Using Bayesian Model Averaging to calibrate meteorogical forecast ensembles: Application to the forecasts of Environment Canada

M. Saïd (1), A.-C. Favre (1), H. Massé (1), V. Fortin (2), L. Perreault (3), N. Evora (3)

(1) Institut National de la Recherche Scientifique - Eau, Terre, Environnement (2) Service météorologique du Canada (3) Institut de recherche d'Hydro-Québec

Introduction

Ensemble meteorological forecasts are often uncalibrated (bias, underdispersion); this is the case for the forecasts from Environment Canada (EC). To calibrate such forecasts, we use the Bayesian Model Averaging (BMA) method proposed by [3]. In the present work, this method is applied to minimum and maximum temperature forecasts from EC, for different forecast ranges and all seasons.

BMA method

BMA method applied to forecast ensembles

If we have a data ensemble of K forecasts $\{f_1, ..., f_K\}$ coming from K models $\{M_1, \ldots, M_K\}$, then the probability density functions (PDF) of the predicted variable y can be written this way:

$$p(y|f_1, ..., f_K) = \sum_{k=1}^{K} w_k \, g_k(y|f_k) \qquad (1)$$

where w_k is the weight of the model M_k and $g_k(y|f_k)$ is the conditional PDF of y, given that f_k is the best forecast in the ensemble.

Choice of the distribution for y

- When the predicted variable y is the temperature, it is reasonable to assume that

$$y \,|\, f_k \sim N\left(f_k, \sigma_k^2\right) = g_k(y|f_k)$$

- The expectation of y given the forecasts is

$$\mathbb{E}(y|f_1, ..., f_K) = \sum_{k=1}^{K} w_k \, f_k$$

Variance decomposition

The variance asociated with the BMA predictive PDF (1) can be decomposed as

$$\mathrm{Var}(y|f_1, \ldots, f_K) = \underbrace{\sum_{k=1}^{K} w_k \left(f_k - \sum_{i=1}^{K} w_i \, f_i\right)^2}_{=\text{ between-forecast variance}} + \underbrace{\sum_{k=1}^{K} w_k \, \sigma_k^2}_{=\text{ within-forecast variance}} \qquad (2)$$

Estimating the variance using only the first term of (2) yields to underestimated variance.

Estimation

General considerations

- The forecasts are corrected by a simple linear regression: $y_t = a_k + b_k f_{kt} + \varepsilon_{kt}$. The corrected forecasts are then: $\tilde{f}_{kt} = \hat{a}_k + \hat{b}_k f_{kt}$.
- We estimate the weights w_k and the variances σ_k^2 by the maximum likelihood method.
- We use the EM algorithm (see [1]), defining the "missing data" $z_{kt} = 1$ if the member k is the best forecast for the day t, and 0 otherwise.
- Let $\hat{z}_{kt}^{(0)}$, $w_k^{(0)}$ and $\sigma_k^{2(0)}$ be the initial values of the unknown z_{kt}, w_k and σ_k^2. The estimates at the jth iteration are given below.

Step E

$$\hat{z}_{kt}^{(j)} = \frac{w_k \, g_k\left(y_t|\tilde{f}_{kt}, \sigma_k^{(j-1)}\right)}{\sum_{i=1}^{K} w_i \, g_i\left(y_t|\tilde{f}_{it}, \sigma_i^{(j-1)}\right)}$$

Step M

$$w_k^{(j)} = \frac{1}{n} \sum_t \hat{z}_{kt}^{(j)} \quad \text{and} \quad \sigma_k^{2(j)} = \frac{\sum_t \hat{z}_{kt}^{(j)} \left(y_t - \tilde{f}_{kt}\right)^2}{\sum_t \hat{z}_{kt}^{(j)}}$$

Data ensemble

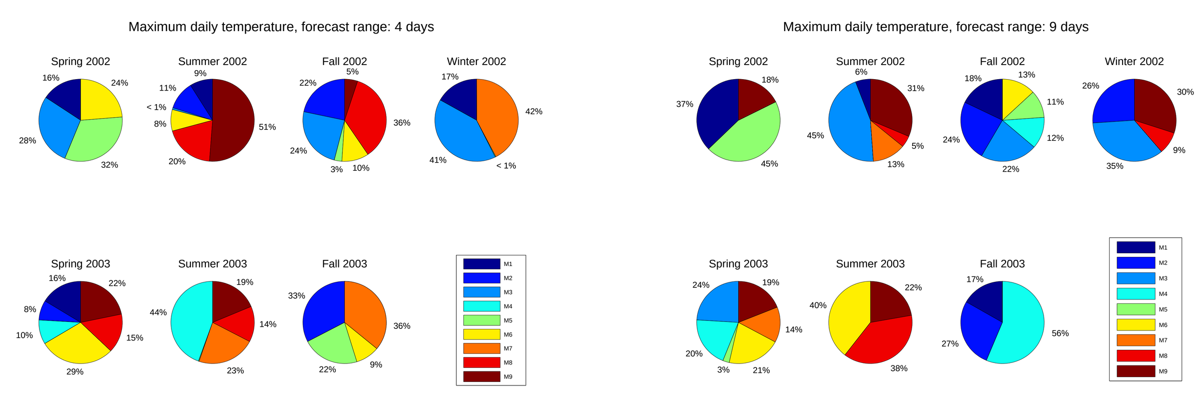
- Global Environmental Multi-scale model (GEM) from Environment Canada
- 8 members and the official forecast for a total of 9 members
- Forecasts on the daily maximum temperature and the daily minimum temperature (maximum and minimum of 8 values)
- Grid point: $(6, 2)$ (see map in [2])
- Forecast range: $k = 0$ to 9 days
- Data from April 2002 to November 2003 (7 seasons)
- Definition of seasons

Spring:	April, May, June	Summer:	July, August and September
Fall:	October, November	Winter:	December, January, February, March

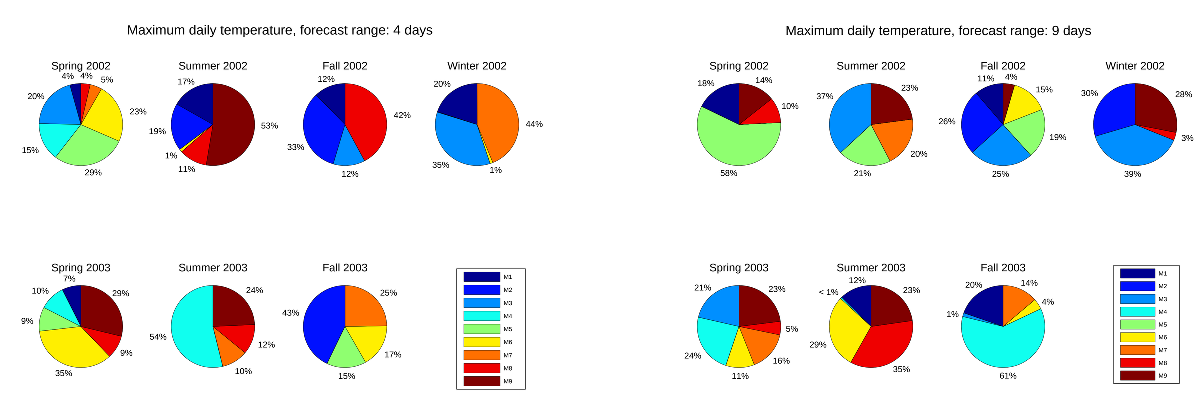
- The analyses using the BMA method concern
  - linear corrected forecasts with distinct slopes b_k and variances σ_k^2
  - linear corrected forecasts with unit slope ($b_k = 1$) and homogeneous variance ($\sigma_k^2 = \sigma^2$)

Weights of different members for maximum daily temperature

σ_k^2, b_k

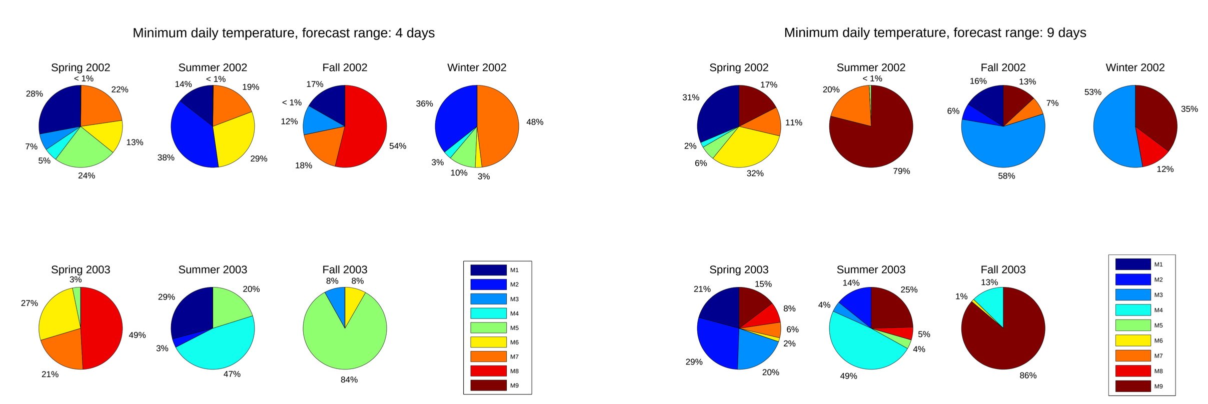


$\sigma_k^2 = \sigma^2$, $b_k = 1$

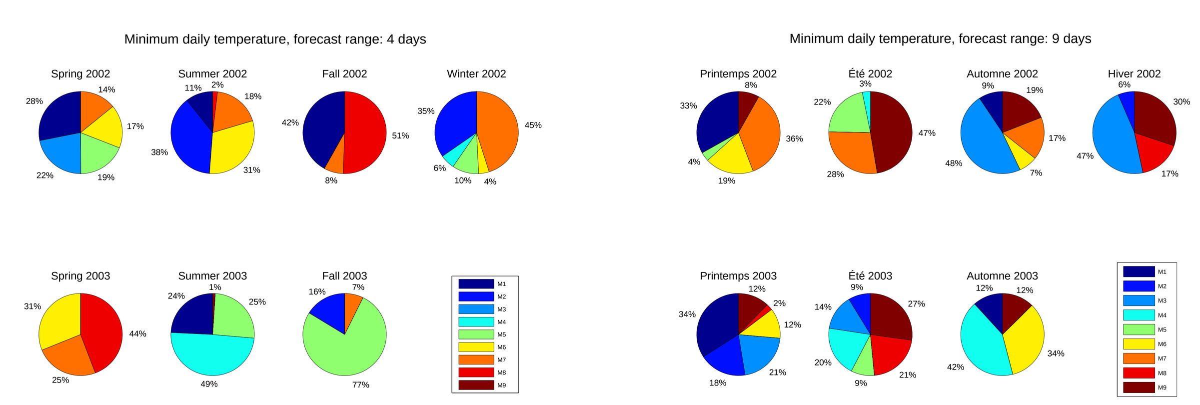


Weights of different members for minimum daily temperature

σ_k^2, b_k

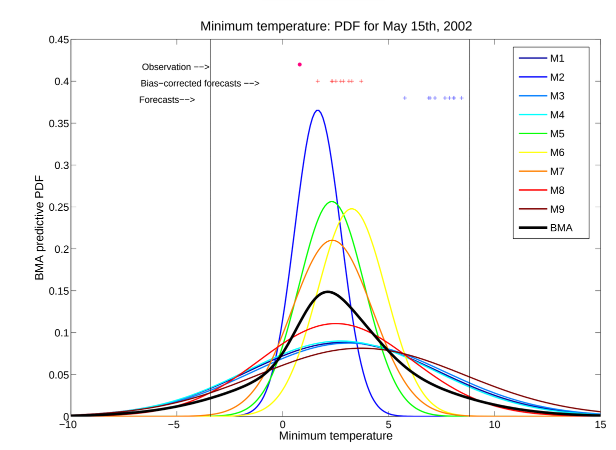


$\sigma_k^2 = \sigma^2$, $b_k = 1$

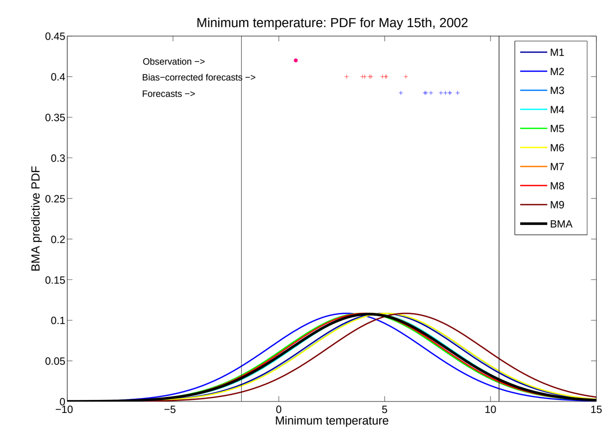


BMA predictive PDF

σ_k^2, b_k

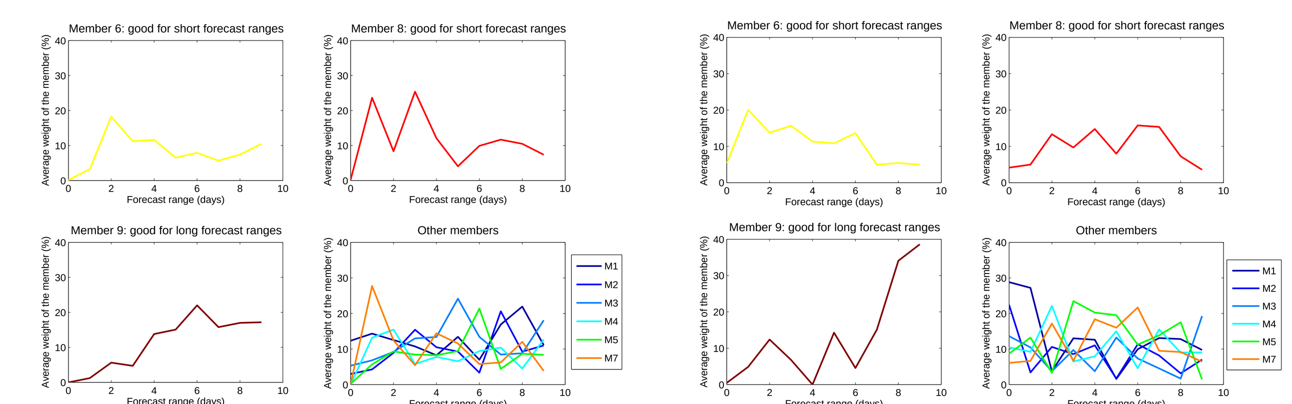


$\sigma_k^2 = \sigma^2, b_k = 1$

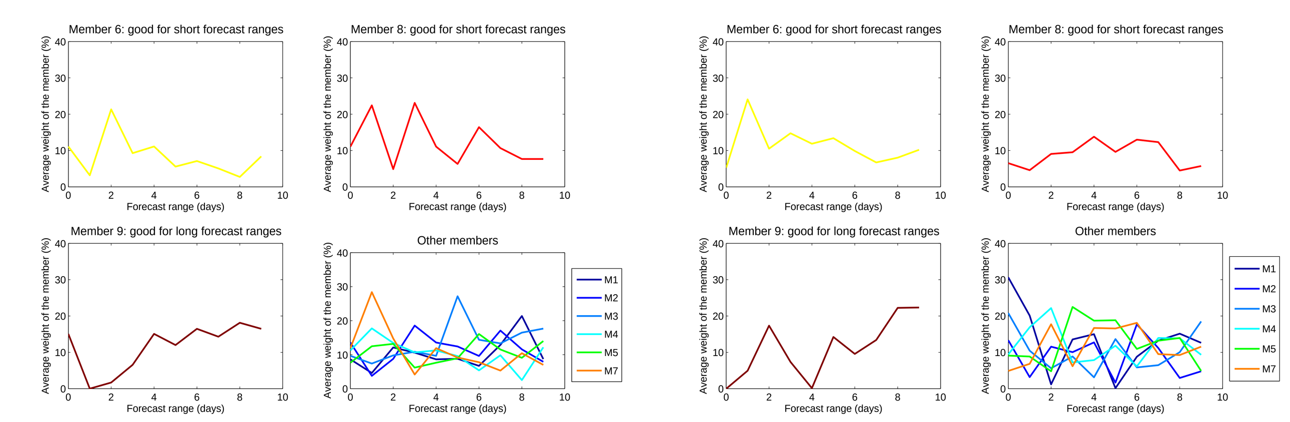


Weights of different members by forecast range

σ_k^2, b_k



$\sigma_k^2 = \sigma^2$
$b_k = 1$



Conclusions and further work

- Members #6 and #8 perform better for short forecast ranges, while member #9 (deterministic forecast) performs better for long forecast ranges.
- The results shown here for linear bias-corrected forecasts with slope 1 and homogeneous variances ($b_k = 1, \sigma_k^2 = \sigma^2$) are not very different from the more general case. We then suggest to use the more parcimonious model.
- The BMA method can be used to calibrate other meteorological variables such as precipitations. However the method should be adapted as the normal distribution is no more appropriate.
- Finally, note that the BMA method supposes the independence of the member forecasts, while these forecasts are often correlated. Additional investigations are needed in order to relax this hypothesis.

References

[1] A. P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B*, 39:1–38, 1977.

[2] N. Evora, L. Perreault, A.-C. Favre, V. Fortin, and H. Massé. *Experimenting a calibration method on canadian meteorological ensemble forecasts for uncertainty assessment of Hydro-Quebec streamflow predictions.* Workshop on the Hydrological Ensemble Prediction Experiment (HEPEX), 2005.

[3] A. E. Raftery, F. Balabdaoui, T. Gneiting, and M. Polakowski. *Using Bayesian Model Aevraging to Calibrate Forecast Ensembles.* Technical Report no. 440, Department of Statistics, University of Washington, 2003.