CMC medium to long-range dynamical forecasts: Method and results

A. Plante[1], P. L. Houtekamer[2], N. Gagnon[1], L. Lefaivre[1], G. Pellerin[1] and R. Verret[1]
[1]Canadian Meteorological Centre
[2]Meteorological Research Branch
Dorval, Canada

Summary: Ensemble forecasts are run operationally at the Canadian Meteorological Centre, with outputs up to 90-days. The ensemble set-up is different for medium-range (1-10 days) than for long-range (30-90 days).

The method of producing the perturbed analyses for the 10-day forecast consists of running independent assimilation cycles that use perturbed sets of observations and are driven by eight different versions of the model obtained from different physical parameterization schemes. Perturbed analyses are doubled by taking opposite pairs. A multi-model approach is then used to obtain the forecasts. The ensemble outputs have been used to generate several products.

For the long-range forecast, analyses lagged by one day are used. A 5 member lagged average is used to produce the 30-day forecasts. The final product consists of a categorical temperature anomaly forecast. Seasonal forecasts are also produced with a lagged average approach, however with a multi-model method, using a forecast model and a climate model. Categorical temperature and precipitation forecasts are produced.

1) INTRODUCTION

The Canadian Meteorological Centre (CMC) operational medium-range ensemble prediction is unique in the way it accounts for the model and the initial state error. The main idea is to perturb the physical parameterizations of the dynamic model that is used in the assimilation cycle and in the production of the 10-day forecasts. Section 2 presents this system which has been running at CMC for almost 4 years. In August 1999, a second dynamical model was added to the system in order to increase the ensemble size to 16.

Dynamical models are also used operationally at CMC to make the 90-day temperature and precipitation forecasts. A multi-model approach, including a climate model and a numerical weather prediction (NWP) model, is used. The climate of both models were computed from a 26 year hindcast. This hindcast serves to remove the models systematic error and provide a verification database. Forecast skill at such range can be modest compared to short term weather forecasts, so it is very important to convey this to users. For this purpose, CMC Web site provides verification maps from the hincast along with the forecast maps. The 90-day forecasts, issued every 3 months since fall of 1996, will be described in Section 3 along with the 30-day forecasts.

CMC forecasts are available at the following Web site: http://www.cmc.ec.gc.ca/cmc/htmls/forecasts.html

2) THE CMC 10-DAY ENSEMBLE PREDICTION SYSTEM

2.1) Description of the method

The ensemble prediction system used at CMC is described in Houtekamer et al (1996) and is summarised in Figure 1. The basis of the method is to produce perturbed analyses through data assimilation procedures.

Eight parallel analysis cycles are run quasi-independently of the high resolution operational analysis cycle (dark grey boxes on fig. 1). Each of the 8 models used in the assimilation cycles, have different physical schemes.

In addition, random perturbations are added to some physical parameters (horizontal diffusion, minimal roughness length over sea and time filter). Perturbations are also introduced in the surface forcing via the sea surface temperature albedo and roughness length. Each observation used to feed the analysis is perturbed (aireps, buoys, radiosondes, satems, satobs, synops, ships, acars/amdars). Each eigenvector of the covariance matrix for the observational error is multiplied with a random value. The resulting perturbation vector is then added to the observations. The random values are different for each piece of information and are different from one perturbed cycle to the other. The analysis scheme used is the Optimal Interpolation (O/I) technique (Rutherford 1972), which has the advantage of being efficient computerwise (the weights for the innovations are calculated once for the control, and used for the *n* ensemble members as well). These perturbations are represented by the white boxes on fig. 1.

The details of the perturbation done to the ensemble prediction system can be found in Houtekamer and Lefaivre (1997). The dynamic core of the 8 models used in the assimilation cycle is based on the spectral model (SEF) described in Ritchie (1991). It has a horizontal resolution of T_L95 and a horizontal diffusion in Δ^4.

To obtain 8 additional analyses to produce medium-range forecasts, pairs of opposing initial conditions are used: the mean of the analyses is subtracted to the high resolution operational analysis and a fraction of this difference is added to the original perturbed analyses.

Once a day, at 00 UTC, 10-day forecasts are produced using:

- 8 perturbed analyses, half of them obtained by taking the opposite addition to the high resolution operational analysis, using the same models as the ones used to produce the trial fields;
- the control run, obtained from an analysis with unperturbed observations and with intermediate model options;
- 8 perturbed analyses, half of them obtained by taking the opposite addition to the high resolution operational analysis, using this time the Global Environmental Multi-scale (GEM) model (Côté *et al* 1998) with its own set of perturbations.

The forecast part of the system is represented by the light grey boxes on fig. 1.

2.2) Experimental precipitation forecasts

The ensemble approach is a natural tool to forecasts the probability of precipitation (POP). Initially, for the eight member ensemble, the classes were defined as follows, for thresholds of 2, 5 10 and 25 mm:

below 19%, if one member or less forecasts a precipitation amount above the threshold;
between 20% and 44%, if 2 to 3 members forecast a precipitation amount above the threshold;
between 45% and 69%, if 4 to 5 members forecast a precipitation amount above the threshold;
above 70%, if more than 6 members forecast a precipitation amount above the threshold.

However, reliability diagrams show that our ensemble method demonstrates a lot of sharpness, but not enough reliability. Because of that, the POP are now calibrated to correct for this bias. An example of the calibrated product can be seen at Figure 2. Validation of these forecasts are performed systematically for a selection of 20 stations over Canada at the end of each season, and the calibration is thus applied to the same season the following year, with a different calibration for thresholds (2, 5, and 10 mm) and lead times. The calibration for the 10 mm threshold is also used for the 25 mm threshold.

The relative operating characteristic (ROC) curves have been proposed by Mason (1982) as a verification measure for probabilistic forecasts. In a ROC curve, the hit rate is shown as a function of the false alarm rate. Figure 3 shows the results for the 5 mm threshold for 4 lead times (day 1, day 4, day 7 and day 10) and for 4 seasons:

(a): December 1998 - January 1999 - February 1999;

(b): March - April - May 1999;
(c): June- July - August 1999;
(d): September - October - November 1999.

The surface under the curve gives a good indication of the performance of the system. The best performance, based on a one year verification, is obtained by the winter (Fig.3 a) forecasts, which shows a relatively large area under the curve at least up to day 7. The performance of the 2 and 10 mm thresholds forecasts is similar to the one on figure 3. The 25 mm threshold is not verified because of the sample being too small.

3) THE CMC 30-DAY AND 90-DAY ENSEMBLE PREDICTION SYSTEM

3.1) Description of the method

A multi-model approach is used to produce the 90-day forecasts. This includes the Canadian second generation General Circulation Model (GCM) (McFarlane et al., 1992) climate model, and the SEF forecast model. The biases are removed from both models by subtracting their climatology. A hindcast project, call the Historical Forecasting Project (HFP), was conducted in order to compute the models climatology and to assess the skill of CMC 90-day forecasts. The HFP set-up almost is the same as the CMC operational set-up and is described below.

The models use as input sea surface temperature (SST), sea ice extent (ICE) and snow. The SST and ICE data are taken from the GISST2.2 data set (Parker et al., 1995). The SST anomaly of the month prior to the start of the integrations is persisted through the 3 month forecasts. The ICE is initialised with a 30 year climatology. The snow line is specified from weekly satellite observations (from NCEP). The SEF model uses a persisted 10 day snow anomaly for the first month and climatology afterwards. The GCM is initialised with the observed snow line and then uses a prognostic scheme. In operational, mode all these analysed surface fields come from the CMC operational analysis.

Each model is integrated throughout the season in an ensemble of 6 members. These 6 runs differ in their starting time that is lagged by 6 hours (24 in the operational mode). This leads to 12 90-day forecasts per season. The CMC issues forecast for 2 variables: the surface air temperature anomalies and the seasonal accumulated precipitation anomalies. The surface air temperature anomaly forecast is done using the 500-1000hPa thickness (DZ) anomaly. The DZ variable of the model runs are output every 12 hours and averaged over the season. The 2 ensembles of 6 forecasts are averaged separately for both models. Then a hybridization of the 2 DZ forecasts is done using the BLUE method (Derome et al., 2000). This method gives better or equivalent results than a normalised average of the 2 model outputs for every season. It is currently used at CMC operations (since spring 1999). The hybridised DZ field is then related to the surface temperature anomalies T by the following "perfect prog" technique :

$$T = b\,DZ. \tag{1}$$

The coefficient b in (1) was derived at Canadian stations from analysed DZ (NCEP reanalysis, Kalnay et al., 1996) and observed T (Vincent and Gullett, 1999) for years 1969 to 1994. There is a different b for every selected station and season. The values of b range from about 0.3 to 0.5 [°C/dam]. The temperature forecast is then compared to the model climatology in order to produce a 3 category forecast (below normal, normal and above normal temperature). The threshold to be different from normal is ±0.43 times the model inter-annual standard deviation. By design all categories have the same probability (1/3) to occur, so that a random forecast would be correct one third of the time in average. Using a contingency table, the Percent Correct was calculated to verify the categorical forecast.

The percent correct (PC) of the HFP surface air temperature anomaly forecasts on 50 km grid is presented for summer (JJA) and winter (DJF) on figure 4. The values at 210 Canadian stations are shown. In theory, the PC has to be higher than 33% to be better than chance. But since there are only 26 years in the verifications it is easy to get score higher than 33% just by chance. With 26 trials a score has to be greater or equal to about 46% to be considered statistically better than the chance (according to the binomial distribution with a 10% confidence level). The areas where the PC is higher than 46% are shaded.

It can be seen from figure 4 that there is good skill in summer (JJA) over the centre of Canada. The winter (DJF) skill cover most of the western and central parts of the country. In spring, the system has good scores in British Columbia, Yukon and Nunavut (not shown). In fall, the skill is mainly found in Quebec and western Ontario (not shown).

The precipitation forecast made using a normalised average of the 2 model outputs. The categorical forecast fields were compared to the observations at more than 340 Canadian stations described by Mekis and Hogg (1999). The performance of the models in forecasting precipitation is much lower than for temperature. The best results are found in winter and spring, but the area covered by significant skill is very close to 10% over Canada, which is what one expect from pure chance. Therefore, there is little skill in the current dynamical precipitation forecasts over Canada.

3.2) 30-day forecasts

The SEF climatology obtained from the HFP was used to remove the systematic error from the March, June, September and December forecasts. Also, the HFP interannual standard deviation for these months were used to get a better balance between the 3 classes of the categorical forecast (below, normal and above). Figure 5 presents the score of the old method (without HFP data), for which the model bias were present, and the new method using the HFP data to remove the model bias and class the forecast (with HFP data). The score of the persistence (previous month) is given for comparison. The score is the average percent correct over Canada for the HFP period (1969-1994). One can see that the HFP data improved the score for all verified months (March, June and September) except for December. The June forecast, which was comparable to the persistence, is much better when produce with the HFP model climate.

4) CONCLUSIONS AND FUTURE WORK

The first priority is to make use of the 10-day forecast 16 member ensemble to add products in terms of probabilistic forecasts. Probabilities thus obtained will be used in the production of worded forecasts up to day 10. New verification methods will also be used to verify the Probability Distribution Function, as proposed by Wilson *et al.* 1999. The generation of perturbed analyses will be produced using the GEM model as the driving model and an adaptive Ensemble Kalman filter as the analysis scheme (Mitchell and Houtekamer, 2000). Exchanges with NCEP ensemble members will be set in order to test "grand ensemble" products.

CMC 30-day and 90-day forecasts were verified over the year 1969-1994 for MAM, JJA, SON and DJF. The best performance of 90-day temperature forecasts over Canada were found in winter (DJF) and summer (JJA). The 90-day precipitation forecasts show little skill for all verified periods. CMC 30-day forecasts was verified for the months of March, June, September and December 1969-1994 over Canada and showed to be much better than persistence. It is planned to run the 30-day forecast with the same set-up as the seasonal one (2 models, 12 members total) and to produce a categorical precipitation forecast.

The HFP is currently extended to all 3 month periods (12) in order to produce the 90-day outlooks every month.

New generation models (GEM and GCMIII) will be tested in a new seasonal Historical Forecasting Project. These models have more comprehensive physical parameterizations and higher resolution than the current ones (GCMII and SEF).

5) REFERENCES

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch 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.

Derome J., G. Brunet, A. Plante, N. Gagnon, G. J. Boer, F.W. Zwiers, S. Lambert, and H. Ritchie, 2000: Seasonal Predictions Based on Two Dynamical Models. In preparation.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, Roy Jenne, and Dennis Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project". *Bull. Amer. Meteor. Soc.*, 77, 437-471.

Houtekamer, P. L, L. Lefaivre, J. Derome, H. Ritchie and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225-1242.

Houtekamer, P. L and L. Lefaivre, 1997. Using Ensemble Forecasts for Model Validation. *Mon. Wea. Rev.*, **125**, 2416-2426

Mason, I., 1982: A model for the assessment of weather forecasts. *Aust. Meteor. Mag.*, **30**, 291-303

McFarlane, N.A., G.J. Boer, J.-P. Blanchet and M. Lazare, 1992: The Canadian Climate Centre second-generation general circulation model and its equilibrium climate. *J. Climate*, **5**, 1013-1044.

Mekis E. and W. D. Hogg, 1999: Rehabilitation and Analysis of Canadian Daily Precipitation. *Atmo.-Ocean*, **37**, 53-85.

Mitchell, H. L. and P. L. Houtekamer, 2000. An Adaptive Ensemble Kalman Filter, *Mon. Wea. Rev.*, **128**, 416-433.

Parker D.E., M. Jackson and E.B. Horton. 1995. The 1961-1990 GISST2.2 sea surface temperature and sea-ice climatology. Technical Note no. 63, Hadley Centre, UKMO, Bracknell, UK.

Ritchie, H. 1991: Application in the semi-Lagrangian method to a multi-level spectral primitive-equations model. *Quart. J. Roy. Meteor. Soc.*, **117**, 91-106

Rutherford, I. D., 1972: Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.*, **29**, 809-815

Vincent, L.A., and D. W. Gullett, 1999: Canadian historical and homogeneous temperature datasets for climate change analyses. *Int. J. Climatol*, **19**, 1375-1388.

Wilson, L. J., W. R. Burroughs and A. Lanzinger, 1999: A Strategy for Verification of Weather Element Forecasts from an Ensemble Prediction System, *Mon. Wea. Rev.*, **127**, 956-970

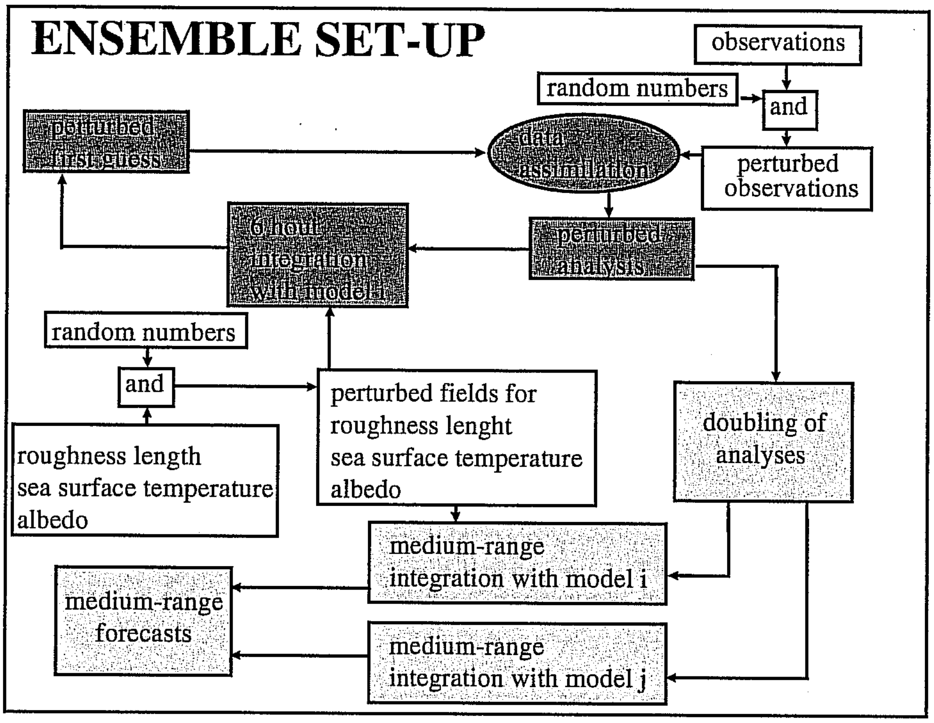

Figure 1: Organisation chart of the 10-day ensemble forecasts set-up.

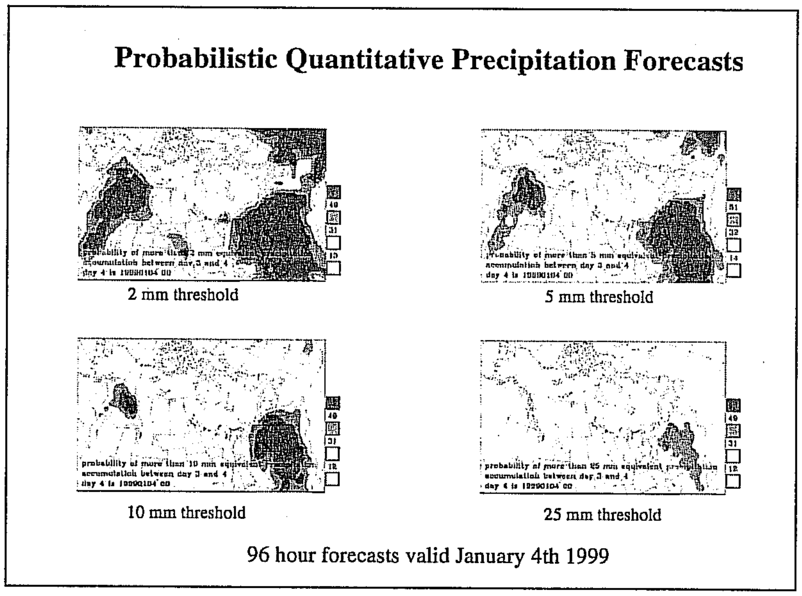

Figure 2: Example of a calibrated probabilistic QPF using the 8 member ensemble set-up. The maximum QPF probability for the 96 hour lead time is 51%. The top left chart shows the probability of getting 2 mm or more of precipitation (equivalent water) in 24 hours ending January 4th 1999 at 00 UTC. The other charts shows the 5, 10 and 25 mm thresholds. This product and other ensemble products are available on the following Web site: http://www.cmc.ec.gc.ca/~cmsw/ensemble

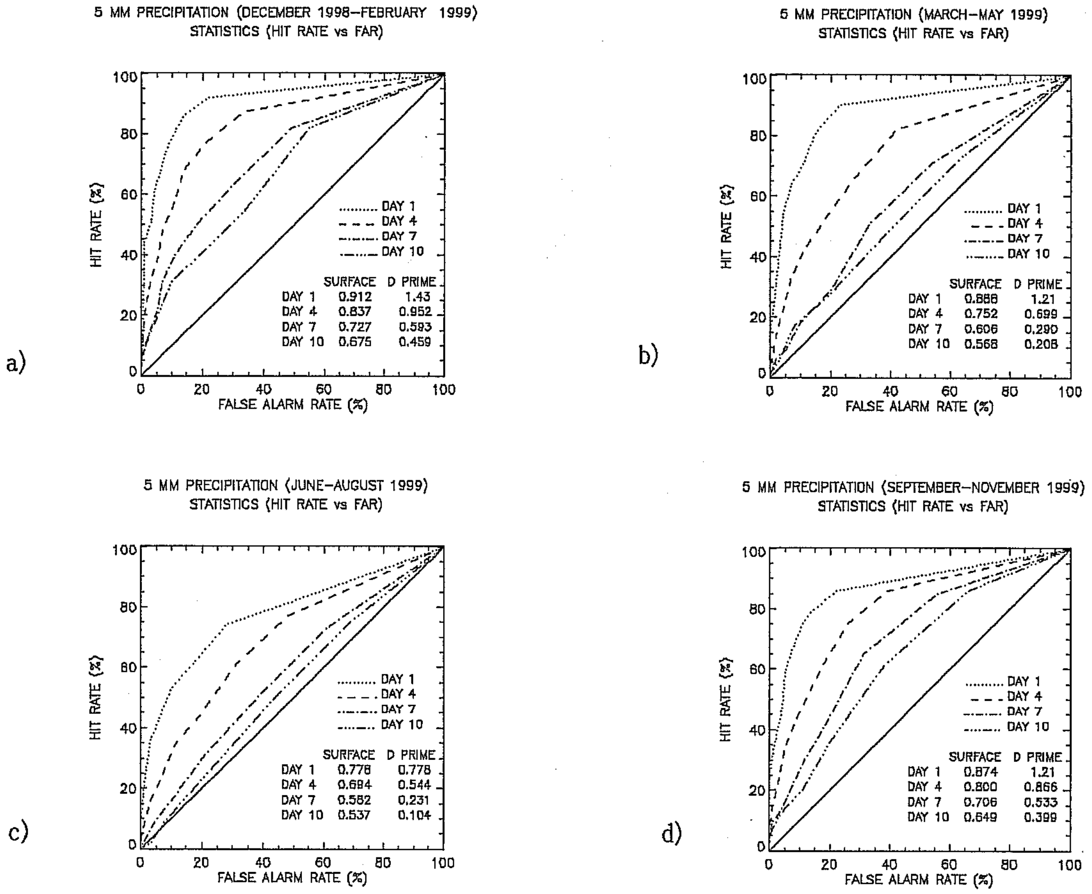

Figure 3: Relative Operating Characteristics verifications of the 5 mm threshold calibrated probabilistic QPF using the 8 member ensemble set-up. The verification are for selected stations in Canada and for a): DJF 1989; b): MAM 1999; c): JJA 1999 and d) SON 1999.

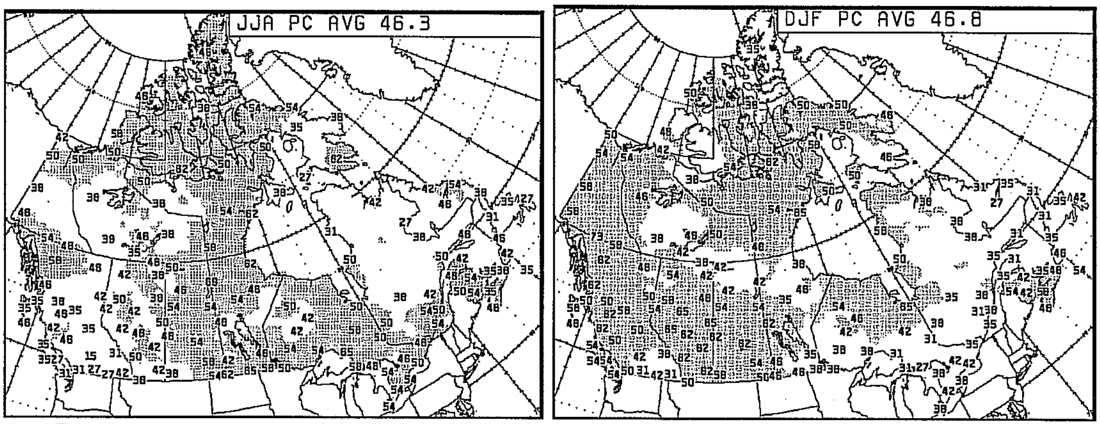

Figure 4: Percentage correct of the categorical temperature seasonal forecasts (Summer (JJA) and winter (DJF)) over the Historical Forecasting project years (1969-1994). With 26 trials a score has to be greater or equal to about 46% to be considered statistically better than the chance (according to the binomial distribution with a 10% confidence level). The areas where the PC is higher than 46% are shaded. CMC seasonal forecasts are available at the following Web site: http://www.cmc.ec.gc.ca/~cmcdev/saisons/seasons.html

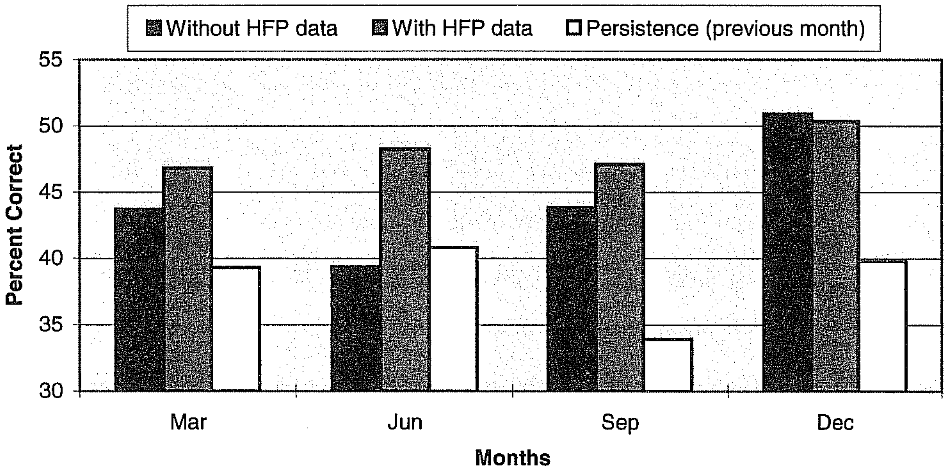

Figure 5: Percent correct (PC) for the 30-day categorical temperature forecasts over the Historical Forecasting Project years (1969-1994) averaged over Canada. The black bars shows the PC obtained by the forecasts produced from the model outputs and the observed climatology (model bias present). The gray bars shows the PC of the forecasts produced from the model outputs and the model climatology (model bias removed). Finally, the white bar shows the PC obtained by the persistence forecasts (previous month observed temperature category).