MONTHLY AND SEASONAL FORECASTS

by Louis Lefaivre¹, André Plante¹ and Normand Gagnon² ¹Numerical Weather Prediction Division, Canadian Meteorological Centre ²Department of Atmospheric and Oceanic Sciences, McGill University

TABLE OF CONTENTS

- 1. Introduction
- 2. <u>Nature of the problem</u>
 - 2.1 <u>Time averaging</u>
 - 2.2 Anomalies
 - 2.3 Ensemble mean
- 3. <u>Validating the operational procedure</u>
 - 3.1 Description of the operational procedure
 - 3.1.1 Surface fields
 - 3.1.2 Monthly forecast
 - 3.1.3 Seasonal forecast with zero lead-time
 - 3.1.4 Seasonal forecast with lead-times of 3, 6, and 9 months
 - 3.2 Verification of the procedure
- 4. Verifications
 - 4.1 Monthly forecasts
 - 4.2 Seasonal forecasts
- 5. Conclusions and future work
- 6. References

1. INTRODUCTION

Long confined to 15-day forecasts based on analogue techniques, and to monthly and seasonal forecasts based on subjective techniques, long-range forecasts at the Canadian Meteorological Centre (CMC) have evolved dramatically over the last two years. Since June 1995, monthly forecasts have been based entirely on numerical techniques, and the seasonal forecasts began using this method in September of the same year. In December 1996, seasonal forecasts were extended to lead-times of 3, 6 and 9 months.

This document discusses the nature of the problem of long-range forecasts (section 2), validation of the operational procedure (section 3) and verifications (section 4). Section 5 draws the necessary conclusions and discusses the future of this endeavour.

2. NATURE OF THE PROBLEM

2.1 **Time averaging**

Time average forms the basis of long-range forecasts based on dynamical techniques. Consider the example of the time series illustrated in Figure 1. Two series are shown: the solid line indicates the height analysed at 500 hPa over Dorval (YUL) from day 1 to day 90 of the year 1993: the dashed line shows the height forecast coming from a general circulation model for the same period. Note in this figure that both curves superposes quite well during the first week of the series. However, beyond this period, the troughs and ridges are not only out of phase, but their amplitudes also differ considerably. This is a well known behaviour of models: their predictability period is about 10 days and their variance is less than that observed in nature. The fact that the results shown come from a general circulation model rather than a forecast model has absolutely no effect on the conclusions. Using a higher resolution model would only increase the forecast period slightly, but the rest of the integration would remain the same. Such results can easily lead one to believe that seasonal forecasts are simply impossible.

Let us go back to <u>Figure 1</u>, and rather than examining the behaviour of the geopotential height from day to day, let us take the mean for the whole period. For the analysis, we obtain a value of 531 decametres with a standard deviation of 18 decametres. For the forecast, we obtain 532 decametres with a standard deviation of 12 decametres. All of the long-range forecasting technique is based on this conclusion: it is impossible to forecast the weather on the 47th day of a seasonal forecast, for example, but the mean can however be predicted. Even the standard deviations are in the same order of magnitude. It is fortunate that the averages of the two time series are almost identical, but does this result, which can be associated with a seasonal mean, vary from year to year? This will be discussed in the next section.

2.2 Anomalies

The inter-annual variations in long-range runs have been studied extensively using general circulation models. Unlike forecast models, the circulation models concentrate on reproducing certain atmospheric climate characteristics. Consequently, they are used to produce long integrations (several years) for which the initial conditions are long forgotten. The inter-annual climate variations are simulated by modifying sea-surface temperatures.

Kumar et al., 1996 (KHLS), shows the ability of a general circulation model to forecast height anomalies. In long-range forecasts, it is the anomalies rather than the fields themselves that one wants to predict. In fact, when the numerical models are run for long periods, they have a tendency to drift away from the initial atmospheric conditions and tend towards the climate of the model. If we want to ensure that the anomalies predicted resemble the ones observed, the model's climate must be subtracted from the forecasts provided. Doing this eliminates the model's drift, which could introduce an artificial anomaly caused by systematic errors of the models.

The top of Figure 2, taken from KHLS, shows the geopotential height anomalies analysed at 200 hPa for two five-month periods (November to March): in 1986/1987 (image a) and 1988/1989 (image b). These are, respectively, El Niño and La Niña conditions. Image a shows negative anomalies over eastern Pacific and northern Mexico, and positive anomalies over Hudson's Bay (this is a positive PNA pattern). Image b shows the inverse response (negative PNA pattern). The order of magnitude of anomalies that a numerical model must be able to predict is typical of the anomalies observed over Hudson's Bay. They vary from +10 decametres in 86/87 to -12 decametres in 88/89. There is a large inter-annual variation in heights, caused mainly by a variation in sea-surface temperatures. Forecasting anomalies at a seasonal scale is thus possible only if the oceans are taken into consideration.

The other two series of images in <u>Figure 2</u> show the results of the two atmospheric models: the one in the middle (images c and d) shows the MRF8 model, which did not reproduce reality very well, and the one at the bottom (images e and f), the MRF9 model, for which a modification in the parameterisation was introduced, which shows an amazing result with very similar predicted and observed anomalies.

The image at the top of <u>Figure 3</u> shows the response of the MRF9 model for all the seasons from 1982 to 1993 in terms of correlation anomalies (CA) at 200 hPa. If .5 CA is considered to be a success criterion, then we can conclude that in climate mode, such a model is capable of accurately predicting height anomalies in winters in which sea-surface temperature anomalies are more pronounced. The second series of images in <u>Figure 3</u> shows the response of height anomalies at 500 hPa in the GCMII model (McFarlane et al., 1992) that was run under the same conditions as the MRF9, and the response is similar. The difference between the two can be explained, among other things, by the fact that the MRF9 model has a higher resolution (T40 versus T32).

2.3 Ensemble mean

We have seen that the numerical models are quite capable of forecasting the mean heights for a seasonal period. We have also seen that these heights are not static from year to year, or at least not over our latitudes. We will now discuss the notion of ensemble mean.

All numerical forecasts contain noise, which is due to the initial conditions (imperfect analysis because of an inadequate observation network) and the errors in the model itself. A theoretical example of this notion is shown in <u>Figure 4</u>. Part a) of the figure shows a series of curves, each of which could represent a numerical integration. At first glance, no conclusion can be drawn from this disparate set of lines. However, by calculating the mean, we obtain a coherent series (part b) in which the noise has been filtered out.

Long-range forecasts use a similar technique, meaning that series of forecasts are issued from analyses lagged by 24 hours. By taking the mean of the series obtained in this manner, we hope to filter out the noise coming from random errors in the model. Nevertheless, we know very well that the models are subject to systematic errors that can be eliminated in part by subtracting the integrations from the model's climatology. We will come back to this point in section 3.1.3.

3. VALIDATING THE OPERATIONAL PROCEDURE

3.1 **Description of the operational procedure**

Several studies have shown the response of models in climate mode (Kumar et al. 1996, Zwiers 1996 and Zwiers et al. 1997). In these studies, the initial atmospheric conditions no longer have an influence because models are run over long periods of time (years). However, oceanic forcing is prescribed: the model is provided with the sea-surface temperature anomalies as they have been observed.

In operational mode, the context is completely different: the initial atmospheric conditions are well known because they come from the operational assimilation cycle. The initial oceanic conditions are known at the beginning, but a method must be applied to forecast their evolution until coupled atmosphere-ocean models are perfected.

3.1.1 Surface fields

<u>Section 2.2</u> of this document showed that surface forcing is extremely important if we want to be able to accurately predict the height anomalies, and ultimately the surface temperature anomalies. Of all the types of forcing, the most significant is no doubt the one involving sea-surface temperatures (SST) (Kumar and Hoerling, 1995). An alternate solution had to be found in operational mode because the SST values are unknown and the CMC has no SST anomaly forecast model at this time. Drawing from the results published by Cane and Zebiak, 1995 (CZ), the persistence of SST anomalies proved to be an excellent compromise.

Figure 5, taken from the article by CZ, shows the performance of two SST forecast models for the equatorial Pacific for different periods (1972 to 1992, 1982 to 1992, and 1972 to 1981) and for different scores (correlation and RMS). In all cases, whatever the period or model, the persistence seems to be the best indication of the evolution of the SSTs for the first three months. At CMC, because we need only the first three months, using the persistence of anomalies is a logical choice. Operationally speaking, the SST anomaly observed over the last 30 days (as shown in Figure 6) is added to the climatology for the next three months.

For the other surface fields such as ice and soil moisture, the climatological values are used. Snow is handled differently depending on the model used: for the SEF, the anomaly for the 30 days is maintained in the first month, with a return to climatology for the

subsequent months; for the GCM, the model is started with the last snow analysis available, after which the model treats the snow as a prognostic variable.

3.1.2 Monthly forecast

The SEF forecast model (Ritchie 1991) provides the basis for monthly forecasts. It is the same model as the one used every day for medium-range forecasts (up to 10 days), but with some dynamic and physical modifications so that it can work in climate mode (Desautels et al., 1996). Among other things, a module has been added so that the surface geophysical fields can be read as the integration progresses. Moreover, its spectral truncation (T63 instead of T199) has been reduced, but more levels have been added in the stratosphere (L23 instead of L21). This model has also been used in sensitivity studies of sea-surface temperature forcing (Peng et al., 1995, Dugas et al., 1995).

The monthly forecasts are temperature anomaly outlooks divided into three classes: under, over and near the normal (see <u>Figure 7</u>). The normal is defined as being within .43 times the value of the standard deviations observed during the period extending from 1961 to 1990 at each of some 240 stations across the country. This definition makes it possible to define three equiprobable classes.

The monthly forecasts are issued twice a month, on the first and on the fifteenth. The SEF model is run up to 35 days every day during the five days preceding the forecast issue date. The ensemble mean is then calculated for the five integrations for the corresponding 30-day periods. Next, the mean for the 1000-500 hPa thicknesses for the 30 days covering the forecast period is calculated, followed by a calculation of their anomalies by subtracting the forecast value from the climatology (30-year analyses). These values are then interpolated for the 240 stations across Canada. Finally, the temperature analyses are estimated using the equation below:

 $T_a = b * DZ_a$ (1) where T_a is the temp and DZ_a is the thick

where T_a is the temperature anomaly and DZ_a is the thickness anomaly.

3.1.3 Seasonal forecast with zero lead-time

Two models are used to produce the seasonal forecasts: the SEF forecast model as described in <u>section 3.1.2</u>, and the GCMII general circulation model (McFarlane et al., 1992). The general circulation model is also a spectral model, with a lower resolution (T32L10) and different physics. Using two models for the seasonal forecast makes it possible to generate a wider range of solutions than would be possible with only one model.

Figure 8 illustrates the strategy used to launch the dynamic seasonal forecasts. For both models, the 96-day forecasts are started every day during the six days before the outlook is issued. These are called zero lead-time forecasts. If longer lead-times are desired (see section 3.1.4), other forecast techniques must be used. Contrary to studies such as those carried out by KHLS, the models in our operational configuration benefit from the initial atmospheric conditions and make use of the predictable part of the seasonal integration. In

addition, surface (SST) forcing must be forecast, whereas it is prescribed in the KHLS-type studies.

Unlike monthly forecasts, seasonal forecasts are issued only four times a year. The models are run every day during the six days before the outlook is issued. In addition, a precipitation anomaly outlook is also produced for this forecast.

The flowchart describing the blending procedure of the two models for temperatures is shown in Figure 9. Like the monthly forecasts, it is a perfect prog approach. The "b" coefficients, described in equation (1), are nonetheless different because of different averaging periods. The other difference is due to the fact that the climatology used to calculate the thickness anomaly forecasts come from the models climatologies, which were obtained from historical forecast runs (see section 3.2). The precipitation anomaly outlook is done following a direct approach (Figure 10) in which the direct outputs of the models are used and from which the climatology of the model is subtracted. Once again, forecasts are produced in three equiprobable classes, using the threshold of .43 times the standard deviation provided by the model's climatology.

3.1.4 Seasonal forecasts with lead-times of 3, 6, and 9 months

The good correlations between the SST anomalies and the surface temperature anomalies made it possible to develop purely statistical forecast methods: the "Canonical Correlation Analysis" (CCA) technique (Barnston, 1993), the "Optimal Climate Normals" (OCN) technique (Zhang et al., 1996), and the "Space-Time Principal Component" (STPC) technique (Vautard et al., 1997). The main advantage of these methods is that they can be calculated at very low computational costs. A 30-year data validation can be done in a few days.

In December 1996, the CCA technique, based on the study by Shabbar and Barnston, 1996 (SB), was introduced to forecast the temperature and precipitation anomalies over Canada for lead-times of 3, 6 and 9 months (note that our definition of lead-time corresponds to lead-times of 3 months in SB). By construction, in order to forecast the surface temperature anomalies over Canada, the technique examines the SST anomalies observed over the last 12 months. The skill, which is measured using a cross-validation method, does not vary much with the lead-time: the quality of the seasonal temperature anomalies forecast over Canada in one year will be similar to that of the anomalies forecast in the following season (see Figure 11, taken from SB). This surprising observation needs to be quantified by the fact that the quality of these forecasts is still low: average correlation of the order of .3 (a variance explained of the order of 10%).

3.2 Verification of the procedure

The operational procedure for the seasonal forecasts has been validated over 16 years (1979 to 1994). This verification was made possible thanks to the recent efforts of NCEP, where a complete reanalysis of the observations was carried out using the same analysis system for the entire period (Kalnay et al. 1996). These new, high-quality coherent

analyses were made available to us and were used as initial conditions for the models and for the validation of the results.

A total of 768 seasonal forecasts were thus run: six forecasts per model for each season of the 16 years of the NCEP reanalyses available. Figure 12 shows the configuration used to start these runs. The procedure used for the surface fields is described in Figure 13 (for the SST), figure14 (for ice) and figure15 (for snow). Once the integrations were done, the operational procedure was applied to obtain the forecast temperatures at the stations.

The verification was carried out for some 50 stations distributed relatively evenly across Canada. The three classes for which the percent correct (PC) was calculated were verified as explained in Figure 16. Considering that the forecasts are done in three equiprobable classes, a PC of less than 33% means that there was no improvement compared with climatology. However, with a sampling of 16 realisations, the PC should be greater than 45% in order for the value to be statistically significant. The results for the temperatures are presented in Figure 17. This figure shows the geographic distribution of the PC for each of the four seasons. The season with the most accurate forecasts is clearly summer, for which not only the national PC is the greatest, but for which most of the regions also show PCs greater than 33%. The worst seasons are the intermediate ones (spring and fall).

The precipitation verifications are only preliminary at this time and are presented in Figure 18 (for the summer and fall) and figure 19 (for the winter and spring). The national mean has not been calculated, but the colour has been heightened for those regions with statistically significant PCs (greater than 45%). One can see that not many regions meet these criteria. These disappointing results will be re-evaluated once a more complete database of observations has been obtained.

This effort has made it possible to validate the operational procedure in order to obtain seasonal forecasts. The climatology of the models has been recalculated, the method of preparing the surface field forcing has been redone, and a more suitable combination (normalized anomalies) has been introduced.

4. VERIFICATIONS

The forecasts have also been verified in real time to ensure that their quality meets the quality obtained in <u>section 3.2</u>. The verifications shown in this section are taken from Verret et al. (1998).

4.1 **Monthly forecasts**

Since the monthly forecasts are issued twice a month, statistics on their performance can be accumulated quickly. Figure 20 shows the percent correct score for the temperature outlooks since June 1996. It also shows the performance of the persistence. On average, the persistence had a lower score (38% compared with 42%), but this difference is not statistically significant. Furthermore, during stable periods, such as from November 1996 to January 1997, the persistence can be a good predictor.

4.2 Seasonal forecasts

The seasonal forecast verifications are summarised in <u>Table 1</u> below. Without comparing the characteristics of such a small sample in detail, we must mention that the blending method of the two models has evolved over time. The method described in <u>section 3.1.3</u> was not introduced until the fall of 1997, which clearly explains the disappointing results for the summer of 1997.

Seasons	PC (%) for temperature anomalies	PC (%) for precipitation anomalies
Summer 1996	42.44	n/d
Fall 1996	31.74	n/d
Winter 1997	60.57	38.89
Spring 1997	43.68	28.76
Summer 1997	35.03	26.20
Fall 1997	58.29	24.22

Table 1: Percent correct (PC) for temperature and precipitation anomalies forecast. The verifications were calculated for all the Canadian stations.

5. CONCLUSIONS AND FUTURE WORK

The first conclusion that can be drawn is that long-range forecasts with modest skill are actually possible. Nevertheless, the work continues in order to improve this accuracy.

First, the 15 years of historical forecasts (described in <u>section 3.2</u>) will be extended by another 10 years (1969-1978), using the continued work by NCEP. The 26 years that will eventually be at our disposal will make it possible to perfect the blending method for the dynamic models. The anticipated technique, the "Best Linear Unbiased Estimators" (BLUE), is based on objective analysis techniques (Daley 1991). Using past performances of models, this method makes it possible to weight the outputs of the two models automatically in order to optimise the final result. Preliminary results of this technique have shown a noticeable improvement in temperature forecasts. The BLUE technique will eventually be expanded to blend dynamical forecasts with statistical methods such as CCA or STPC.

The precipitation outlooks will be evaluated in more details, using a more complete database for the observations. The BLUE technique will also be tested on this weather element.

Monthly forecasts will eventually be generated like the seasonal forecasts: use of the seasonal forecast outputs to evaluate the monthly climatologies of the model will be tested. We thus hope to reduce the systematic bias of the model, obtain more realistic thickness forecast anomalies, and obtain temperature anomalies that provide better verification. Similarly, we will be able to provide outlooks of precipitation anomalies.

In the longer term, the dynamic forecasts will be extended to provide seasonal forecasts with lead-times of 3, 6 and 9 months. To do this, an SST forecast method will have to be introduced, at least over the eastern Pacific.

6. **REFERENCES**

Barnston, A. G., 1993. Linear Statistical Short-term Climate Predictive Skill in the Northern Hemisphere, *J. Climate*, **7**, 1513-1564

Daley, R. 1991, Atmospheric Data Analysis, *Cambridge Atmospheric Sciences Series*, 457 pp

Desautels, G, L. Lefaivre, H. Ritchie and Bernard Dugas, 1996. Monthly forecasts using a dynamic technique. *CMC Review*, Vol. 3, No. 1 and 2, 2-16

Dugas, B., G. Brunet and H. Ritchie, 1995. Progress on AMIP integrations with the RPN spectral forecast model. *Proceedings of the AMIP Scientific Conference*, Monterey, California

Cane and Zebiak, 1995 Science, 269, 1699-1702

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 Reanalysis Project. *B. A. M. S.*, **77**, 437-490

Kumar, A., M. Hoerling, M. Ji, A. Leetma and P. Sardeshmukh, 1996. Assessing a GCM's Suitability for Making Seasonal Prediction. *J. Climate*, **9**, 115-129

Kumar, A. and M. Hoerling, 1995. Prospects and Limitations of Seasonal Atmospheric GCM Predictions. *B. A. M. S.*, **76**, 335-345

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

Peng, S., L. A. Mysak, H. Ritchie, J. Derome and B. Dugas, 1995. The differences between early and midwinter atmospheric responses to sea surface temperature anomalies in the Northwest Atlantic. *J. Climate*, **8**, 137-157

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

Shabbar, A. and A. G. Barnston, 1996. Skill of Seasonal Forecasts in Canada Using Canonical Correlation Analysis, *Mon. Wea. Rev.*, **124**, 2370-2385

Vautard, R., G. Plaut, R. Wang et G. Brunet, 1997. Seasonal Prediction of North American Surface Air Temperatures using Space-Time Principal Components, submitted to *J. Climate*

Verret, R., A. Bergeron, L. Lefaivre and André Plante, 1998. Surface temperature anomaly forecasts over periods ranging from five to ninety days. *Proceedings of the 14th Conference on Probability and Statistics*, Phoenix, Arizona.

Zhang, X., A. Shabbar and W. D. Hogg, 1996. Seasonal prediction of Canadian surface climate using optimal climate normals. *Proc.* 21st Annual Climate Diagnostics Workshop, Huntsville, Alabama

Zwiers, F. W., 1996. Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCMII, *Climate Dynamics*, **12**, 825-847

Zwiers, F. W., X. L. Wang and J. Sheng, 1997. Low frequency variability in an ensemble of GISST simulations, submitted.



Figure 1



FIG. 4. Departure from zonal symmetry of the 200-mb height anomalies during winter (December through April) for 1986/87 (left panels) and 1988/89 (right panels). Observations shown at top, MRF8 simulations in middle, and MRF9 simulations at bottom. The model results are the ensemble average of nine integrations with the observed SSTs. Contour interval is 10 m, and negative values are dashed. Anomalies computed with respect to the corresponding dataset's 1982–1993 climatology. Map projection as in Fig. 1.

· · .





Figure 4

-2

 $^{-4}$



(right). Correlations and rms errors between predicted and observed NINO3 SST anomalies for three different

time periods. In each panel, results from the standard, new, and persistence forecasts are shown for comparison. Persistence recasts were obtained by assuming initial SST anomalies remained constant.

Figure 5



Figure 6



Figure 7

analyses	Prévisions	Prévisions	Prévisions	Prévisions
décalées	saisonnières	saisonnières	saisonnières	saisonnières
de 24	avec préavis	avec préavis	avec préavis	avec préavis
heures	nul	de 3 mois	de 6 mois	9 mois
	3 mois domaine de la prévision numérique	3 mois domaine des p	3 mois prévisions statist	3 mois t iques

Figure 8: Schema explaining the strategy used between dynamic forecasts (with no advance notice) and statistical forecasts (longer advance notice)

Analyses staggered by 24 hours Seasonal forecasts with no advance notice Seasonal forecasts with advance notice of 3 months Seasonal forecasts with advance notice of 6 months Seasonal forecasts with advance notice of 9 months 3 months 3 months 3 months Numerical forecasting domain/ Statistical forecasting domain

ESTIMÉ DES TEMPÉRATURES (APPROCHE PROG PARFAIT)

SEF

GCM



Figure 9



Figure 10

Snow for months 0, 1, 2 & 3

• <u>GCM</u>

For initial time: GCM snow climate [kg/m2] is taken where NCEP snow mask has snow.

Then the snow amount is forecasted.

• <u>SEF</u>

For initial time: NCDC snow climate [cm] is taken where NCEP snow mask has snow.

For month 1: snow depth anomaly is persisted.



For months 2&3: NCDC climatology.

Figure 11



Figure 12

SST for months 0, 1, 2 & 3

Same procedure for both models

1) SST background is a 30 average of GISST data: 1951-1980 for HFP years 1979-1990 1961-1990 for HFP years 1991-1994

2) SST anomaly is computed for the start month: SSTA = SST(start month) - SST(30 year average)

3) SSTA persisted through month 1 to 3.



Figure 13



Figure 14





Figure 15. CCA forecast skill averaged over Canada for 3-month mean temperature except the SST field is weighted double its natural value.

Example of contingency table





Figure 16



Figure 17



Figure 18



Figure 19



Figure 20: Percent correct (PC) for the monthly temperature anomaly forecasts. The score is presented relative to the expected value of 33.3% for stochastic forecasts. The dynamical forecasts are in black and persistence in white. The verification was done from June 1996 (J96) to December 1997. The forecasts issued at mid-month are also shown.