Post-processing multimodel seasonal climate forecasts: When can we expect an improvement from using unequal weighting, and how big is the expected improvement?

Stefan Siegert

August 16, 2016

Contents

1 Introduction

multiple ensembles for the same prediction target; ensemble of ensembles; superensemble

MMEs are highly structured data with a (potentially) very rich correlation structure (fill in some details and examples)

It has been pointed out that equal weighting is often preferable, even when we have reason to believe that the weights should be different. (fill in details, literature, examples relevant to seasonal)

Questions:

How to optimally combine the MME into a single prediction?

When can we expect unequal weighting to be beneficial?

How big is the expected improvement?

1.1 Literature

DelSole (2013): Is unequal weighting significantly better than equal weighting for multi-model forecasting? (QJ 139, 176–183)

2 Seasonal forecast data: El Nino Southern Oscillation and the North Atlantic Oscillation

We use seasonal forecasts of ENSO and NAO. (fill in model details)

2.1 Summary statistics

ENSO ensemble means and observations: time means, variances (using $1/n$ instead of $1/(n-1)$), and correlations

MEANS :

```
   obs   cfs   cmc  gfdl    mf  nasa    ec
26.70 25.64 25.77 24.89 27.05 25.29 24.82
```

STANDARD DEVIATIONS:
```
 obs  cfs  cmc gfdl   mf nasa   ec
1.21 1.41 1.54 1.30 0.67 1.19 1.24
```

CORRELATIONS:
```
        obs  cfs  cmc gfdl   mf nasa   ec
obs   1.00 0.81 0.89 0.85 0.87 0.88 0.91
cfs   0.81 1.00 0.78 0.89 0.93 0.89 0.86
cmc   0.89 0.78 1.00 0.82 0.83 0.90 0.91
gfdl 0.85 0.89 0.82 1.00 0.87 0.93 0.92
mf    0.87 0.93 0.83 0.87 1.00 0.90 0.90
nasa 0.88 0.89 0.90 0.93 0.90 1.00 0.94
ec    0.91 0.86 0.91 0.92 0.90 0.94 1.00
```

NAO ensemble means and observations: time means, variances, and correlations

MEANS:
```
  obs  ecmwf  lodyn  metfr    mpi   ukmo
  0.6  746.4  704.7 1029.6  740.1  596.2
```

STDEVS:
```
  obs ecmwf lodyn metfr   mpi  ukmo
  1.4 140.1 118.2 158.5 104.1 136.4
```

CORRELATIONS:
```
      obs  ecmwf lodyn metfr   mpi  ukmo
```

```
obs    1.00 -0.070  0.03  0.18  0.019 -0.14
ecmwf -0.07  1.000  0.47  0.23 -0.006  0.14
lodyn  0.03  0.466  1.00  0.16  0.051  0.06
metfr  0.18  0.230  0.16  1.00 -0.130  0.09
mpi    0.02 -0.006  0.05 -0.13  1.000 -0.10
ukmo  -0.14  0.140  0.06  0.09 -0.101  1.00
```

3   Optimal weighting by multivariate analysis

Suppose we knew exactly that the observations and ensemble means in each year were independent draws from multivariate Normal distributions with parameters given by the sample means and correlations above. Then it would be straightforward to derive the optimal weights. (fill in details)

If the observation y_t and vector of ensemble means $\boldsymbol{x}_t = (x_t^{(1)}, \ldots, x_t^{(M)})'$ at time t are draws from a multivariate Normal distribution

$$
\begin{bmatrix} y_t \\ \boldsymbol{x}_t \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_y \\ \boldsymbol{\mu}_x \end{bmatrix}, \begin{bmatrix} \Sigma_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{bmatrix} \right) \tag{1}
$$

then the conditional distribution of the observation, given a values of the ensembles $\boldsymbol{m}_t$ is given by

$$
p(y_t | \boldsymbol{x}_t = \boldsymbol{m}_t) = \mathcal{N}(\mu_{y|x}, \Sigma_{y|x}) \tag{2}
$$

where

$$
\mu_{y|x} = \mu_y + \boldsymbol{\Sigma}_{yx} \boldsymbol{\Sigma}_{xx}^{-1} (\boldsymbol{m}_t - \boldsymbol{\mu}_x) \tag{3}
$$

$$
\Sigma_{y|x} = \Sigma_{yy} - \boldsymbol{\Sigma}_{yx} \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\Sigma}_{xy} \tag{4}
$$

(c.f. Mardia et al). Therefore, the optimal weighting scheme of the M ensemble means is the multiple regression given by

$$\hat{y}_t = a + \sum_{m=1}^{M} w_m x_t^{(m)} \qquad (5)$$

where

$$a = \mu_y - \boldsymbol{\Sigma}_{yx} \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\mu}_x \qquad (6)$$

$$w_m = \boldsymbol{\Sigma}_{yx} (\boldsymbol{\Sigma}_{xx}^{-1})_m \qquad (7)$$

Assuming the sample means, variances, and correlations above are the correct values, the overall mean s, weights w_m, and residual variance $\Sigma_{y|x}$ for ENSO are given by

```
    a  wcfs  wcmc wgfdl   wmf wnasa   wec     v
 -4.07 -0.10  0.33  0.17  0.66 -0.16  0.28  0.19
```

and the values for NAO are

```
     a  wecmwf  wlodyn  wmetfr   wmpi   wukmo      v
-0.36920 -0.00127  0.00074  0.00190  0.00038 -0.00141  1.57664
```

The weights become more interpretable by performing a scale free analysis, i.e., by normalising the observations and time series of ensemble means. Then the weights for ENSO and NAO are given by

```
ENSO:
   a  wcfs  wcmc wgfdl   wmf wnasa   wec     v
 0.00 -0.12  0.43  0.18  0.37 -0.16  0.28  0.13
NAO:
   a wecmwf wlodyn wmetfr   wmpi  wukmo      v
 0.000 -0.130  0.063  0.219  0.029 -0.139  0.929
```

Given the above correlation structures, what would be the optimal post-processing scheme of the multi-model mean (MMM)? If the vector $\boldsymbol{x}$ is a random variable $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then the random variable $\boldsymbol{Ax} \sim \mathcal{N}(\boldsymbol{A\mu}, \boldsymbol{A\Sigma A'})$. Setting

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & \frac{1}{M} & \ldots & \frac{1}{M} \end{pmatrix} \tag{8}$$

yields the joint distribution of the observation and the MMM from which the forecast distribution $p(y_t | \bar{x}_t = m_t)$ can be derived using eq. 2:

$$p(y_t | \bar{x}_t = \bar{m}_t) = \mathcal{N}\left(\mu_y + \frac{\Sigma_{y\bar{x}}}{\Sigma_{\bar{x}\bar{x}}}(\bar{m}_t - \mu_{\bar{x}}), \Sigma_{yy} - \frac{\Sigma_{\bar{x}y}^2}{\Sigma_{\bar{x}\bar{x}}}\right) \tag{9}$$

where $\Sigma_{y\bar{x}}$, $\Sigma_{\bar{x}\bar{x}}$ and $\mu_{\bar{x}}$ are given by ... (fill in details here).

We can now ask what is the potential benefit of using unequal weighting vs equal weighting for MMEs with a known correlation structure. We consider the expected squared difference between forecast mean and observation, as well as the expected Ignorance score, and the expected CRPS of the forecast distribution. We calculate these measures for the optimally weighted forecasts, for the equally-weighted forecasts and for the climatological forecast (where all weights are zero).

If the observations and ensemble means are multivariate Normal, the climatological distribution is simply the marginal distribution of the observation

$$p_{clim}(y_t) \sim \mathcal{N}(\mu_y, \Sigma_{yy}) \tag{10}$$

The expected mean squared error of this forecast is simply

$$E[(y - \mu_y)^2] = \Sigma_{yy} \tag{11}$$

The expected Ignorance score of the climatological distribution is equal to

$$E[-\log(p_{clim}(y))] = \log\sqrt{2\pi\Sigma_{yy}} + \frac{1}{2}\frac{E[(y-\mu_y)^2]}{\Sigma_{yy}} \quad (12)$$

$$= \frac{1}{2}\left(\log 2\pi + \frac{1}{2}\log\Sigma_{yy} + 1\right) \quad (13)$$

i.e., the entropy of the climatological distribution. The expected CRPS of the climatological distribution is equal to

$$E[crps(\mathcal{N}(\mu_y, \Sigma_{yy}), y)] = \sqrt{\frac{\Sigma_{yy}}{\pi}} \quad (14)$$

which follows from integration by parts. Note that, for the climatological Normal distribution, the RMSE and the CRPS are related by a multiplicative factor of $\sqrt{\pi}$.

In fact, the expectation of $crps(y, \mathcal{N}(\mu, \sigma^\in))$, where $y \sim \mathcal{N}(\mu, \sigma^2)$ is always equal to $\sqrt{\sigma^2/\pi}$, independent of the values of μ and σ. Therefore, the expected CRPS of the conditional forecast distribution in a multivariate Normal setting is always equal to $\sqrt{\Sigma_{y|x}/\pi}$.

The expected MSE of the conditional forecast distribution is given by

$$E[(y - \mu_y - \Sigma_{yx}\Sigma_{xx}^{-1}(x - \mu_x))^2] = \Sigma_{y|x} \quad (15)$$

The expected Ignorance score is given by

$$E[-\log p_{y|x}(y)] = \frac{1}{2}(\log 2\pi + \log\Sigma_{y|x} + 1). \quad (16)$$

The respective values for the ENSO and NAO correlation structures are:

```
ENSO
         mse crps  ign
clim    1.51 0.69 1.63
equal   0.24 0.28 0.71
unequal 0.19 0.25 0.59


NAO
          mse     crps     ign
clim    1.9636 0.79058 1.7563
equal   1.9630 0.79046 1.7562
unequal 1.8244 0.76205 1.7196
```

For both ENSO and NAO, the unequal weighting performs better than equal weighting. This result is expected. Since the unequal weighting scheme was based on knowledge of the full joint distribution of the data, it uses the most complete knowledge that we can possibly have about the observation given the forecasts. On the other hand, the equal weighting scheme assumes a simplified covariance structure and is therefore based on imcomplete information.

The analysis also shows the magnitude of improvements we might hope to achieve by using unequal weighting. This improvement can be quantified by relative scores (or skill scores):

$$Skill = \frac{S_{clim} - S_{equal}}{S_{clim} - S_{unequal}} \tag{17}$$

```
skill scores:
ENSO
 mse crps  ign
0.96 0.93 0.89
NAO
```

```
   mse   crps    ign
0.0043 0.0042 0.0041
```

(should state the ignorance difference, not the ratio) The relative improvements are vastly different. For ENSO, the scores achieved by equal weighting is relatively close the scores achieved by unequal weighting, but for NAO the equal weighting scores are (relatively) rarely better than the climatological scores.

However, in general we do not know the population parameters exactly, as was assumed in all the analyses above. In general all we have is a finite sample of forecasts and observations. We can make an assumption of joint Normality and independence and estimate the parameters. In this case, the quality of our post-processed forecasts will also be affected by random errors due to the finiteness of the data set that is used for estimating the parameters.

4   Skill degradation due to parameter estimation errors