Status of the Global EPS at Environment Canada

Peter Houtekamer, Martin Charron, Herschel Mitchell, and Gérard Pellerin

Meteorological Research Division
Dorval, Québec, Canada
peter.houtekamer@ec.gc.ca, martin.charron@ec.gc.ca,
herschel.mitchell@ec.gc.ca

ABSTRACT

A Monte-Carlo strategy is followed in the Canadian ensemble prediction system (EPS) to arrive at ensembles in which the differences between ensemble members are representative of the difference between the ensemble mean and the unknown true state of the atmosphere. The main difficulty is how to sample the unknown defects of the assimilation and prediction system. This paper reviews the variety of methods to sample this "model error" that have been developed and implemented in the Canadian EPS.

1 Introduction

The Canadian operational medium-range Ensemble Prediction System (EPS) has been designed to obtain a representative ensemble by comprehensively simulating the behavior of errors in the forecasting system. Sources of uncertainty that are deemed to be significant are sampled by means of random perturbations that are different for each member of the ensemble. Arguments for the use of nonselective, purely random, ensemble perturbations are presented in Anderson (1997) and Descamps and Talagrand (2007). The only deviation in the EPS from a pure Monte Carlo approach is, as we will see, that different versions of the model are selected following a pre-arranged non-random procedure to sample part of the model error component.

In an ensemble Kalman filter (EnKF) (Evensen 1994; Houtekamer and Mitchell 1998) a sufficiently large representative ensemble of first-guess fields provides flow-dependent statistics to the data-assimilation procedure. Since 12 January 2005, an EnKF is used operationally to provide an ensemble of initial conditions to the medium-range EPS (Houtekamer et al. 2005; Houtekamer and Mitchell 2005). Having an ensemble with reasonably representative ensemble statistics from the early short-range onwards, in principle, permits new applications of the global EPS such as piloting a short-range regional EPS.

A successful EPS should simulate the effect of both initial and model-related uncertainties on forecast errors (Buizza et al. 2005). Recently, a major development effort went into the implementation of the stochastic perturbation of physical tendencies (first proposed by Buizza et al. 1999) and Stochastic Kinetic Energy Backscatter (SKEB, first proposed by Shutts 2005). Together with the addition of isotropic random errors and the use of different model versions, the Canadian EPS thus uses four different complementary methods to simulate and parameterize model error.

Section 2 gives an overview of the global EPS. Section 3 describes some critical aspects of the EnKF implementation. Section 4 describes how model error is handled in the EnKF. Section 5 describes how initial conditions for the medium-range forecasts are obtained from the EnKF data-assimilation cycle. Section 6 describes how model error is accounted for in the medium-range ensemble forecasts. Section 7 gives some results and finally section 8 summarizes the current status and suggests areas where future work is needed.

2 The global EPS

Twice daily, the Meteorological Service of Canada (MSC) produces a 16-day 20-member medium-range ensemble forecast. Since 10 July 2007, both the data-assimilation and the forecasting components of the EPS use the gridpoint Global Environmental Multiscale model (GEM; Côté et al. 1998) for all members. This model is used with a uniform 400×200 global grid, with 28 η-levels and with the model top at 10 hPa.

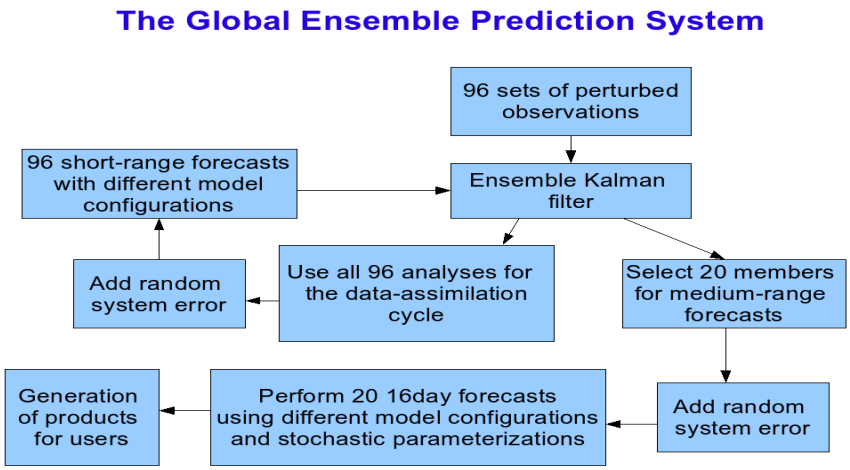

Figure 1: Components of the Global Ensemble Prediction System

Figure 1 shows how the data-assimilation and forecasting components are linked together. We notice that 96 members are used in the EnKF whereas only 20 members are used for the medium-range ensemble forecast. Essentially, we need a large number of members in the EnKF in order to obtain a high-dimensional analysis increment based on the ensemble statistics for the background error and based on the O(200 000) observations. On the other hand, the users of the EPS medium-range products are mostly interested in probability distributions for scalars, and providing such distributions involves only low-dimensional estimation problems. In the sequel, the data-assimilation and forecasting components, as well as how they are connected together, will be described in more detail.

3 The ensemble Kalman filter

The EnKF approximates the optimal Kalman filter (Kalman 1960) using a finite size representative ensemble. Due to the limited ensemble size, it is necessary to introduce additional terms that are not suggested by the Kalman filter equations. In particular, it is necessary to reduce the impact of noisy low-amplitude correlations. The multiplication of ensemble-based correlations by a finite-support fifth-order piecewise rational function (Gaspari and Cohn 1999, Eq. (4.10)) is used for this purpose (Houtekamer and Mitchell 2001; Whitaker and Hamill 2002). In the operational EnKF, the impact of observations is smoothly forced to zero over a horizontal distance of 2800 km and over two units of ln(pressure) in the vertical (Houtekamer et al. 2005).

To reduce the cost of the matrix operations that are required to compute the Kalman gain matrix, observations are assimilated in a batch-wise sequential manner (Anderson and Moore, 1979, pp. 142-146; Houtekamer and Mitchell 2001). The currently operational EnKF uses at most 300 observations per batch, however with the forthcoming assimilation of profiler observations, which are available at high spatial and temporal density over the American plains (St-James and Laroche 2005), it was found preferable to increase the maximum batch size

to 900.

The MSC recently (Gauthier et al. 2007) replaced the 3d-variational data-assimilation system that was used in the high-resolution deterministic forecasting system with a 4d-variational algorithm. In the 4d algorithm, the forward interpolation operator includes a temporal interpolation of the model trajectory to the observation times. With this more accurate forward operator, it is possible for the 4d algorithm to benefit from observations near the edges of the assimilation window. In the previous 3d algorithm, it was necessary to remove observations far from the central analysis time. The EnKF uses a quality controlled observation dataset that is supplied by the variational analysis. The big positive impact of the change from 3d to 4d-var and the corresponding availability of quality controlled observations near the edges of the assimilation window, motivated the implementation of time interpolation also into the EnKF algorithm.

Time-interpolation has been implemented following Houtekamer and Mitchell (2005, sect. 4e). The ensemble of model trajectories is sampled at five time levels (at 3.0, 4.5, 6.0, 7.5 and 9.0 h). The model states at these five time levels are subsequently used to evaluate the forward interpolation operator $\mathscr{H}(x^f(t_{obs}))$ at the time of validity t_{obs} of each observation. Subsequently an ensemble of joint state-observation vectors $(x(t = 6\text{ h}), \mathscr{H}(x^f(t_{obs})))$ is formed as suggested by Tarantola (1987) and Anderson (2001). Observations are still assimilated in batches and during this sequential algorithm the background ensemble of joint state-observation vectors gradually evolves into an ensemble of analyses. A digital filter finalization technique (Fillion et al. 1995) is used to impose balance from 3 h onwards in the 9-h integrations that are required for the EnKF (Houtekamer and Mitchell 2005).

As mentioned in Candille et al. (2007), the EnKF uses a configuration with 4 subensembles of 24 members. Here, for the assimilation of observations into one subensemble, the Kalman gain is computed from the 72 guess fields of the other 3 subensembles. Formerly, a configuration with two subensembles of 48 members had been used as in Houtekamer and Mitchell (1998).

4 The representation of model error in the EnKF

In standard data-assimilation terminology, the term "model error" is used to refer to all errors of unknown origin. These errors can be in both the forecast model and the data-assimilation algorithm. It is generally not possible to obtain a complete statistical description of the model error (Dee 1995). Typically, the addition of a specific covariance matrix, defined by a few parameters for amplitudes and length-scales, is employed to account for all these errors (Dee 1995; Mitchell and Houtekamer 2000). In this paper, we use the term "system error" for the unexplained parameterized "model error". In addition, as discussed below, we have various ways to actively simulate specific uncertainties in the model.

4.1 The addition of parameterized system error

As shown in Figure 1, an ensemble of random perturbations is added to the ensemble of analyses, to reflect unknown error sources that degrade the quality of the ensemble mean analysis, but whose effects are not otherwise simulated in the Monte Carlo procedure. Formerly (Houtekamer et al. 2005), such an error term was added to the ensemble of guess fields (at 6 h), however, in order to permit time interpolation in the EnKF, we had to locate this term after the analysis (Houtekamer and Mitchell 2005, Eqs. 10 and 24).

The random perturbation fields are the sum of a (balanced) streamfunction field and an unbalanced temperature field. The random streamfunction field is isotropic, with a specified length-scale and with the same amplitude for all model levels. The standard deviation of the field is $(5.5m)g/f_0$ where the constants $g = 9.81ms^{-2}$ and $f_0 = 1.03\,10^{-4}s^{-1}$ are used to convert the 5.5 geopotential meters into streamfunction. The unbalanced temperature field is also isotropic and has a standard deviation of about $0.5K$ near the surface and near the top of the model and of about $0.13K$ in the mid-troposphere. For the length-scale, we used for both variables

a third-order autoregressive function (Mitchell et al. 1990). For streamfunction and unbalanced temperature, correlations of respectively 0.58 and 0.46 are obtained at 640 km. Attempts to add a humidity component to this simple formulation have not been successful in spite of a clear under-dispersion for humidity.

The above description of system error is considerably simpler than current background error descriptions in variational data-assimilation algorithms (Rabier et al. 1998). Likely we could obtain better results by using a more complex and appropriate formulation. It is, however, our long-term strategy to reduce the importance and amplitude of the system error component by using a better model and data-assimilation system and by actively simulating the impact of known sources of error (Mitchell and Houtekamer 2000).

4.2 The use of different parameterizations for different members

In the EnKF, we have 4 subensembles of 24 members. Each subensemble uses the same set of 24 configurations. Within a subensemble, different members can use different combinations of physical parameterizations. For deep convection, we have a choice of four different parameterizations: Kuo (Kuo 1974), symmetric Kuo (Wagneur 1991), Relaxed Arakawa Schubert (Moorthi and Suarez 1992) and Kain and Fritsch (1993). For the surface, we can use either a force-restore type algorithm (Deardorff 1978) or the more comprehensive Interaction Soil-Biosphere-Atmosphere (ISBA; Noilhan and Planton 1989) package. For the computation of the mixing length, we use either the Blackadar (1962) or the Bougeault and Lacarrère (1989) formulation. Finally, we use different values for the Prandtl number (Pr). The configuration of each member with the above parameterizations is shown in table 1.

Model error simulation in the EnKF (4 x 24)

#	Deep convection	Surface Scheme	Mixing Length	Prandtl number
1	Kain & Fritsch	ISBA	Bougeault	1.0
2	Kuo	ISBA	Blackadar	0.85
3	Relaxed Arakawa Schubert	force-restore	Bougeault	0.85
4	Symmetric Kuo	force-restore	Blackadar	1.0
5	Kuo	force-restore	Bougeault	1.0
6	Kain & Fritsch	force-restore	Blackadar	0.85
7	Symmetric Kuo	ISBA	Bougeault	0.85
8	Relaxed Arakawa Schubert	ISBA	Blackadar	1.0
9	Kain & Fritsch	ISBA	Blackadar	0.85
10	Kuo	ISBA	Bougeault	1.0
11	Relaxed Arakawa Schubert	force-restore	Blackadar	1.0
12	Symmetric Kuo	force-restore	Bougeault	0.85
13	Kuo	force-restore	Blackadar	0.85
14	Kain & Fritsch	force-restore	Bougeault	1.0
15	Symmetric Kuo	ISBA	Blackadar	1.0
16	Relaxed Arakawa Schubert	ISBA	Bougeault	0.85
17	Symmetric Kuo	force-restore	Bougeault	1.0
18	Kain & Fritsch	ISBA	Blackadar	0.85
19	Kuo	ISBA	Bougeault	0.85
20	Relaxed Arakawa Schubert	force-restore	Blackadar	1.0
21	Relaxed Arakawa Schubert	ISBA	Blackadar	0.85
22	Kuo	force-restore	Bougeault	1.0
23	Kain & Fritsch	force-restore	Blackadar	1.0
24	Symmetric Kuo	ISBA	Bougeault	0.85

Table 1: The configurations that are selected for each member in a 24-member subensemble.

Note that the configurations of members 5 and 7 are reused for members 22 and 24 respectively. This undesirable property is a consequence of our not specifying in the program that generates the model configurations that different members must use different configurations. The choice of different convective parameterizations leads to a desired significant increase of the ensemble spread where the model is uncertain and contributes to an improved quality of the ensemble mean. The use of different land-surface algorithms has an impact near

the surface for in particular the amplitude of the diurnal cycle. The other changes to the model have relatively smaller impacts. A variety of other parameters in the namelist of the model has been modified in research mode, but each time the impact was insufficient to justify using different values of the parameter in the EPS.

5 The link between the EnKF and the medium-range forecasts

In principle, one would like to have a seamless link between the data-assimilation and forecasting components of the EPS. Historically, we used various ad-hoc inflation procedures to increase the initial spread of the medium-range ensemble. In the absence of sufficiently rapid error growth, this was deemed necessary to arrive at a realistic spread in the medium range. These procedures, however, led to various adjustment problems and have now been abandoned.

Since it is not a priori clear that all four subensembles will always have the same statistics, it was decided to select an equal number of initial conditions from each subensemble. For the collaboration in the North American Ensemble Forecast System (NAEFS), we have a requirement for 20 medium-range ensemble forecasts. We therefore need to select 5 members from each subensemble (figure 2). To minimize adjustment problems, we maintain the same model configuration that was used in the EnKF for the medium-range forecast. As we shall see in section 6, however, there are some extensions.

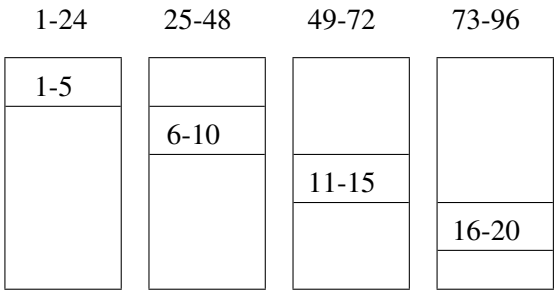

Figure 2: The correspondence between members of the 96-member EnKF ensemble and the 20-member medium-range forecast ensemble.

Since the 96-member mean analysis is our best estimate of the initial state, we translate the initial conditions of the 20-member ensemble so that their mean becomes identical to this best estimate. Any resulting negative or oversaturated humidity values are clipped to zero and 100 % relative humidity respectively.

6 Accounting for model error in the medium-range forecasts

As in the EnKF, parameterized system error is added at the beginning of each model integration. However (see below), for the medium-range forecasts, the amplitude of the parameterized error is almost twice as large.

The set of basic changes to the model parameterizations that is used in the EnKF has been extended with a change in the intensity of the gravity wave drag (that was already described in Houtekamer and Lefaivre (1997)).

Since we decided not to add parameterized model error every six hours during the medium-range integrations, we found it necessary to add stochastic algorithms to simulate model error. Currently we use both the stochastic perturbation of physical tendencies and SKEB as described below.

6.1 The addition of parameterized system error

For the parameterized system error, we use the same correlations in the applications for the EnKF and for the medium-range forecasts. We have, however, increased the amplitudes by a factor of 1.8. The standard deviation of the added streamfunction field is $(10.0m)g/f_0$ and the added unbalanced temperature field has a standard deviation of about $0.9K$ near the surface and near the top of the model and of about $0.23K$ in the mid-troposphere. We need these larger amplitudes because we found it difficult to obtain rapid initial error growth in the EPS. Essentially the amplitude of the system error can be adjusted to obtain a realistic amount of spread in the early medium-range. Unlike previously used inflation procedures, the addition of parameterized system error does not lead to significant balance problems during the first few days of the forecast. The isotropic system error does, however, effectively blur the interesting flow-dependent statistics that are developed by the EnKF and we would therefore like to keep it as small as possible.

6.2 The use of different parameterizations for different members

As mentioned in section 5, we would like to have a seamless link between the data-assimilation and forecasting components of the EPS. It is hard to justify using different physical parameterizations in the two systems. Mostly for historical reasons, the medium-range EPS uses different intensities of the upper-level gravity wave drag.

In contrast to the EnKF, the medium-range EPS does have an unperturbed control member. This is to permit a straightforward comparison of the quality of the EPS with the high-resolution deterministic forecast. In the short-range, we need to ascertain that the quality of the EPS is not too far behind.

Table 2 shows the current configuration of the EPS with 20 perturbed members and one unperturbed control member. Since we need only 20 members, all members use different configurations.

Model configurations for the medium range

#	Deep convection, surface scheme, mixing length, Pr	Intensity of gravity wave drag	Backscatter used	Stochastic physics used
0	As member 1	standard	no	no
1	As in the EnKF for members 1-20. Member 1 has a configuration that is almost identical to the configuration of the high-resolution deterministic model of our centre.	weak	yes	yes
2		strong	yes	yes
3		weak	yes	yes
4		strong	yes	yes
5		weak	yes	yes
6		strong	yes	yes
7		weak	yes	yes
8		strong	yes	yes
9		weak	yes	yes
10		strong	yes	yes
11		weak	yes	yes
12		strong	yes	yes
13		weak	yes	yes
14		strong	yes	yes
15		weak	yes	yes
16		strong	yes	yes
17		weak	yes	yes
18		strong	yes	yes
19		weak	yes	yes
20		strong	yes	yes

Table 2: The configuration of each of the members in the medium-range forecast ensemble.

For medium-range forecasts, one may note that the use of different parameterizations will cause the different members to have different bias properties, which has the disadvantage of requiring additional post-processing steps for the users.

6.3 The stochastic perturbation of physical tendencies

Buizza et al. (1999) proposed accounting for uncertainty in the parameterized model physics in a comprehensive manner by multiplying the tendency vector that is output from the model subgrid-scale parameterizations by a random field. The random field changes in space, time and between members, which should minimize undesired model and ensemble drifts.

Presently, we use Markov chains in spectral space to obtain random numbers which are, as in Buizza et al. (1999), between 0.5 and 1.5. The random fields are on total wavenumbers 1 to 10 and have mean 1.0 and standard deviation 0.23. The decorrelation time scale for the Markov chains is 3 h.

Currently, the stochastic perturbation of physical tendencies forms an essential component of the medium-range EPS which adds significantly to the error growth. On occasion, however, when the physics tendencies have locally been multiplied by values near the maximum of 1.5 over several time steps, we start noticing severe $2\Delta t$ oscillations near the surface. This problem is currently under investigation.

6.4 Stochastic Kinetic Energy Backscatter

In comparison to the real atmosphere, numerical weather prediction models overdissipate kinetic energy near their truncation scale. In the GEM model, a significant fraction of the numerical diffusion comes from the off-centered time stepping and the semi-Lagrangian advection schemes. A non-negligible part of the turbulent inverse cascade process is likely inhibited because of this overdissipation (Shutts 2005). Following Shutts (2005), we implemented a Stochastic Kinetic Energy Backscatter (SKEB) scheme to compensate for the overdissipation.

In our implementation, rotational perturbation modes are introduced between wave numbers 40 and 128. In addition, we have a stochastic temperature forcing that does not assume a specific balance. The decorrelation time scale for the Markov chains is 36 h.

It may be noted that, because the SKEB parameterization can be justified with reference to specific weaknesses of the model physics and dynamics, significant changes to the forecast model may require a corresponding readjustment of the parameters for the SKEB algorithm.

7 Results

A development goal for the Canadian EPS has been to arrive at reliable medium-range ensemble forecasts. Recently, various ways of measuring reliability, have suggested nearly perfect reliability for the mid-tropospheric extra-tropical ensemble forecasts.

As an example, we show the rank histogram (Anderson 1996) for October 2006 for the Northern Extratropics for 850 hPa Geopotential Height (Fig. 3). This diagram no longer shows the typical U-shape of an under-dispersive ensemble that we obtained with previous configurations of the EPS. Since we did not account for the impact of the uncertainty of the verifying analyses on the rank histogram, one could argue that the ensemble is somewhat overdispersive in the short-range where the analysis error is relatively large (Anderson 1996). Scores of the same type have been obtained for various variables with the reliability component of the Continuous Ranked Probability Score (CRPS; Hersbach 2000). Here, again, due to the impact of observational errors of often unknown magnitude, it is often not clear if the ensemble is under- or over-dispersive. Given the verification

tools that we have and having control (i) over the magnitude of the initial system error, (ii) over the range of the random numbers in the stochastic perturbation of physical tendencies and (iii) over many parameters in the SKEB algorithm, it is possible to adjust our EPS to obtain a nearly perfect reliability for many variables. Evident non-trivial problems remain with the amount of dispersion and with bias near the surface, near the top of the model and in the tropics.

To give an impression of the overall quality of the EPS, we give a verification of 500 hPa geopotential height against the global radiosonde network in Fig. 4. It can be seen that the American and Canadian ensembles have a similar overall quality. A significant gain, of the order of 1-day, is obtained from combining the two ensembles as is done operationally. Somewhat different results can be obtained for different variables or different periods because the two systems have different strengths and weaknesses and because improvements are regularly made in both systems.

In general, for extra-tropical areas, verifications suggest that further improvements to the EPS must come from the resolution component of the scores. In practice, we can either try to better assimilate more data with the EnKF or we can move to improved - often higher-resolution - versions of the GEM forecast model.

8 Future work

A variety of methods is used in the different parts of the EPS to account for model error. The current EPS uses two different methods to account for model error in the EnKF and four different methods to account for it in the medium-range forecast. It would seem that sometimes different methods each in their own manner sample the same uncertainty. In the future, we will attempt to arrive at a coherent treatment of model error in the data-assimilation and forecasting components.

In the global EPS, we have long-standing problems with the error dynamics near the top of the model (e.g. Houtekamer et al. 2005, Figs. 2,9). In the near future, we hope to migrate to a version of the GEM model with improved stratospheric dynamics and at the same time we want to raise the top of the model beyond 10 hPa.

In the future, we would like to run a regional ensemble over the North American area. Possibly, we would use the global ensemble to pilot the regional ensemble.

Finally, we are undertaking various intercomparison projects at our centre to exploit our unique position with both in-house ensemble-based and adjoint-based algorithms.

Acknowledgements

We thank our many colleagues at Environment Canada for their help, suggestions and encouragement during the entire development period of the global ensemble prediction system.

9 References

Anderson, J.L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. Journal of Climate, 9, 1518–1530.

Anderson, J.L., 1997: The impact of dynamical constraints on the selection of initial conditions for ensemble predictions: Low-order perfect model results. Mon. Wea. Rev., 125, 2969–2983.

Anderson, J.L., 2001: An ensemble adjustment Kalman filter for data assimilation. Mon. Wea. Rev., 129, 2884–2903.

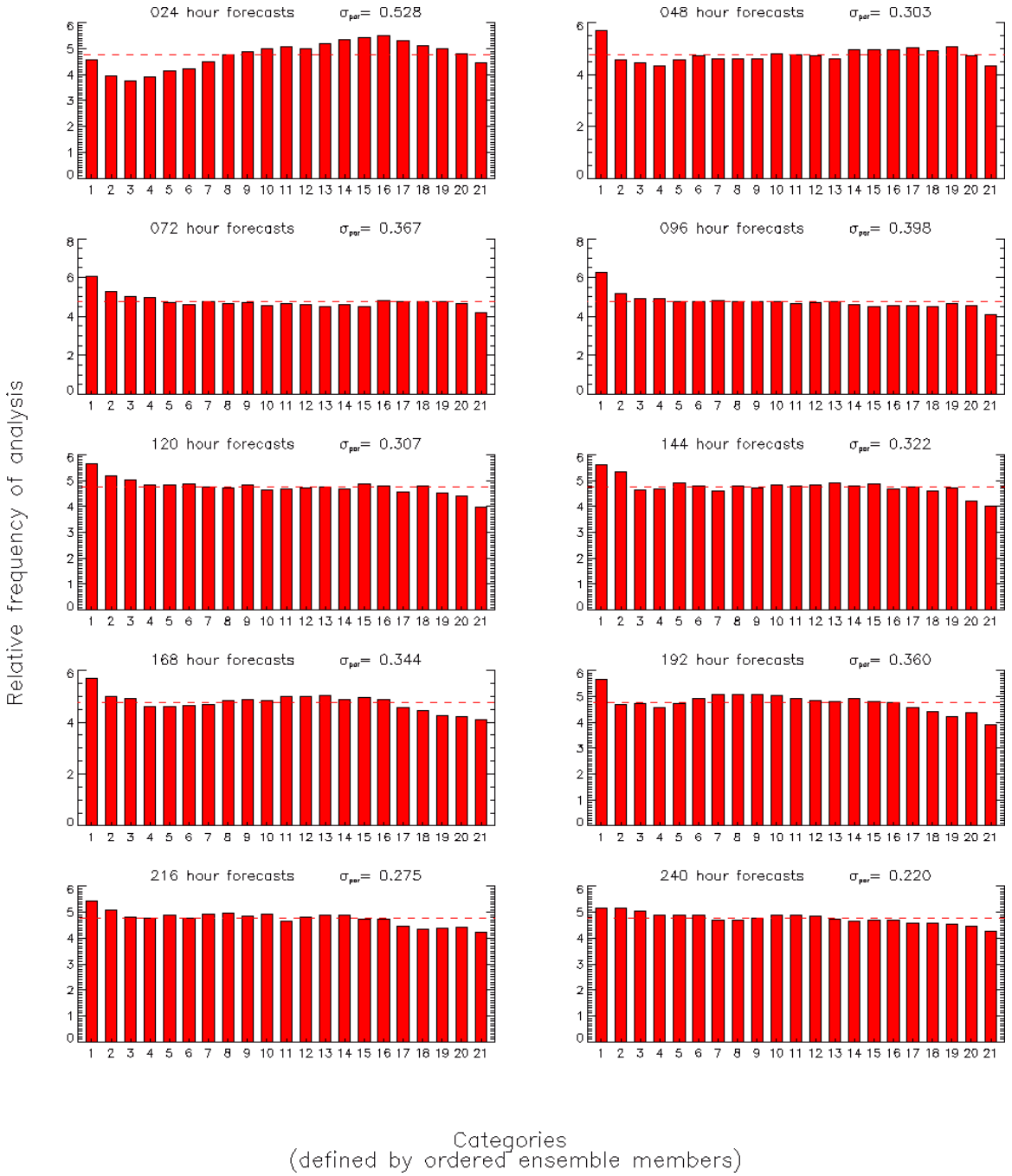

Figure 3: Rank Histogram for October 2006, Northern Extratropics, 850 hPa Geopotential Height, pre-implementation experiments. The printed standard deviation measures the distance to the optimal dashed horizontal line.

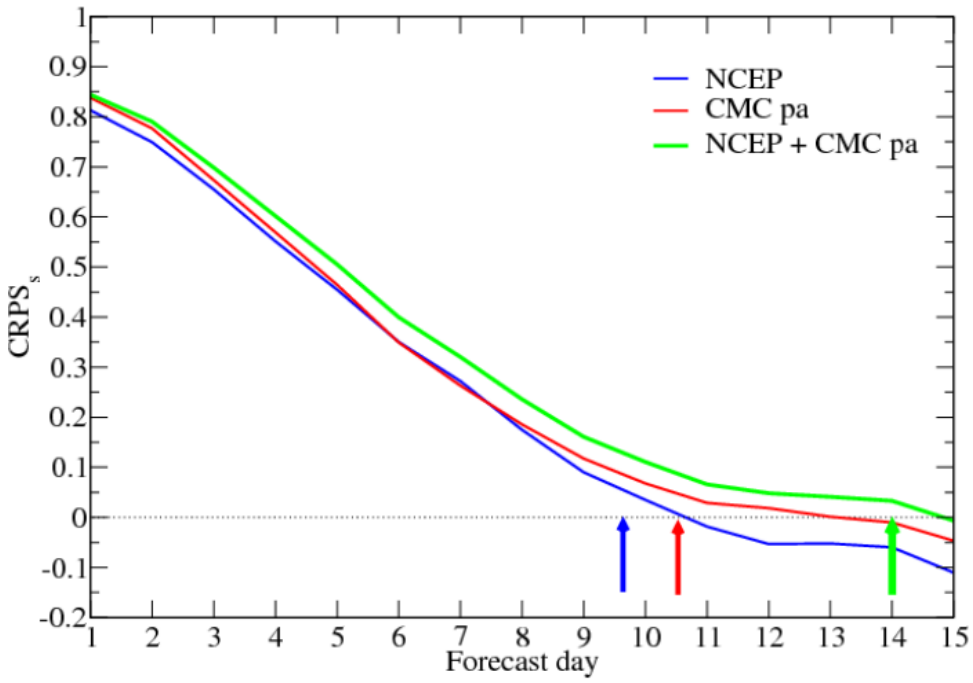

Figure 4: The CRPS Skill score for 500 hPa Geopotential Height for October 2006. The verification is against radiosondes. For the Canadian ensemble (red), pre-implementation experiments were used. The American operational ensemble, of that time, is in blue and the score for the combined NAEFS ensemble is in green. The arrows are at score 0.05.

Anderson, B.D.O., and J.B. Moore, 1979: Optimal Filtering, Prentice-Hall, 357 pp.

Blackadar, A. K. 1962: The vertical distribution of wind and turbulent exchange in a neutral atmosphere, Journal of Geophysical Research, Vol. 67, No. 8, 3095–3102.

Bougeault, P., and P. Lacarrère, 1989: Parameterization of orography-induced turbulence in a meso-beta-scale model. Mon. Wea. Rev., 117, 1872–1890.

Buizza, R., Miller, M., and T.N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. Quart. J. Roy. Meteor. Soc., 125, 2887-2908.

Buizza, R., Houtekamer, P.L., Toth, Z., Pellerin, G., Wei, M., and Y. Zhu, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. Mon. Wea. Rev., 133, 1076–1097.

Candille, G., Côté, C., Houtekamer, P.L., and G. Pellerin, 2007: Verification of an ensemble prediction system against observations. Mon. Wea. Rev., 135, 2688–2699.

Côté, J., Gravel, S., Méthot, A., Patoine,A., Roch, M., and A. Staniforth, 1998: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. Mon. Wea. Rev., 126, 1373–1395.

Deardorff, J. W., 1978: Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation, J. Geophys. Res., 83, 1889–1903.

Dee, D.P., 1995, On-line estimation of error covariance parameters for atmospheric data assimilation. Mon. Wea. Rev., 123, 1128–1145.

Descamps, L., and O. Talagrand, 2007: On some aspects of the definition of initial conditions for ensemble prediction. Mon. Wea. Rev., 135, 3260–3272.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo

methods to forecast error statistics. J. Geophys. Res., 99 (C5), 10143–10162.

Fillion, L., Mitchell, H.L., Ritchie, H., and A. Staniforth, 1995: The impact of a digital filter finalization technique in a global data assimilation system. Tellus, 47A, 304–323.

Gaspari, G., and S.E. Cohn, 1999: Construction of correlation functions in two and three dimensions. Quart. J. Roy. Meteor. Soc., 125, 723–757.

Gauthier, P., Tanguay, M., Laroche, S., Pellerin, S., and J. Morneau, 2007: Extension of 3DVAR to 4DVAR: implementation of 4DVAR at the Meteorological Service of Canada. Mon. Wea. Rev., 135, 2339–2354.

Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. Wea. Forecasting, 15, 559–570.

Houtekamer, P.L., and L. Lefaivre, 1997: Using ensemble forecasts for model validation. Mon. Wea. Rev., 125, 2416–2426.

Houtekamer, P.L., and H.L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. Mon. Wea. Rev., 126, 796–811.

Houtekamer, P.L., and H.L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Wea. Rev., 129, 123–137.

Houtekamer, P.L., and H.L. Mitchell, 2005: Ensemble Kalman filtering. Quart. J. Roy. Meteor. Soc., 131, 3269–3289.

Houtekamer, P.L., Mitchell, H.L., Pellerin, G., Buehner, M., Charron, M., Spacek, L., and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: results with real observations. Mon. Wea. Rev., 133, 604–620.

Kain, J.S., and J.M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain- Fritsch scheme. in *The representation of cumulus convection in numerical models*. Meteor. Monogr., No. 46, Amer. Meteor. Soc., 165–170.

Kalman, R.E., 1960: A new approach to linear filtering and prediction problems. Trans. ASME, Ser. D, J. Basic Eng., 82, 35–45.

Kuo, H.L., 1974: Further studies of the parameterization of the influence of cumulus convection on large-scale flow. J. Atmos. Sci., 31, 1232–1240.

Mitchell, H.L, Charette, C., Chouinard, C., and B. Brasnett, 1990: Revised interpolation statistics for the Canadian data assimilation procedure: their derivation and application. Mon. Wea. Rev., 118, 1591–1614.

Mitchell, H.L, and P.L. Houtekamer, 2000: An adaptive ensemble Kalman filter. Mon. Wea. Rev., 128, 416–433.

Moorthi, S., and M.J. Suarez, 1992: Relaxed Arakawa-Schubert. A parameterization of moist convection for general circulation models. Mon. Wea. Rev., 120, 978–1002.

Noilhan, J., and S. Planton, 1989: A simple parameterization of land surface processes for meteorological models. Mon. Wea. Rev., 117, 536-549.

Rabier, F., McNally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A., and F. Bouttier, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. Quart. J. Roy. Meteor. Soc., 124, 1809–1829.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. Quart. J. Roy. Meteor. Soc., 131, 3079–3102.

St-James, J.S., and S. Laroche, 2005: Assimilation of wind profiler data in the Canadian Meteorological Cen-

tre's analysis systems. J. Atmos. Oceanic Technol., 22, 1181–1194.

Tarantola, A., 1987: Inverse Problem Theory, Elsevier, 613 pp.

Wagneur, N., 1991: Une évaluation des schémas de type Kuo pour le paramétrage de la convection, Msc Thesis, UQAM, 76 pp.

Whitaker, J.S., and T.M. Hamill, 2002: Ensemble data assimilation without perturbed observations. Mon. Wea. Rev., 130, 1913–1924.