13A.2 TOWARDS THE ASSIMILATION OF PRECIPITATION DATA FROM GROUND-BASED RADARS IN ECMWF 4D-VAR

Philippe Lopez *
ECMWF, Reading, UK

1. INTRODUCTION

The purpose of data assimilation is to produce a three-dimensional representation of the atmospheric state that is as close as possible to a set of available observations and that can be used to initialize numerical weather forecasts. In practice, this optimal state (called the *analysis*) is obtained by combining observational information with some *a priori* information coming from the numerical forecast model itself (*background* or *first-guess*). For this purpose, most operational centers for weather prediction around the world have opted for the variational assimilation method (usually 3D-Var or 4D-Var). In such approaches, the analysis is found by iteratively minimizing a cost function that measures the misfit between the unknown model state and both observations and background information, weighted with their respective errors.

After the assimilation of satellite microwave brightness temperatures in cloudy and rainy situations over ocean became operational in ECMWF's 4D-Var system in June 2005 (Bauer *et al.* 2006a,b), the feasibility and potential benefits of assimilating precipitation data from ground-based radars started to be investigated at ECMWF. The hope is that the assimilation of such data could lead to an improvement of analyses, in particular of the hydrological budget (precipitation analysis) over land.

Ground-based radar precipitation observations have the advantage of being complementary to satellite microwave brightness temperatures (TBs), which are currently assimilated over oceans only, because of large uncertainties in surface emissivities. At the same time, radar data can offer an excellent temporal sampling, in contrast to polar orbiting satellites.

However, it is now well recognized that the assimilation of precipitation observations is far from easy and full of uncertainties.

The purpose of this paper is twofold: first, results will be presented from experimental attempts to assimilate precipitation observations from the U.S.A. ground-based radar network in the ECMWF system using either a two-step 1D+4D-var approach or direct 4D-Var. Secondly, the main issues and uncertainties that affect the assimilation of rain measurements will be summarized.

* *Corresponding author address*: Philippe Lopez, ECMWF, Reading, RG2 9AX, UK; e-mail: **philippe.lopez@ecmwf.int**.

2. RAIN ASSIMILATION METHODS

Two different approaches have been tested to assimilate ground-based radar precipitation data: a two-step 1D+4D-Var method and direct 4D-Var.

2.1 1D+4D-Var

1D+4D-Var was originally developed by Marécal and Mahfouf (2003) and subsequently implemented in operations at ECMWF in June 2005 (Bauer *et al.* 2006a,b). First, rainy observations are passed to a 1D-Var procedure that retrieves individual profiles of temperature and moisture increments. Temperature increments are disregarded since they are often significantly smaller than moisture ones (true with ECMWF's physics at least). Moisture increments are then vertically integrated to create a pseudo-observation of Total Column Water Vapor (TCWV), which is then assimilated in the full 4D-Var system (as described in section 2.2), together with all other observation types.

In 1D+4D-Var experiments, ground-based radar precipitation rates (RR expressed in mm h^{-1}) were first converted to $\log_{10}$(RR) so as to make the statistical distribution of obs−model departures closer to Gaussian, as required in variational data assimilation.

2.2. Direct 4D-Var

As an alternative to the latter indirect approach, direct 4D-Var assimilation of radar rain rates has recently been tested. In this case, radar precipitation observations are assimilated together with all other observation types. ECMWF's 4D-Var uses an incremental formulation which is crudely summarized in Fig.1. Two main steps are to be distinguished:

1) the *trajectory*, which is a high resolution integration with the full forecast model over the 4D-Var window (12 hours, here), and

2) the *minimization*, which determines the analysis increments needed to improve the fit of the model evolution to the observations. The minimization is run at lower resolution and uses linearized simplified versions of the model physical parameterizations so as to cut computational cost and satisfy the linearity assumption of 4D-Var, which makes the cost function quadratic.

In practice, weak nonlinearities can be taken into account by completing several successive trajectory/minimization loops (3 in the present case). The minimization resolution is gradually increased (T95 ≈ 200 km, T159 ≈ 130 km and T255 ≈ 80 km) so that synoptic scales are adjusted first. In this study, trajectories are run at T511 (≈ 40 km) resolution.

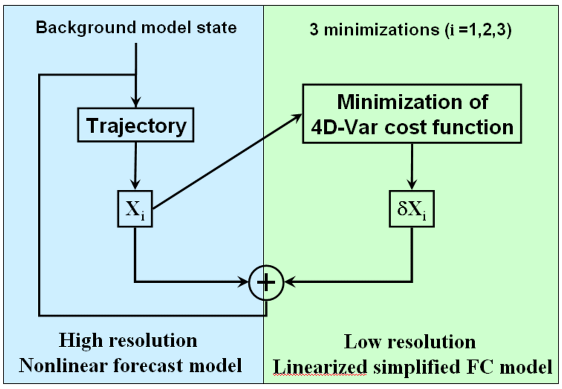

Fig.1. Schematic diagram of incremental 4D-Var.

In the preliminary direct 4D-Var experiments presented in this paper, NCEP Stage IV rain rates were converted to $\ln(RR_{mm/h} + 1)$ prior to assimilation to make background departures more Gaussian, as recommended by Mahfouf et al. (2007).

3. OBSERVATIONS

The observations used in this study are the NCEP* Stage IV surface precipitation analyses over the conterminous U.S.A.. These data combine the advantages of a relatively wide spatial coverage, spatial homogeneity (in terms of production and quality control) and straightforward availability. Rain rates retrieved from the NEXRAD (NEXt generation RADar) ground-based network are merged with rain gauge observations for calibration purposes (Fulton *et al.* 1998). Original hourly rain rates are obtained on a 4-km resolution polar stereographic grid; however, they are averaged onto ECMWF's model Gaussian grid before being assimilated, to ensure consistency in resolution between model and observations in the 4D-Var trajectory calculations.

4. EXPERIMENTAL RESULTS

4.1 1D+4D-Var

1D+4D-Var assimilation experiments of NCEP Stage IV precipitation data were run using ECMWF's 4D-Var system (12h window) at T511 spectral resolution and with 60 model vertical levels, for the period 20 May to 20 June 2005. Results will only be summarized here and more details can be found in Lopez and Bauer (2007).

It was found that assimilating NCEP Stage IV precipitation observations together with all other "usual" observations led to a modest improvement in standard forecast scores for geopotential, temperature, wind and moisture. A significant improvement was found over North America during the first three days of the forecast, but also downstream over Europe around days 7 and 8 of the forecast. Precipitation forecasts over the U.S.A. also got slightly better, but only for ranges shorter than 24 hours. This relative modest impact was at least partly attributed to the fact that the additional 1200 or so precipitation observations assimilated in each 12h 4D-Var cycle had to compete with comparatively low-error conventional observations, such as radiosondes or SYNOP measurements.

This was confirmed by a set of denial experiments in which all humidity-sensitive measurements from radiosondes, surface stations and satellite were not assimilated over the U.S.A.. Figure 2 shows how the mean TCWV from 00Z analyses is affected (a) by the removal of all conventional moisture-related observations over the domain and (b) by the assimilation of NCEP Stage IV rain data as the only source of humidity information over the U.S.A.. It is clear that the precipitation observations alone allow to recover from the strong dry bias caused by the withdrawal of humidity observations over the U.S.A..

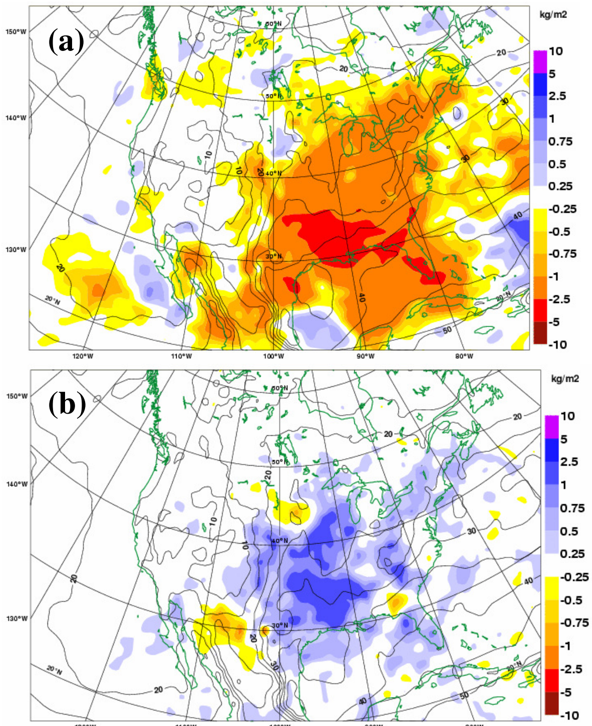

Fig.2. Mean impact on 4D-Var TCWV analyses of (a) discarding all moisture-sensitive observations over the U.S.A. in 4D-Var and (b) of assimilating NCEP Stage IV rain rate observations through 1D+4D-Var as the only source of moisture information over the domain. Units are in kg m^{-2}.

---

* National Centers for Environmental Prediction (U.S.A.).

4.2. Direct 4D-Var

An interesting feature of direct 4D-Var is that precipitation observations can be accumulated in time before being assimilated. Mahfouf and Bilodeau (2007) showed that the linearity assumption, central to 4D-Var, is better verified for rainfall accumulated over several hours than for instantaneous rain rates. This is confirmed in Fig.3 that displays a Taylor diagram of the statistical match in observation space between T95 linear observation–model departures produced by the first minimization and corresponding departures computed from the following T799 (≈ 25 km) nonlinear trajectory.

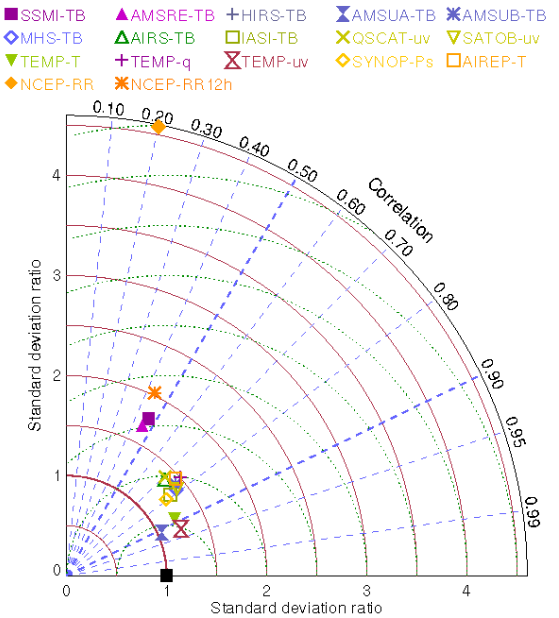

Fig.3. Taylor diagram displaying the statistical match in observation space between T95 linear model-observation departures from the first minimization and corresponding departures from the following T799 nonlinear trajectory (single 4D-Var cycle). Each symbol shows a given observation type and variable (see top legend).

In Fig.3, each symbol corresponds to a given observation type and variable. For the linearity assumption to be valid, both the standard deviation ratio (radius) of low-resolution linear to high-resolution nonlinear departures and the corresponding correlation coefficient (azimuth) should be close to 1 (black square). It is clear that for all observations that are not related to moisture, the linearity assumption is not perfect but not too wrong either. On the other hand, for observations directly affected by clouds and precipitation (SSM/I and AMSR-E TBs), the validity of the linear hypothesis is somewhat degraded. The latter becomes even poorer for "instantaneous" NCEP Stage IV rain data ("NCEP-RR"), but when the same observations are accumulated over 12 hours prior to assimilation ("NCEP-RR12h"), linearity is substantially improved.

First experiments with direct 4D-Var assimilation of NCEP Stage IV 12h accumulated precipitation data have recently been run with a spectral truncation of T511 and 91 vertical levels, over the period 1 to 30 April 2009. One should note that only NCEP Stage IV observations located east of 105ºW and over flat terrain were assimilated in these experiments, to avoid problems associated to the use of radars in mountainous regions. Besides, observations were also rejected in conditions of low-level freezing or favorable to anomalous propagation (as diagnosed from model fields according to Lopez 2009). On average, about 500 rain observations were assimilated every 12 hours.

Results show that observation–model precipitation departures significantly decrease after assimilation, as illustrated in Fig.4. In particular, the standard deviation of the departures drops from 0.265 to 0.209 (logarithmic space). However, there is also some indication of a slight over-drying of the model (positive departures after analysis). It is also nice to note that the distributions are roughly Gaussian, which is important to ensure 4D-Var optimality.

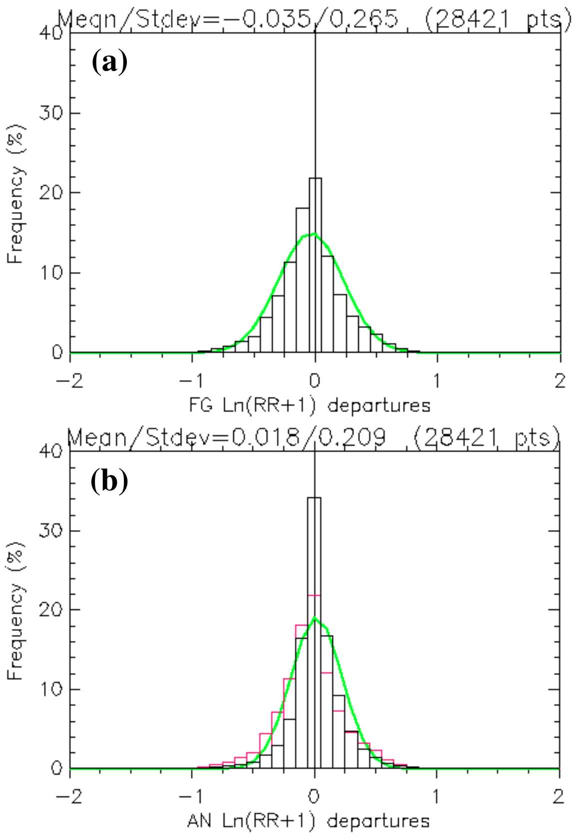

Fig.4. Frequency distribution of (a) first-guess and (b) analysis observation–model departures from a month-long experiment with direct 4D-Var assimilation of NCEP Stage IV 12h-accumulated rain amounts (April 2009). Fitted Gaussian distributions are plotted in green, and in panel (b) first-guess histogram is superimposed in red, for comparison purposes. Departures are in $ln(RR_{mm/h} + 1)$ space.

Furthermore, Table 1 indicates that the assimilation of NCEP Stage IV data bring the monthly averaged precipitation computed from

24h model forecasts closer not only to the NCEP Stage IV estimates, as expected, but also to independent PRISM* 4-km resolution high-density rain gauge analyses. Mean biases and root mean square (RMS) differences are reduced by 50-60% and 6-8%, respectively.

MODEL	OBS	Mean		Bias	RMS
		Model	Obs		
CTRL	NCEP	3.38	3.01	0.37	1.19
	PRISM		2.92	0.46	1.04
EXPER	NCEP	3.16	3.01	0.15	1.11
	PRISM		2.92	0.24	0.95

Table 1. Statistical comparison for April 2009 of daily precipitation from 24h model forecasts with NCEP Stage IV and PRISM observations. CTRL = control 4D-Var; EXPER = 4D-Var with assimilation of NCEP Stage IV 12h accumulated RR. Units are mm day^{-1}.

However, standard forecast scores for temperature, geopotential and moisture are slightly degraded (not shown), which suggests that several issues will require further investigation before operational application can be envisaged.

5. ISSUES

One of the main shortcomings of 1D+4D-Var in general is the fact that the model background is used twice (once in 1D-Var, once in 4D-Var), which can reduce the impact of precipitation observations in the final 4D-Var analyses. Two other limitations, specific to our implementation of 1D+4D-Var, stems from the loss of information in the vertical due to the use of TCWV pseudo-observations and from the discarding of temperature increments in 1D-Var. Also, 1D+4D-Var cannot properly deal with time-accumulated precipitation measurements, unlike direct 4D-Var.

In addition, several major issues exist that can hinder the successful assimilation of precipitation observations, most of them common to 1D+4D-Var and to the preferred direct 4D-Var method. These issues can be summarized as follows.

5.1. Nonlinearities

In 4D-Var, tangent-linear and adjoint computations of the minimization are performed using a set of linearized simplified physical parameterizations, designed in such way as to achieve a compromise between realism, nearness to their full nonlinear counterparts, computational efficiency and linearity. Therefore, by construction, the strong nonlinearities sometimes involved in condensation and microphysical processes cannot be properly represented during the minimization, leading to sub-optimality of the 4D-Var analyses. The effects of mismatching parameterizations between minimizations and trajectories was already seen in Fig.3 for observations affected by clouds or (worse) precipitation such as SSM/I and AMSR-E TBs and NCEP Stage IV rain data. However, Fig.3 also showed that accumulating precipitation observations can substantially improve the validity of the linearity assumption. Tremendous efforts are devoted to keeping linearized simplified parameterizations as close as possible to the full nonlinear ones, but the linear framework imposed by 4D-Var will always be a serious limitation.

5.2 Resolution differences

Another potential source of sub-optimality in 4D-Var analyses arises from the fact that the horizontal resolution is always lower in the minimizations than in the trajectories (see Fig.1), mainly to decrease computational cost. Such differences can be significant for precipitation which often exhibits a high spatial variability and a strong dependence on resolution. Again, assimilating accumulated precipitation can be beneficial.

5.3 Structure functions

The shape of the 4D-Var analysis increments is largely determined by the so-called *structure functions*, or in other words, by the model background error covariances (i.e. matrix **B**, with usual notation). It is suspected that the current structure functions used in ECMWF's 4D-Var might be too wide, too isotropic and not flow-dependent enough to ensure an optimal extraction of precipitation information. This means that only observations located in widespread regions with coherent positive or negative background departures can have a significant impact in 4D-Var.

5.4 "0-rain" issue

Whenever the model background is non-rainy, the adjoint sensitivity of the simulated precipitation to the input model variables becomes zero and consequently the associated precipitation observations have no impact on the analysis.

Symmetrically, whenever the observations are non-rainy, there is a strong ambiguity about the corresponding atmospheric state in terms of temperature and moisture profiles, in particular. This ambiguity can only be alleviated if other types of observations are available in the vicinity (e.g. radiosoundings).

As a consequence, only points where there is simultaneously precipitation in the model background and in the observations are currently assimilated. However, it should be noted that assimilating time-accumulated instead of instantaneous observations does reduce the overall occurrence of non-rainy

* Parameter-elevation Regressions on Independent Slopes Model (http://www.prismclimate.org).

points (due to the propagation of precipitating systems).

5.5 Asymmetry of increments

In earlier assimilation experiments with NCEP Stage IV precipitation data as well as in the operational assimilation of satellite microwave TBs (Geer *et al.* 2008), it was noticed that it was always easier to reduce model precipitation than to increase it during the analysis. In more recent tests, such asymmetry could be successfully reduced by switching off the systematic clipping of humidity increments above saturation implemented in ECMWF's standard 4D-Var. The remaining asymmetry might originate from differences in the sensitivities of moist physical parameterizations and from the use of a normalized relative humidity control variable in 4D-Var.

5.6 Observation and representativity errors

The specification of error statistics for precipitation observations and their associated representativity is still uncertain. Currently, the error standard deviation is arbitrarily set to roughly 20% of the actual observation amount and horizontal correlations have been neglected so far. These assumptions will deserve further attention in the near future. In particular, error spatial correlations might not be negligible for precipitation estimates derived from radars.

6. SUMMARY AND CONCLUSIONS

The assimilation of NCEP Stage IV precipitation data over conterminous U.S.A. have been tested at ECMWF in month-long global experiments, using both a two-step 1D+4D-Var technique and a direct 4D-Var approach. 1D+4D-Var experiments with hourly NCEP Stage IV hourly observations led to a slight improvement in precipitation forecasts and in temperature, geopotential and wind forecast scores mainly over North America and for ranges shorter than 24 hours. Denial 1D+4D-Var experiments with no operational moisture-sensitive observations over mainland U.S.A. exhibited a large improvement in the moisture field and significantly better forecast scores of other variables over North America in the first 4-5 days when radar rain rates were assimilated as the sole source of information about humidity. This suggests that the full benefit of the precipitation data in full 4D-Var experiments might not be obtained because of the competition with other more conventional measurements.

As a preferred alternative, direct 4D-Var assimilation of NCEP Stage IV rain data has been developed and has started to be tested. This approach makes it possible to assimilate accumulated precipitation over several hours, which improves the validity of the linearity assumption and reduces the occurrence of problematic non-rainy points in model and observations. First direct 4D-Var experiments suggest that model precipitation can be significantly improved, but traditional forecast scores are still slightly degraded, especially at short ranges.

Many issues still remain to be addressed to improve the overall performance of direct 4D-Var. These include the question of the discrepancies between 4D-Var low-resolution linear minimizations and high resolution nonlinear trajectories, the possible inadequacy of structure functions, the specification of observation errors (including spatial correlations), the "0-rain" issue and the asymmetry of increments (depending on the sign of background observation–model departures).

Once all these practical problems have been alleviated, it should become possible to assimilate precipitation data from ground-based radar networks in ECMWF's operational system, provided of course that data quality and availability requirements are satisfied.

Acknowledgments

NCEP and JOSS/UCAR* should be acknowledged for processing and providing, respectively, the Stage-IV precipitation observations used in this study. PRISM monthly precipitation data were provided by the Climate Group at Oregon State University (http://www.prismclimate.org).

References

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

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, S. Saarinen, and M. Bonazzola, 2006b: Implementation of 1D+4D-Var assimilation of precipitation affected microwave radiances at ECMWF, Part II: 4D-Var, *Q. J. R. Meteorol. Soc.*, 132, 2307-2332.

Fulton, R. A., J. P. Breidenbach, D. J. Seo, D. A. Miller and T. O'Bannon, 1998: The WSR-88D rainfall algorithm, *Weather Forecast.*, 13, 377-395.

Geer, A., P. Bauer and P. Lopez, 2008: Lessons learnt from the operational 1D+4D-Var assimilation of rain- and cloud-affected SSM/I observations at ECMWF., *Q. J. R. Meteorol. Soc.*, 134, 1513-1525.

* Joint Office for Science Support / University Corporation for Atmospheric Research (U.S.A.).

Lopez, P., and. P. Bauer, 2007: "1D+4D-Var" Assimilation of NCEP Stage IV Radar and Gauge Hourly Precipitation Data at ECMWF, *Mon. Weather Rev.*, 135, 2506-2524.

Lopez, P., 2009: A 5-year 40-km Resolution Global Climatology of Super-refraction for Ground-based Weather Radars, *J. Appl. Meteor.*, 48, 89-110.

Mahfouf , J.-F., B. Brasnett, and S. Gagnon, 2007: A Canadian Precipitation Analysis (CaPA) Project: Description and Preliminary Results, *Atmos. Ocean.*, 45, 1-17.

Mahfouf , J.-F., and B. Bilodeau, 2007: Adjoint Sensitivity of Surface Precipitation to Initial Conditions, *Mon. Weather Rev.*, 135, 2879-2896.

Marécal, V., and J.-F. Mahfouf, 2003: Experiments on 4D-Var assimilation of rainfall data using an incremental formulation, *Q. J. R. Meteorol. Soc.*, 129, 3137-3160.