Implementation of Deterministic Weather Forecasting Systems Based on Ensemble–Variational Data Assimilation at Environment Canada. Part II: The Regional System

JEAN-FRANÇOIS CARON,* THOMAS MILEWSKI,† MARK BUEHNER,* LUC FILLION,* MATEUSZ RESZKA,† STEPHEN MACPHERSON,* AND JUDY ST-JAMES†

* Data Assimilation and Satellite Meteorology Research Section, Environment Canada, Dorval, Quebec, Canada
† Data Assimilation and Quality Control Development Section, Environment Canada, Dorval, Quebec, Canada

(Manuscript received 14 October 2014, in final form 6 March 2015)

ABSTRACT

The modifications to the data assimilation component of the Regional Deterministic Prediction System (RDPS) implemented at Environment Canada operations during the fall of 2014 are described. The main change is the replacement of the limited-area four-dimensional variational data assimilation (4DVar) algorithm for the limited-area analysis and the associated three-dimensional variational data assimilation (3DVar) scheme for the synchronous global driver analysis by the four-dimensional ensemble–variational data assimilation (4DEnVar) scheme presented in the first part of this study. It is shown that a 4DEnVar scheme using global background-error covariances can provide RDPS forecasts that are slightly improved compared to the previous operational approach, particularly during the first 24 h of the forecasts and in the summertime convective regime. Further forecast improvements were also made possible by upgrades in the assimilated observational data and by introducing the improved global analysis presented in the first part of this study in the RDPS intermittent cycling strategy. The computational savings brought by the 4DEnVar approach are also discussed.

1. Introduction

In the first part of this study (Buehner et al. 2015, hereafter Part I), the latest modifications to the Global Deterministic Prediction System (GDPS) implemented operationally at Environment Canada (EC) in the fall of 2014 were described. The most notable change is the replacement of the four-dimensional variational data assimilation (4DVar) scheme by a four-dimensional ensemble–variational data assimilation (4DEnVar) scheme in which the background-error covariances are represented by a blend of homogeneous and isotropic covariances and 4D flow-dependent covariances derived from a global ensemble Kalman filter (EnKF). In this paper we report on the implementation of the same 4DEnVar scheme in the Regional Deterministic Prediction System (RDPS), a forecasting system based on a limited-area version of the Global Environmental Multiscale (GEM) model (Côté et al. 1998) covering North America (Fig. 1) with an approximate 10-km horizontal grid spacing. The previous RDPS relied on a limited-area 4DVar scheme (Tanguay et al. 2012, hereafter T12) for the limited-area analysis and a three-dimensional variational data assimilation (3DVar) scheme for the synchronous global driver analysis at EC operations between October 2012 and the fall of 2014. The forecast impacts resulting from improvements to the assimilated observational data and the introduction of the improved global analysis presented in Part I in the RDPS intermittent cycling strategy are also presented.

Similarly to global data assimilation applications (see Part I and references therein), considerable effort has been devoted in recent years to both comparing and combining ensemble- and variational-based limited-area data assimilation approaches (Wang et al. 2008a,b; Wang 2011; Li et al. 2012; Zhang and Zhang 2012; Zhang et al. 2013; Liu and Xiao 2013; Schwartz et al. 2013; Schwartz and Liu 2014; Gustafsson and Bojarova 2014; Pan et al. 2014). As reported by Zhang and Zhang...
using 4DVar in combination with an ensemble-derived covariance matrix (hybrid 4DVar) can outperform a stand-alone EnKF or a 4DVar in a limited-area forecasting system, in agreement with the results from Buehner et al. (2010b) in a global configuration. However, hybrid 4DVar still requires the use of tangent-linear (TL) and adjoint (AD) versions of the forecast model, whose time integrations dominate the cost of the analysis step and that requires, moreover, significant development and maintenance effort. On the other hand, as explained in Buehner et al. (2013, hereafter B13), 4DEnVar uses 4D ensemble covariances in a way that essentially replaces the use of TL and AD versions of the forecast model and has been shown to produce forecasts with similar or improved accuracy at short lead times in a global context by B13 and Part I. Recently Gustafsson and Bojarova (2014) reported that 4DEnVar can outperform both 4DVar and hybrid 4DVar in the context of the limited-area HIRLAM forecasting system. Therefore, the 4DEnVar approach described in Part I seemed appropriate for the RDPS in order to 1) improve, or at least maintain, the RDPS forecast accuracy obtained using the operational limited-area 4DVar scheme and 2) make more efficient use of limited resources at EC by moving toward a more unified data assimilation approach for deterministic and ensemble forecasting.

Since there is currently no operational equivalent to the global EnKF (Houtekamer et al. 2014) at the regional scale at EC, we simply based our 4DEnVar scheme for the RDPS limited-area analysis on the use of 4D ensemble covariances derived from the global EnKF as in the GDPS configuration described in Part I. Our approach is thus similar to the National Centers for Environmental Prediction (NCEP), which recently replaced the 3DVar scheme in the North American Mesoscale Forecast System (NAM) and the Rapid Refresh (RAP) regional forecasting system by a 3DEnVar scheme based on their global EnKF system (see National Weather Service 2014a,b).

In the next section, the forecast component and the different data assimilation approaches are presented. In section 3, it will be shown that a 4DEnVar scheme using global background-error covariances can provide RDPS
forecasts with either similar or slightly improved accuracy compared to the previous operational approach. In section 4, the forecast impacts resulting from various improvements to the observational data assimilated are presented while section 5 presents a thorough forecast performance evaluation of the configuration implemented at EC operations in the fall of 2014. Finally, a summary of the results as well as future directions are presented in section 6.

2. Design of the RDPS

a. Forecast model configurations and cycling strategies

The RDPS represents the main NWP guidance for forecasters of the Meteorological Service of Canada for days 1 and 2. It produces forecasts over North America (see again Fig. 1) up to $T + 48\ h$ ($T + 54\ h$) at 0000 and 1200 UTC (0600 and 1800 UTC) using a limited-area version of the GEM model [limited-area model (LAM); Côté et al. 1998] on a cylindrical equidistant grid with a horizontal grid spacing of 0.09° (approximately 10 km) and 80 vertical levels (model lid at 0.1 hPa, as in the GDPS). The GEM model is a two-time-step implicit, semi-Lagrangian, gridpoint model, with a latitude–longitude C grid that is staggered in the horizontal direction. Unlike the GDPS, the GEM model configuration in the RDPS still uses a hydrostatic-pressure coordinate [defined in Charron et al. (2012)] and a discretization on a regular (unstaggered) vertical grid due to compatibility issues between some physical process parameterizations used in the RDPS and the log-hydrostatic-pressure vertical coordinate defined on a staggered grid proposed by Girard et al. (2014).

This forecasting system employs an intermittent upper-air cycling strategy (Fig. 2) where, 6 h before the analysis time $T$, the (approximately 25 km) analysis from the GDPS is interpolated onto the LAM grid and initializes a short-term forecast (named LB in Fig. 2). This short-term forecast serves as the background state for the LAM analysis step at time $T$ (named LF in Fig. 2). In parallel, the same procedure is applied to a GEM driving model with a horizontal grid spacing of approximately 33 km. The resulting synchronous 33-km global driving analysis (named DF in Fig. 2) and forecast, as proposed and tested by Fillion et al. (2010, hereafter F10), allow the observations outside the LAM analysis domain to influence the LAM forecasts through the lateral boundary conditions (LBCs).1

The 33-km driving model uses the same vertical levels as the LAM in order to minimize interpolations in the prescription of the LBCs (updated every hour), whose

---

1 The addition of a 33-km global driver in the RDPS was necessary because the GDPS is executed after the RDPS in EC’s operational schedule.
formulation in the limited-area version of GEM is based on Thomas et al. (1998). In each forecast, a digital filter procedure (Fillion et al. 1995) is applied to eliminate the spurious gravity waves triggered by imbalances in the initial conditions. Interested readers are referred to Vaillancourt et al. (2012) and references therein for further details on the model configuration used in the RDPS, especially regarding the parameterization of the physical processes.

Both the LAM and the 33-km driving component use the Interactions between Soil, Biosphere, and Atmosphere (ISBA) land surface scheme, which has its own data assimilation and cycling strategy [see Bélair et al. (2003a,b) for complete details]. In the 33-km driver, the surface variables from the GDPS at $T - 6 \text{ h}$ are used to initialize the DB forecast (see Fig. 2) and no update of these variables is made in DF at time $T$ (i.e., DF starts from 6-h forecast surface variables). The LAM component uses a continuous data assimilation cycling strategy where surface temperatures and soil moisture contents are updated every 24 h at 0000 UTC. LAM forecasts (LB and LF) initialized at 0600, 1200, and 1800 UTC simply rely on the forecast surface variables from the most recent 0000 UTC surface analysis.

b. Data assimilation approaches

1) LIMITED-AREA 4DVAR AND GLOBAL 3DVAR

The previously operational 4DVar scheme used to perform the LAM analysis is based on the formulation described in T12 (to our knowledge, unique among operational centers) and is a temporal extension of the 3DVar scheme proposed by F10. It operates with an incremental formulation (Courtier et al. 1994) to produce analysis increments on a global 400 × 200 Gaussian grid (i.e., with a horizontal grid spacing of approximately 100 km). The homogeneous and isotropic background-error covariances used at the start of the assimilation window ($T - 3 \text{ h}$) are defined with a global spherical-harmonics spectral representation, but the TL and AD models are defined over a limited-area domain exceeding (by 30% in each direction) the horizontal grid of the high-resolution ($\sim 10 \text{ km}$) nonlinear model (see Fig. 5 in T12). Because the TL and AD model grid chosen is an exact subdomain of the global Gaussian analysis increment grid, the communication of information (i.e., analysis increments and adjoint sensitivities) is direct, with no spatial interpolations required. Note that only observations over the high-resolution nonlinear model domain are assimilated and that the nonlinear trajectory for the extended TL and AD limited-area models is obtained from an additional nonlinear forecast performed on the extended limited-area horizontal grid ($\sim 100 \text{ km}$) and initialized with the GDPS analysis. The analysis increments generated on the global Gaussian grid are then interpolated to the high-resolution limited-area grid to produce the LAM analysis.

This configuration mimics the solution of a global 4DVar (but at a reduced computational cost) in order to prevent distortions in the analysis increments near the lateral boundaries of the nonlinear model and to optimize the analysis of the large scales in this continental-scale LAM. The TL and AD models used here include the same dynamical processes as the high-resolution forecast model but with only two physical processes: simplified vertical diffusion and simplified grid-scale condensation. Only one outer loop with 45 iterations is performed and the resulting global analysis increments at low resolution ($\sim 100 \text{ km}$) are then interpolated to the high-resolution LAM grid ($\sim 10 \text{ km}$) and added to the background state at $T - 3 \text{ h}$. A 3-h nonlinear forecast is then needed to carry this analysis to time $T$. See T12 and F10 for further details. We remark that this 4DVar system presents a significant difference in horizontal grid spacing between the nonlinear model and the analysis increments ($\sim 10 \text{ vs} \sim 100 \text{ km}$). This is due to the high computational cost of 4DVar as well as to the tight operational constraints imposed on the RDPS at EC.

The 33-km global driver analysis, on the other hand, is obtained from a 3DVar system since a global 4DVar system would require too much time to meet EC’s operational constraints for RDPS products. The analysis increments are also produced on a global 400 × 200 ($\sim 100 \text{ km}$) Gaussian grid but using the same background-error covariances employed by the global 4DVar in the previously operational version of the GDPS (see Charron et al. 2012) and with all the observational data available over the entire globe.

2) 4DEnVAR USING GLOBAL BACKGROUND-ERROR COVARIANCES

The 4DEnVar scheme tested and adopted in the RDPS is identical to the system implemented in the GDPS and presented in section 2a of Part I and in B13. Using an analysis increment grid that matches the horizontal grid spacing of the 4D background-error covariances obtained from the global EnKF ($\sim 66 \text{ km}$ in sections 3 and 4; $\sim 50 \text{ km}$ in the final configuration presented in section 5), it is possible to obtain analysis increments at a higher resolution than the limited-area 4DVar scheme presented above. Because no TL or AD model integrations are necessary in 4DEnVar, the computational cost of performing an LAM analysis at a higher resolution using the ensemble covariances from the global EnKF is still lower than the limited-area
4DEnVar scheme. Moreover, the computational efficiency of the 4DEnVar makes it possible to use this approach for the 33-km global driver analysis. A comparison of timings and computational resources between the different data assimilation approaches is presented in section 5d.

The ensemble-derived background-error covariances are blended with the same homogeneous and isotropic background-error covariances (so-called hybrid approach) used in the GDPS for each analysis (LAM and 33-km global driver). That is, we use the same vertical profile of weighting factors (equal 50/50 weights in the troposphere; see Fig. 1 in B13), the same spatial localization design [see section 2b in Buehner et al. (2010a)], and the same time resolution for the ensemble covariances (1h) employed in the GDPS described in Part I. As in the 4DVar scheme, the LAM analysis increments from this 4DEnVar scheme are thus generated on a global grid and then interpolated to the high-resolution limited-area grid to produce the LAM analysis. Again, only observations over the LAM domain are assimilated in the LAM analysis step.

Unlike the GDPS, the RDPS based on 4DEnVar still uses a cold start strategy (i.e., without cycling certain physics variables, such as clouds) and a nonincremental digital filter initialization procedure in forecasts initialized at time $T$ simply because the incremental analysis update and the physics recycling (warm start) capability now used in the GDPS (see section 2g in Part I) are not available in the older version of the GEM model used in the LAM and the driver components. Therefore, it was decided to select the 4DEnVar analysis increment valid at time $T$ obtained after 50 (70) iterations in the LAM (driver) analysis, which further reduces the computational cost compared to 4DVar, since no forecast step is needed to carry the analysis from $T-3h$ to $T$.

3. Comparison between 4DEnVar and 4DVar (3DVar)

a. Experiment configuration

To examine the feasibility of an operational implementation of the 4DEnVar approach in the RDPS, we conducted two data assimilation experiments similar to B13, where the only difference is the replacement of the 4DVar (3DVar) scheme in the LAM (driver) analysis by the 4DEnVar scheme (hereafter referred to as experiments 4D1 and EN1, respectively). For each winter (February and March) and summer (July and August) period of 2011, 118 (in total) 48-h forecasts were initialized at 0000 and 1200 UTC. The same periods of investigation were used for every experiment in this paper. Table 1 in Part I lists some of the observations assimilated in each experiment. The global ensemble of background states in EN1 was taken from the EnKF configuration using 192 members with a ~66-km horizontal grid spacing, as reported in Houtekamer et al. (2014). The GDPS analyses at $T-6h$ were taken from a retrospective experiment using the system configuration that was operational at EC between 13 February 2013 and the fall of 2014 and that used a ~25-km horizontal grid spacing model in combination with a 4DVar scheme producing analysis increments with a ~100-km horizontal grid spacing (see Zadra et al. 2014).

A comparison of 4DEnVar and 4DVar with analysis increments at the same horizontal resolution was outside the scope of this paper. In the context of comparing approaches for potential operational use, adopting a higher horizontal resolution in 4DEnVar is appropriate because of the significantly lower computational cost of 4DEnVar as compared with 4DVar (see section 5d). It was also outside the scope of this paper to compare 4DEnVar and three-dimensional ensemble–variational data assimilation (3DEnVar). Interested readers are referred to B13 and Wang and Lei (2014) for a demonstration of the benefit of time-evolving ensemble-derived background-error covariances.

Unlike the RDPS operational configuration, the surface data assimilation cycle was not activated. Instead, the surface variables needed to initialize the LAM in both experiments were taken from the ~15-km RDPS, which was operational in 2011 and that relied on a 3DVar scheme for the LAM analysis as described in F10. See Table 1 for a summary of 4D1 and EN1 experiments.

b. Impact on forecasts

Figure 3 shows the standard deviation (solid curves) and bias (dashed curves) of forecast minus observation using North American radiosonde observations and $T+24$-h LAM forecasts. The differences in the forecast scores are quite small. For zonal winds, some reductions in the standard deviation around ~250 hPa can be seen for both winter (Fig. 3a) and summer (Fig. 3b) in the EN1 experiment. In summer, the small reductions also extend to the lower troposphere and to the bias in the upper levels. Regarding temperature, the forecasts in EN1 and 4D1 are essentially equivalent in winter (Fig. 3c) but in summer, some reductions occurred in the EN1 standard deviation above 250 hPa and in the EN1
bias at some levels (except above 30 hPa where 4D1 is better). At $T + 48$ h, the differences between EN1 and 4D1 are reduced (Fig. 4). In winter, no statistically significant differences can be seen in the standard deviation of zonal wind (Fig. 4a) or temperature (Fig. 4c), while the biases show either no differences or a mixed signal. However, in summer some marginal reductions in EN1 are maintained for both zonal wind (Fig. 4b) and temperature (Fig. 4d), except for biases above 30 hPa where 4D1 performs better. The fact that our 4DEnVar scheme reverts to a 3DVar scheme above 10 hPa due to the lower model lid used in the EnKF could explain why the forecasts from 4D1 exhibit smaller biases above 30 hPa. The greater improvements from 4DEnVar in summer compared to winter are probably the result of the lack of moist physical processes in the TL and AD models, which impedes the performance of the 4DVar scheme compared to 4DEnVar, in which the analysis increments are derived from a linear combination of nonlinear forecast perturbations. The higher horizontal resolution of the analysis increments in 4DEnVar could also have contributed to further improve the forecasts compared to 4DVar in summer.

Forecast comparison against radiosonde observations shows that using the 4DEnVar approach in the RDPS not only preserves the forecast accuracy relative to the 4DVar and 3DVar analyses but also brings some small improvements at some upper-air levels. This is consistent with the results in the Northern Hemisphere in B13 and with the improvements observed over the eastern North Atlantic Ocean in winter by Gustafsson and Bojarova (2014), although that study reported much larger benefits especially in the lower troposphere. A comparison with observations from surface-based weather stations also led to similar conclusions (not shown). It is important to stress that adopting the 4DEnVar scheme only in the LAM analysis (i.e., keeping the 3DVar scheme for the driver analysis) led to somewhat smaller improvements at $T + 48$ h although no forecast degradation below 20 hPa was observed compared to 4D1 (not shown). Based on the above results, it was decided to adopt the 4DEnVar approach in the RDPS for both the LAM and the global driver analyses.

### 4. Changes to the observational data assimilated

#### a. Summary of changes

The various improvements to the processing and the volume of data assimilated in the GDPS described in section 2 of Part I were also implemented in the RDPS. In short, the most important changes are as follows: 1) coefficients from a revised satellite radiance bias correction scheme were implemented, 2) the processing and the assimilation of radiosonde and aircraft data were upgraded [this includes introducing a bias correction strategy for aircraft temperature data and taking into account the horizontal drift of radiosonde balloons as in Laroche and Sarrazin (2013)], 3) the number of assimilated channels for Atmospheric Infrared Sounder (AIRS) and Infrared Atmospheric Sounding Interferometer (IASI) radiances was increased, and, finally, 4) zenith tropospheric delay (ZTD) data from ground-based GPS (GB-GPS) receivers (sensitive to precipitable water) over North America are now being assimilated (see sections 2b–f and Table 1 in Part I for complete details). Although most changes were tested individually in separate experiments, we present here, for simplicity, the impact from the introduction of all the above changes in an experiment called EN2, based on the configuration of experiment EN1 (see again Table 1). This choice is motivated by the fact that the impact of each individual change was found to be very similar to that in the global system. In other words, the examination of the individual modifications in the context of the RDPS led to no further understanding. This is not surprising given that the two forecasting systems are now based on the same 4DEnVar scheme. Interested readers

---

**Table 1. Summary of experiments.**

<table>
<thead>
<tr>
<th>Analysis (approx increment horizontal grid spacing)</th>
<th>Global background ensemble (approx horizontal grid spacing)</th>
<th>Assimilated obs</th>
<th>Surface variables in LAM</th>
<th>GDPS analysis at $T - 6$ h (approx increment/analysis horizontal grid spacing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4D1</td>
<td>4DVar/3DVar (100 km)</td>
<td>—</td>
<td>Standard</td>
<td>No cycling</td>
</tr>
<tr>
<td>4DF</td>
<td>4DVar/3DVar (100 km)</td>
<td>—</td>
<td>Standard</td>
<td>Cycling</td>
</tr>
<tr>
<td>EN1</td>
<td>4DEnVar (66 km)</td>
<td>192 members (66 km)</td>
<td>Standard</td>
<td>No cycling</td>
</tr>
<tr>
<td>EN2</td>
<td>4DEnVar (66 km)</td>
<td>192 members (66 km)</td>
<td>Upgraded</td>
<td>No cycling</td>
</tr>
<tr>
<td>ENF</td>
<td>4DEnVar (50 km)</td>
<td>256 members (50 km)</td>
<td>Upgraded</td>
<td>Cycling</td>
</tr>
</tbody>
</table>

3 No bias correction coefficients are computed in the RDPS. The coefficients are simply taken from the GDPS.
are, therefore, referred to Part I and references therein for the individual examination of each of the above changes in the context of the GDPS.

b. Impact on forecasts

To be consistent with the assimilation of radiosonde data at their reported or estimated geographical positions during the balloon ascent in EN2, forecasts from both EN2 and EN1 experiments were verified against radiosonde data by taking into account the horizontal drift. This explains why EN1 scores in Fig. 5 are slightly different from the values presented in Fig. 4. However, as in Part I and in Laroche and Sarrazin (2013), all the observations used for the verification were assumed to be valid at the synoptic time (0000 or 1200 UTC) even though the reported (or estimated) observation times were taken into account in the data assimilation step of EN2. This verification strategy will also be used in section 5.

The verifications of $T + 48$-h forecasts show some improvements for temperature and zonal winds in both seasons mostly above 500 hPa in EN2 (Fig. 5). The small
reductions in the standard deviation can be attributed mainly to the improved assimilation of the radiosonde data. As for biases, where larger changes can be seen, the reductions are mainly due to the combined effect of the improved radiance bias correction and the introduction of a bias correction strategy for aircraft temperature data. The former are the main contributor to the reductions of zonal wind (temperature) bias above 100 hPa (above 30 hPa) while the latter contribute mainly to the reduction of the temperature bias between 100 and 300 hPa. Increasing the volume of AIRS and IASI radiance data assimilated had a neutral impact in EN2. Finally, the assimilation of the GB-GPS ZTD data had a significant positive impact on summertime forecasts of precipitable water and precipitation (not shown, but verifications for these variables are presented in the examination of the final configuration in section 5c).

5. Final configuration

a. Changes with respect to EN2

In the last experiment called ENF, the 4DEnVar-based RDPS was configured as closely as possible to the system implemented at EC operations in the fall of 2014. The GDPS analyses at $T = 6\, \text{h}$ were taken from the 4DEnVar-based system presented in Part I and the ensemble covariances were taken from the version of the EnKF using 256 members with a $\sim 50\, \text{km}$ horizontal grid spacing. The analysis increments for both the LAM and the driver were thus computed on a global $800 \times 400$ ($\sim 50\, \text{km}$)
Gaussian grid. To measure the full impact of the changes, the surface cycling strategy in the LAM (see section 2a) was activated in ENF and in a new 4DVar-based control experiment called 4DF (see again Table 1). Finally, as in the GDPS, an error in the formulation of the prognostic equations for the snow canopy density was corrected in the model version used for the LAM and the driving forecasts in ENF, leading to an approximate doubling of the snow density seen by the models. This correction in the model formulation has a significant impact on the near-surface temperature in winter as will be shown later in section 5c together with other changes in the RDPS. Again, the impact from this change in the LAM examined in isolation was quite similar to that in the global model at the corresponding forecast lead times. Readers interested in the sole impact of this modification to the model formulation are, therefore, referred to section 4a of Part I where it is described in the context of the GDPS.

b. Impact on forecasts: ENF versus EN2

We first examined the impact of the changes in ENF with respect to the previous 4DEnVar-based experiment EN2 presented in section 4. As shown in Fig. 6, the improved GDPS analysis and the upgraded ensemble-derived background-error covariances led mostly to small but statistically significant improvements in 148-h forecasts when compared to radiosonde observations. In the winter period, the standard deviation for zonal wind (Fig. 6a) and temperature (Fig. 6c) is reduced in the upper troposphere and in the stratosphere. In
summer, the reductions also extend to the lower troposphere (Figs. 6b,d) although the improvements for zonal winds are smaller in the upper troposphere compared to winter. In terms of biases, the impact is neutral or mixed, except for temperature in winter, where reductions can be seen above 400 hPa but at the expense of some increase within the boundary layer (Fig. 6c).

c. Impact on forecasts: ENF versus 4DF

We now perform a thorough comparison of forecast performance between the new operational configuration (ENF) and the previous one (4DF) for the 2011 two periods again using observations from radiosondes but also from the North American network of surface-based weather stations and GB-GPS receivers.

1) Upper-air verification

Given the improvements from the incremental changes shown in sections 3b, 4b, and 5b, it is not surprising to find that the $T + 48$-h forecasts from ENF are closer to radiosonde observations at most vertical levels than 4DF in both winter and summer periods (Fig. 7). In terms of standard deviation, the reductions to the zonal winds and the temperature, similar in each season, span almost all pressure levels and are maximal in the upper troposphere. As for the biases, reductions for the zonal winds are present only in the stratosphere (i.e., above 150 hPa; see Figs. 7a,b) while the temperature bias reductions also extend to the upper troposphere (i.e., up to 300 hPa; see Figs. 7c,d). Similar improvements can also be seen at shorter lead times (not shown).
2) GROUND-BASED GPS

Forecasts of precipitable water were compared to the values derived from GB-GPS ZTD observations. The confidence intervals for the scores reported here and in the following subsection were estimated from the bootstrap resampling technique described in Candille et al. (2007, see their section 2c).

Figure 8 shows the differences between ENF and 4DF for forecasts in the summer period, initialized at 0000 UTC (Fig. 8a) and 1200 UTC (Fig. 8b). In both cases, a significant reduction of the standard deviation can be seen throughout the forecast period with a maximum reduction of about 1 mm at $T + 6$ h. In terms of biases, the moist bias in 1200 UTC forecasts is also reduced throughout the forecast period (Fig. 8b) whereas no significant differences (except at $T + 18$ h) can be seen in forecasts initialized at 0000 UTC (Fig. 8a). These improvements are largely due to the addition of GB-GPS ZTD observations in the data assimilated in both the RDPS and the GDPS in ENF and are in agreement with the findings of Macpherson et al. (2008). In the winter period, when precipitable water is very low, no change was observed in the verification scores (not shown).

3) SURFACE VERIFICATION

The forecasts were also verified against the aviation routine weather report (METAR) and synoptic observation data available operationally at EC, as in Wilson

---

4 As explained in section 2c of Part I, the ZTD data are currently neither bias corrected in the assimilation step nor the verification step.
and Vallée (2003). Only reports from surface-based stations with an elevation difference with the LAM topography of less than 100 m were used, which led to a total of 1282 stations being considered: 984 in Canada and 210 in the United States. Therefore, the verification scores presented here are more representative of forecast performance over Canada. Statistics were partitioned between forecasts initialized at 0000 and 1200 UTC in order to capture the different behavior during the daytime and nighttime regimes.

We first examine the verification of temperature and dewpoint temperature at screen level. During the summer period, the temperature standard deviation (Figs. 9a,b) in ENF shows some reductions compared to 4DF throughout most of the forecast period, but the differences never exceed 0.1 K, except at the initial time, mainly due to a recurrent increased fit to the observations in 4DEnVar-based analyses. In terms of biases, a small reduction can be seen in the first 12 h of 1200 UTC forecasts (Fig. 9d) while no difference can be
observed in 0000 UTC forecasts (Fig. 9c). The reductions are clearer in the dewpoint temperature standard deviation for forecasts from both 0000 (Fig. 10a) and 1200 UTC (Fig. 10b) with a maximum difference between 0.1 and 0.2 K in the first 12 h. The averaged near-surface humidity is somewhat increased in ENF, but this does not lead to significant changes in forecast bias (Figs. 10c,d).

In the winter period, the forecast differences are larger due to the aforementioned correction to the prognostic equations for snow canopy density in the ENF model formulation. The standard deviation is reduced at short lead times \((T + 3 \text{ h})\) by up to 0.2 K for temperature (Figs. 11a,b) and up to 0.4 K for dewpoint temperature (Figs. 12a,b). The improvements are even larger when considering observations over Canada only (not shown) because of the larger extent of the snow canopy. The correction in the model surface scheme in ENF reduced the variation of the bias with lead time for both temperature and humidity. In terms of temperature, this had a mixed impact, as both ENF and 4DF experiments exhibit a cold bias (Figs. 11c,d) whereas this led to improvements in the verification for dewpoint temperature (Figs. 12c,d) throughout most of the 48-h forecast period and reaching as much as 0.4 K up to \(T + 12 \text{ h}\) for forecasts initialized at 0000 UTC (Fig. 12c).
Precipitation accumulations were also verified using the equitable threat score (ETS; see e.g., Mason 2003) for five thresholds ($0.5, 2, 5, 10,$ and $15$ mm) of 12-h accumulation from synoptic reports. For simplicity, forecasted accumulations were linearly interpolated to the observation locations (i.e., no technique was applied to alleviate the representativeness errors between the forecast model and the observations). Again, some improvements in ENF compared to 4DF can be seen at most of the forecast lead times (Figs. 13 and 14) although many are barely statistically significant. In summer, increases in ETS in ENF forecasts from 0000 UTC were found for the smallest (largest) accumulation thresholds during the nighttime (daytime) for days 1 and 2 (Figs. 13a,c vs Figs. 13c,g) whereas ENF forecasts from 1200 UTC showed only improvements over 4DF for each accumulation threshold in the first 12 h (daytime of day 1; Fig. 13b) and to the smallest accumulation thresholds in the daytime of day 2 (Fig. 13f). In the winter period, the ENF forecasts from 1200 UTC showed more improvement, as higher ETS from ENF can be seen for most of the accumulation thresholds at each of the forecast lead times (Figs. 14b,d,f,h) whereas the improvements in forecasts from 0000 UTC are limited to the largest accumulation thresholds for daytime at days 1 (Fig. 14c) and 2 (Fig. 14g). Comparison of the precipitation biases shows a mostly neutral impact from ENF except in the first 12 h of the forecasts in winter and in a few of the larger accumulation thresholds in summer where an increase in the precipitation led to a small degradation compared to 4DF (not shown).
A comparison against surface pressure observations also showed some small improvements in ENF compared to 4DF in terms of the standard deviation and bias (not shown). Finally, verifications of wind fields (speed and direction) and cloud cover against observations revealed no significant differences in forecast performance between ENF and 4DF (not shown).

d. Computational resources

In the previous section, it was shown that forecasts from the new operational configuration (ENF) are generally improved compared to the previous one (4DF). We show here that these forecast improvements are also associated with a major reduction in the computational cost of the analysis steps in the RDPS. Table 2 shows the average computational resources of the analysis steps (LAM and global driver) for the two aforementioned experiments, where the computational cost is defined in terms of the product between the number of CPUs used and the elapsed wall-clock time. In the LAM component, the replacement of the 4DVar reduces the computational cost by an order of magnitude due to significant reductions in both the required number of CPUs (320 vs 2048) and in the wall-clock time (7 vs 17 min). This is largely the result of replacing the integration of the TL and AD models in 4DVar by the use of an ensemble of nonlinear model states to estimate 4D background-error covariances over the assimilation time window. However, it is important to stress that 4DEnVar analyses are generated using an updated version of the variational analysis program, and benefit from improvements in the optimization of the Fortran code, especially from a better usage of the message passage interface (MPI) protocol. Note that the
6. Summary and discussion

In the second part of the description of the latest changes implemented in the deterministic forecasting systems at EC operations in the fall of 2014, we described the modifications to the Regional Deterministic Prediction System (RDPS), a limited-area system providing NWP guidance up to day 2 using a LAM covering North America with an approximate 10-km horizontal grid spacing. The main modification is the replacement of the limited-area 4DVar algorithm for the LAM analysis and the associated 3DVar scheme for the synchronous global driver analysis by the same 4D ensemble–variational (4DEnVar) algorithm implemented in the GDPS (Part I), in which the background-error covariances are represented by a blend of homogeneous and isotropic covariances and 4D flow-dependent covariances derived from a global EnKF. Through verifications of forecast experiments covering a two-month period in two different seasons (winter and summer) of 2011 against North American radiosonde observations, we first showed that the 4DEnVar scheme
Fig. 13. Verification of forecasts from the ENF (red) and 4DF (blue) experiments against 12-h precipitation accumulation measured in terms of the ETS as a function of accumulation threshold ($\gamma$; mm) for forecasts initialized at (a),(c),(e),(g) 0000 and (b),(d),(f),(h) 1200 UTC for the summer period. Verifications for accumulation from 0 to +12 h are shown in (a),(b); from +12 to +24 h are shown in (c),(d); from +24 to +36 h are shown in (e),(f); and from +36 to +48 h are shown in (g),(h). The scores were obtained from an average over 59 cases. For each comparison, the difference between the two experiments (ENF minus 4DF) is shown below with the gray shaded area representing the 90% confidence interval based on a 2000-member bootstrap analysis.
using global background-error covariances presented in Part 1 can provide RDPS forecasts with either similar or slightly improved accuracy compared to the previously operational 4DVar- and 3DVar-based configuration. The greatest benefits were observed during the summer period and are believed to be an indication of the lack of moist physical processes representation in our TL and AD models, which hampered the performance of the

Fig. 14. As in Fig. 13, but for the winter period.
TABLE 2. Comparison of average computational resources used for the analysis steps of the RDPS in experiments ENF and 4DF on an IBM Power 7 system. The wall clock represents the elapsed time between the start of the variational analysis program and the end of the writing of the analysis increment. The cost is simply defined as the product of the wall-clock time and the number of CPUs used. The cost of the nonlinear forecast from $T - 3\ h$ to $T$ in 4DVar is not included.

<table>
<thead>
<tr>
<th></th>
<th>CPUs</th>
<th>Wall clock (min)</th>
<th>Cost (CPU min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4DVar</td>
<td>2048</td>
<td>17</td>
<td>34816</td>
</tr>
<tr>
<td>4DEnVar</td>
<td>320</td>
<td>7</td>
<td>2240</td>
</tr>
<tr>
<td>Global driver</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3DVar</td>
<td>64</td>
<td>9</td>
<td>576</td>
</tr>
<tr>
<td>4DEnVar</td>
<td>320</td>
<td>9</td>
<td>2880</td>
</tr>
</tbody>
</table>

The cost of the nonlinear forecast from $T - 3\ h$ to $T$ in 4DVar is not included.

4DVar scheme. The higher horizontal resolution of the analysis increments in 4DEnVar could also have contributed further to improve the forecasts compared to 4DVar in summer.

As in the GDPS, several improvements to the processing and the volume of data assimilated in the RDPS were introduced and contributed to further small forecast improvements, mainly in the upper troposphere and stratosphere. The final new configuration was obtained by initializing the RDPS intermittent upper-air cycle with the improved 4DEnVar-based GDPS analyses presented in Part I and by adopting the ensemble of backgrounds from the upgraded global EnKF system, which led to further enhancements in the forecast performance. From an extensive objective comparison with observations from the network of radiosonde, surfaced-based weather stations and ground-based GPS receivers over North America, it was shown that forecasts from the new operational version of the RDPS generally show an improved fit to observations for many variables throughout the 48-h forecast period during each of the two-month periods considered here. The benefits over the previous 4DVar-based operational version were also confirmed by a subjective evaluation from experienced forecasters at EC (not shown). This new version of the RDPS became operational on 18 November 2014 at EC together with the changes to the GDPS described in Part I and to the EnKF. Readers interested in the objective and subjective evaluations of the RDPS parallel run that preceded this operational implementation are referred to Canadian Meteorological Centre (2014).

Replacing the 4DVar scheme by the 4DEnVar approach also led to a 10-fold reduction in the computational cost of the LAM analysis step despite the reduction in horizontal grid spacing of the analysis increments from $\sim 100$ to $\sim 50\ km$. This is largely due to the replacement of the integration of TL and AD models in 4DVar by the use of an ensemble of nonlinear model states to estimate four-dimensional background-error covariances over the assimilation time window. The LAM analysis step now only takes about 7 min compared to 20 min with 4DVar (17 min to generate the analysis increment plus 3 min for its nonlinear propagation from $T - 3\ h$ to $T$), the maximum real-time allocation available at EC operations for the analysis step in the RDPS. Adopting the 4DEnVar approach in both the RDPS and the GDPS will also allow a more efficient use of limited resources at EC by moving toward a more unified data assimilation approach for deterministic and ensemble forecasting.

The RDPS adopted the same 4DEnVar configuration as the GDPS for the LAM analysis only because there is currently no operational equivalent to the global EnKF at the regional scale at EC. However, significant efforts have recently been devoted at EC to developing limited-area EnKF systems (e.g., Chang et al. 2014) and a North American continental-scale configuration is currently being tested for an operational implementation in the near future (see Baek et al. 2014). There are also plans to increase the horizontal resolution of the forecast model in the RDPS to a convection-permitting scale (i.e., with a grid spacing of a few kilometers). Our future work will thus be devoted to developing a limited-area 4DEnVar scheme to improve the analysis and forecast at small scales, where the importance of moist processes and nonlinearities should give further advantage to the 4DEnVar approach relative to 4DVar. We also expect that the computational savings brought by the 4DEnVar scheme in the analysis step will ease the operational implementation of this convection-permitting system since the computational cost of the forecast step is expected to increase dramatically.

Acknowledgments. The authors thank their colleagues, Drs. Stéphane Laroche and Sylvain Heilliette, for their guidance on the upgrades to the assimilated observational data and Marcel Vallée for providing the surface verification package. We also thank Allan Rahill and Charles Creese for their careful subjective evaluation of the forecasts, as well as Nathalie Gauthier for pointing out the error in the surface scheme of the model.

REFERENCES


Bélair, S., L.-P. Crevier, J. Mailhot, B. Bilodeau, and Y. Delage, 2003a: Operational implementation of the ISBA land surface


Zadra, A., and Coauthors, 2014: Improvements to the Global Deterministic Prediction System (GDPS) (from version 2.2.2 to 3.0.0), and related changes to the Regional Deterministic Prediction System (RDPS) (from version 3.0.0 to 3.1.0), Canadian Meteorological Centre Tech. Note, 88 pp. [Available online at http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/op_systems/doc_opchanges/technote_gdps300_20130213_e.pdf.]
