Four-dimensional variational assimilation of SSM/I precipitable water content data

By M.-A. FILIBERTI1†, F. RABIER2, J. N. THÉPAUT3, L. EYMARD1 and P. COURTIER2

1Centre d’étude des Environnements Terrestre et Planétaires, France
2European Centre for Medium-Range Weather Forecasts, UK
3Météo-France, France

(Received 23 October 1996; revised 12 September 1997)

SUMMARY

Satellite microwave radiometers provide measurements of precipitable water content (PWC) over the oceans, with a horizontal resolution of a few tens of kilometres. These data represent the water vapour content integrated over the atmospheric column. We assess the value of this source of information for numerical weather prediction systems. We used the European Centre for Medium-Range Weather Forecasts Integrated Forecasting System with a four-dimensional variational method to assimilate the Special Sensor Microwave/Imager (SSM/I) PWC data over a 24-hour period. Our experiments use an incremental variational method. Comparison with observations is made using a multi-level global primitive-equation T106 spectral model with physical parametrizations. Minimization is performed using a T63 adiabatic dynamics model which includes only simplified physics (horizontal diffusion, a simple surface drag and a vertical diffusion scheme).

Comparing the control and assimilation experiments with aircraft and other data shows that the use of PWC data from SSM/I improves the analysis. We also obtain a slight improvement in short-range forecasts of almost all parameters when SSM/I-PWC data are used in the assimilation.

KEYWORDS: Data assimilation  Numerical weather prediction  Satellite data

1. INTRODUCTION

The impact of humidity data on forecast quality has been the subject of many studies. Smagorinsky et al. (1970) were the first to study the issues for forecasts. They suggested that numerical weather prediction models are rather insensitive to initial moisture conditions. Atkins (1974) showed that information on atmospheric humidity in initialization processes had a positive impact on forecasts, especially of precipitation. Norquist (1987) improved the analysis with non-conventional humidity data in areas without radiosonde data. Illari (1989), Zhang et al. (1989) and McNally and Vesperini (1996) showed that TIROS-N Operational Vertical Sounder (TOVS) satellite-derived water vapour retrievals had a positive impact on both analysis and forecast. The Special Sensor Microwave/Imager (SSM/I) precipitable water content (PWC) should be extremely useful data to define the global moisture field. These microwave data represent the water vapour content integrated over the atmospheric column and are available in almost all weather, contrary to infrared and visible data, with horizontal resolution of a few tens of kilometres. The gradient of PWC allows the accurate location of fronts in all seasons and over all oceans (Katsaros et al. 1989; Filiberti et al. 1994). Eymard et al. (1989), McMurtrie and Katsaros (1991) and Liu et al. (1992) compared values of PWC found by radiometer with European Centre for Medium-Range Weather Forecasts (ECMWF) analyses, and noted systematic errors in the ECMWF model for the humidity fields over some parts of the oceans. Comparisons between measured and analysed values of PWC suggested an overabundance of PWC in the ECMWF model. Phalippou (1992), comparing SSM/I-PWC data with ECMWF model outputs, found good coherence in structure and magnitude except in certain subsidence

* Corresponding author: LODYC, Université Pierre et Marie Curie, Case 100, 4 place Jussieu, 75252 Paris Cedex 05, France. e-mail: maf@lodyc.jussieu.fr.
† Now at LODYC, Université Paris VI, Paris, France.
regions (north-east Pacific Ocean and south-east Atlantic). With the most recent version of the ECMWF model, McNally and Vesperini (1996) and Vesperini (1998) documented an underestimate of humidity in the tropics and excessively moist conditions in the subtropics in the ECMWF operational analyses by comparison with the SSM/I-PWC data.

The first results (to the best of our knowledge) of SSM/I-PWC data (hereafter SSM/I data) assimilation in a mesoscale forecast model (Filiberti et al. 1994) using an optimal interpolation (O.I.) method showed how use of SSM/I data improved analyses and short-range forecasts. The impact was weak, especially in forecasts for fields not directly related to the observations, and was mainly located in areas where dynamics was not preponderant. The analysis of humidity in O.I. (as in the three-dimensional variational (3D-Var) assimilation method) is univariate, which possibly leads to inconsistencies with the model dynamics.

In the present study, we test a four-dimensional variational assimilation method (4D-Var) with the Integrated Forecasting System (IFS)/Action de recherche petite échelle grande échelle (ARPEGE) system developed in collaboration between ECMWF and Météo-France. This approach circumvents some of the practical O.I. weaknesses, since the 4D-Var analysis uses the model’s nonlinear multivariate dynamics to interpret all the available observations. The experiment we perform consists in assimilating SSM/I data over a 24-hour period.

Variational assimilation of meteorological observations using the adjoint technique has been widely studied in recent years (Rabier and Courtier 1992; Thépaut et al. 1993a, b; Andersson et al. 1994 or Thépaut et al. 1996). An earlier study involving SSM/I data assimilation with a variational method was performed to evaluate the ability of the IFS system to extract the information contained in the SSM/I humidity data. This preliminary experiment was carried out using an adiabatic 4D-Var system (Filiberti 1993), except for a horizontal diffusion similar to the one described by Thépaut and Courtier (1991), a surface drag and a simple linear scheme for representing vertical diffusion (Buizza 1994). The results of this experiment indicated that when SSM/I data were used the 4D-Var results for all fields were improved compared to a control run without SSM/I, even in data-dense areas, and that these data could be used consistently with other observations. These preliminary results with limited physics also showed the adiabatic model’s inability to take into account moist processes, which are important for the humidity field. A similar conclusion was reached in 4D-Var experiments without the use of SSM/I data (Rabier et al. 1992).

It is now possible to take physics into account in the assimilation procedure through the incremental approach of 4D-Var described by Courtier et al. (1994). The minimization still uses a simplified model for the evolution of the increments but the differences between the current analysis and the data (‘departures’) are evaluated with the full nonlinear model at regular intervals within the minimization. It was decided to test this approach in the context of assimilating SSM/I data. Contrary to the previous experiment, which used adiabatic dynamics to compare with observations, we now use a model dynamics with physical parametrizations (in particular large-scale and convective precipitation) to compare with observations. The dynamics used to estimate the evolution of errors is still adiabatic but the specific humidity is constrained between zero and its saturation value.

In the following sections, we present the 4-D variational assimilation method and its formulation in terms of increments (section 2). We also describe the model and the data used in this study (section 3). Section 4 describes the different global experiments. We then provide in section 5 the results of a single SSM/I-PWC observation experiment to estimate the structure functions implicitly used in 4D-Var. The results of the global experiments are presented in sections 6 and 7. Some concluding remarks are in section 8.
2. FOUR-DIMENSIONAL VARIATIONAL ASSIMILATION

(a) General formulation

The general problem of assimilation of observations in numerical weather prediction consists in defining the initial conditions of the forecast model, using all the available information on the atmosphere.

4D-Var seeks an optimal balance between observations and the dynamics of the atmosphere by finding a model solution which is as close as possible, in a least-square sense, to the observations available over a given period \([t_0, t_n]\). The misfit to the data and other possible constraints is measured by a cost function which can be written as

\[
J[\mathbf{x}(t_0)] = J^o + J^b + J^c,
\]

where \(\mathbf{x}(t_0)\) is the atmospheric state at time \(t_0\), \(J^o\) the distance of the atmospheric state vector \(\mathbf{x}\) to the observational vector \(\mathbf{y}^o\) over space and time, and \(J^b\) the distance between \(\mathbf{x}(t_0)\) and a background state \(\mathbf{x}^b(t_0)\) (which will be provided, for example, from a previous 6-hour forecast). Other constraints can be added through an additional term \(J^c\). In our experiments, \(J^c\) makes the increments tangent to the slow manifold. \(J^o\) and \(J^b\) take the following forms respectively:

\[
J^o = 1/2 \sum_{i=0}^{n} [H_i \mathbf{x}(t_i) - \mathbf{y}_i^o]^T \mathbf{R}_i^{-1} [H_i \mathbf{x}(t_i) - \mathbf{y}_i^o] \quad (2)
\]

\[
J^b = 1/2 [\mathbf{x}(t_0) - \mathbf{x}^b(t_0)]^T \mathbf{B}_0^{-1} [\mathbf{x}(t_0) - \mathbf{x}^b(t_0)], \quad (3)
\]

where at any time \(t_i\), \(\mathbf{y}_i^o\) is the vector of observations, \(H_i\) the operator providing the equivalent of the data from the model variable \(\mathbf{x}(t_i)\) (\(H_i\) provides the interpolation to the observation locations, the calculation of the observation variables (as in 3D-Var) and also the action of the forecast model), \(\mathbf{R}_i\) the observation error covariance matrix (including measurement and representativeness errors (Lorenc 1986)), and \(\mathbf{B}_0\) the background error covariance matrix. The background information valid for time \(t_0\) summarizes all information used before time \(t_0\). Superscripts −1 and T denote respectively matrix inverse and transpose.

The atmospheric state \(\mathbf{x}(t_i)\) at time \(t_i\) is defined as

\[
\mathbf{x}(t_i) = M(t_i, t_0) \mathbf{x}(t_0),
\]

where \(M\) is the forecast model integrated from time \(t_0\) to \(t_i\).

The cost function \(J\) is minimized through a descent algorithm (in our case MIQN3 from INRIA (Gilbert and Lemaréchal 1989)) which requires several computations of the gradient of \(J\) with respect to the initial state \(\mathbf{x}(t_0)\). These computations are made using the adjoint technique.

(b) Incremental approach

The quality of the 4D-Var assimilation depends on the quality of the assimilating forecast model (because of the underlying hypothesis of a perfect model). However, the better the model is (in particular if it contains sophisticated physical parametrizations), the less regular it becomes because of the many thresholds in the current parametrizations. The strong nonlinearities included in the model can lead to convergence problems during the minimization.

The incremental 4D-Var (Courtier et al. 1994) consists in computing the background trajectory and the departures (observations minus model) using the full model (including
physics), and minimizing the cost function in the space of increments at the initial time using a linear model $M$ and its adjoint. It allows a compromise between the two previous issues: the forecast model can be improved and the nonlinearities are not taken into account for the minimization.

Writing $x(t_0) = x^b(t_0) + \delta x(t_0)$ and $H_i[x(t_i)] = H_i[x^b(t_i)] + H_i \delta x(t_i)$, with $H_i$ being the linearisation of $H_i$ in the vicinity of the background $x^b$ and $\delta x(t_i) = M(t_i, t_0) \delta x(t_0)$, the new cost function becomes:

$$\frac{1}{2} \sum_{i=0}^{n} [H_i \delta x(t_i) - d_i]^T R_i^{-1} [H_i \delta x(t_i) - d_i] \quad \text{for } J^a$$

$$+ \frac{1}{2} [\delta x(t_0)]^T B_0^{-1} [\delta x(t_0)] \quad \text{for } J^b,$$

where $d_i = y_i^o - H_i x^b(t_i)$ is the innovation vector.

The analysis is given by $x^a(t_0) = x^b(t_0) + \delta x^a(t_0)$ where $\delta x^a(t_0)$ is the minimizing solution of (Eq. 5).

Courtier et al. (1994) made two important remarks about this method: it is possible to introduce some nonlinearities by updating iteratively the trajectory about which the full model is linearized (outer loop); this formulation allows for simplification of the tangent linear model and reduction of the control variable (inner loop).

In practice, $M$ is taken as the tangent linear model of the adiabatic part of the forecast model $M$, and the increments are computed at a lower resolution. Twenty-five iterations for each inner-loop minimization and four external updates (outer loop) of the trajectory were performed (which gives a total of one hundred iterations).

3. The model and the data

(a) The IFS model

We use the IFS primitive equation model (cycle 12r1) developed in cooperation between Météo-France (ARPEGE project) and ECMWF (Courtier et al. 1991). The features of the IFS/ARPEGE system which are of interest for this study are a global multilevel spectral model, a three-dimensional variational analysis and its four-dimensional extension. The model uses the spectral representation of the vorticity, divergence, temperature, specific humidity and logarithm of surface pressure fields based on spherical harmonics with a triangular truncation. In this study we chose to use a triangular truncation at wave number 106, and 19 vertical levels (T106L19), with a reduced gaussian grid—implying a quasi-uniform horizontal resolution of about 120 km. The evolution of the specific humidity $q$ is determined by

$$\frac{dq}{dt} = \frac{\partial q}{\partial t} + u \frac{\partial q}{\partial x} + v \frac{\partial q}{\partial y} + \eta \frac{\partial q}{\partial \eta} + S,$$

where $u$ and $v$ are the components of wind field, $\eta$ is the vertical hybrid coordinate based on pressure and following the orography, $\dot{\eta} = (\partial \eta / \partial t)$ is the vertical pseudo-velocity and $S$ is the source term. This source term corresponds to a truncation for negative and supersaturation values:

$q = q_{\text{sat}}$ for $q + dq > q_{\text{sat}}$ (where $q_{\text{sat}}$ is the specific humidity at saturation) and $q$ is reset to 0 for $q + dq < 0$.

Near zero and $q_{\text{sat}}$, we specified a smooth transition zone to obtain differentiable corrections.
The tangent-linear and adjoint versions of the adiabatic part of the full model have been developed for the investigation of 4D-Var (3D-Var is in essence coded as the particular case of a zero time-step of the model). The minimization problem involved in 4D-Var experiments can be considered as "large scale", since the number of degrees of freedom in the control variable is more than $3 \times 10^5$ at truncation T63 with 19 levels (horizontal resolution of about 210 km).

(b) The SSM/I data

The first SSM/I was launched on 19 June 1987. It was mounted aboard the Defence Meteorological Satellite Program (DMSP) satellite of the US Navy, flying in a circular, sun-synchronous near-polar orbit at an altitude of 833 km, with an orbital period of 102 min.

The SSM/I is a passive, scanning multichannel microwave radiometer sensor (Hollinger et al. 1990). The SSM/I radiometer measures brightness temperatures at 19.3, 22.2, 37.0 and 85.5 GHz. All frequencies are received in dual polarization, except 22 GHz, which is received in vertical polarization only. The active portion of the conical scan covers a swath about 1400 km broad, and the horizontal sampling of processed brightness temperatures is 25 km (except at 85.5 GHz).

Several algorithms have been developed to retrieve the precipitable water from the brightness temperatures (Wentz 1988; Alishouse et al. 1990; Lojou et al. 1994). In this study we use precipitable water contents directly obtained from the WetNet Data Base (Dodge and Goodman 1994). The PWC is obtained using the statistical algorithm of Alishouse et al. (1990). The accuracy of the retrieved precipitable water content decreases when strong precipitation occurs. The root-mean-square (r.m.s.) error and the biases compared to radiosonde data depend on the value of the observation. The bias ranges from $-1.6$ to $1.2$ kg m$^{-2}$ and the r.m.s. error from 1.9 to 4.3 kg m$^{-2}$ as shown by an internal ECMWF study (L. Phalippou, personal communication). The 37 GHz signal depolarization is used to reject the data when there is strong precipitation.

To include SSM/I data in the variational assimilation, an additional term is included in the cost function $J^o$. The cost function measuring the misfit of the model to SSM/I data is written in the simple form

$$J^o_{\text{SSM/I}} = 1/2 \sum_{i=0}^{n} [y_{\text{mod}} - y^o_{\text{SSM/I},i}]^T R_i^{-1} [y_{\text{mod}} - y^o_{\text{SSM/I},i}],$$

(7)

where $y^o_{\text{SSM/I}}$ is the SSM/I observation and $y_{\text{mod}} = H(t_i)$. In this case, the observation operator $H$ contains:

— a forecast model from $t_0$ to the observation time;
— an inverse spectral transform to go from spectral to grid-point space;
— a horizontal interpolation to go from grid point to the observation location;
— a vertical integration to evaluate the PWC from the specific humidity ($q$) at the model levels.

The observation operator here is trivial:

$$y_{\text{mod}} = \int_{p_s}^{0} q \, dp,$$

(8)

where $p_s$ is the surface pressure at the observation location, $q$ is the specific humidity and $p$ is the pressure at the same level. The adjoint of all these operators is also needed to compute the gradient of $J^o_{\text{SSM/I}}$ with respect to the control variable.
Figure 1. Trajectories of the centre of the surface low of the storm of 15/16 October 1987: (a) forecast by the ECMWF model (pecked line); (b) forecast by the Météo-France EMERAUDE (EME) model (dotted line); (c) trajectories based on the operational analyses of Météo-France (solid line). Times and dates are shown in the form hh/dd, where hh is the hour and dd the day. Pressures at the centre of the surface low are in hPa. (Diagram taken from Jarraud, M., Goas, J. and Deyts, C. (1989) Weather and Forecasting, 4, 517–536.)

4. ASSIMILATION EXPERIMENTS

In this study, data are assimilated over a 24-hour period from 15 UTC 13 October 1987 to 15 UTC 14 October 1987, and a 36-hour forecast is started at 12 UTC 14 October valid for 00 UTC 16 October.

During the period chosen, a great storm struck the coasts of France and the south coast of England. This case was badly forecasted by some meteorological models, especially for 36-hour forecasts. Figure 1 (taken from Jarraud et al. 1989) shows the four-day forecasts of the track of the low pressure centre as forecasted from the ECMWF model and from the Météo-France model, EMERAUDE. The forecast available from Météo-France is more consistent with the observations than the ECMWF one for the location of the low, even if the trough intensity is underestimated.

The meteorological situation has been extensively documented (Jarraud et al. 1989; Gadd and Morris 1988; Lorenc et al. 1988) and we shall not go into details. The storm was most developed at 00 UTC 16 October 1987.

As the storm was present in the medium-range forecasts, the atmospheric models were able to represent the phenomenon. One of the reasons advocated for the bad performance of most operational forecasting centres was the non-representation of realistic structure functions within O.I. in such meteorological situations. Isotropic and barotropic covariance functions are poorly adapted for such an exceptional situation. Regarding the specific problem of the structure functions in the case of a rapidly evolving system and in the presence of strong baroclinic instability, Thépaut et al. (1993b) and Rabier and Courtier (1992) studied the behaviour and the performance of 4D-Var. Thépaut et al. (1993b) showed that the structure functions implicitly used in 4D-Var are flow dependent, contrary to those usually specified in traditional sequential assimilation systems. In the case of the
exceptional storm of 16 October 1987, the baroclinic information could then be passed onto the forecast error covariances. A better model resolution and the increase of the number of observational data should also improve the results. However, observations are very sparse over the ocean, particularly in cloudy areas where the TOVS data cannot be used. Lorenc et al. (1988) investigated the behaviour of the United Kingdom Meteorological Office data assimilation schemes for this period. They found that the inclusion of late observations, particularly aircraft reports (AIREPs) and satellite soundings (SATEMs), made a significant impact, as did increasing the weight given to observations. They also noticed that some structures of the storm development were not resolved by the available observations.

This extreme event is used at ECMWF as a test case to evaluate the improvements stemming from refinements in the four-dimensional assimilation system (such as the introduction of physics and an increase in resolution).

Our purpose is to try out 4D-Var in this situation using the SSM/I data and to compare short-range forecasts coming from 4D-Var and 3D-Var analyses (produced in similar conditions) and a 3D-Var analysis.

(a) Quality control of SSM/I data

The SSM/I PWC is not available over land. We assumed that the SSM/I observation errors at adjacent places (in the scan and transversely) were uncorrelated. (We could not find a significant correlation between adjacent points.) We take one SSM/I observation out of five in both directions (across and along the track), to ensure horizontal consistency with the spatial resolution of the model. The final sampling is then for points 125 km apart. This separation removes any spatial correlation between observations and so the need to take it into account. These data are distributed over all oceans. To avoid contamination of the data by the effects of ice we eliminate data southward of 60°S and northward of 60°N between 20°W and 60°W (Greenland). Data for points too close to land are also rejected. The observational error of these data is taken as 20% of the observed value if the value is smaller than 25 kg m⁻², or 10% if the value is greater than 25 kg m⁻², as used in internal ECMWF studies (L. Phalippou, personal communication). To take into account the possibility that individual observations in data-sparse areas might be biased, we assumed an ‘observational error standard deviation’ which was greater than the calculated standard deviation of the observed values. We then determined the difference between the background value and each observed value. An observed value was used when this difference was less than four times the observational error standard deviation. Such differences depend on model resolution, because of variations in the accuracy of the landsea mask. This method allows the elimination of anomalous observation points but also of points for which the large difference may be due to a model error. This could, however, reduce the impact of SSM/I data.

(b) Experimental framework

All assimilation experiments presented here have been performed at truncation T106L19 and increments are calculated at T63L19. The same observations (except SSM/I) and the same quality control are used for 3D-Var and 4D-Var experiments. For the 3D-Var and 4D-Var background the model of errors used is 15% of the relative humidity value.

A 3D-Var experiment without SSM/I data is used as a reference run for 4D-Var experiments. The 3D-Var experiment includes conventional observations, by which we mean those observations with errors which are not correlated in the horizontal, such as temperature, humidity and wind measurements from radiosondes (TEMP), wind measurements
from ascending balloon (PILOT), aircraft wind data (AIREP) and surface observations from land (SYNOP). In addition, cloud drift wind (SATOB from METEOSAT and GOES) measurements, drifting-buoy surface-pressure observations (DRIBU) and satellite thicknesses (SATEM from TOVS/NOAA11) are included. (As TOVS humidity data are not available for this period at ECMWF, 1D-Var is not used.) In the vertical, the number of radiosonde and pilot-balloon observations is greatest at 850 hPa, decreasing slowly through the troposphere and the low stratosphere, whereas aircraft reports are concentrated around 250 hPa. The 3D-Var analysis experiments were performed from 1200 UTC 12 October 1987 to 00 UTC 16 October 1987, in a cycle of analyses and 6-hour forecasts.

Two sets of 4D-Var experiments have been made: one with (hereafter denoted SSM1) and the other without SSM/I data (control run, hereafter denoted by CONTROL), both including the same observations as 3D-Var. In our experiments, the SSM/I values are the most numerous data, especially at times other than the four synoptic hours, but they represent only an integrated quantity.

The 4D-Var assimilation period spans 24 hours from 15 UTC 13 October 1987 to 15 UTC 14 October 1987. Then we make a 36-hour forecast starting at 12 UTC 14 October valid for 00 UTC 16 October. All these experiments (3D-Var and 4D-Var) are better understood by referring to Fig. 2. The period from 12 UTC 12 October to 15 UTC 13 October was the warm-up period of the assimilation using 3D-Var assimilation. We can compare the results of the 3D-Var and 4D-Var analyses from 15 UTC 13 October to 15 UTC 14 October. We continued the 3D-Var analysis after 15 UTC 14 October, in order to have a verification for the forecasts.
5. Structure functions

In order to estimate the impact of SSM/I data, we performed a single observation experiment in 3D-Var and 4D-Var. The shape of the analysis increments provides a three-dimensional picture of the covariances of the background error (modified by the dynamics in the case of 4D-Var).

(a) Structure functions specification in 3D-Var

In 3D-Var, the analysis increments, \( x^a - x^b \), are a linear combination of the observation increments, \( y^o - H x^b \), (Lorenc 1986)

\[
x^a - x^b = B H^T (H B H^T + R)^{-1} [y^o - H x^b].
\]

(9)

For a single observation \( H^T \) becomes a vector of the same dimension as the model state vector. Denoting this vector by \( k^T \) the previous expression becomes

\[
x^a - x^b = B k^T \left( \frac{y^o - H x^b}{\sigma_b^2 + \sigma_o^2} \right),
\]

(10)

where \( \sigma_b^2 \) is the effective background error variance of the model equivalent to the observation (\( \sigma_b^2 = k B k^T \)) and \( \sigma_o^2 \) the observation error variance. Introducing only one PWC observation at 15 UTC 14 October, the 3D-Var analysis differs from the background only at this time (and not before). The observation increment is 10 kg m\(^{-2}\) with an expected standard deviation of 4 kg m\(^{-2}\). Figure 3(a) presents the specific humidity increments (analysis minus background) at 850 hPa for a single PWC observation at 40°N, 15°W. They have an almost circular shape and are located around the position of the observation, in agreement with the isotropic specific humidity covariances. The vertical cross-section of the same parameter (Fig. 3(b)) is symmetric about the observation location. Since PWC is integrated in the vertical, the details in the vertical of the analysis increments are imposed by the specified prediction error covariance. It is larger at 1000 hPa and almost negligible at 500 hPa with a ‘barotropic’ structure, which corresponds to the specified structure functions for specific humidity.

(b) Structure functions in 4D-Var

Unlike static schemes such as 3D-Var, 4D-Var is able to specify flow-dependent structure functions. The analysis increments from a single datum are proportional to the error covariance of the background. Single observation experiments allow us to visualize the forecast-error covariances of the Kalman filter equivalent to 4D-Var (Thépaut et al. 1996). In the results presented here, the background-error covariances used at \( t_0 \) are identical to the 3D-Var background-error covariances.

If we consider the analysis increment at \( t_n \) we find:

\[
x^a(t_n) - x^b(t_n) = M B M^T k^T \left( \frac{y^o - H x^b}{\sigma_b^2 + \sigma_o^2} \right).
\]

(11)

The assimilation period spans 24 hours from 15 UTC 13 October to 15 UTC 14 October. We included (as in 3D-Var) a single PWC observation located at 40°N, 15°W at the end of the assimilation period (15 UTC 14 October).

Figure 4(a) shows the specific humidity increments (analysis minus background) at the end of the assimilation period (15 UTC 14 October) at 850 hPa. The spatial propagation of
Figure 3. Increments of specific humidity (analysis minus background) in 3D-Var at 850 hPa for 15 UTC 14 October 1987: (a) corresponding to a single observation of precipitable water content at 40°N, 15°W; (b) a vertical cross-section along the 40°N parallel of latitude. Contour intervals, with zero contour suppressed: (a) 0.2 g kg⁻¹; (b) 0.1 g kg⁻¹.

Figure 4. Increments of specific humidity (analysis minus background) in 4D-Var at 850 hPa for the end of the assimilation period, 15 UTC 14 October 1987: (a) corresponding to a single observation of precipitable water content at 40°N, 15°W (marked by a cross); (b) a vertical cross-section along the NW–SE solid line in (a). Contour interval: 0.2 k g⁻¹, with the zero contour suppressed.
the information within 4D-Var is significantly different from that in 3D-Var. The increments are clearly anisotropic and asymmetric with respect to the observation location. Another feature of 4D-Var is the existence of increments with opposite signs. We also notice south-west–east elongation of the system consistent with the dynamics of the forecast model for this situation and a south-east–north-west propagation of the information, leading to a second extremum (with opposite sign) near 45°N, 18°W. Figure 4(b) shows a vertical cross-section for the specific humidity increments. The cross-section location is depicted in Fig. 4(a). The information of the observation is more distributed in the vertical than in the 3D-Var experiment. The propagation with altitude seems to be north-westwards for the positive values and south-eastwards for the negative values. The humidity gradient is increased in the area of maximum humidity when using the observation.

For the wind and temperature increments (produced by the same observation) at the end of the assimilation period, the impact is also anisotropic (not shown). In the tangent linear model, advection is the dominant process in the evolution of the specific humidity (Eq. (6)). The wind evolution can therefore be inferred from specific humidity. Other equations of the dynamics explain the evolution of mass fields (temperature and geopotential fields). It should be mentioned that the relation between humidity change and temperature change is not included in the minimization process.

In 4D-Var, the structure functions are modified by the dynamics used to propagate the errors (Thépaut et al. 1996). This dynamics is a priori correct in non-saturated areas, but not in saturated ones and thus within precipitation areas. However, there is no major problem in the dynamics used in our case, even for the specific humidity (as suggested by the convergence of the incremental method shown in section 6).

6. Results of the assimilation

(a) Numerical results: the cost function

As shown in Fig. 5, the distance between the model solution and the observations decreased by 50% after 100 iterations (25 inner × 4 outer loops). The jumps in Fig. 5

![Cost function graph](image)

Figure 5. Cost functions for the SSM/I run (pecked line) and CONTROL run (solid line), normalized by their respective initial values.
correspond to the update of the trajectory at T106 resolution. Both experiments have the same behaviour. To go more deeply into this result, we determine the convergence of the incremental method. Following Courtier (1995) we can estimate an accuracy criterion: we compute the observation departure at point \( x \) for the full \( \mathbf{d}(x) \) and simplified \( \mathbf{d}s(x) \) problem, and at the end of the minimization both are computed at the point \( x + dx \), with the following definitions:

\[
\begin{align*}
\mathbf{d}(x) &= y^o - H[M(x)] \\
\mathbf{d}s(x) &= y^o - G[L(x)] \\
\mathbf{d}(x + dx) &= y^o - H[M(x + dx)] \\
\mathbf{d}s(x + dx) &= y^o - G[L(x)] - G[L(dx)],
\end{align*}
\]

where \( G \) is the simplified observation operator, \( L \) the simplified model, and \( G \) and \( L \) the linearized operators.

The accuracy criterion is then approximated by:

\[
\frac{a}{b} \approx \frac{|\mathbf{d}(x + dx) - \mathbf{d}(x)| - |\mathbf{d}s(x + dx) - \mathbf{d}s(x)|}{|\mathbf{d}s(x + dx) - \mathbf{d}s(x)|}.
\]

The numerator represents the level of approximation of the assimilation and the denominator the part taken into account. The observations for which this number is higher than one are rejected by the accuracy criterion. This could occur, for example, if the dynamics used is insufficiently accurate. This selection is made a posteriori. The data rejected by the criterion are still used in the system.

This criterion is evaluated at each update of the trajectory for SSM/I observations, radiosonde specific humidity (\( Q \)) data (TEMP) and AIREP wind (\( U \) and \( V \)) and temperature (\( T \)) data. Table 1 presents the percentage of observations rejected per update. The first column describes the observation type, the second column the middle of the time window (observations made during the same six-hour period are grouped together), the third column shows the number of observations, and the four following columns show the percentage of rejected observations at each update. SSM/I data do not have an atypical behaviour as compared to other data. The number of SSM/I data rejected is generally smaller than for TEMP and AIREP data. For example, for the first update and at the beginning of the assimilation period (\( \pm 3 \) h around 18 UTC), we obtain 15% of rejected observations for SSM/I, 21% for TEMP humidity data (but for a small number of observations) and 21% for AIREP data (average for \( U \), \( V \) and \( T \)). We cannot determine a specific area for rejected points. Higher percentages of observations are rejected the more times the trajectory is updated and the more that time passes during an assimilation. At the beginning we have a rough structure, but the more we increase the resolution the less the incremental approach is valid. At the end of the assimilation period (\( \pm 3 \) h around 12 UTC), for the first update, we obtain 31% of rejected observations for SSM/I data, 41% for TEMP humidity data (average for different levels) and 37% for AIREP data (average for \( U \), \( V \) and \( T \)). This indicates that SSM/I data do not damage the convergence and can be assimilated using an adiabatic dynamics to propagate the increments. We do not make the same estimation for the CONTROL run as we expect negligible differences because global convergence in the two experiments is similar. In Table 2, a global diagnostic of the different assimilation fits to the observations is presented. Panels (a), (b) and (c) represent the value of the observational cost function for different types of conventional observations: (a) for the northern hemisphere; (b) for the tropics and (c) for the southern hemisphere. The first column describes the observation type, the second the number of observations,
### TABLE 1. Percentage of Rejected Observations Per Update for Different Types of Observations

<table>
<thead>
<tr>
<th>Observation Type</th>
<th>Time (UTC)</th>
<th>Number of Observations</th>
<th>Update 1 (%)</th>
<th>Update 2 (%)</th>
<th>Update 3 (%)</th>
<th>Update 4 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIREP (V)</td>
<td>18</td>
<td>585</td>
<td>23</td>
<td>30</td>
<td>38</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>678</td>
<td>25</td>
<td>37</td>
<td>48</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>673</td>
<td>36</td>
<td>38</td>
<td>43</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>695</td>
<td>42</td>
<td>47</td>
<td>57</td>
<td>46</td>
</tr>
<tr>
<td>AIREP (U)</td>
<td>18</td>
<td>585</td>
<td>22</td>
<td>31</td>
<td>48</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>678</td>
<td>34</td>
<td>39</td>
<td>45</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>673</td>
<td>41</td>
<td>38</td>
<td>47</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>695</td>
<td>40</td>
<td>49</td>
<td>48</td>
<td>52</td>
</tr>
<tr>
<td>AIREP (T)</td>
<td>18</td>
<td>605</td>
<td>20</td>
<td>26</td>
<td>37</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>699</td>
<td>34</td>
<td>39</td>
<td>40</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>683</td>
<td>34</td>
<td>44</td>
<td>45</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>722</td>
<td>29</td>
<td>44</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>TEMP (Q) at 1000 hPa</td>
<td>18</td>
<td>33</td>
<td>21</td>
<td>36</td>
<td>26</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>107</td>
<td>30</td>
<td>32</td>
<td>47</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>98</td>
<td>42</td>
<td>56</td>
<td>46</td>
<td>47</td>
</tr>
<tr>
<td>TEMP (Q) at 700 hPa</td>
<td>0</td>
<td>575</td>
<td>36</td>
<td>44</td>
<td>48</td>
<td>50</td>
</tr>
<tr>
<td>TEMP (Q) at 500 hPa</td>
<td>12</td>
<td>607</td>
<td>40</td>
<td>48</td>
<td>49</td>
<td>56</td>
</tr>
<tr>
<td>SSM/I</td>
<td>18</td>
<td>7755</td>
<td>15</td>
<td>29</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>6004</td>
<td>20</td>
<td>32</td>
<td>40</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>7015</td>
<td>25</td>
<td>38</td>
<td>45</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>5037</td>
<td>31</td>
<td>39</td>
<td>46</td>
<td>47</td>
</tr>
</tbody>
</table>

The third the initial cost function before minimization, the fourth the final cost function after minimization when no SSM/I data are used and the fifth the final cost function when SSM/I data are used.

The fit is equivalent for the SSMI* run and the CONTROL run for almost all types of observations and all geographical areas. The impact of SSM/I data is neutral (if we take into account the statistical significance of the results, calculated at the 95% confidence interval of a $\chi^2$ test) in the southern hemisphere (panel (c)) but slightly positive in the northern hemisphere (panel (a)) and in the tropics (panel (b)), and particularly for AIREP data. We also present in Table 2 (panel (d)) the fit for different variables; this is in order to allow us to investigate the impact of SSM/I data on other parameters. Except for the humidity (relative or specific) parameter for which the results are slightly negative but not significant, the results are improved by the presence of SSM/I data. In other words, the extra data have been used consistently with the conventional ones. This is a good indication that the data are compatible with the usual ones and that 4D-Var is able to extract the additional information properly. If we compare these results with those obtained with earlier experiments performed with a simpler assimilating model (Filibreri 1993), they are slightly worse: the model has improved, consequently the SSM/I data have a smaller impact. In the previous case, the fit was better for all types of observations and all geographical areas.

* SSMI is the experiment (compared to CONTROL); SSM/I is the data name.
TABLE 2. CONTRIBUTION OF THE OBSERVATIONAL COST FUNCTION FOR VARIOUS OBSERVATION TYPES AND GEOGRAPHICAL AREAS, AND FOR SEVERAL PARAMETERS.

(a) Northern hemisphere, conventional observations

<table>
<thead>
<tr>
<th>Observation type</th>
<th>Number of observations</th>
<th>Initial cost function</th>
<th>Final cost function (no SSM/I)</th>
<th>Final cost function (SSM/I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNOP+DRIBU</td>
<td>14493</td>
<td>26009</td>
<td>12514</td>
<td>12523</td>
</tr>
<tr>
<td>AIREP</td>
<td>6748</td>
<td>14409</td>
<td>8034</td>
<td>7939</td>
</tr>
<tr>
<td>TEMP+PILOT</td>
<td>49415</td>
<td>78030</td>
<td>44153</td>
<td>43830</td>
</tr>
</tbody>
</table>

(b) Tropics, conventional observations

<table>
<thead>
<tr>
<th>Observation type</th>
<th>Number of observations</th>
<th>Initial cost function</th>
<th>Final cost function (no SSM/I)</th>
<th>Final cost function (SSM/I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNOP+DRIBU</td>
<td>3117</td>
<td>5968</td>
<td>2609</td>
<td>2690</td>
</tr>
<tr>
<td>AIREP</td>
<td>906</td>
<td>1932</td>
<td>1321</td>
<td>1229</td>
</tr>
<tr>
<td>TEMP+PILOT</td>
<td>6511</td>
<td>12491</td>
<td>6735</td>
<td>6773</td>
</tr>
</tbody>
</table>

(c) Southern hemisphere, conventional observations

<table>
<thead>
<tr>
<th>Observation type</th>
<th>Number of observations</th>
<th>Initial cost function</th>
<th>Final cost function (no SSM/I)</th>
<th>Final cost function (SSM/I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNOP+DRIBU</td>
<td>2139</td>
<td>12157</td>
<td>2280</td>
<td>2371</td>
</tr>
<tr>
<td>AIREP</td>
<td>314</td>
<td>1109</td>
<td>558</td>
<td>557</td>
</tr>
<tr>
<td>TEMP+PILOT</td>
<td>5570</td>
<td>8544</td>
<td>6518</td>
<td>6535</td>
</tr>
</tbody>
</table>

(d) Globe, cost function evolution for different parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number of observations</th>
<th>Initial cost function</th>
<th>Final cost function (no SSM/I)</th>
<th>Final cost function (SSM/I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U + V</td>
<td>54598</td>
<td>89990</td>
<td>50811</td>
<td>50600</td>
</tr>
<tr>
<td>T + Z + Ps + DZ</td>
<td>45159</td>
<td>75212</td>
<td>35205</td>
<td>34944</td>
</tr>
<tr>
<td>RH + Q</td>
<td>8327</td>
<td>17312</td>
<td>6515</td>
<td>6734</td>
</tr>
<tr>
<td>PWC</td>
<td>28178</td>
<td>24718</td>
<td>25977</td>
<td>11678</td>
</tr>
</tbody>
</table>

Statistically significant results are in bold type.

(b) Analyses for 12 UTC 14 October 1987

We present the result of the assimilation at 12 UTC 14 October 1987 because it is a synoptic time (near the end of the assimilation) so that analyses (operational or 3D-Var) are available and can help for validation. For the purpose of qualitative validation, the SSM/I PWC is used, even if these data are included in the analysis, because this parameter is a good indication of frontal structures.

In order to estimate the global evolution of the specific humidity, vertical cross-sections of the zonal mean of this parameter are presented (Fig. 6)*. The 4D-Var with SSM/I is shown in Fig. 6(a). We notice that the difference between 3D-Var and 4D-Var with SSM/I (Fig. 6(b)) is always positive, with largest values in the tropics (30°N–20°S) below 800 hPa (reaching 1 kg kg⁻¹). The 4D-Var with SSM/I has the effect of drying the

* Note that Figs. 6(d), 8(d) and 14(d) give the difference between the two 4D-Var as CONTROL minus SSMI, whereas Figs. 9(b), 10(b), 11(b), 12(d), 13(b) and 16(b) give the negative value of this. Because of a computer change at ECMWF it has not been possible to show Figures with the comparisons all in the same sense.
Figure 6. Vertical cross-sections of zonal-mean specific humidity for 14 October 1987: (a) analysis using 4D-Var SSMI; (b) difference between the analyses using 3D-Var and 4D-Var SSMI (3D-Var minus 4D-Var SSMI); (c) difference between the two 4D-Var analyses (CONTROL minus SSMI). Contour intervals: (a) 2 g kg$^{-1}$; (b) 0.2 g kg$^{-1}$, with zero contours suppressed; (c) 0.1 g kg$^{-1}$, with zero contours suppressed.
model results compared with 3D-Var. If we compare both 4D-Var experiments (Fig. 6(c)), the difference is less significant (reaching 0.3 g kg$^{-1}$) and presents a vertical structure. Around 20°N–0° and around 20°S the SSM/I run is wetter than the CONTROL run (SSMI is more coherent with 3D-Var); elsewhere it is the opposite. Compared to 3D-Var, 4D-Var is dryer with or without SSM/I. A possible interpretation is that the spin-up for 4D-Var is more important than for other systems. It increases the precipitation but dries up more than 3D-Var. We can make the hypothesis that time constants in 4D-Var lead to precipitation more rapidly than evaporation, than do those in 3D-Var. The hypothesis can be compared with results obtained for gravity waves by Courtier and Talagrand (1990) as a transient state to the attractor.

In the following, results for the eastern Atlantic region and for a region in the southern hemisphere (near Madagascar) are presented. Figure 7 represents the SSM/I PWC for 18 h local time 14 October (when the satellite crosses the equator for each track) which is sufficiently coherent in time with the analysis at 12 UTC in the area located near Madagascar (43°E–50°E). For the area located in the northern hemisphere, to estimate the PWC value at 12 UTC we also consider the PWC at 6 h local time (not shown). Figure 7 allows us to estimate the coverage of SSM/I data for 24 hours using only southbound passes of the satellite.

(i) The Atlantic zone. In Fig. 7 we notice a very wet tongue (reaching 50 kg m$^{-2}$) around 40°N–45°N. Figure 8 represents (a) the PWC field obtained from 3D-Var, (b) 4D-Var without SSM/I data, (c) 4D-Var using SSM/I data and (d) the difference (CONTROL
Figure 8. Precipitable water content at 12 UTC 14 October 1987 over the North Atlantic: (a) 3D-Var analysis; (b) 4D-Var analysis without SSM/I data; (c) 4D-Var analysis with SSM/I data; (d) difference between the two 4D-Var analyses (CONTROL minus SSMI). Contour intervals: (a), (b) and (c) 5 kg m$^{-2}$; (d) 2 kg m$^{-2}$, with zero contour suppressed.

minus SSMI) between both 4D-Var. In comparison with SSM/I data (Fig. 7) for almost the same period, the PWC shown in Fig. 8(a) is correct in the frontal area in both location and spatial extension, but the magnitude is slightly too high (almost 5 kg m$^{-2}$). Comparing panels (a) and (b) of Fig. 8 it is clear that 4D-Var without SSM/I (panel (b)) overestimates PWC, especially in the frontal zone. The introduction of SSM/I data improves the quality of the PWC analysis (panel (c)), since the magnitude and shape then become comparable to the 3D-Var one (panel (a)). The comparison between the two 4D-Var (panel (d)) shows PWC differences which can reach 10 kg m$^{-2}$ in the frontal area and near the African coasts. We observe a drying effect in the south-western region and an increase of PWC near the coasts. The positive impact of SSM/I (as it is more coherent with 3D-Var) is important because in such precipitable areas the 4D-Var dynamics is known to be incorrect. SSM/I data compensate for some 4D-Var deficiencies.

It is also interesting to see the impact of these data on the wind field (Fig. 9) in a strong and quickly developing system. Figure 9(a) represents the 250 hPa wind field without SSM/I and Figure 9(b) the difference between the two 4D-Var experiments. Both analyses exhibit an area of high winds around 45°N, 40°W with a maximum exceeding 80 m s$^{-1}$, corresponding to the upper-level jet and consistent with the ECMWF operational analysis.
One can notice an impact of SSM/I data over the Atlantic (panel (b)). The inclusion of SSM/I data reduces the activity of the system in the diffuent zone of the jet (around 45°N, 10°W), which is compatible with the shape of the PWC field (Fig. 7). Figure 10 presents the analysed sea-level pressure for the two 4D-Var experiments. Panel (a) represents results with SSM/I and panel (b) the difference between the two 4D-Var experiments. The main low is located at 55°N, 15°W with a value of 980 hPa in both analyses. This low-pressure area is extended by two troughs, one towards the tip of Brittany and the other towards the mid-Atlantic. The western one will be the one responsible for the storm mentioned in section 4. The associated low is located at approximately 43°N, 30°W with a pressure of 1000 hPa. This low is well located (compare its position with that shown in Fig. 1 for 12 UTC 14 October on the real trajectory (solid line)), although the pressure is overestimated. The difference between the two analyses can reach 2 hPa at the same location of maximum impact in PWC in the frontal area and over France. Especially in the region of the second trough, the structure is shifted. Figure 11 represents the temperature field at 850 hPa for the 4D-Var experiment with SSM/I and the difference between the two 4D-Var experiments. The lower-level baroclinic area is present in both analyses, as can be seen from the strong gradient around 45°N, 40°W. The difference between the two analyses can reach 2 K.
(ii) The southern hemisphere. An example of the impact of SSM/I data in the southern hemisphere (east of Madagascar) is given in Fig. 12, which represents the PWC field obtained from the 3D-Var analysis (a), the 4D-Var assimilation without SSM/I data (b), the 4D-Var experiment using SSM/I data (c) and the difference between the two 4D-Var experiments (d). For this zone we have a good spatial coherence at 12 UTC with the SSM/I observation at 18 h local time (Fig. 7). The observations exhibit two wet zones, oriented from north-west to south-east, the eastern one being wetter than the other. In comparison with SSM/I data, the PWC structures shown in Fig. 12(a) are correct in both location and magnitude. The comparison between 4D-Var experiments without and with SSM/I shows small differences (4 kg m$^{-2}$) in the south and larger differences in the north, reaching 9 kg m$^{-2}$ around 30$^\circ$S near the tropics. We obtain results closer to those of the 3D-Var experiment from the 4D-Var experiment with SSM/I. For the wind at 250 hPa, the biggest wind increment over the selected area corresponds to an area of a strong PWC gradient (not shown). In 4D-Var, the observation of PWC gives information on the advective part.
of the wind. For the surface pressure (Fig. 13) the structures are almost identical, though the difference can reach 3 hPa in the low.

To summarize, we can say that although slightly different, both 4D-Var analyses exhibit the basic features of the atmospheric flow but results with SSM/I are closer to observations and to 3D-Var results.

7. RESULTS OF THE FORECASTS

A 36-hour forecast has been calculated from the analysis of 1200 UTC 14 October 1987. We have archived results every 12 hours. We can compare results of 3D-Var and 4D-Var with and without SSM/I at 00 UTC and 12 UTC 15 October and at 00 UTC 16 October, and also with the 3D-Var analysis and with SSM/I observations for the same period (Fig. 2).
ASSIMILATION OF WATER CONTENTS

Figure 12. Precipitable water content at 12 UTC 14 October 1987 over the south Indian Ocean, south and east of South Africa: (a) 3D-Var analysis; (b) 4D-Var analysis without SSM/I data; (c) 4D-Var analysis with SSM/I data; (d) difference between the two 4D-Var analyses (SSMI minus CONTROL). Contour intervals: (a), (b) and (c) 5 kg m$^{-2}$; (d) 2 kg m$^{-2}$, with the zero contour suppressed.

(a) **Numerical results: fit to observations**

The observations are obviously not used in these forecast experiments but are available for comparison. If we consider the differences after a 12-hour run (00 UTC 15 October) between model state and observations for different parameters (Table 3(a)), we find that the results are better for mass fields (temperature, pressure, thickness and geopotential) in the case with SSM/I data and neutral for wind and humidity data.

The results after a 24-hour run (12 UTC 15 October) are better (Table 3(b)) for the run with SSM/I data in the assimilation part than those after the 12-hour run, except for humidity where they are neutral.

The results after a 36-hour run (00 UTC 16 October) are still good (Table 3(c)) except for the mass fields. Good results at this time for most parameters show that SSM/I data still have an impact after a 36-hour forecast. It is very important to notice that, after a 36-hour run, the SSM/I data have a positive impact on humidity and wind fields.

Firstly, we improve the dynamic fields, particularly the wind fields. As humidity is a passive tracer, the results for humidity are then improved. For the forecasted humidity field, the results for the wind fields are more important than the humidity data.

(b) **Forecasts for 00 UTC 16 October**

If we compare the vertical cross-sections of the zonal mean of specific humidity (Fig. 14), we find both 3D-Var and 4D-Var CONTROL forecasts are drier than the 3D-Var
analysis in the tropics below 700 hPa while the 4D-Var SSMI forecast is wetter than the analysis. Note that the 3D-Var experiment is known to have a small underestimation of the humidity in the tropics (at least for the IFS cycle used in this study). Outside the tropical regions and above 700 hPa in the tropics, the 3D-Var and 4D-Var CONTROL forecasts are more consistent with the analysis than the 4D-Var SSMI.

We now consider the 36-hour forecast valid 00 UTC 16 October for the North Atlantic region. Figure 15 displays the surface pressure for (a) the operational analysis (the O.I. used in 1987) at this date and time, (b) the 3D-Var forecast, (c) the 4D-Var forecasts without SSM/I data and (d) with SSM/I data. The main low (965 hPa) in the operational analysis is located slightly to the south of Cornwall, with a strong pressure gradient over Brittany. This value is located well enough but is not as low as the observed value (Fig. 1). The lows produced by all 36-hour forecasts are not deep enough and are located too far north. On the one hand, the value of 964 hPa in the forecast started from the 4D-Var without
Figure 14. Vertical cross-sections of zonal-mean specific humidity for 00 UTC 16 October 1987: (a) 36-hour forecast using 4D-Var SSMI; (b) difference between the 3D-Var analysis and the 36-hour forecast using 4D-Var SSMI (3D-Var minus 4D-Var SSMI); (c) difference between the 3D-Var 36-hour forecast and the 4D-Var SSMI 36-hour forecast (3D-Var minus 4D-Var SSMI); (d) difference between the two 4D-Var 36-hour forecasts (CONTROL minus SSMI). Contour intervals; (a) 2 g kg$^{-1}$; (b), (c) and (d) 0.2 g kg$^{-1}$, with zero contour suppressed.
### TABLE 3. DIFFERENCE BETWEEN THE MODEL FORECASTS AND THE OBSERVATIONS

(a) 00 UTC 15 October 1987

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number of observations</th>
<th>Final cost function (no SSM/I)</th>
<th>Final cost function (SSM/I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U + V</td>
<td>19088</td>
<td>30056</td>
<td>29847</td>
</tr>
<tr>
<td>T + Z + Ps + DZ</td>
<td>19885</td>
<td>27928</td>
<td>27747</td>
</tr>
<tr>
<td>RH + Q</td>
<td>6648</td>
<td>10980</td>
<td>11089</td>
</tr>
</tbody>
</table>

(b) 12 UTC 15 October 1987

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number of observations</th>
<th>Final cost function (no SSM/I)</th>
<th>Final cost function (SSM/I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U + V</td>
<td>19230</td>
<td>39982</td>
<td>38848</td>
</tr>
<tr>
<td>T + Z + Ps + DZ</td>
<td>20616</td>
<td>38245</td>
<td>36462</td>
</tr>
<tr>
<td>RH + Q</td>
<td>6521</td>
<td>12948</td>
<td>13024</td>
</tr>
</tbody>
</table>

(b) 00 UTC 16 October 1987

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number of observations</th>
<th>Final cost function (no SSM/I)</th>
<th>Final cost function (SSM/I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U + V</td>
<td>18600</td>
<td>48141</td>
<td>47279</td>
</tr>
<tr>
<td>T + Z + Ps + DZ</td>
<td>19451</td>
<td>57876</td>
<td>59853</td>
</tr>
<tr>
<td>RH + Q</td>
<td>6569</td>
<td>13738</td>
<td>13439</td>
</tr>
</tbody>
</table>

Statistically representative results are in bold type.

SSM/I (panel c) is slightly better than the 967 hPa in panels (b) and (d). On the other hand, the front is more marked in the 4D-Var (and slightly more with SSM/I) as in the 3D-Var forecast. For the winds at 250 hPa, the difference between both 4D-Var is small (not shown). For the winds at 850 hPa, the structures are almost identical. The difference between the two 4D-Var experiments reaches 8 m s⁻¹ in the frontal area (Fig. 16(a)). The 4D-Var experiment with SSM/I is in better agreement with the 3D-Var results (not shown). For the specific humidity at 850 hPa, Fig. 16(b) shows the difference between the two 4D-Var. The forecasts are very close. They locate the maximum wet tongue a little farther south than the analysis (not shown). In the case of 4D-Var with SSM/I, the structure is shifted slightly north, so that it is closer to the analysis. However, it is important to notice that none of the simulations represented the storm, even though the forecast with SSM/I was slightly better.

Finally, we have performed some hydrological diagnostics during the forecasts, to determine the spin-up of the model. The 4D-Var experiment with SSM/I has less water at the beginning of the forecast than the 4D-Var without SSM/I. This characteristic is still valid at the end of the forecast, whereas the water content increases more rapidly in the 4D-Var with SSM/I experiment. Thus, the spin-up is more important in the run with SSM/I. We have almost the same latent heat in both forecasts but the loss by precipitation (convective and stratiform) is less important in the SSM/I run. Therefore the experiment including SSM/I data has less activity. The solar radiation at height is less important in the SSM/I run: we have fewer lower clouds (which is consistent with the reduction in water content). The thermal radiation at height is also less important in the SSM/I run which indicates that we have higher clouds.
ASSIMILATION OF WATER CONTENTS

8. DISCUSSION AND CONCLUDING REMARKS

We have exploited the ability of 4-D variational assimilation to make consistent use of the information coming from both the observations and the model dynamics. Obviously there are many factors which show some evidence of the impact of SSM/I data in the 4D-Var.

Preliminary results (Rabier et al. 1992; Filiberti 1993) suggested the necessity of using physics in the model and not only in the forecast part. However even with a model without physics, we obtained a global improvement of the analysis in terms of fit to observations.

In the present study, we have shown that SSM/I data can be used consistently with other observations and model dynamics. We have also shown that 4D-Var copes nicely with the introduction of asynoptic high-frequency data without degrading the fit to other conventional observations. The results point to the fact that SSM/I data are neutral or even slightly improve the fit to other observations. SSM/I data slightly improve the convergence of the minimization. 4D-Var can benefit from a high-density data coverage, which makes SSM/I a good candidate for being operationally assimilated.

Another strong point in favour of 4D-Var is its intrinsic ability to allow us to infer information about non-directly observed components of the flow. This theoretical result can be of paramount importance for the improvement of wind analyses in the tropics for
instance, providing that some simplified moist physics is introduced in the tangent-linear model. The SSM/I observations are regularly spaced in time, but if one looks at the bias between model and observations, it is positive at the beginning and negative at the end of the assimilation period. Due to the absence of some physics, we put in more humidity at the beginning than at the end. If the SSM/I data are to be exploited fully, there are still some other processes which should also be included.

The results of this study of a specific meteorological situation show impacts on wind fields that are significant locally. The 4D-Var decreases the model humidity more than the 3D-Var, but it seems that SSM/I PWC can improve this in some areas.

We also notice the robustness of the impact of SSM/I data in a short-range forecast and that there is still a significant impact after 36 hours for humidity and wind parameters. SSM/I data are treated more like a passive tracer in the model, which can explain why the impact is more important for wind than for humidity.

The incremental approach can decrease the errors caused by the tangent-linear assumption, but can help little on the adiabatic assumption in the minimization process. For humidity information to be assimilated properly, adiabatic processes should be included.
The impact of SSM/I data in the forecast is rather neutral for this situation, but should be improved when the tangent-linear model physics is progressively upgraded.

ACKNOWLEDGEMENTS

For his useful comments and suggestions on this manuscript, we are very grateful to A. Hollingsworth (at ECMWF). We should also like to thank D. Vasiljevic and L. Phalippou (both also at ECMWF) for helpful discussions and advice during this study.

REFERENCES

Alishouse, J. C., Snyder, S. A., Vongsathorn, J. and Ferraro, R. R.

Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J., McNally, A. P., Kelly, G. and Courtier, P.

Atkins, M. J.

Buizza, R.

Courtier, P.

Courtier, P. and Talagrand, O.
1990 Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. Tellus, 42A, 531–549

Courtier, P., Freydiére, C., Geleyn, J.-F., Rabier, F. and Rochas, M.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A.

Dodge, J. C. and Goodman, H. M.
1994 The WeiNet project, Remote Sensing Reviews, 11, 1–4

Eymard, L., Klapaiz, C. and Bernard, R.
1989 Comparison between Nimbus-7 SSMR and ECMWF model analyses: The problem of the surface latent-heat flux. J. Atmos. Oceanic Technol., 6, 866–881

Filiberti, M.-A.
1993 Assimilation dans des modèles météorologiques de données de vapeur d’eau intégrée mesurées par radiométrie hyperfréquence spatiale. Thèse de Doctorat, University of Paris 7, France

Filiberti, M. A., Eymard, L. and Urban, B.

Gadd, A. J. and Morris, R. M.

Gilbert, J. Ch. and Lemaître, C.

Hollinger, J. P., Perce, J. L. and Poe, G. A.

Illari, L.
1989 The quality of satellite precipitable water vapour content data and their impact on analysis moisture fields. Tellus, 41A, 319–337

Jarraud, M., Goas, J. and Deyts, C.

Katsaros, K. B., Bhatti, I., McMurtrie, L. A. and Petty, G. W.
1989 Identification of atmospheric fronts over the ocean with microwave measurements of water vapour and rain. Weather and Forecasting, 4, 449–460

Liu, W. T., Tang, W. and Wentz, F. J.


Smagorinsky, J., Miyakoda, K. and Strickler, R. F. 1970 The relative importance of variables in initial conditions for dynamical weather prediction. Tellus, 22, 141–157


