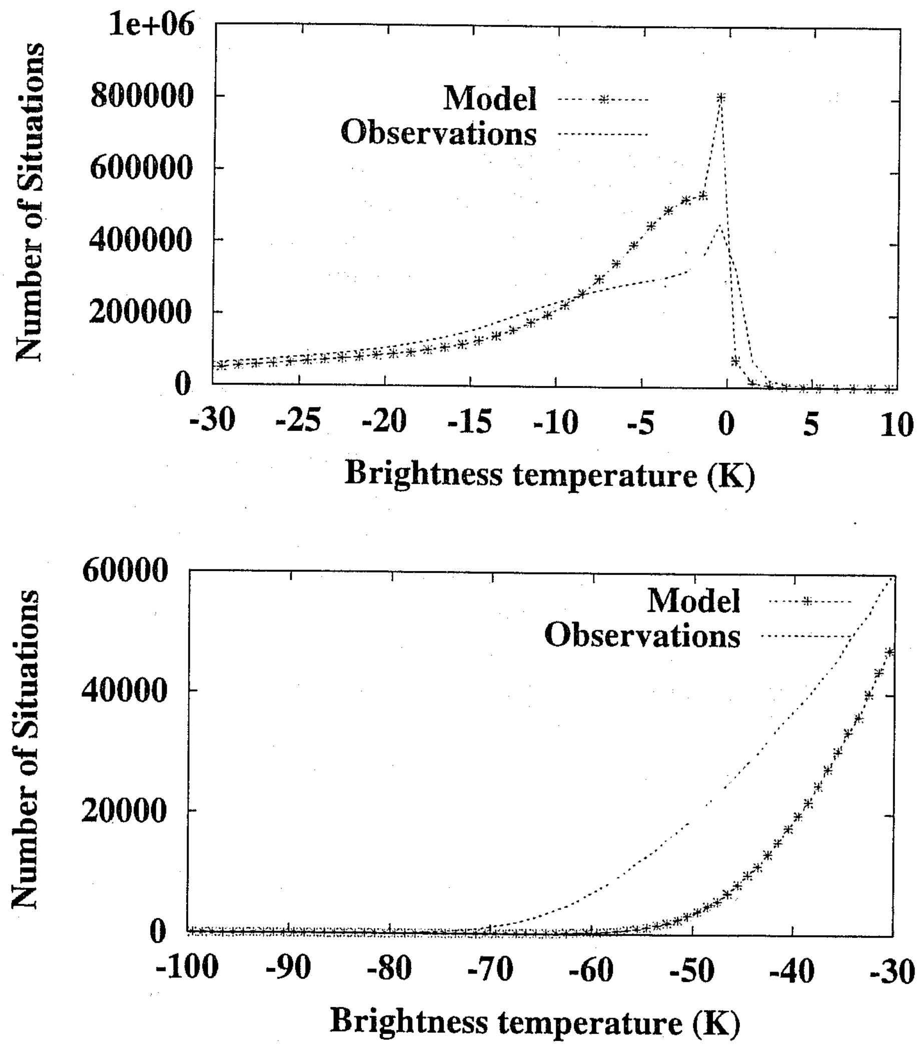


Figure 11: Histograms of differences between total sky radiances (observed and simulated) for HIRS Channel 8 and simulated clear sky radiances for a 3-month period (DJF 86-87) over mid-latitude oceans [30N- 60N and 30S-60N]

condensed water (liquid/ice) and cloud fraction from the forecast model. Figure 11 shows a histogram of cloudy radiances from HIRS-8 where observed and simulated brightness temperatures are compared in terms of differences with clear sky simulated values (for a three-month period over mid-latitude oceans). The signature of clouds is given by colder values than the clear sky ones: high clouds are then characterized by strong negative differences. The general cloud structures are well located by the ECMWF model (the two histograms have a similar shape). Some weaknesses are identified such as an underestimation of high-top cloud condensate in these regions (not enough situations with large negative brightness temperatures). Simulated and observed cloudy radiances have also been processed for the retrieval of cloud properties (cloud top pressure and effective cloud amount) using the CO_2 slicing method (Wylie *et al.*, 1994). These results are discussed in Chevallier *et al.* (2001)

8.3 1D-Var assimilation of simulated radiative fluxes

Preliminary 1D-Var experiments have been performed at ECMWF using a methodology similar to the one presented for the assimilation of rainfall rates, but for simulated observations of outgoing longwave radiation (OLR) at the top-of-the-atmosphere. The parametrization of moist processes is replaced by the ECMWF longwave radiation scheme (Morcrette, 1991) and by a simple diagnostic cloud scheme (Slingo, 1987). This cloud scheme produces fractional cover and liquid/ice water contents mostly from temperature and humidity information (there are other dependencies with convective precipitation and vertical velocity which are not taken into account in the present study). Considering a clear-sky observed value of OLR and having clouds in the first-guess profile, is it possible for 1D-Var to remove clouds ? The answer is yes as depicted in Figure 12. The initial value of 186 W/m^2 has been increased to 271 W/m^2, a value very close to the observation (285 W/m^2). Such agreement is possible because the observation is supposed to be very accurate (with 2 W/m^2 error). This is reasonable for simulated observations, but the error should be increased when real observations are considered. The reduction of cloud cover has been done through a warming and a drying of the corresponding atmospheric layer. The local drying at high level is associated with a much larger drying around 850 hPa, which reflects the vertical structure of the humidity background errors. Another exercise has been tried. If the cloud amount in the model is too low, can 1D-Var increase cloud cover ? The answer again is yes. Starting from a cloud fraction of 20% and using a simulated OLR observation associated with a cloud cover of 100%, the 1D-Var is able to reconstruct a cloud structure very similar to the expected one (Figure 13). The analyzed cloud fraction at the top is smaller than the observed one because this region is associated with very small amounts of ice water. The conclusion from these preliminary studies is that, at least for some cases, a 1D-Var assimilation of radiative fluxes can improve the description of the vertical cloud structure produced by the model, through an adjustment of temperature and humidity profiles. This approach has also been tested with success using simulated cloudy radiances from HIRS (Chevallier *et al.*, 2001). However the vertical structure of the increments is strongly influenced by the statistics of background errors. The importance of this a-priori information in the 1D-Var retrieval could be reduced by including additional observations on the vertical such as those provided by radars or lidars.

9. Future developments and remaining issues

Current satellite already provide information on clouds and precipitation that is not used by operational data assimilation systems. These data can be important particularly for small scale phenomena not well described by the synoptic network and over tropical regions where data are sparce and where the circulation is strongly linked to the latent heat release in convective clouds. New satellites will provide more accurate (increased spectral resolution) and less ambiguous (e.g. vertical profiles) information on clouds and precipitation. Variational data assimilation techniques currently run in a number of operational centres (ECMWF, UKMO, Météo-France, NCEP, CMC) can be used to extract information contained in cloudy and rainy radiances. These techniques require accurate forward modelling of the observations as well as useful observation operators (moist physics, radiation). Several encouraging

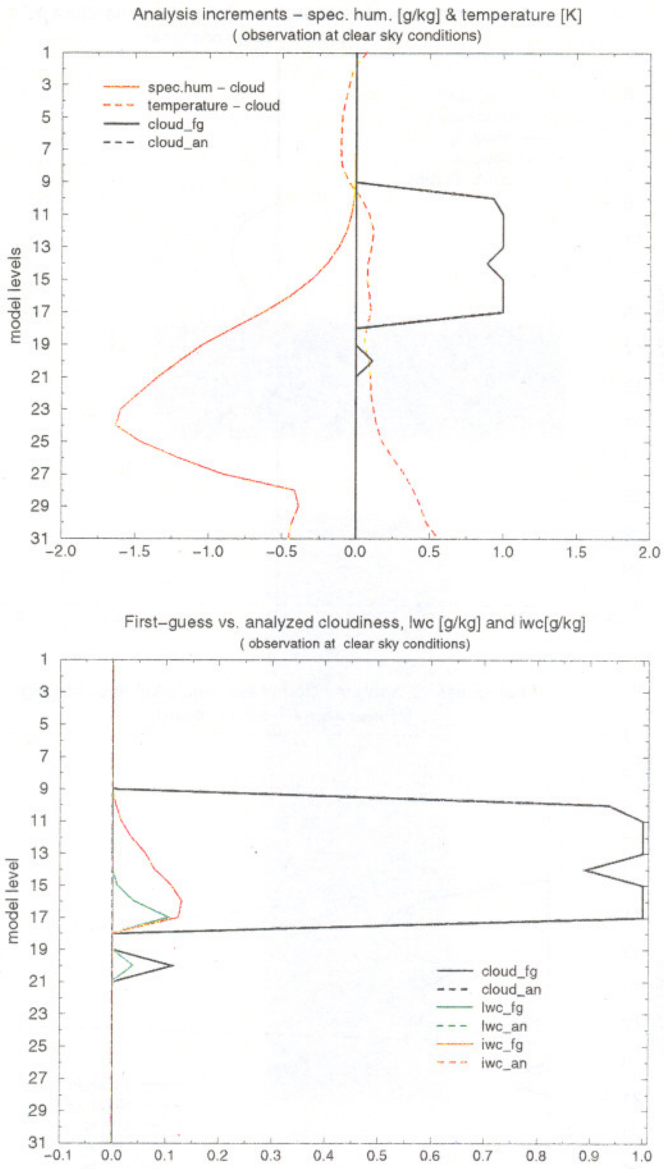


Figure 12: Analysis increments of temperature and humidity (top panel) produced by a 1D-Var assimilation of clear sky longwave flux at the top of the atmosphere. A cloud layer is present in the background profile between level 18 and level 9 as shown on the lower panel together with liquid and ice water contents.

results obtained at ECMWF have been illustrated in this paper for both TMI rainfall rates and simulated cloudy radiation fluxes.

There are still issues that need to be addressed in order to assimilate efficiently cloudy and rainy radiances. The spatial resolution of NWP models should be increased in order to reduce representativeness errors when comparing model to observations. This is particularly true for the horizontal resolution used to solve the minimization in 4D-Var (given the fact that physical parametrization schemes are more sensitive to resolution changes than spatial interpolation schemes). Efficient and/or simplified observation operators for convection, clouds and radiation need to be defined with their corresponding linearized versions. For example, since ECMWF longwave radiation scheme is expensive, an approach based on neural networks and mean Jacobians has been defined by Janiskova *et al.* (2000). Similar methodology should be used for clouds and deep convection. Both model and observation errors need to be assessed through systematic comparisons in order to derive reliable statistics. More specifically, the quantification of observation errors for satellite derived products can

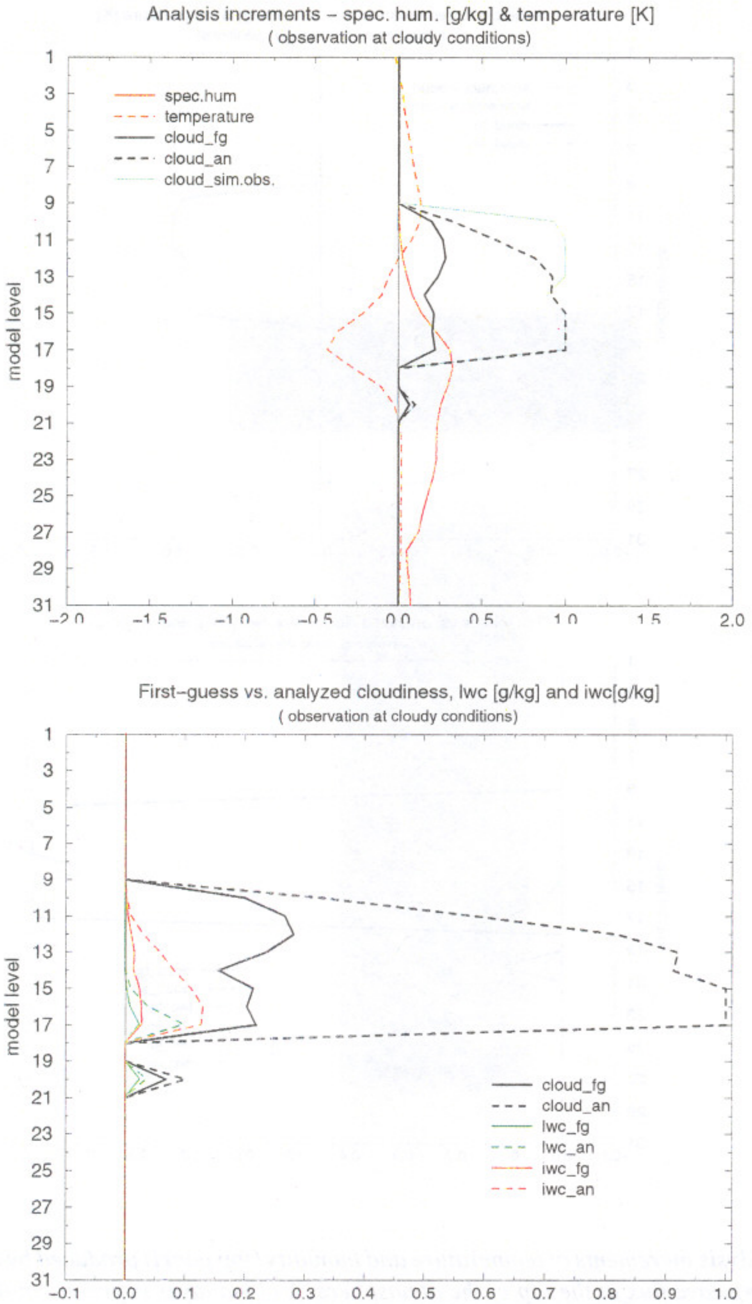


Figure 13 : Analysis increments of temperature and humidity (top panel) produced by a 1D-Var assimilation of cloudy longwave flux at the top of the atmosphere. The initial cloud layer in the background profile between level 18 and level 9 is shown on the lower panel together with liquid and ice water contents (solid lines). The analysed cloud cover and water contents are also presented on the lower panel (dashed lines).

be estimated from more accurate collocated data (such as the precipitation radar on board TRMM to calibrate TMI retrievals). The statistics of background errors for humidity also need improvements. They are currently univariate and have no geographical dependence. For example, it is likely that correlation length scales are much smaller in rainy and cloudy systems than the current "averaged" values. In such systems, the link between temperature, humidity and the dynamics is strong. This implies error correlations which currently do not exist. Some specific limitations of the 4D-Var approach for the assimilation of clouds and rain should also be addressed (non-linearities, thresholds, assumption of Gaussian errors). In current global data assimilation systems, cloud variables are not explicitly modified by the analysis (they are assumed to adjust rapidly to the dynamic and thermodynamic forcings). To which extent such assumption is valid may depend upon the application (short-range versus medium range forecasts).

Given the current status of forward modelling for both radiation and cloud processes it has been suggested that cloudy radiances could be assimilated directly whereas the complexity of the description of rainy radiances requires a rain rate retrieval before assimilation. However, considering low frequency channels in the microwave (10 and 19 GHz) which are weakly affected by ice scattering it may be possible to evaluate model rain profiles and examine if simulated radiances have any resemblance with observed radiances.

These very exciting developments will require important collaboration between research groups developing data assimilation systems, physical parametrization schemes and satellite retrievals. These three branches should benefit each other by a better usage of satellite data for assimilation and also by an improved evaluation of parametrization schemes. All these developments should lead to improved forecast skills and to a better understanding of the hydrological cycle of the atmosphere.

References

Andersson, E., J. Pailleux, J.-N. Thépaut, J.R. Eyre, A.P. McNally, G.A. Kelly and P. Courtier (1994): Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. *Quart. J. Roy. Meteor. Soc.*, **120**, 627-653

Arkin, P.A. and B.N. Meiner (1987): The relationship between large scale convective rainfall and cold cloud over the western hemisphere during 1982-84. *Mon. Wea. Rev.*, **115**, 51-74

Buizza, R. (1994): Sensitivity of optimal unstable structures. *Quart. J. Roy. Meteor. Soc.*, **120**, 429-451

Chevallier, F. and J.-J. Morcrette (2000): Comparison of model fluxes with surface and top-of-the-atmosphere observations. *Mon. Wea. Rev.*, **128**, 3839-3852

Chevallier, F., P. Bauer, G. Kelly, J.-F. Mahfouf, C. Jakob and T. McNally (2001): Requirements for the assimilation of cloudy radiances. Proceedings of the EuroTRMM/ECMWF Workshop on the assimilation of clouds and precipitation, 6-9 November 2000, Reading (U.K.)

Deirmendjian, D. (1975) : Far-infrared and submillimeter wave attenuation by clouds and rain. *J. Appl. Meteor.*, **14**, 1584-1593

Ebert, E.E. and J.A. Curry (1992): A parametrization of ice optical properties for climate models. *J. Geophys. Res.*, **97D**, 3831-3836

Falkovich, A., E. Kalnay, S. Lord and M.B. Mathur (2000): A new method of observed rainfall assimilation in forecast models. *J. Appl. Meteor.*, **39**, 1282-1298

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

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

Hogan, R.J., C. Jakob, and A.J. Illingworth (2001): Comparison of ECMWF cloud fraction with radar derived values. *J. Appl. Meteor.*, **40**, 513-525.

Kasahara, A., J.I. Tsutsui and H. Hirakuchi (1996): Inversion methods of three cumulus parameterizations for diabatic initialization of a tropical cyclone model. *Mon. Wea. Rev.*, **124**, 2304-2321

Klein, S.A. and C. Jakob (1999): Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.*, **127**, 2514-2531

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

Krishnamurti, T.N., H.S. Bedi and K. Ingles (1993): Physical initialization using SSM/I rain rates. *Tellus*, **45A**, 247-269

Janiskova, M., J.-F. Mahfouf, J.-J. Morcrette and F. Chevallier (2000): Development of linearized radiation and cloud schemes for the assimilation of cloud properties. *ECMWF Technical Memorandum No 301*, 31 pp.

Le Dimet, F.X. and O. Talagrand (1986): Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, **38A**, 97-110

Liebe H.J. (1989) : MPM-an atmospheric millimeter-wave propagation model. *Int. J. Infrared and Millimeter Waves*, **10**, 631-650

Mahfouf, J.-F. (1999) : Influence of physical processes on the tangent-linear approximation. *Tellus*, **51A**, 147-166

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

Marécal, V. and J.-F. Mahfouf (2000b) : Four dimensional variational assimilation of total column water vapour in rainy areas. *ECMWF Technical Memorandum No 314*, 29 pp.

Manobianco, J., S. Koch, V Mohan Karyampudi and A.J. Negri (1994) : The impact of assimilating satellite-derived precipitation rates on numerical simulations of the ERICA IOP 4 cyclone. *Mon. Wea. Rev.*, **122**, 341-365

Miller, S.D., G.L. Stephens and A.C.M Beljaars (1999) : A validation survey of the ECMWF prognostic cloud scheme using LITE. *Geophys. Res. Lett.*, **26**, 1417-1420

Molteni, F., R. Buizza, T.N. Palmer and T. Petroliagis (1996) : The ECMWF Ensemble Prediction System : methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73-119

Morcrette, J.-J. (1991) : Radiation and cloud radiative properties in the ECMWF operational forecast model. *J. Geophys. Res.*, **96D**, 9121-9132

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

Puri K., J. Barkmeijer and T.N. Palmer (2001): Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Quart. J. Roy. Meteor. Soc.*, **126**, 709-732.

Raisanen, P. (1998): Effective longwave cloud fraction and maximum-random overlap clouds- a problem and a solution. *Mon. Wea. Rev.,*. **126**, 3336-3340

Rossow, W.B (1996): Properties of clouds and cloud systems. *Radiation and Water in the Climate System: Remote Measurements*, NATO ASI Series 45, Springler-Verlag, 249-276

Simmer, C. (1996): Retrieval of precipitation from satellites. *Radiation and Water in the Climate System: Remote Measurements*, NATO ASI Series 45, Springler-Verlag, 249-276

Saunders, R., M. Matricardi and P. Brunel (1998): An improved fast radiative transfer model for the assimilation of satellite radiance observations. *Quart. J. Roy. Meteor. Soc.*, **124**, 1407-1426

Slingo, J.M. (1987) : The development and verification of cloud prediction scheme in the ECMWF model. *Quart. J. Roy. Meteor. Soc.*, **13**, 899-927

Smith, E.A. and L. Shi (1992): Surface forcing of the infrared cooling profile over the Tibetan plateau. Part I: Influence of relative longwave radiative heating at high altitude. *J. Atmos. Sci*, **49**, 805-822

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

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

Wylie, D.P., W.P. Mezel, H.M. Woolf and K.I. Strabala (1994) : Four years of global cirrus cloud statistics using HIRS. *J. Climate*, **7**, 1972-1986