TECHNICAL MEMORANDUM

627

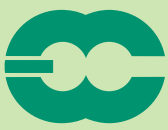
Direct 4D-Var assimilation of NCEP Stage IV radar and gauge precipitation data at ECMWF

Philippe Lopez

Research Department

To be submitted to Monthly Weather Review

July 2010



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

Series: ECMWF Technical Memoranda

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

Contact: library@ecmwf.int

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

Abstract

Direct four-dimensional variational (4D-Var) data assimilation of NCEP Stage IV radar and gauge precipitation observations over the Eastern USA have been developed and tested in ECMWF's Integrated Forecasting System. This is the natural extension of earlier work using a two-step 1D+4D-Var approach. Major aspects of the implementation are described and discussed in this paper. In particular, it is found that assimilating 6-hour precipitation accumulations instead of the original hourly data substantially improves the behaviour of 4D-Var, especially as regards the validity of the tangent-linear assumption.

The comparison of background and analysis precipitation departures demonstrate that most of the information contained in the new precipitation observations is properly assimilated. Experiments run over the periods April-May and September-October 2009 also show that local precipitation forecasts become significantly better for ranges up to 12 hours, which indicates that a genuine precipitation analysis can now be obtained over the Eastern USA. Geopotential, temperature, moisture and wind forecast scores are generally neutral or slightly positive for all regions of the globe and at all ranges, which is consistent with previous 1D+4D-Var results.

The most crucial issue that remains unsolved is the treatment of non-precipitating model background occurrences because of the corresponding absence of sensitivity in the linearized moist physics. For the moment, only points where both model background and observations are rainy are assimilated. Operational implementation using American data is planned in 2011 and one can hope that new networks of radars (and maybe rain-gauges) can be added in the 4D-Var assimilation process in the future.

1 Introduction

Atmospheric moist processes which govern the life cycle of clouds and precipitation and the Earth's hydrological cycle, currently remain one of the most uncertain components of any numerical weather prediction (NWP) model. The main reason for this lies in the extreme complexity, diversity, strong natural variability and lack of predictability of

(1) moist processes themselves (condensation/evaporation, heterogeneous/homogeneous nucleation, collection/aggregation, phase changes,...),

(2) the characteristics of particles involved (droplets of various sizes, ice aggregates with various shapes, sizes, densities and fall velocities),

(3) the concentration distributions of hydrometeors.

At the same time, global-scale observations that are routinely available from satellite imagers, surface spaceborne radars/lidars, ground-based radars and rain-gauge networks can only provide partial and generally temporally and spatially integrated information about microphysical processes. Besides, our knowledge on clouds and precipitation from field experiments is restricted to specific meteorological situations and regions of the globe. As a result of all these uncertainties, current parameterizations of moist processes used in operational NWP models remain simplified representations of the truth, based on the explicit prediction of a few categories of particles (e.g. cloud liquid water, cloud ice, snow and rain) and with a few microphysical processes accounted for. As a consequence, the forecast skill for clouds and precipitation is often much poorer than the skill for temperature, wind or water vapour, even after a single day of forecast. This is particularly true and systematic over certain regions of the globe such as West Africa during the summer monsoon (Agustí-Panareda *et al.* 2010) or in specific meteorological situations such as those characterized by unorganized convection or trade-wind stratocumuli.

Another significant source of inaccuracy in NWP outputs lies in the uncertainty in the three-dimensional (3D) atmospheric states that are provided as initial conditions to forecast models. In the last four decades, various

data assimilation (DA) techniques were developed and successfully used in an operational context to constrain the initial model state towards a set of reliable observations that are available within a few hours of the analysis time. In the variational DA approach (e.g. Le Dimet and Talagrand 1986), this optimal state (or *analysis*) is obtained by searching for the model state that best fits all the observations and some a priori (or *background*) information from the model, in a least-square sense. Temperature, wind and surface pressure information from radiosondes, synoptic stations and satellites which have now been operationally assimilated for several decades, were followed by water vapour measurements and only recently by observations affected by clouds and precipitation. The assimilation of observations that are strongly affected by moist atmospheric processes has turned out to be extremely challenging for two main reasons. First, adequate observation operators had to be developed to link the observable quantity (e.g. a radiance or a radar reflectivity) with the assimilation state vector which typically consists of temperature, specific humidity and wind profiles and surface pressure. Such an operator would usually combine a parameterization of moist processes and a radiative transfer model. Secondly, since linearity is one of the main underlying assumptions of the now widespread variational DA methods (such as 3D or 4D-Var), the strong non-linearities which often characterize atmospheric moist processes have to be overcome through the design of purpose-built sets of moist physical parameterizations for both convective (subgrid-scale) and stratiform (large-scale) processes. Indeed, any parameterization meant to be employed in the minimization of the variational cost function has to be differentiable, smooth, computationally efficient and at the same time as realistic as possible (e.g. Mahfouf 1999; Janisková *et al.* 1999), a hard-to-reach compromise.

These issues started to be successfully (but still partially) addressed during the last decade, so that operational 3D or 4D-Var systems are now able to assimilate polar-orbiting satellite all-skies microwave radiances from SSM/I, AMSR-E and TMI, for instance (see Appendix 1 for abbreviations). Bauer *et al.* (2006a, 2006b) gave an example of such operational implementation at ECMWF, using the two-step 1D+4D-Var approach originally designed by Marécal and Mahfouf (2003; hereafter MM03). The latter was recently replaced by the direct assimilation of all-sky microwave radiances in 4D-Var, as described in Bauer *et al.* (2010) and Geer *et al.* (2010). On-going work is also aiming at the assimilation of infrared radiances from geostationary satellites in cloudy regions (e.g. Vukicevic *et al.* 2006).

As far as precipitation-related observations are concerned, some operational weather services already assimilate information from ground-based weather radar networks: Macpherson (2001) implemented a latent heat nudging technique at the UK Met Office, while Treadon *et al.* (2002) used NCEP's 3D-Var and Tsuyuki *et al.* (2002) JMA's 4D-Var. Ducrocq *et al.* (2002) used a diabatic initialization technique based on Meteosat infrared imagery and ground-based precipitation radar data to improve short-range high-resolution forecasts of convective storms over France. The Ensemble Kalman Filter technique was tested by Tong and Xue (2005) and Caya *et al.* (2005) as an alternative to 4D-Var for the assimilation of simulated ground-based radar volumetric data on the convective scale. More recently, Caumont *et al.* (2010) presented some promising results from their pre-operational 1D+3D-Var assimilation of ground-based radar reflectivity profiles at 2.5 km resolution. In their approach, a Bayesian retrieval first yields relative humidity profiles which are then assimilated as pseudo-observations in their 3D-Var system. One advantage of this method is to avoid the problematic coding of tangent-linear and adjoint versions of their complex observation operator which includes detailed microphysics.

Lopez and Bauer (2007; LB07 hereafter) employed MM03's 1D+4D-Var method to experimentally assimilate hourly combined radar and gauge surface precipitation estimates from the NCEP Stage IV archive (Baldwin and Mitchell 1996; Lin and Mitchell 2005) over the conterminous USA. Rain rates were first assimilated through 1D-Var to produce total column water vapour (TCWV) retrievals, which were then passed as pseudo-observations to the 4D-Var system. LB07 found that the main impact of the additional observations was to reduce errors in short-range precipitation forecasts (up to 24 hours) over the USA. Short-range forecasts of temperature, wind and geopotential were also improved, to a lesser extent though. A hint of an eastward

propagation of this positive impact over the North Atlantic and Europe during the forecast was also identified. The relatively weak impact of the assimilated precipitation data was attributed to their competition with other observation types such as radiosondes and synoptic station data and to shortcomings specific to the 1D+4D-Var approach (e.g. double use of the model background). Finally, denial 1D+4D-Var experiments, in which NCEP Stage IV precipitation estimates were the only source of moisture-related information assimilated over mainland USA, highlighted the strong potential positive impact of these data.

As a natural evolution of LB07's indirect 1D+4D-Var approach, the present paper describes results from new experiments in which NCEP Stage IV precipitation data have been *directly* assimilated in ECMWF's 4D-Var system, which offers more consistency with the way all other observations are treated in daily operations. Direct 4D-Var assimilation of NCEP Stage IV data is expected to become operational at ECMWF in 2011.

The NCEP Stage IV observations to be assimilated are described in section 2. Section 3 introduces the 4D-Var assimilation method in general terms while section 4 provides details specific to the assimilation of NCEP Stage IV data as implemented in this study. Results from direct 4D-Var assimilation experiments are presented in section 5, while remaining issues are discussed in section 6. Section 7 summarizes the main findings of this study and gives an outlook on the future assimilation of ground-based precipitation observations at ECMWF.

2 NCEP Stage IV precipitation data

The new observations to be assimilated in this study are NCEP Stage IV precipitation data which combine precipitation estimates from about 150 Doppler NEXt-generation RADars (NEXRAD) with about 5,500 hourly rain-gauge measurements over the conterminous USA (Baldwin and Mitchell 1996; Lin and Mitchell 2005). Technically speaking, NEXRAD corresponds to the so-called WSR-88D (Weather Surveillance Radar, 1988, Doppler) (Fulton *et al.* 1998). Each NCEP Stage IV precipitation analysis is initiated 35 min after the end of each hourly collection period and may be updated over a period of several hours with new data coming from the twelve USA regional centres. A first inflow of automatically generated precipitation data is available within a few hours after the accumulation time, while a second inflow of updated manually-quality-controlled data becomes available later (with a delay of up to 12 hours). The spatial coverage of the early release is usually not far away from its maximum extent. In this work, 4D-Var experiments have been performed using manually quality controlled data obtained from the JOSS/UCAR archive (website: *http://www.joss.ucar.edu/codiac/*). Even though original precipitation data are available on a 4-km resolution polar-stereographic grid, they are averaged on the ECMWF model's Gaussian grid prior to assimilation. Besides, although hourly precipitation accumulations are obtained from the archive, it will be shown in section 4 that it is preferable to assimilate 6 hourly accumulations. In the following, the observations used in this study will be referred to as "NEXRAD" observations for simplicity.

3 The 4D-Var method

The aim of 4D-Var assimilation is to find the optimal initial 3D atmospheric state (the *analysis*) that leads to a short-range model forecast that best fits a set of observations and some a priori information from the model (the so-called *model background* or *trajectory*) over a certain time window (typically up to 12 hours). Formally, the analysis corresponds to the initial model state vector, $\mathbf{x}_0$, which minimizes the following cost function, J,

$$J = \underbrace{\frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b)}_{J_b} + \underbrace{\frac{1}{2}\sum_t (H_t(\mathbf{x}_0) - \mathbf{y}_t)^T \mathbf{R}^{-1}(H_t(\mathbf{x}_0) - \mathbf{y}_t)}_{J_o} + J_c \quad (1)$$

where subscript t denotes the model time step and $\mathbf{x}_0^b$ is the model background state at initial time. The term J_c corresponds to an additional weak constraint for the control of fast gravity waves using the digital filter approach developed by Gauthier and Thépaut (2001). In ECMWF's 4D-Var system, the model state consists of temperature, humidity, vorticity, divergence and surface pressure. H_t is the often non-linear observation operator used for converting the initial model state into observed equivalents at time t. All observations available in the assimilation window are gathered in vector $\mathbf{y}_t$. $\mathbf{R}$ and $\mathbf{B}$ are respectively the observation and model background error covariance matrices. $\mathbf{B}$ is made flow-dependent through a wavelet formulation (Fisher 2004).

In practice at ECMWF, J is re-formulated using an incremental approach (Courtier *et al.* 1994) as

$$J = \frac{1}{2}\delta\mathbf{x}_0^T\mathbf{B}^{-1}\delta\mathbf{x}_0 + \frac{1}{2}\sum_t(\mathbf{H}_t\delta\mathbf{x}_0 - \mathbf{d}_t)^T\mathbf{R}^{-1}(\mathbf{H}_t\delta\mathbf{x}_0 - \mathbf{d}_t) + J_c \tag{2}$$

where $\delta\mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_0^b$ are increments relative to the model background state, $\mathbf{d}_t = \mathbf{y}_t - H_t(\mathbf{x}_0^b)$ is the so-called innovation vector and $\mathbf{H}_t$ is the tangent-linear version of H_t (i.e. the matrix of local derivatives of the observation operator with respect to each variable of the model state vector). It is also essential to note that in each 4D-Var cycle three successive minimizations are performed at T95 (≈200 km), T159 (≈130 km) and finally T255 (≈80 km) horizontal resolution. After each minimization, the model trajectory is recomputed at high resolution (T511≈40 km in this study; T1279≈15 km in current ECMWF operations) to update innovation vectors, $\mathbf{d}_t$. Starting with the lowest resolution ensures that larger scales are adjusted first, reduces the computational cost of 4D-Var and permits the inclusion of weak non-linearities such as those in H_t. The latter advantage allows to partially overcome the rather constraining underlying assumption of linearity inherent in the variational method.

Of particular interest for the assimilation of precipitation observations, simplified paramaterizations of convection (Lopez and Moreau 2005) and large-scale moist processes (Tompkins and Janisková 2004) are used in each minimization. Other processes also represented in $\mathbf{H}_t$ include radiation (Janisková *et al.* 2002), vertical diffusion and gravity wave drag (Mahfouf 1999).

4 Implementation details of NEXRAD assimilation in 4D-Var

4.1 Change of variable

To better satisfy the requirement of Gaussian distributions of observation errors in 4D-Var and to avoid the sub-optimality of the 4D-Var analysis (Errico *et al.* 2000), a logarithmic transform, namely $\ln(RR+1)$, is applied to observed and model equivalent precipitation amounts (RR, expressed in mm h^{-1}) before the assimilation. Such transform was successfully used by Mahfouf *et al.* (2007) to produce precipitation analyses over Canada. Besides, in order to better satisfy the 4D-Var linearity assumption, as discussed in more details in section 6.1, it was decided to consider 6-hour precipitation accumulations ($RR6h$, hereafter), still expressed in mm h^{-1}, instead of the original hourly precipitation data. This means that all precipitation departures involved in J_o (see Eq. (2)) are expressed in terms of $\ln(RR6h+1)$.

4.2 Screening of observations

As far as screening is concerned, NEXRAD precipitation data are rejected over steep or rugged orography to account for the possible degradation of radar and rain-gauge measurement quality due to ground clutter, radar beam blocking, local precipitation enhancement or poorer representativity of rain gauges. In practice, all observations located west of 105°W are discarded to avoid the Rocky Mountains area. Over the remaining eastern

half of the USA, NEXRAD observations are also rejected if either the orography is higher than 1500 m or if the standard deviation of the orography is higher than 100 m (on the original NEXRAD 4-km grid). This eliminates data over the highest terrain of the Appalachian Mountains. NEXRAD observations are also discarded if snowfall is expected at ground level (namely if model background 2-meter temperature is below $+1^{\circ}$C) to account for the fact that radar and rain gauge measurements can become less accurate in such situations.

Despite the quality control applied by NCEP when creating the original precipitation dataset, an additional screening is applied to points that are likely to be affected by the occurrence of ducting in the lower troposphere. Ducting corresponds to the anomalous propagation of electromagnetic waves in the atmosphere, which can lead to the mis-interpretation as precipitation of ground echoes returned towards ground-based radars receivers. In this work, ducting is assumed to be present whenever atmospheric refractivity, N, sharply decreases with height ($dN/dz < -0.157$ m^{-1}). N is diagnosed from model background profiles of temperature, T (in K), water vapour partial pressure, e (in Pa), and total atmospheric pressure, P (in Pa), using

$$N = \frac{0.776P}{T} + \frac{3730e}{T^2} \tag{3}$$

Therefore, conditions favourable to ducting require the existence of either temperature inversions and/or sharp negative vertical gradients of humidity. Such situations are more likely to be encountered in the lower troposphere, during nighttime or over water surfaces much colder than the air above or in the outflow region of thunderstorms. More details on this diagnostic and its applications can be found in Lopez (2009).

Another major restriction is related to what can be called the ”0-rain” issue: wherever the model background has no precipitation while observations have, 4D-Var will be unable to correct the model towards the observations because the adjoint sensitivity of model precipitation to the 4D-Var control variables is zero. Conversely, wherever there is no precipitation in the observations while the model background is rainy, 4D-Var can substantially reduce model precipitation but a significant ambiguity remains regarding what the actual temperature and moisture profiles should be, unless actual measurements of these are available in the vicinity (from radiosondes, for instance). It has therefore seemed wise to restrict the assimilation of NEXRAD data to locations where precipitation is simultaneously higher than a small threshold (0.001 mm h^{-1}) in the model background and in the observations. This issue will be further discussed in section 6.2.

To illustrate the result of the screening process, an example of NEXRAD 6-hourly accumulated precipitation data coverage is displayed in Fig. 1 for a single 6-hour period ending at 0300 UTC 3 April 2009. Note the data-void area over the Appalachian Mountains after application of the rugged terrain criteria. Also note the limited number of points affected by surface snowfall and anomalous propagation (black and grey square symbols, respectively).

4.3 Observation errors

Matrix **R** in Eq. (2) is supposed to describe observation errors in terms of variances (diagonal terms) and covariances (off-diagonal terms). Theoretically, these should account not only for instrumental uncertainties but also for errors associated to the observation operator used to convert the model state into observed equivalent, to representativity and to the mapping of observations onto the model grid (spatial averaging for NEXRAD observations). Since little is known about NEXRAD error statistics, several assumptions have been made here: matrix **R** is supposed to be block-diagonal (i.e. no spatial correlation among NEXRAD observations) and a constant value $\sigma_o = 0.18$ has been assigned to the error standard deviation of NEXRAD 6-hourly accumulated precipitation amounts in $\ln(RR6h+1)$ space. This constant value of σ_o implies that higher precipitation amounts are deemed to be more reliable than smaller ones in a relative sense. Earlier attempts to express σ_o as a function of precipitation itself did not lead to any improvement in experiments scores and section 6.4 will

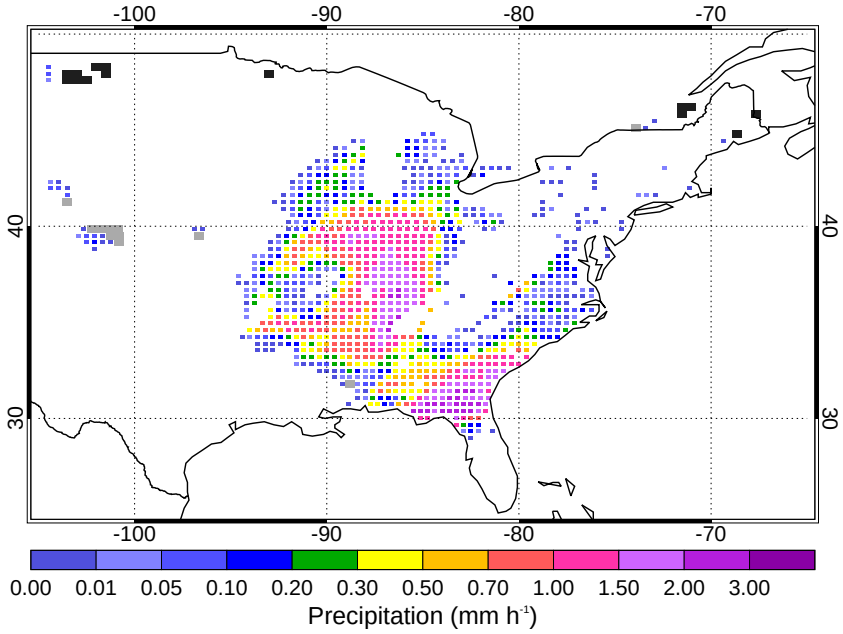


Figure 1: Example of NEXRAD 6-hourly accumulated precipitation data coverage at 0300 UTC 3 April 2009. Colour squares show precipitation observations (in mm h^{-1}) actually used in 4D-Var. Light grey (resp. black) squares indicate points that are rejected because they are likely to be affected by anomalous propagation (resp. surface snowfall).

provide some a posteriori verification of the above choice.

4.4 Bias-correction

Since 4D-Var relies on the assumption that both observations and model background are unbiased, statistics of observation minus model background departures expressed in $\ln(RR6h+1)$ space were computed from two-month long passive monitoring 4D-Var experiments. Several bias correction (*BC* hereafter) formulations were tested:

(1) *BC*=constant,

(2) $BC = \sum_{i=0}^{2} \alpha_i \, \overline{\ln(RR6h+1)}^{i}$ with α_i =constant,

(3) $BC = \sum_{i=0}^{2} \alpha_i(t_{obs}) \, \overline{\ln(RR6h+1)}^{i}$ where α_i is a function of observation local time, t_{obs}.

The quantity $\overline{\ln(RR6h+1)}$ corresponds to the average between model background and observation. Using the average of model and observation avoids undesirable spurious asymmetries in the BC, as demonstrated in Geer and Bauer (2010). Preliminary experimentation suggested that option (2) leads to the largest improvements in analyses and forecasts. Therefore, all results presented in this paper are based on this BC formulation. Coefficients α_i of the polynomial fit were derived from statistics computed over the two-month long passive monitoring experiments mentioned earlier. As an illustration, Fig. 2 displays observation minus model background departures in $\ln(RR6h+1)$ space as a function of $\overline{\ln(RR6h+1)}$ values for the two 2-month periods of

2009 considered in this study (more details in section 5.1). The polynomial fit used to define the bias correction is also shown for both period. To account for the gradual reduction of class population, n, with increasing precipitation values, a weight equal to $\left(\log(n_{max}/n)+1\right)^{-1}$ was assigned to each class in the definition of the fit, where n_{max} is the maximum class population. Figure 2 demonstrates that the curves are similar for both periods with an overestimation (resp. underestimation) in the model for 6-hour precipitation accumulations lower (resp. higher) than 1 mm h^{-1} (i.e. $\ln(RR6h+1) \approx 0.7$). The overall mean bias is rather small: -0.008 and -0.016 for April-May and September-October 2009, respectively.

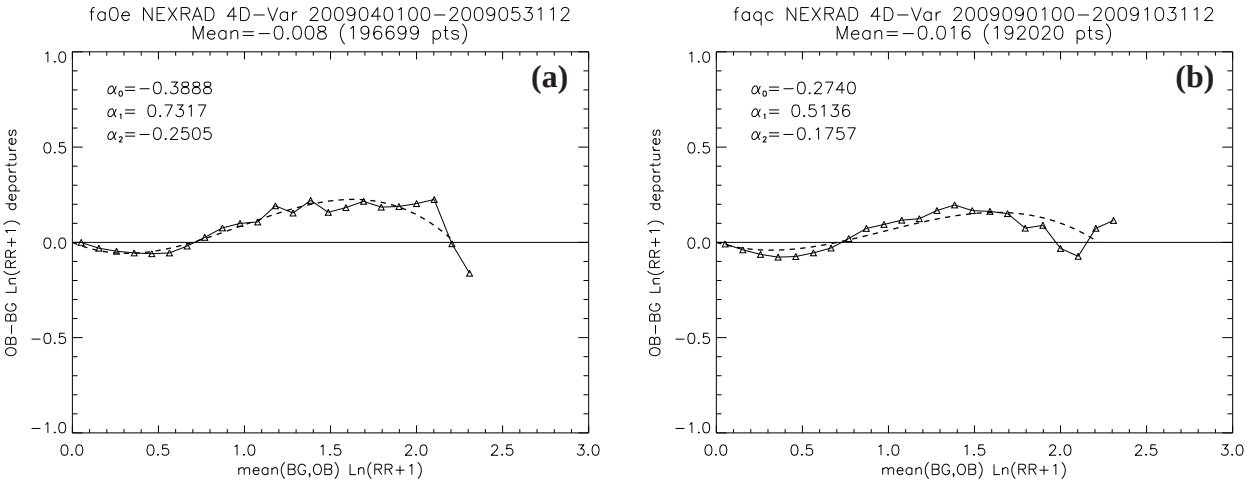


Figure 2: Average observation minus model background departures in $\ln(RR6h+1)$ *space (solid line) as a function of* $\overline{\ln(RR6h+1)}$ *for (a) April-May 2009 and (b) September-October 2009. The dash line shows the second degree polynomial fit chosen to define the bias-correction for the direct 4D-Var NEXRAD assimilation experiments:* $BC = \sum_{i=0}^{2} \alpha_i \overline{\ln(RR6h+1)}^{i}$. *Values of fit coefficients,* α_i, *are also given.*

4.5 First-guess check and variational quality control

As already implemented for all other observation types in 4D-Var, NEXRAD observations are subjected to an a-priori first-guess check which rejects measurements that depart too much from the model background. In the present case, a NEXRAD observation is rejected if $|\mathbf{y}_t - H_t(\mathbf{x}_0^b)| > 4\sqrt{\sigma_o^2+\sigma_b^2}$, where σ_o and σ_b are the observation and background error standard deviations in $\ln(RR6h+1)$ space, respectively, and are both set equal to 0.18. In other words, all observations leading to absolute values of background departures larger than 1.02 in $\ln(RR6h+1)$ space are screened out. This first-guess check for NEXRAD data is only applied in the first trajectory of each 4D-Var cycle, as for all other observation types.

In addition, in the course of each minimization, the variational quality control (VarQC; Andersson and Järvinen 1999) already applied to all other observation types is also applied to NEXRAD data. Any observation which leads to large departures that are deemed inconsistent with neighbouring measurements, has its weight in the cost function of Eq. (2) artificially inflated, in the current minimization only.

5 Experiments

5.1 Set-up

Two 2-month long periods have been selected for running global experiments of direct 4D-Var assimilation with NEXRAD precipitation observations over the eastern half of USA. The version of the ECWMF Integrated Forecasting System (IFS) employed in this study is cycle 35r2, which was used in operations between 10 March and 7 September 2009. The first period runs from 0000 UTC 1 April 2009 until 1200 UTC 31 May 2009 and the second one from 0000 UTC 1 September 2009 to 1200 UTC 31 October 2009. These two periods were chosen because they were characterized by numerous precipitating events of both convective and stratiform nature. For each period, two 4D-Var experiments were performed: a control run (CTRL) that uses all standard observations available as in ECMWF's operations and an experiment (NEW) with NEXRAD 6-hourly precipitation accumulations also included in the 4D-Var assimilation process. All experiments were run with 91 vertical levels and with a horizontal spectral resolution of T511 ($\approx$40 km) in the 4D-Var trajectory computations as well as in the subsequent 10-day forecasts started from 4D-Var analyses. The coarser resolutions used in the three successive minimizations of each 4D-Var 12-hour cycle are described in section 3.

5.2 Results and validation

5.2.1 Comparison of NEXRAD with PRISM observations

A first verification of precipitation fields is provided in Fig. 3 which compares NEXRAD with PRISM observations over the Eastern USA and over the two periods April-May and September-October 2009. The PRISM dataset consists of monthly precipitation 4-km gridded data generated from about 7,000 rain-gauges over mainland USA by the PRISM Climate Group (Oregon State University, http://www.prismclimate.org). No temporal resolution better than monthly was available for PRISM data at the time of this study. More details on the PRISM dataset can be found in Di Luzio *et al.* (2008). Note that both NEXRAD and PRISM were averaged onto the same T511 model reduced Gaussian grid prior to their comparison. Even though some rain-gauge observations are used in the production of NCEP Stage IV precipitation analyses, PRISM data can be considered as the only nearly-independent high-quality high-resolution precipitation dataset currently available over the USA. Figure 3 shows that NEXRAD and PRISM are remarkably close to one another for both selected periods, even though NEXRAD have a slight tendency to exhibit finer scale features.

Table 1 summarizes the spatial mean values of NEXRAD$-$PRISM differences, root-mean-square (RMS) differences and correlation between the maps displayed in Fig. 3. Table 1 confirms the excellent level of agreement

	Apr-May 2009	Sept-Oct 2009
PRISM Mean	3.18	3.75
Mean NEXRAD$-$PRISM diff.	0.03	$-$0.19
RMS NEXRAD$-$PRISM diff.	0.49	0.56
Correlation NEXRAD/PRISM	0.96	0.97

Table 1: Statistical comparison of two-month averaged precipitation accumulations (in mm day^{-1}) from NEXRAD and PRISM in April-May and September-October 2009 over the domain shaded in Fig. 3

.

between NEXRAD and PRISM, with very small mean differences, modest RMS values (as compared for instance to those found in Table 4 when comparing model with PRISM) and correlations close to unity.

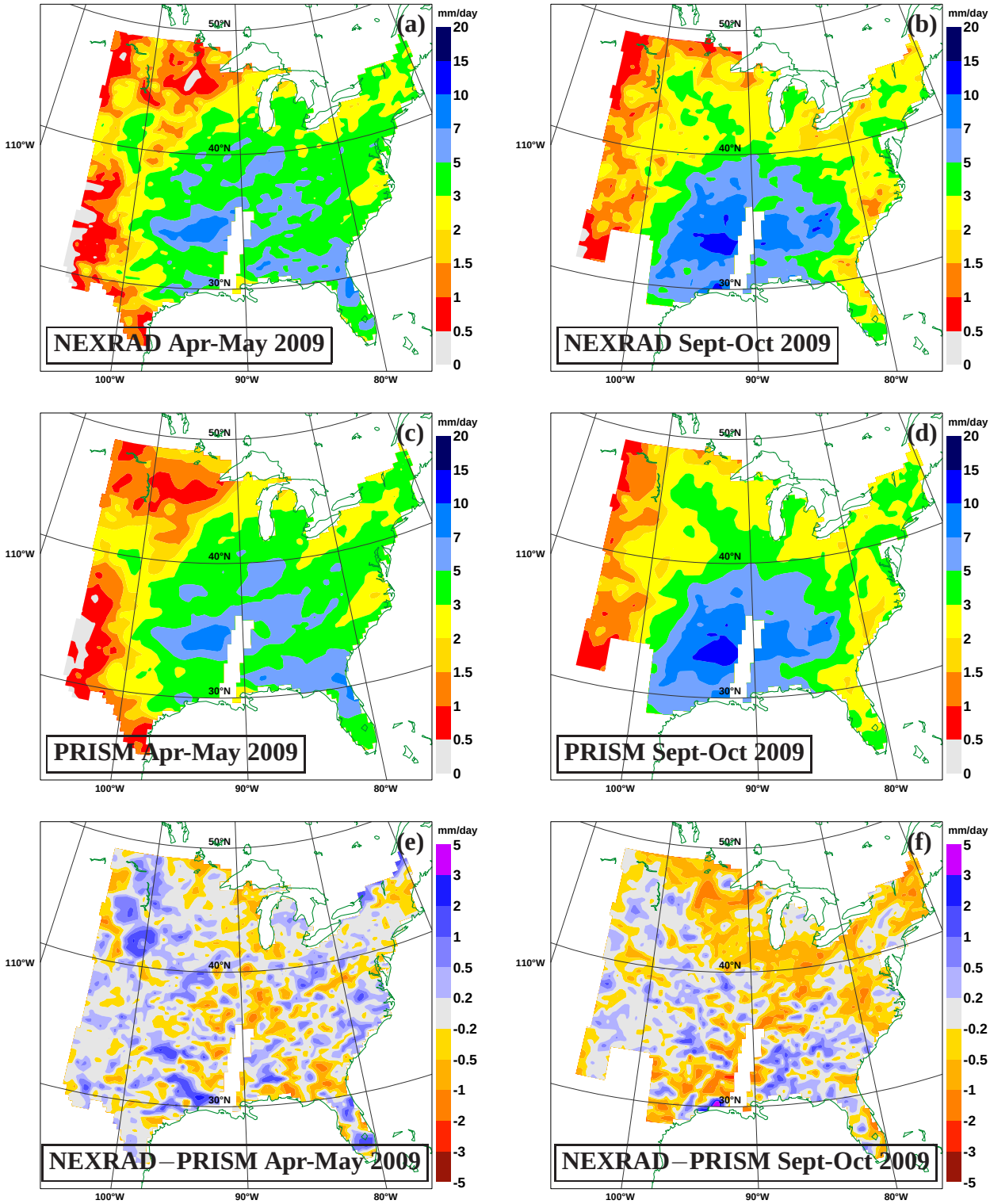


Figure 3: Comparison of NEXRAD (top) with PRISM (middle) average precipitation amounts (in mm day^{-1}) over the periods April-May 2009 (left) and September-October 2009 (right). Bottom panels show time-averaged NEXRAD−PRISM differences. The white stripe over the Mississipi Valley is due to the presence of a few missing hourly NEXRAD data which, if not rejected, might have led to a spurious mismatch between NEXRAD and PRISM accumulations.

5.2.2 *Background and analysis precipitation departures*

When assimilating a new type of observations such as NEXRAD data, it is essential to check how background and analysis departures expressed in observation space are statistically distributed. Their probability density functions (PDF) in $\ln(RR6h+1)$ space are plotted in Fig. 4 for each of the two-month periods considered in this paper. Samples size is 86,135 in April-May 2009 (resp. 99,298 in September-October 2009), which means that on average 706 (resp. 814) NEXRAD observations were assimilated in each 4D-Var 12h cycle. Panels (a) and (c) show that the distribution of observation minus background departures after bias correction (see section 4.4) is not too far from Gaussian, which is desirable in the 4D-Var context. The residual bias is rather small for both periods (-0.037 and -0.024) and reduced compared to that of uncorrected background departures (red histogram). One should note that these biases are of course different from those derived from the passive monitoring experiments to construct the bias correction procedure. Panels (b) and (d) demonstrate that the observation minus analysis departures PDF becomes much narrower than the background departures PDF. For both periods, the standard deviation of analysis departures is reduced by a good 33% compared to background departures (going down from 0.342/0.323 to 0.227/0.218). The mean bias is also reduced in the analyses. One can also remark that analysis departures are also nearly normally distributed. All this indicates that 4D-Var succeeded in bringing the model precipitation closer to the observations through the changes in temperature, moisture, wind and surface pressure imposed in the analyses.

5.2.3 *Precipitation scores against NEXRAD*

The NEXRAD precipitation observations that are assimilated in 4D-Var are expected to modify temperature, specific humidity, wind and surface pressure analyses. It seems natural to verify how these 4D-Var analyses changes feedback on subsequent precipitation forecasts. First, short-range precipitation forecasts from experiments CTRL and NEW have been validated against NEXRAD observations themselves, even though these cannot be considered as independent validation data. Tables 2 and 3 summarize the statistical results of this comparison computed over periods April-May 2009 and September-October 2009, respectively, and over the Eastern half of the USA. Precipitation accumulations over the first 6h, 12h and 24h of the forecasts started at 00Z are considered here. Displayed statistics are mean model minus NEXRAD bias, root-mean-square errors (RMSE) and correlation coefficient, computed over all points (rainy and non-rainy). Note that RMSE and correlation were obtained by averaging daily values over each two-month periods. The mean value of NEXRAD precipitation is also given on the first row, in mm day^{-1}.

Apr-May 2009	*RR*6h		*RR*12h		*RR*24h	
NEXRAD Mean	3.12		2.90		2.99	
	CTRL	NEW	CTRL	NEW	CTRL	NEW
Bias	0.15	$-$**0.11**	0.28	**0.11**	0.48	**0.42**
RMSE	10.03	**9.09**	7.41	**6.81**	5.73	**5.55**
Correlation	0.59	**0.63**	0.63	**0.67**	0.66	**0.68**

Table 2: Mean statistics of 6, 12 and 24-hour precipitation accumulations from short-range forecasts from experiments CTRL and NEW versus NEXRAD observations in April-May 2009 over the Eastern USA. Mean NEXRAD observations, model minus NEXRAD bias, RMS and correlation coefficient are computed over domain 25°N-50°N and 105°E-70°E. Mean, bias and RMS are in mm day^{-1}.

As a preliminary remark, it is noteworthy that the over-estimation of RR6h found in CTRL is consistent with

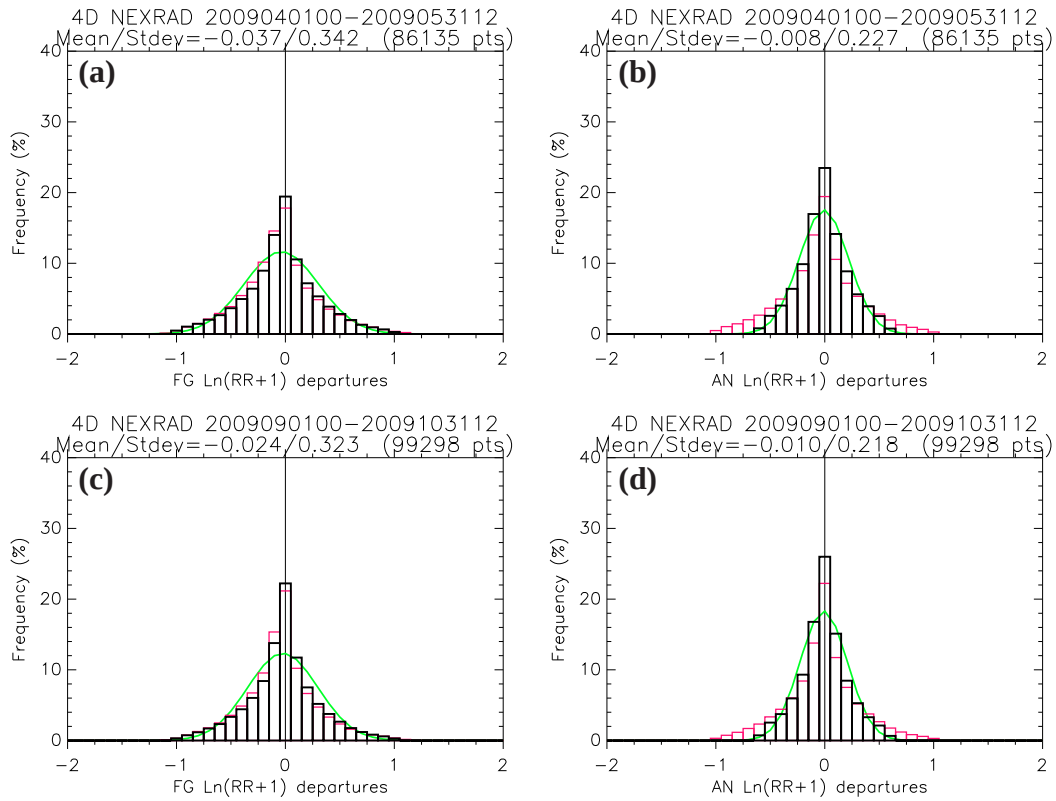


Figure 4: Histogram of observation minus background (left) and observation minus analysis (right) precipitation departures (in $\ln(RR6h+1)$ *space) from 4D-Var assimilation experiments with NEXRAD data for April-May 2009 (top) and September-October 2009 (bottom). Black lines display the histogram of bias-corrected background departures (left panels) and analysis departures (right panels). Red curves show the uncorrected background departures (left panels) and the bias-corrected background departures (right panels; i.e. a copy of the black curve from the left panel). Green curves show the Gaussian distribution with the same mean and standard deviation as the black histogram. Frequency (y-axis) are in % and the mean, standard deviation and total population of the black histogram are shown at the top of each panel.*

Sept-Oct 2009	*RR*6h		*RR*12h		*RR*24h	
NEXRAD Mean	3.01		3.15		3.19	
	CTRL	NEW	CTRL	NEW	CTRL	NEW
Bias	0.52	**0.22**	0.46	**0.21**	0.63	**0.53**
RMSE	9.82	**8.76**	7.78	**6.97**	6.12	**5.74**
Correlation	0.57	**0.61**	0.62	**0.66**	0.65	**0.67**

Table 3: Same as in Table 2, but for September-October 2009.

the results obtained in the definition of the bias-correction procedure in section 4.4. More importantly, both tables give clear evidence that short-range precipitation forecasts are better in experiment NEW than in CTRL, for both two-month periods. The absolute magnitude of the mean bias is substantially reduced. In particular, for the second period (Table 3), the bias is more than halved for forecast precipitation accumulated over 6h and 12h. RMSE also go down by roughly 10% for 6h and 12h accumulations and by around 4% for 24h accumulations. Similarly, the correlation is increased by 0.04 and 0.02, respectively. It is also worth stressing that bias values in experiment NEW but also CTRL are surprisingly modest. Tables 2 and 3 also suggest that the improvement in precipitation forecasts brought by the assimilation of NEXRAD data tends to wane with forecast range. This is confirmed by Fig. 5 which displays similar statistics for forecast ranges between 6 and 72 hours (all forecasts are started at 0000 UTC). Note that precipitation accumulations considered in this plot only run between successive selected forecast ranges. For instance, statistics shown at time 36h apply to precipitation accumulations between 24h and 36h. To convert to CST or EST (USA local time zones), one should substract 5 or 6 hours, respectively. As a first remark, panels (a) and (e) indicate that the model tends to systematically overestimate daytime precipitation. The precipitation diurnal cycle appears to be more pronounced in both experiments than in NEXRAD observations. This could indicate that convection is triggered too early in the model over the Eastern USA, as already identified in LB07. More interestingly, panels (b-d) and (f-h) confirm that the assimilation of NEXRAD data brings a significant improvement in bias, RMS difference and correlation between model and observations in the first 12 hours of the forecast and that this improvement vanishes for ranges beyond 24 hours. Such short-term positive impact on precipitation forecast is consistent with previous findings from 1D+4D-Var experiments of LB07. However, this apparent short life of the precipitation improvement might just be an artefact of its eastward propagation out of the Eastern USA domain considered here. Unfortunately, this assumption cannot be verified since there are currently no reliable quantitative precipitation estimates over the North Atlantic Ocean.

A last verification against NEXRAD observations has been performed by calculating Equitable Threat Scores (ETS) and False Alarm Rates (FAR), as defined in Appendix 3. The higher ETS and the lower FAR are, the better. Both scores are plotted in Fig. 6 for 12h forecasts started at 0000 UTC from experiments CTRL and NEW as a function of various precipitation thresholds and for the two selected two-month periods.

Figure 6 clearly demonstrates that NEXRAD observations are successfully assimilated in 4D-Var since ETS is systematically increased and FAR reduced, especially for precipitation rates higher than 3 mm day^{-1}. Beyond the 12-hour range, the positive impact of NEXRAD assimilation on threat scores quickly vanishes to become neutral beyond day 1 (not shown), consistent with previous findings.

5.2.4 Precipitation scores against PRISM

To complement the validation against NEXRAD data, independent monthly precipitation gridded data generated from 7,000 rain-gauges over mainland USA by the PRISM Climate Group (Oregon State University,

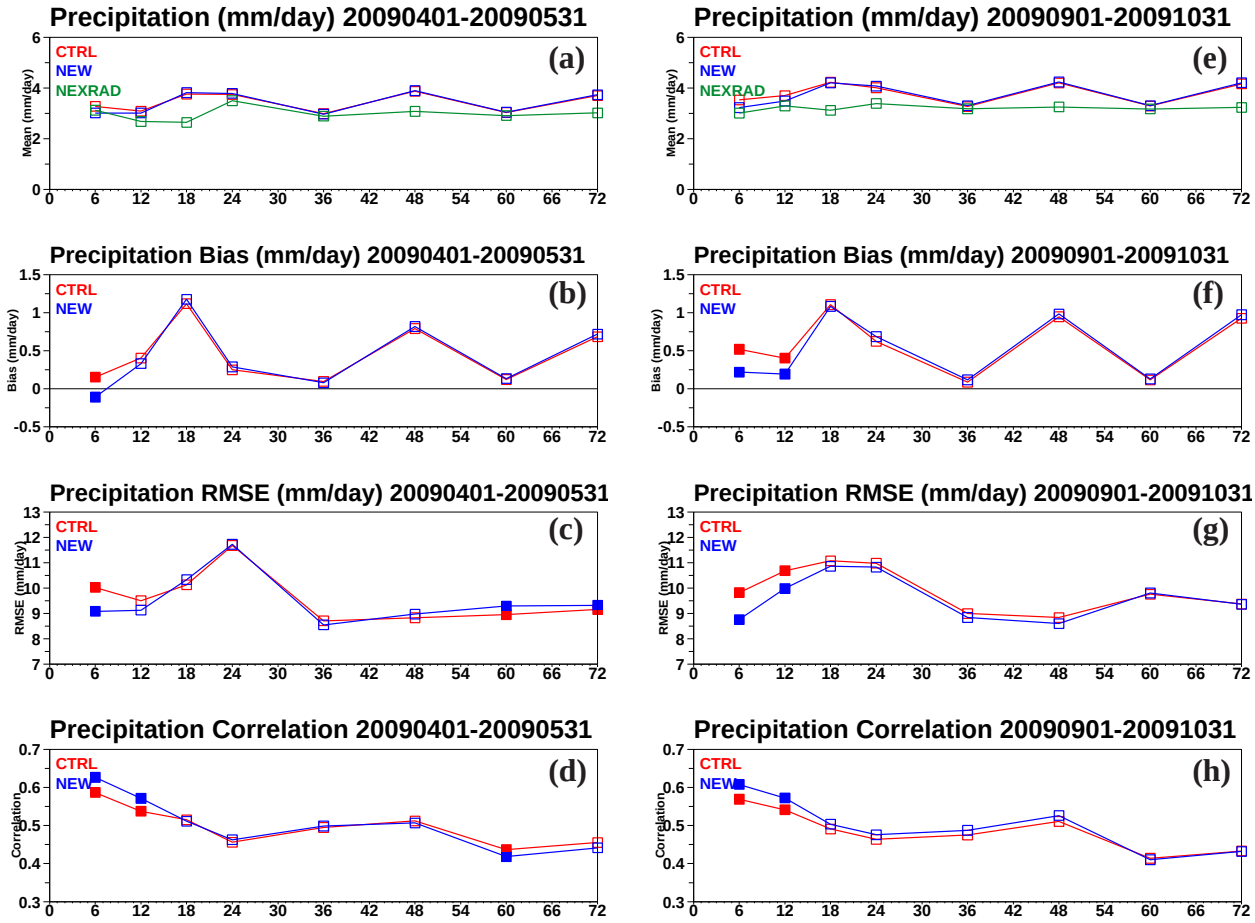


Figure 5: Statistics of precipitation forecasts against NEXRAD observations for various forecast ranges from 6 to 72 hours (x-axis) and computed over the periods April-May 2009 (left) and September-October 2009 (right). Statistics include (a,e) mean precipitation, (b,f) mean model minus NEXRAD bias, (c,g) RMSE and (d,h) correlation coefficient. Units are mm day^{-1} for the first three statistical quantities. Each symbol corresponds to the end of the accumulation period running since the previous symbol. For example, a symbol plotted at 36h corresponds to precipitation accumulated between 24h and 36h of the forecasts. Red (resp. blue) curve is for CTRL (resp. NEW) experiment. Filled symbols indicate that the difference seen between CTRL and NEW is significant (t-test with 95% confidence level).

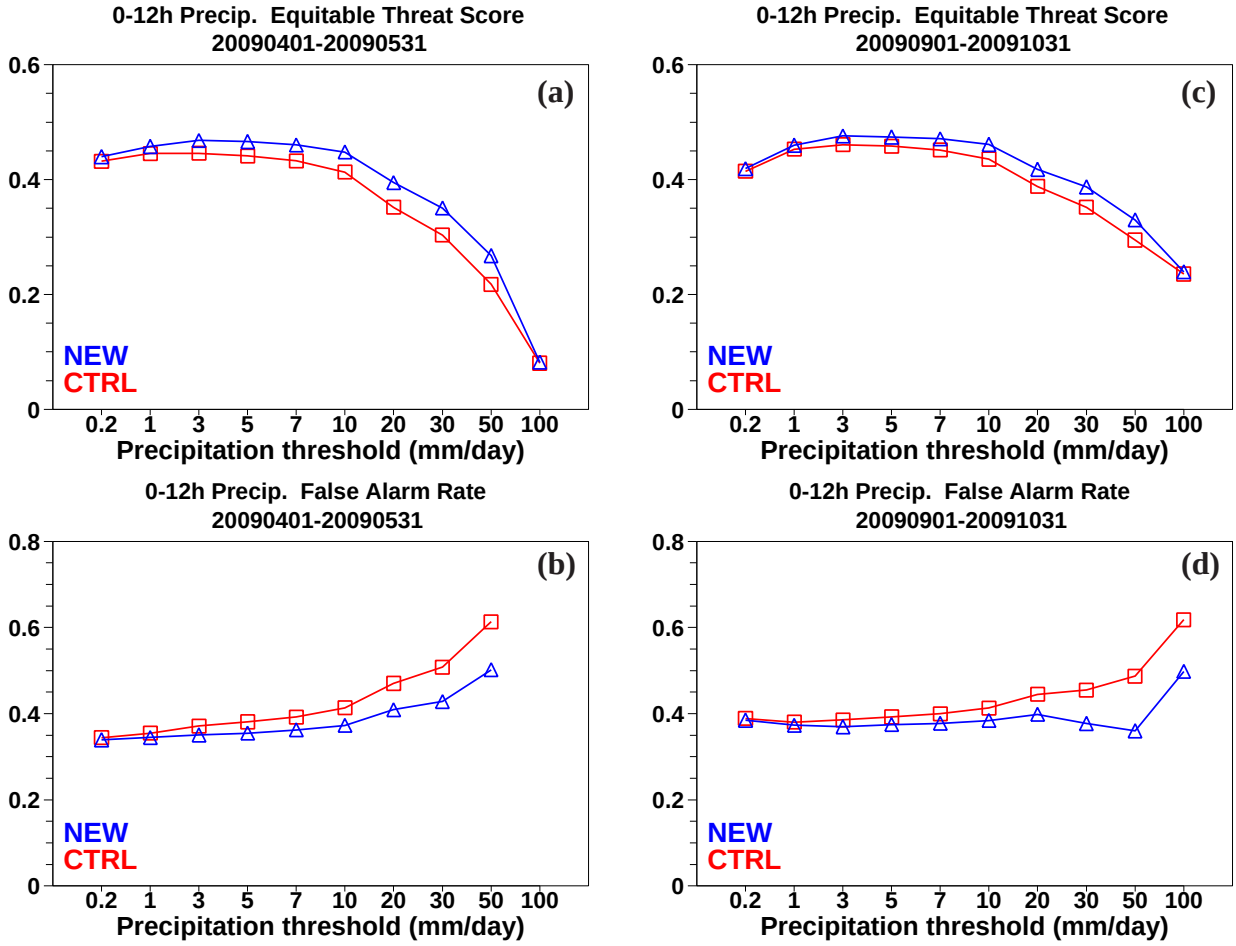


Figure 6: Equitable threat scores (top) and false alarm rates (bottom) computed from 12h precipitation forecasts from experiments CTRL (red line) and NEW (blue line) for various precipitation thresholds (in mm day^{-1}). Left (resp. right) panels are for April-May 2009 (resp. September-October 2009).

http://www.prismclimate.org) have been compared to experiments CTRL and NEW. More details on the PRISM dataset can be found in Di Luzio *et al.* (2008). Although the original horizontal resolution of PRISM data is 4 km, these data were averaged over the T511 model Gaussian grid prior to comparison. Table 4 summarizes mean, bias, RMSE and correlation of two-month precipitation accumulations computed from 24h forecasts (started at 00Z) and from PRISM monthly data. Note that small differences in observation coverage between

	Apr-May 2009		Sept-Oct 2009	
PRISM Mean	3.21		3.74	
	CTRL	NEW	CTRL	NEW
Bias	0.42	**0.35**	0.30	**0.17**
RMSE	0.91	**0.85**	0.91	**0.81**
Correlation	0.89	**0.89**	0.93	**0.94**

Table 4: Mean statistics of two-month precipitation accumulations from 24h forecasts from experiments CTRL and NEW versus PRISM gridded rain-gauge observations in April-May and September-October 2009. Mean PRISM observations, model minus PRISM bias, RMSE and correlation coefficient are computed over domain 25°N-50°N and 105°E-70°E. Mean, bias and RMSE are in mm day^{-1}.

NEXRAD and PRISM observations over the Eastern USA explain the discrepancies between the statistics reported in Tables 1 and 4. Also note that RMSE (resp. correlations) are much lower (resp. higher) for PRISM than for NEXRAD since here the statistics were calculated from two-month averages instead of daily data. Table 4 confirms that the assimilation of NEXRAD rain data improves the agreement of 24h precipitation forecasts with PRISM data in terms of mean bias, RMSE and only marginally correlation.

5.2.5 Atmospheric scores against verifying analyses

To verify how atmospheric variables are affected by the NEXRAD assimilation, anomaly correlations of 10-day forecasts against verifying analyses were computed for geopotential height, temperature, wind vector and relative humidity on different pressure levels. Only the most prominent changes in scores brought by the 4D-Var assimilation of NEXRAD observations shall be summarized here. Besides, to reduce further the number of plots, the two periods 1 April-21 May 2009 and 1 September-21 October 2009 were grouped together in the statistical calculations, yielding a total population of 102 cases. The last 10 days of each experiment were not included in the statistics since scores were computed against the respective analyses of experiments CTRL and NEW.

Figure 7 displays the impact on geopotential height (Z), temperature (T) and wind vector (WV) forecast anomaly correlation for selected pressure levels (as mentioned on each panel in hPa) over the Northern and Southern Hemispheres, North America, North Pacific and Tropics, as a function of forecast range from 12 hours to 10 days (all forecasts started at 00Z). Note that on all plots, score changes are normalized by the statistics from experiment CTRL and that positive (resp. negative) values on the y-axis indicate an improvement (resp. a degradation) of the selected score. Furthermore, purple bars highlight the level of significance of the changes, based on a 95% confidence level. An impact value is significant only if the associated bar does not cross the y=0 line. Figure 7 shows that scores over the Northern Hemisphere are slightly improved for all variables up to day 5 or 6, in a significant way up to day 3 of the forecast. They are degraded beyond day 6, significantly for $T850$ around day 8. Over the southern Hemisphere, the impact of assimilating NEXRAD data is positive, especially during the first three days of the forecast. It was checked that this rather unexpected impact gradually develops during the 4D-Var cycling. Over North America, changes in scores tend to be slightly neu-

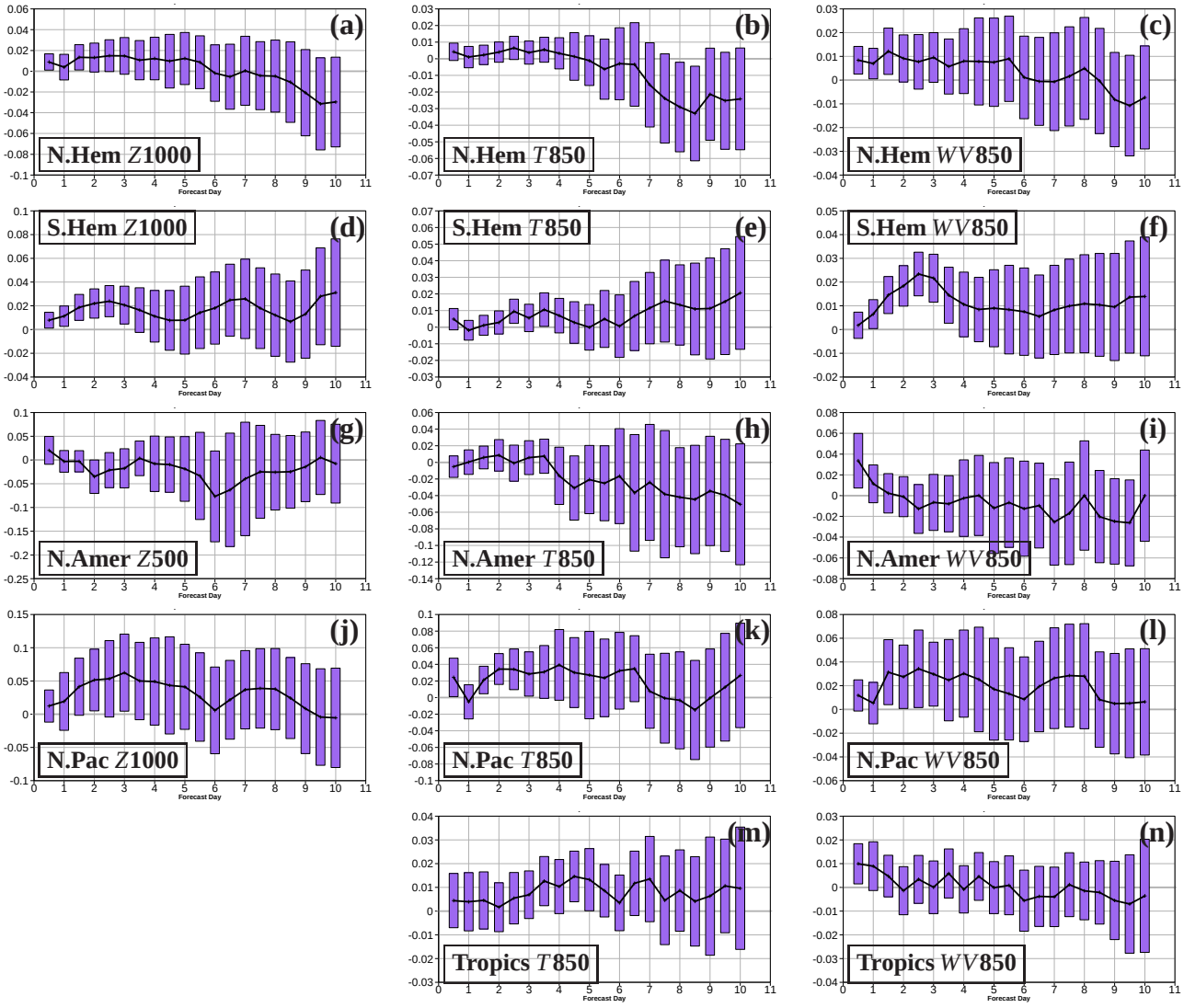


Figure 7: Impact of 4D-Var assimilation of NEXRAD precipitation observations on forecast anomaly correlation against own analyses as a function of forecast range (in days). Score changes for 1000 hPa geopotential height (Z1000; left panels), 850 hPa temperature (T850; middle panels) and 850 hPa wind vector (WV850; right panels) are plotted for different regions: (a)-(c) Northern Hemisphere (N.Hem), (d)-(f) Southern Hemisphere (S.Hem), (g)-(i) North America (N.Amer), (j)-(l) North Pacific (N.Pac) and (m)-(n) Tropics. Note that 500 hPa Z is shown instead of 1000 hPa Z over North America, since many points in this regions are situated well above sea level. Changes shown on y-axis are normalized by the score in experiment CTRL and positive (resp. negative) values indicate an improvement (resp. degradation) of the score. Purple bars indicate significance at the 95% confidence level.

tral (for Z and T) over the first 4 days, but become negative later, though not significantly. Wind is particularly improved close to analysis time (panel (i)) at 850 hPa but also at other levels, especially close to the tropopause (not shown). North Pacific scores are mostly positively affected at all forecats ranges, with good confidence after day 1 and until day 3 or 4. Similarly, tropical scores tend to become better especially for temperature around days 4-5 (panel (m)). Over all these regions, scores at higher levels were found to be either unaffected or modified in a way consistent with Fig. 7. Over all other sub-regions, including Europe, North Atlantic and Asia, score changes turned out to be either neutral or not significant (not shown). Relative humidity scores, which are now known to be more difficult to interpret when assimilating new moisture-related observations for reasons explained in Geer *et al.* (2010), did not exhibit any significant changes either. Overall the assimilation of NEXRAD precipitation observations in 4D-Var is either neutral or slightly positive on standard atmospheric verification scores, which is not really surprising given the limited spatial coverage of NEXRAD observations and their competition with other observation types with small errors, such as radiosoundings or synoptic station measurements. These results are consistent with those from 1D+4D-Var experiments in LB07, who also demonstrated through denial experiments that NEXRAD data could have a much larger positive impact if they were assimilated as the sole source of moisture information over the USA.

5.2.6 Verification against other observation types

Consistent with the overall small impact of NEXRAD observations that was found when looking at atmospheric scores against own analyses in section 5.2.5, no significant changes appear either when considering observation-minus-background and observation-minus-analysis departures statistics for all types of observations assimilated in 4D-Var (using OBSTAT software). This remains true even when statistics are restricted to the Eastern USA, as illustrated in Fig. 8 which compares profiles of background and analysis departure standard deviations from experiments CTRL and NEW against all radiosonde measurements of temperature, relative humidity and zonal wind assimilated in 4D-Var over this sub-domain. These plots are for period September-October 2009. If anything, the standard deviation of the departures is slightly reduced at all levels in NEW compared to CTRL. Similar conclusions can be drawn about period April-May 2009 for all other observation types and for mean bias as well (not shown).

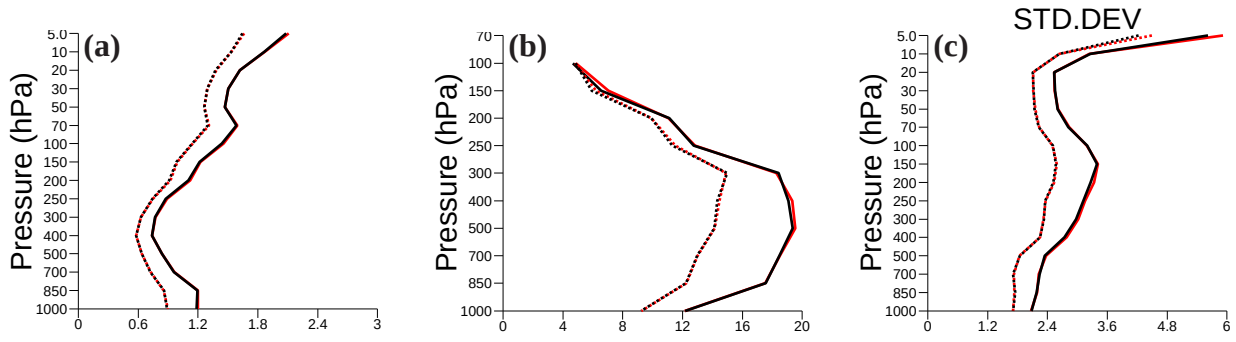


Figure 8: Background departure (solid line) and analysis departure (dotted line) standard deviation profiles with respect to radiosonde observations of (a) temperature (K), (b) relative humidity (%) and (c) zonal wind ($m\,s^{-1}$) from experiments CTRL (red) and NEW (black). Statistics were computed over about 7,500 radiosoundings over the Eastern USA in September-October 2009. Vertical axis shows pressure in hPa.

5.2.7 Validation against GOES-12 infrared brightness temperatures

An attempt was also made to verify the impact of NEXRAD data assimilation on cloud analyses and forecasts through a comparison to geostationary infrared imagery (10.7 μm) from the GOES-12 satellite, overlooking the Americas (full-disk data obtained through NOAA/CLASS). To perform this validation in 10.7-μm brightness temperature space, simulated images were computed from the ECMWF model temperature, moisture, cloud and ozone fields using the RTTOV-9 radiative transfer model (http://research.metoffice.gov.uk/research/interproj/-nwpsaf/rtm/rtm_rttov9.html). Model minus GOES mean statistics computed over two months (not shown) did not exhibit any significant change, even when focusing on the eastern half of the USA and on short-range forecasts. This may be explained by the fact that infrared TBs are mainly sensitive to cloud top temperatures only, which are not necessarily strongly correlated with precipitation amounts, especially in stratiform cloud systems. However, substantial positive and negative local impact on TBs can be found when considering individual images, as illustrated in Fig. 9 for 12h forecasts of two frontal situations. Here, impact is defined as $|\text{NEW}-\text{GOES}| - |\text{CTRL}-\text{GOES}|$. Most of the positive and negative changes are located inside cloudy regions, which is understandable since NEXRAD data were only assimilated at points with precipitation both in the model background and in the observations. Therefore, cloud water is simply redistributed inside the cloud system. But again, on a monthly time scale, these positive and negative changes cancel out and the impact of the NEXRAD assimilation on 10.7-μm TBs is neutral.

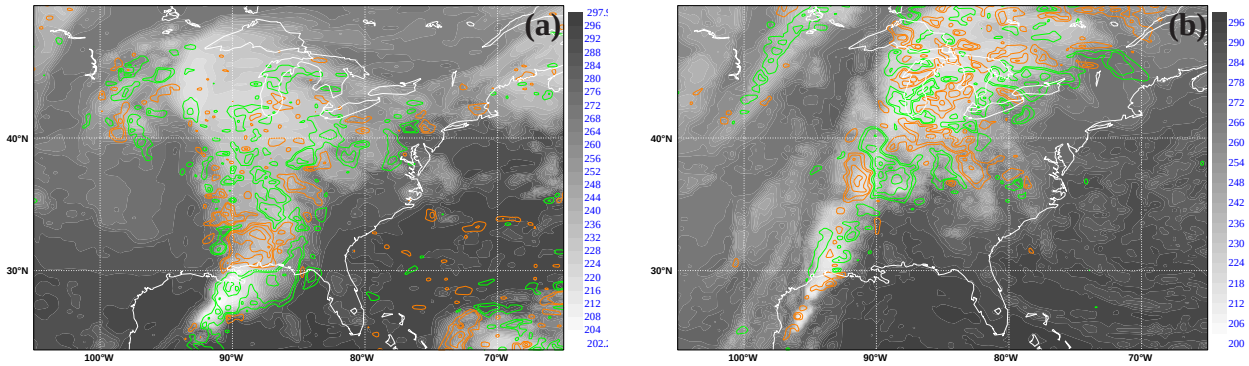


Figure 9: Two examples of impact of NEXRAD data assimilation on 10.7 μm simulated TBs from 12h forecasts compared to GOES-12 observations (shown with black and white shading, in K) (a) at 0000 UTC 23 October 2009 and (b) at 1200 UTC 30 October 2009. Positive (resp. negative) impact is shown with green (resp. orange) contour isolines (5, 10, 20, 30 K).

6 Discussion

6.1 Minimization versus trajectory

In the course of each 4D-Var assimilation cycle, model minus observation departures evolved at lower resolution using the linearized model in each minimization (i.e. $D_{minim} = \mathbf{H}_t \delta\mathbf{x}_0 - \mathbf{d}_t = H_t[\mathbf{x}_0^b] + \mathbf{H}_t \delta\mathbf{x}_0 - y_t$) should not, in theory, differ too much from model minus observation departures evolved at high resolution with the non-linear model in the subsequent trajectory (i.e. $D_{traj} = H_t[\mathbf{x}_0^b + \delta\mathbf{x}_0] - y_t$).

In practice, however, discrepancies between D_{minim} and D_{traj} can arise from:

(1) the fact that the linearized model, **H**, used in the minimizations is only a simplified version of the fully-fledged non-linear model, H, used in trajectory computations (for reasons mentioned in the introduction),

(2) strong non-linearities in H, especially if moist processes are to be represented, which implies that 4D-Var's underlying assumption of linearity is breached,

(3) resolution differences between minimizations and trajectories (see section 3).

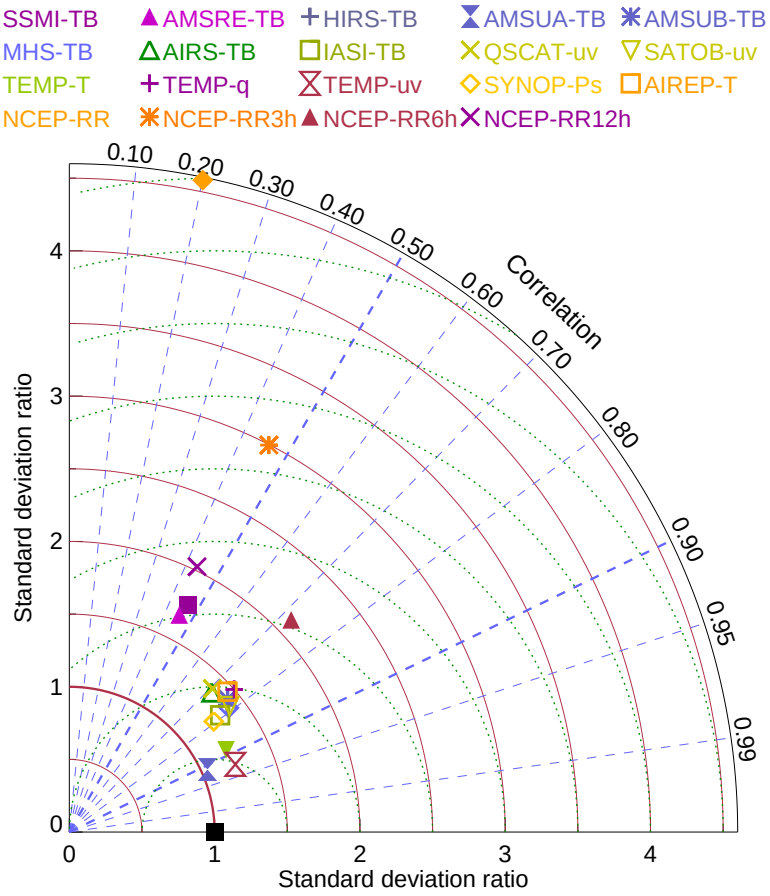


Figure 10: Taylor diagram showing the level of agreement between model minus observation departures (D_{minim}) computed in the first minimization using the simplified linearized model at T95 resolution with those (D_{traj}) evolved in the following trajectory using the fully-fledged non-linear forecast model at T799 resolution. Azimuth shows correlation between D_{traj} and D_{min} while radial distance measures the ratio of their standard deviations (SDR). The black square indicates a perfect agreement, when correlation and SDR are both equal to unity. Each symbol corresponds to a given observation type assimilated in 4D-Var, as shown in the legend at the top. Statistics are based on a single 4D-Var 12-hour cycle valid at 0000 UTC 1 April 2009.

The level of matching between D_{minim} and D_{traj} was assessed over a single 4D-Var cycle for each observation type currently assimilated in the ECMWF operational system as well as for NEXRAD observations. Note that in this particular experiment, the trajectory was run at T799 (≈25 km) resolution and that statistics were computed over the whole 4D-Var 12-hour assimilation window. Results for the first (T95 resolution) of the three minimizations are summarized on the Taylor diagram in Fig. 10 which displays the correlation coefficient between D_{traj} and D_{minim} (azimuth) and the associated standard deviation ratio, *SDR*, (radial distance). Each symbol corresponds to a given observation type as indicated in the top legend (see Appendices 1 and 2 for

the meaning of abbreviations). The black square symbol identifies a perfect match between D_{traj} and D_{minim}. Figure 10 shows that for observations that are not directly affected by cloud and precipitation processes, *SDR* remains reasonably close to unity and the correlation higher than 0.7. For microwave brightness temperatures from SSM/I and AMSR-E which are highly sensitive to clouds and precipitation, *SDR* reaches 1.7 and the correlation drops below 0.5, consistent with Bauer *et al.* (2010). This points towards a degradation in the validity of the linearity hypothesis. In the case of hourly NEXRAD precipitation data, the breach of linearity becomes terrifying since *SDR* reaches 5.2 and the correlation plummets to 0.2, which offers a daunting prospect for their optimal assimilation in 4D-Var.

However, it was found that the linearity could be much improved by assimilating precipitation observations accumulated over several hours instead of the original hourly data. Linearity results for accumulation lengths of 3, 6 and 12 hours are plotted in Fig. 10. Accumulating NEXRAD observations over 6 hours brings *SDR* down to 2 (not far from SSM/I or AMSR-E) and correlation up to 0.72 (better than for SSM/I or AMSR-E), indicating a strong improvement in the level of linearity. Similar conclusions can be drawn for the second and third minimizations (not shown), although the gap between the various NEXRAD accumulation periods is slightly reduced and the overall agreement between minimization and trajectory increments better than in the first minimization.

In view of these results, it was decided to assimilate NEXRAD 6-hourly accumulations ($RR6h$), which provides the best compromise between linearity and observation usage over the 4D-Var 12-hour window. Therefore, the quantity actually assimilated in the 4D-Var experiments is $\ln(RR6h+1)$, where $RR6h$ is expressed in mm h^{-1}. The positive impact of the temporal accumulation on linearity is entirely consistent with recent findings of Mahfouf and Bilodeau (2007), Fabry and Sun (2010) and Fabry (2010). In terms of practical implementation, it is worth noting that the contributions of each NEXRAD hourly observation to the adjoint sensitivities ($\mathbf{H}^T$) have to be summed up over 6-hour time slots. One can also stress that the assimilation of temporally accumulated observations has the advantage of reducing the occurrence of "0-rain" points (see sections 4.2 and 6.2) as a result of the propagation of weather systems over longer periods of time.

6.2 "0"-rain issue

One of the major issue that remains unsolved is the so-called "0-rain" issue, as already introduced in section 4.2. When the model background precipitation is zero, a rainy NEXRAD observation will be unable to produce any increment of the assimilation control vector (temperature, humidity, wind and surface pressure) as a result of the absence of sensitivity in the adjoint moist physics. With a two-step assimilation method such as the 1D+3D-Var of Caumont *et al.* (2010), it is possible to use a first-guess which is artificially forced towards precipitation through a moistening of the background vertical profile. This approach may give satisfactory results at kilometric horizontal resolutions since atmospheric profiles with precipitation can be assumed to be saturated at some vertical levels. However, even in this case, some uncertainty remains about the exact location of those levels, especially if only 2D precipitation observations are assimilated. For coarser resolutions (say 5 km or more), there is some additional uncertainty about which relative humidity values should be specified in the modified first-guess profile, since by construction the model can produce precipitation even if the grid box is far from saturation. In the case of direct 4D-Var, using a first-guess which is different from the model background is theoretically possible but would require further investigation with probably some additional technical developements. Furthermore, the inclusion of "0-rain" cases would also mean to address the issues that (1) the occurrence of non-rainy events is currently underestimated in the ECMWF model (too many grid points with small precipitation amounts) and (2) NEXRAD observations might not detect weaker precipitation because of the minimum detection threshold of the radar.

6.3 Asymmetry of analysis departures

It was found in earlier experimentation that 4D-Var could more easily dry the model state (when the background has more precipitation than in the observations) than moisten it (when the background is less rainy than the observations), consistent with the findings of Geer *et al.* (2010) for the assimilation of all-sky microwave radiances. Removing the limitation of humidity analysis increments above saturation implemented in the standard model code (Hólm 2002) helped to reduce this asymmetry in all experiments presented here. Although radical, this measure helps to overcome the limiting effect of the saturation threshold, especially in very moist regions, as already identified in Hou *et al.* (2000), for instance. However, some asymmetry remains as illustrated in Fig. 11 which displays 4D-Var analysis-minus-observation versus background-minus-observation departures in $\ln(RR6h+1)$ space from experiment NEW in September-October 2009. Figure 11 clearly shows

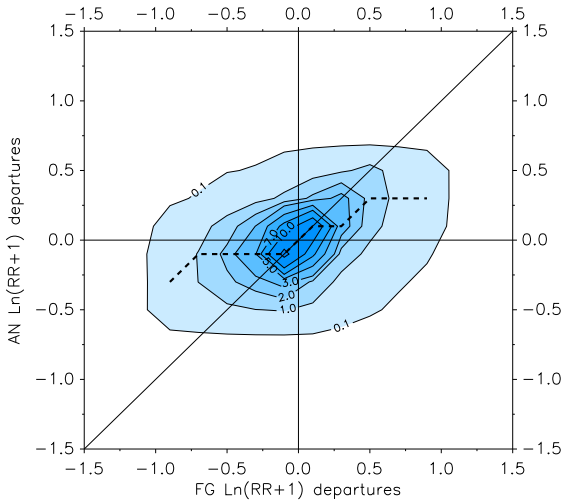


Figure 11: 2D PDF of analysis-minus-observation versus background-minus-observation departures from experiment NEW for September-October 2009. All departures are expressed in $\ln(RR6h+1)$ *space. Colour shading indicate frequency (labelled 0.1, 1, 2, 3, 5, 7, 10 and 15%) while the black dashed line corresponds to the mode of the analysis departure distribution for each background departure class.*

that the analysis gets closer to the observations when model background is too rainy (left half of the plot) than in the opposite case (right half of the plot). This might be the consequence of the existence of an asymmetry in the sensitivities of the precipitation produced by the moist physics to its input variables, in particular temperature and moisture.

6.4 Observation error statistics

it was mentioned in section 4.3 that the error standard deviation for NEXRAD observations was set to a relatively arbitrary constant value of 0.18. In order to assess the validity of this specification, the diagnostic approach proposed by Desroziers *et al.* (2005; DZ05 hereafter) was applied to the outputs from the 4D-Var assimilation experiments (NEW) presented in this paper. Among other things, this method allows the a posteriori estimation of observation error statistics (matrix **R**) from background and analysis departures computed in 4D-Var. DZ05's Eq.(3) states that

$$\mathbf{R} \quad = \quad E[\mathbf{d}_a^o(\mathbf{d}_b^o)^T] \tag{4}$$

where $\mathbf{d}_b^o$ and $\mathbf{d}_a^o$ are respectively the observation-minus-background and observation-minus-analysis departures and E is the statistical expectation operator. This computation can be applied to a selected subset of observations, namely here all assimilated NEXRAD observations, to diagnose the a posteriori observation error standard deviation as

$$\sigma_o \quad = \quad \Big(\sum_{j=1}^{p} (\mathbf{d}_a^o)_j (\mathbf{d}_b^o)_j / p \Big)^{\frac{1}{2}} \tag{5}$$

where p is the number of NEXRAD observations (indexed j). Figure 12 displays the values of σ_o diagnosed from Eq. (5) as a function of $\ln(RR6h+1)$. It appears that the diagnosed observation errors do not differ too much from the constant value of 0.18 prescribed in this study (dashed line in Fig. 12). There is an indication that σ_o might be slightly overestimated in this study for precipitation rates lower than 0.2 mm h^{-1} and underestimated for higher rain rates. Even though previous experimentation showed that results from direct 4D-Var assimilation of NEXRAD data were not very sensitive to σ_o variations of such amplitude, a slight retuning of observation errors will be considered before operational implementation.

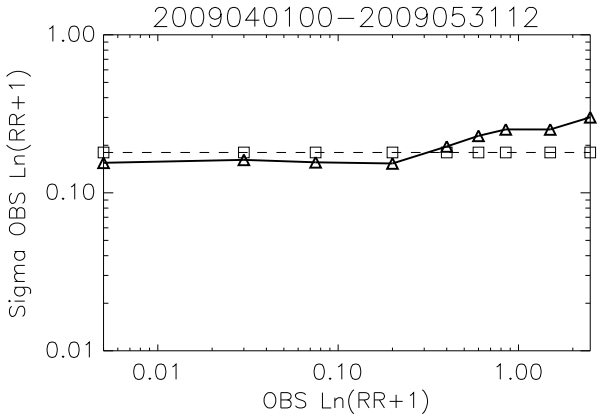


Figure 12: Observation error standard deviations obtained from Desroziers et al. (2005) a posteriori diagnostics (solid line) as a function of precipitation rate. All values are in $\ln(RR6h+1)$ *space. The dashed line corresponds to the constant value of* σ_o *=0.18 used in this study. Diagnostics are valid for April-May 2009.*

7 Conclusions

Direct 4D-Var assimilation experiments using 6-hour precipitation accumulations from NCEP Stage IV combined ground-based radar (NEXRAD) and rain-gauge observations over the Eastern half of the USA have been run over two 2-month long periods in 2009. The motivation was to assess the impact of these additional data on both 4D-Var analyses and subsequent forecasts. Precipitation observations averaged over 6 hours were preferred to the original hourly data since this helped to reduce the discrepancies between the minimizations run at low resolution with simplified linearized physics and subsequent trajectories run at high resolution with full non-linear physics. Although observation minus background mean biases turned out to be relatively small, a simple bias correction procedure was defined as a second-order polynomial function of precipitation amount.

As expected from a well behaved assimilation system, analysis departures turn out to be significantly reduced compared background departures. Besides, the assimilation of the new rain observations leads to a significant improvement on precipitation forecasts for ranges up to 12 hours, with respect to NEXRAD observations but

also to independent monthly PRISM rain-gauge measurements. This improvement quickly vanishes, which is consistent with previous 1D+4D-Var experiments of LB07. Atmospheric forecast scores are not significantly affected in general. Even though a degradation of mid-level temperature scores beyond day 7 appears over the Northern Hemisphere, a significant positive impact can be identified on temperature, wind and gepotential scores over the North Pacific and the Southern Hemisphere up to day 4, and for tropospheric wind over North America during the first day of the forecasts. The limited impact of NEXRAD observations can be explained by their competition with all other observation types available over North America.

Efforts in the near future will be devoted to the transfer of all changes required in the IFS (129 files, including scripts and SQL requests) to the latest cycle available and to run new tests at higher resolution (T799 or T1279), prior to operational implementation in 2011 (as planned). The screening of snowy situations ought to be revisited to try to increase the coverage of actually assimilated NEXRAD observations in the wintertime. Also, to avoid the need for constantly retuning the bias-correction coefficients everytime the model is changed, one might consider the inclusion of NEXRAD observations in the variational bias correction framework implemented in ECMWF's operational 4D-Var system (Dee and Uppala 2009). The planned implementation of a new prognostic variable for cloud condensate in the 4D-Var assimilation control vector might also improve the performance of precipitation assimilation in general. Further work should also aim at overcoming the "0-rain" issue which currently restricts NEXRAD observation usage to the situations where both model background and observations are rainy. Using a first-guess modified from the background at the beginning of each 4D-Var cycle could theoretically be a solution, but this was never tested in ECMWF's 4D-Var.

In the longer term, one could consider the assimilation of more radar networks (e.g. Europe, China, Canada,...), once problems of data availability and homogeneity are solved (e.g. Lopez 2008 on OPERA data). Besides, the new ability to assimilate accumulated precipitation measurements should make it possible to consider 4D-Var assimilation of rain-gauge observations from synoptic stations, which are currently used for verification purposes only, even though this would mean to address the issue of their representativity, first.

Despite the rather neutral or slightly positive impact found on traditional atmospheric scores so far, the clear improvement of short-range precipitation forecasts suggests that genuine precipitation analyses can now be obtained over the Eastern USA, which should be beneficial to the quality of the surface analysis, in particular of soil moisture contents. More generally, it is undeniable that obtaining better high-resolution precipitation analyses over land is vital for nowcasting, hydrological and climatological applications.

Acknowledgements

Special thanks should go to my colleagues Anne Fouilloux, Milan Dragosavac and Deborah Salmond for their technical assistance throughout this work. Marta Janisková, Anton Beljaars, Alan Geer, Peter Bauer, Martin Miller and Jean-Noël Thépaut should be acknowlegded for their early review of this paper. NCEP should be acknowledged for producing and JOSS/UCAR for providing the Stage IV precipitation observations used in this study. I am also grateful to the PRISM Climate Group (Oregon State University, USA) for giving access to their monthly precipitation gridded data. NOAA should be acknowledged for supplying GOES-12 satellite data through CLASS.

APPENDIX 1

List of abbreviations used in the text:

NOAA	=	National Oceanic And Atmospheric Administration (USA)
NCEP	=	National Centers for Environmental Prediction (USA)
JOSS	=	Joint Office for Science Support (USA)
UCAR	=	University Corporation for Atmospheric Research (USA)
JMA	=	Japan Meteorological Administration
ECMWF	=	European Centre for Medium-range Weather Forecasts
PRISM	=	Parameter-elevation Regressions on Independent Slopes Model
CLASS	=	Comprehensive Large Array-data Stewardship System
OPERA	=	Operational Programme for the Exchange of weather RAdar information (Europe)
SSM/I	=	Special Sensor Microwave Imager
AMSR-E	=	Atmospheric Microwave Scanning Radiometer - Earth Observing System
TMI	=	TRMM Microwave Imager
TRMM	=	Tropical Rainfall Measuring Mission
GOES	=	Geostationary Operations Environmental Satellite

APPENDIX 2

List of abbreviations used in Fig. 10:

TB	=	Brightness temperature
HIRS	=	High-resolution Infrared Radiation Sounder
AMSU	=	Advanced Microwave Sounding Unit
AIRS	=	Atmospheric Infrared Sounder
IASI	=	Infrared Atmospheric Sounding Interferometer
MHS	=	Microwave Humidity Sounder
QuikSCAT	=	Quick Scatterometer
QSCAT-uv	=	QuikSCAT winds
SATOB-uv	=	Geostationary satellite motion vectors
TEMP-T	=	Radiosonde temperature
TEMP-q	=	Radiosonde specific humidity
TEMP-uv	=	Radiosonde wind components
SYNOP-Ps	=	Synoptic station surface pressure
AIREP-T	=	Aircraft temperature reports
NCEP-RR	=	NCEP Stage IV hourly precipitation amounts
NCEP-RR12h	=	NCEP Stage IV 12-hour accumulated precipitation amounts
NCEP-RR6h	=	NCEP Stage IV 6-hour accumulated precipitation amounts
NCEP-RR3h	=	NCEP Stage IV 3-hour accumulated precipitation amounts

APPENDIX 3

Precipitation scores used in this study are the Equitable Threat Score (ETS) and the False Alarm Rate (FAR), defined as follows

$$ETS = \frac{H - H_e}{H + M + F - H_e} \quad (6)$$

$$FAR = \frac{F}{H + F} \quad (7)$$

where H is the number of correct hits, M is the number of misses and F is the number of false alarms. H_e is the number of correct hits purely due to random chance and is computed as

$$H_e = \frac{(H + F)(H + M)}{N} \quad (8)$$

where N is the sample size.

References

Agustí-Panareda, A., Beljaars, A., Cardinali, C., Genkova, I., and Thorncroft, C. (2010). Impact of assimilating AMMA soundings on ECMWF analyses and forecasts. *Weather Forecast.* in press.

Andersson, E. and Järvinen, H. (1999). Variational quality control. *Q. J. R. Meteorol. Soc.*, 125:697–722.

Baldwin, M. E. and Mitchell, K. E. (1996). The NCEP hourly multi-sensor U.S. precipitation analysis. In *Preprints 11th Conf. on Numerical Weather Prediction, Norfolk, VA (USA), 19–23 August 1996*, pages J95–J96.

Bauer, P., Geer, A. J., Lopez, P., and Salmond, D. (2010). Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation. *Q. J. R. Meteorol. Soc.* Accepted.

Bauer, P., Lopez, P., Benedetti, A., Salmond, D., and Moreau, E. (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., Lopez, P., Benedetti, A., Salmond, D., Saarinen, S., and Bonazzola, M. (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.

Caumont, O., Ducrocq, V., Wattrelot, E., Jaubert, G., and Pradier-Vabre, S. (2010). 1D+3D-Var assimilation of radar reflectivity data: a proof of concept. *Tellus*, 62A:173–187.

Caya, A., Sun, J., and Snyder, C. (2005). A Comparison between the 4DVAR and the Ensemble Kalman Filter Techniques for Radar Data Assimilation. *Mon. Weather Rev.*, 133:3081–3094.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteorol. Soc.*, 120:1367–1388.

Dee, D. P. and Uppala, S. (2009). Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Q. J. R. Meteorol. Soc.*, 135:1830–1841.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.*, 131:3385–3396.

Di Luzio, M., Johnson, G. L., Daly, C., Eischeid, J. K., and Arnold, J. G. (2008). Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. *J. Appl. Meteor.*, 47:475–497.

Ducrocq, V., Ricard, D., Lafore, J.-P., and Orain, F. (2002). Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Weather Forecast.*, 17:1236–1256.

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

Fabry, F. (2010). For How Long Should What Data Be Assimilated for the Mesoscale Forecasting of Convection and Why? Part II: On the Observation Signal from Different Sensors. *Mon. Weather Rev.*, 138:256–264.

Fabry, F. and Sun, J. (2010). For How Long Should What Data Be Assimilated for the Mesoscale Forecasting of Convection and Why? Part I: On the Propagation of Initial Condition Errors and Their Implications for Data Assimilation. *Mon. Weather Rev.*, 138:242–255.

Fisher, M. (2004). Generalized frames on the sphere, with application to the background error covariance modelling. In *Proceedings of the ECMWF Seminar on recent developments in numerical methods for atmospheric and ocean modelling, 6-10 September 2004*, pages 87–102. Available from ECMWF, Reading, UK.

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

Gauthier, P. and Thépaut, J.-N. (2001). Impact of the Digital Filter as a Weak Constraint in the Preoperational 4DVAR Assimilation System of Météo-France. *Mon. Weather Rev.*, 129:2089–2102.

Geer, A. J. and Bauer, P. (2010). Enhanced use of all-sky microwave observations sensitive to water vapour, cloud and precipitation . Technical report. ECMWF Technical Memorandum No. 620, 41 pages, available from ECMWF, Reading, UK.

Geer, A. J., Bauer, P., and Lopez, P. (2010). Direct 4D-Var assimilation of all-sky radiances. Part II: Assessment. *Q. J. R. Meteorol. Soc.* Accepted.

Hólm, E., Andersson, E., Beljaars, A., Lopez, P., Mahfouf, J.-F., Simmons, A. J., and Thépaut, J.-N. (2002). Assimilation and Modelling of the Hydrological Cycle: ECMWF's Status and Plans. Technical report. ECMWF Technical Memorandum No. 383, 55 pages, available from ECMWF, Reading, UK.

Hou, A. Y., Ledvina, D. V., Silva, A. M. D., Zhang, S. Q., Joiner, J., Atlas, R. M., Huffman, G. J., and Kummerow, C. D. (2000). Assimilation of SSM/I-Derived Surface Rainfall and Total Precipitable Water for Improving the GEOS Analysis for Climate Studies. *Mon. Weather Rev.*, 128:509–537.

Janisková, M., Mahfouf, J.-F., Morcrette, J.-J., and Chevallier, F. (2002). Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation. *Q. J. R. Meteorol. Soc.*, 128:1505–1527.

Janisková, M., Thépaut, J.-N., and Geleyn, J.-F. (1999). Simplified and Regular Physical Parametrizations for Incremental Four-Dimensional Variational Assimilation. *Mon. Weather Rev.*, 127:26–45.

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

Lin, Y. and Mitchell, K. E. (2005). The NCEP Stage II/IV Hourly Precipitation Analyses: Development and Applications. In *Proceedings of the 19th AMS Conference on Hydrology, San Diego, CA (USA), 5–14 January 2005*.

Lopez, P. (2008). Comparison of OPERA precipitation radar composites to CMORPH, SYNOP and ECMWF model data. Technical report. ECMWF Technical Memorandum No. 569, 22 pages, Available from ECMWF, Reading, UK.

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.

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

Macpherson, B. (2001). Operational experience with assimilation of rainfall data in the Met.Office mesoscale model. *Meteorol. Atmos. Phys.*, 76:3–8.

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

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

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

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

Tompkins, A. M. and Janisková, M. (2004). A cloud scheme for data assimilation: Description and initial tests. *Q. J. R. Meteorol. Soc.*, 130:2495–2518.

Tong, M. and Xue, M. (2005). Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: OSS Experiments. *Mon. Weather Rev.*, 133:1789–1807.

Treadon, R. E., Pan, H.-L., Wu, W.-S., Lin, Y., Olson, W. S., and Kuligowski, R. J. (2002). Global and Regional Moisture Analyses at NCEP. In *Proceedings of the ECMWF Workshop on Humidity Analysis, 8-11 July 2002*, pages 33–47. Available from ECMWF, Reading, UK.

Tsuyuki, T., Koizumi, K., and Ishikawa, Y. (2002). The JMA Mesoscale 4D-Var System and Assimilation of Precipitation and Moisture Data. In *Proceedings of the ECMWF Workshop on Humidity Analysis, 8-11 July 2002*, pages 59–67. Available from ECMWF, Reading, UK.

Vukicevic, T., Sengupta, M., Jones, A. S., and Haar, T. V. (2006). Cloud-Resolving Satellite Data Assimilation: Information Content of IR Window Observations and Uncertainties in Estimation. *J. Atmos. Sci.*, 63:901–919.