Climate model dependence and the replicate Earth paradigm

Author:
Bishop, C; Abramowitz, Gabriel

Publication details:
Climate Dynamics
v. 41
Chapter No. 3-4
pp. 885-900
0930-7575 (ISSN)

Publication Date:
2013

Publisher DOI:
http://dx.doi.org/10.1007/s00382-012-1610-y

Electronic Supplementary Material

for

Climate model dependence and the replicate Earth paradigm

Craig H. Bishop
Naval Research Laboratory,
Marine Meteorology Division
7 Grace Hopper Avenue, Monterey, California 93943-5502
craig.bishop@nrlmry.navy.mil

Gab Abramowitz
Climate Change Research Centre
University of New South Wales
Kensington, Sydney
gabriel@unsw.edu.au

19th September, 2012

Corresponding Author:

Gab Abramowitz
gabriel@unsw.edu.au
Climate Change Research Centre
University of New South Wales, NSW 2052
Australia
Ph: +61 2 9385 8958
Fax: +61 2 9385 8969

ESM A: Optimal weights for the minimum MSD estimate

We seek the vector of coefficients $\mathbf{w}^T = \left[w_1, w_2, ..., w_K \right]$ that minimises

$$\sum_{j=1}^{J} \left(\mu_e^j - y^j \right)^2 \quad \text{where} \quad \mu_e^j = \mathbf{w}^T \mathbf{x}^j = \sum_{k=1}^{K} w_k x_k^j \tag{A1}$$

with the additional constraint that $\sum_{k=1}^{K} w_k = 1$. We should be clear that the x_k^j represent bias-corrected model time series (i.e. they have zero mean error). This requires minimising the function

$$F(\mathbf{w}, \lambda) = \frac{1}{2} \left[\frac{1}{(J-1)} \sum_{j=1}^{J} \left(\mu_e^j - y^j \right)^2 \right] - \lambda \left(\left(\sum_{k=1}^{K} w_k \right) - 1 \right). \tag{A2}$$

Note that the first term in this cost function measures the distance between μ_e^j and the observations y^j and the second term is a constraint term associated with the Lagrange multiplier λ that ensures that the sum of the weights is equal to one. To simplify (A2), we define the *K*-vector $\mathbf{1}^T = \underbrace{\left[1, 1, ..., 1 \right]}_{K-elements}$ and define $\left(\mathbf{y}^j \right)^T = y^j \mathbf{1}^T$ so that $\mathbf{w}^T \mathbf{y}^j = \mathbf{w}^T \mathbf{1} y^j = y^j \left(\sum_{k=1}^{K} w_k \right) = y_j$ provided that $\sum_{k=1}^{K} w_k = 1$. Using $\mathbf{w}^T \mathbf{y}^j = y^j$ and (A1) and (A2) gives

$$\begin{aligned}
F(\mathbf{w}, \lambda) &= \frac{1}{2} \left[\frac{1}{(J-1)} \sum_{j=1}^{J} \left(\mathbf{w}^T \mathbf{x}^j - \mathbf{w}^T \mathbf{y}^j \right)^2 \right] - \lambda \left(\mathbf{w}^T \mathbf{1} - 1 \right) \\
&= \frac{1}{2} \left[\frac{1}{(J-1)} \sum_{j=1}^{J} \mathbf{w}^T \left(\mathbf{x}^j - \mathbf{y}^j \right) \left(\mathbf{x}^j - \mathbf{y}^j \right)^T \mathbf{w} \right] - \lambda \left(\mathbf{w}^T \mathbf{1} - 1 \right) \\
&= \frac{1}{2} \left[\mathbf{w}^T \left[\frac{\sum_{j=1}^{J} \left(\mathbf{x}^j - \mathbf{y}^j \right) \left(\mathbf{x}^j - \mathbf{y}^j \right)^T}{J-1} \right] \mathbf{w} \right] - \lambda \left(\mathbf{w}^T \mathbf{1} - 1 \right) \\
&= \frac{1}{2} \mathbf{w}^T \mathbf{A} \mathbf{w} - \lambda \left(\mathbf{w}^T \mathbf{1} - 1 \right)
\end{aligned} \tag{A3}$$

where $\mathbf{A}$ is the sample-based estimate of the covariance of the bias-corrected errors between all of the ensemble members

$$\mathbf{A} = \frac{\sum_{j=1}^{J}\left(\mathbf{x}^j - \mathbf{y}^j\right)\left(\mathbf{x}^j - \mathbf{y}^j\right)^T}{J-1}. \tag{A4}$$

The cost function F is minimized at the value of the weight vector $\mathbf{w}$ and Lagrange multiplier λ that make the gradients of F with respect to λ and each element of $\mathbf{w}$ zero. The expressions for these gradients are given by

$$\frac{\partial F}{\partial \mathbf{w}} = \begin{bmatrix} \frac{\partial F}{\partial w_1} \\ \vdots \\ \frac{\partial F}{\partial w_K} \end{bmatrix} = \mathbf{Aw} - \lambda\mathbf{1} = \mathbf{0} \quad \text{and} \quad \frac{\partial F}{\partial \lambda} = 1 - \mathbf{w}^T\mathbf{1} = 0 \tag{A5}$$

Setting $\frac{\partial F}{\partial \mathbf{w}}$ to zero gives

$$\mathbf{w} = \lambda\mathbf{A}^{-1}\mathbf{1}. \tag{A6}$$

Using (A6) for $\mathbf{w}$ in the expression for $\frac{\partial F}{\partial \lambda}$ gives $\lambda = \frac{1}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}}$ and hence both of the derivatives in (A5) are simultaneously satisfied when

$$\mathbf{w} = \frac{\mathbf{A}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}}. \tag{A7}$$

Note then that

$$\mu_e^j = \mathbf{w}^T\mathbf{x}^j = \frac{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{x}^j}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}} \tag{A8}$$

defines the minimum error variance estimate. While we noted in Section 2 that performance-only weights can be constructed by ignoring error correlation between models, that is

$$\mathbf{A} = \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_K^2 \end{pmatrix} \quad \text{so that} \quad w_k = \left(\frac{\frac{1}{\sigma_k^2}}{\sum_{j=1}^{K}\frac{1}{\sigma_k^2}} \right), \tag{A9}$$

note also that dependence-only weights can be constructed by scaling the error variance of all models to be equal when constructing $\mathbf{A}$.

The expected error variance of the estimate obtained from (A8) may also be deduced. First note that since $\frac{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{y}^j}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}} = \frac{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}} y^j = y^j$, it follows that one can subtract y^j from both sides of (A8) to obtain

$$\mu_e^j - y^j = \frac{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{x}^j}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}} - \frac{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{y}^j}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}} = \frac{\mathbf{1}^T\mathbf{A}^{-1}}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}}\left(\mathbf{x}^j - \mathbf{y}^j\right) \tag{A10}$$

The average squared error (or distance from observations) over J realizations is then given by

$$\begin{aligned} s_m^2 &= \frac{\sum_{j=1}^{J}\left(\mu_e^j - y^j\right)^2}{J-1} \\ &= \frac{\mathbf{1}^T\mathbf{A}^{-1}}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}}\left[\frac{\sum_{j=1}^{J}\left(\mathbf{x}^j - \mathbf{y}^j\right)\left(\mathbf{x}^j - \mathbf{y}^j\right)^T}{J-1}\right]\frac{\mathbf{A}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}} \\ &= \frac{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{A}\mathbf{A}^{-1}\mathbf{1}}{\left(\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}\right)^2} \\ &= \frac{1}{\left(\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}\right)}. \end{aligned} \tag{A11}$$

As the minimum error variance estimate of the observations, μ_e will be used as our CPDF mean estimate and s_m^2 used to constrain our replicate Earth-like ensemble variance. This is discussed in Section 5.

ESM B: Properties of Earth replicates

Here we deduce $\mathbf{A}_r$, the ensemble error covariance matrix $\mathbf{A}$ that would be obtained if each member of the ensemble $\left(\mathbf{x}^j\right)^T = \left(x_1^j, x_2^j, \ldots, x_K^j\right)$ was a forecast from an Earth replicate and the number of observations tended to infinity so that

$$\mathbf{A}_r = \lim J \to \infty \left[\frac{\sum_{j=1}^{J}\left(\mathbf{x}^j - \mathbf{y}^j\right)\left(\mathbf{x}^j - \mathbf{y}^j\right)^T}{J-1}\right] = \left\langle\left(\mathbf{x} - \mathbf{y}\right)\left(\mathbf{x} - \mathbf{y}\right)^T\right\rangle \tag{B1}$$

where the angle brackets indicate the expectation operator over an infinite time series of verifying observations and ensemble forecasts. Expanding (B1) gives

$$
\begin{aligned}
\mathbf{A}_r &= \left\langle \begin{bmatrix} x_1 - y \\ x_2 - y \\ \vdots \\ x_K - y \end{bmatrix} \left[x_1 - y, x_2 - y, \ldots, x_K - y \right] \right\rangle \\
&= \left\langle \begin{bmatrix} (x_1 - \mu) - (y - \mu) \\ (x_2 - \mu) - (y - \mu) \\ \vdots \\ (x_K - \mu) - (y - \mu) \end{bmatrix} \left[(x_1 - \mu) - (y - \mu), \ldots, (x_K - \mu) - (y - \mu) \right] \right\rangle
\end{aligned}
\tag{B2}
$$

where μ is the true instantaneous mean of the true instantaneous CPDF. Note that

$$\left\langle \left[(x_m - \mu) - (y - \mu) \right] \left[(x_n - \mu) - (y - \mu) \right] \right\rangle = \overline{\sigma_r^2} \left(\delta_{mn} + 1 \right) \tag{B3}$$

because $\left\langle (x_m - \mu)(y - \mu) \right\rangle$=0 and $\left\langle (x_m - \mu)(x_n - \mu) \right\rangle = \delta_{mn} \overline{\sigma_r^2}$ where δ_{mn} is 1 for m=n but zero for $m \neq n$ and where $\overline{\sigma_r^2} = \left\langle (y - \mu)^2 \right\rangle$. Note that because the expectation operator is over time and the true mean μ is evolving in time, $\overline{\sigma_r^2}$ represents a time average of the variance of the time evolving CPDF. It is not the instantaneous variance of the time evolving CPDF. Using (B3) in (B2) gives

$$
\begin{aligned}
\mathbf{A}_r &= \begin{bmatrix} 2\overline{\sigma_r^2} & \cdots & \overline{\sigma_r^2} \\ \vdots & \ddots & \vdots \\ \overline{\sigma_r^2} & \cdots & 2\overline{\sigma_r^2} \end{bmatrix} = \overline{\sigma_r^2} \mathbf{1}\mathbf{1}^T + \overline{\sigma_r^2} \mathbf{I} = \left(2\overline{\sigma_r^2} \right) \mathbf{1}\mathbf{1}^T + \left(\mathbf{I} - \mathbf{1}\mathbf{1}^T \right) \overline{\sigma_r^2} \\
&= \left(2\overline{\sigma_r^2} \right) \mathbf{1}\mathbf{1}^T + \left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}^T}{K} \right) \overline{\sigma_r^2} - \frac{(K-1)\mathbf{1}\mathbf{1}^T}{K} \overline{\sigma_r^2} \\
&= \frac{\left(2K - (K-1) \right) \overline{\sigma_r^2}}{K} \mathbf{1}\mathbf{1}^T + \left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}^T}{K} \right) \overline{\sigma_r^2} \\
&= (K+1) \overline{\sigma_r^2} \frac{\mathbf{1}\mathbf{1}^T}{K} + \left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}^T}{K} \right) \overline{\sigma_r^2}.
\end{aligned}
\tag{B4}
$$

Equation (B4) shows that with a perfect ensemble of Earth replicates (a) the error variance of each of the members is equal to twice the average time averaged climatological variance $\overline{\sigma_r^2}$, and (b) the covariance of the errors of one ensemble member with another member is equal to the time averaged climatological variance. That is, if we agree that the best we can expect from our climate models is to be a perfect replicate Earth, “independence” is not defined by zero error correlation, but rather error covariance equal to $\overline{\sigma_r^2}$. This then implies that the expected error correlation of independent models is $\overline{\sigma_r^2}/\left(\sqrt{2\overline{\sigma_r}}\cdot\sqrt{2\overline{\sigma_r}}\right)=1/2$.

As a check, we will compute the weights **w** (from (A7)) for members of a replicate Earth ensemble. To do this, we require the inverse of $\mathbf{A}_r$, given by

$$\mathbf{A}_r^{-1}=\frac{1}{\left(\left(K+1\right)\overline{\sigma_r^2}\right)}\frac{\mathbf{11}^T}{K}+\frac{1}{\overline{\sigma_r^2}}\left(\mathbf{I}-\frac{\mathbf{11}^T}{K}\right) \qquad \text{(B5)}$$

To prove that (B5) gives the needed inverse, note that using (B4) and (B5)

$$\begin{aligned}\mathbf{A}_r\mathbf{A}_r^{-1}&=\left[\overline{\sigma_r^2}\mathbf{11}^T+\overline{\sigma_r^2}\mathbf{I}\right]\left[\frac{1}{\left(\left(K+1\right)\overline{\sigma_r^2}\right)}\frac{\mathbf{11}^T}{K}+\frac{1}{\overline{\sigma_r^2}}\left(\mathbf{I}-\frac{\mathbf{11}^T}{K}\right)\right]\\&=\frac{K}{K+1}\frac{\mathbf{11}^T}{K}+\frac{1}{K+1}\frac{\mathbf{11}^T}{K}+\left(\mathbf{I}-\frac{\mathbf{11}^T}{K}\right) \qquad \text{(B6)}\\&=\frac{\mathbf{11}^T}{K}+\left(\mathbf{I}-\frac{\mathbf{11}^T}{K}\right)=\mathbf{I}, \quad \text{as was required.}\end{aligned}$$

Using (B5) in (3 or A7) gives the weights

$$\mathbf{w}_r=\frac{\mathbf{A}_r^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{A}_r^{-1}\mathbf{1}}=\frac{\left[\frac{1}{\left(\left(K+1\right)\overline{\sigma_r^2}\right)}\frac{\mathbf{11}^T}{K}+\frac{1}{\overline{\sigma_r^2}}\left(\mathbf{I}-\frac{\mathbf{11}^T}{K}\right)\right]\mathbf{1}}{\mathbf{1}^T\left[\frac{1}{\left(\left(K+1\right)\overline{\sigma_r^2}\right)}\frac{\mathbf{11}^T}{K}+\frac{1}{\overline{\sigma_r^2}}\left(\mathbf{I}-\frac{\mathbf{11}^T}{K}\right)\right]\mathbf{1}}=\frac{\frac{\mathbf{1}}{\left(\left(K+1\right)\overline{\sigma_r^2}\right)}}{\frac{K}{\left(\left(K+1\right)\overline{\sigma_r^2}\right)}}=\frac{1}{K}\mathbf{1} \qquad \text{(B7)}$$

so each of the perfect Earth replicate ensemble members would be given an equal weight of 1/*K*, as one would be expect given the equivalent skill of each member of the ensemble.

The average square difference s_r^2 between the estimate of the mean of the CPDF obtained from the (perfect) replicate Earth ensemble and any particular replicate Earth is obtained by using $\mathbf{A}_r^{-1}$ in (A11) to obtain

$$s_r^2=\frac{1}{\mathbf{1}^T\mathbf{A}_r^{-1}\mathbf{1}}=\frac{K+1}{K}\sigma_r^2=\sigma_r^2+\frac{\sigma_r^2}{K} \qquad \text{(B8)}$$

Equation (B8) shows that, for example, the time averaged squared error of the mean of a perfect ensemble of Earth replicates decreases from $2\sigma_r^2$ to $1.5\sigma_r^2$ as the ensemble size is increased from K=1 to K=2 while it only decreases from $1.033\sigma_r^2$ to $1.017\sigma_r^2$ as the ensemble size is increased from K=30 to K=60. Hence, extremely large ensemble sizes should not be necessary to estimate the time evolving mean of the CPDF. However, one should recognize that the time evolving variance of the CPDF is also of interest and that for this quantity ensemble sizes larger than 60 would probably be required.

ESM C: The ensemble transformation process

If the k^{th} preliminary weights w_k of the weight vector $\mathbf{w}$ gave the relative probability that the k^{th} ensemble member was a member of the CPDF then the instantaneous mean μ_e^j would be as in (A8) while the instantaneous variance σ_e^{2j} of the CPDF would be given by

$$\sigma_e^{2j} = \sum_{k=1}^{K} w_k \left(x_k^j - \mu_e^j \right)^2 . \qquad \text{(C1)}$$

Assuming that climate change is relatively slow, the instantaneous CPDF variance, averaged over time, will approximate the variance of the observations about the CPDF mean (i.e. the error variance of μ_e^j). That is,

$$\frac{1}{J} \sum_{j=1}^{J} \sigma_e^{2j} \approx s_e^2 \qquad \text{(C2)}$$

holds. However, the minimization of F in (A2) does nothing to ensure that (C1) would satisfy (C2). In particular, if any of the weights w_k are negative then they cannot be interpreted as probabilities and the definition of variance given by (C1) does not make sense. To obtain a transformed ensemble that has mean μ_e but which also has a meaningful version of (C1) that satisfies (C2), we first note that the sum of the ensemble perturbations is zero $\mathbf{1}^T \mathbf{x}'^j = 0$ where $\mathbf{x}^j = \overline{x}^j + \mathbf{x}'^j$. Hence,

$$\begin{aligned} \mu_e^j &= \mathbf{w}^T \mathbf{x}^j = \mathbf{w}^T \overline{x}^j + \mathbf{w}^T \mathbf{x}'^j = \overline{x}^j + \mathbf{w}^T \mathbf{x}'^j \\ &= \overline{x}^j + \left(\mathbf{w}^T + (\alpha - 1) \frac{\mathbf{1}^T}{K} \right) \mathbf{x}'^j \end{aligned} \qquad \text{(C3)}$$

where α is *any* scalar. However, the sum of the elements of the row vector $\left(\mathbf{w}^T + (\alpha - 1) \frac{\mathbf{1}^T}{K} \right)$ is not unity. Their sum is given by

$$\sum_{k=1}^{K}\left(w_k + \frac{(\alpha-1)}{K} \right) = \alpha. \qquad \text{(C4)}$$

Ensemble variance is not equal to error variance in general. One way to address this mismatch is to alter the magnitude of the ensemble perturbations. If we adjust the size of the ensemble perturbations by the factor α to obtain $\mathbf{z}'^{j} = \alpha \mathbf{x}'^{j}$, (C3) can be rewritten in the form

$$\begin{aligned} \mu_e^j &= \bar{x}^j + \left[\frac{\left(\mathbf{w}^T + (\alpha-1)\dfrac{\mathbf{1}^T}{K} \right)}{\alpha} \right] \mathbf{z}'^{j} \\ &= \bar{x}^j + \tilde{\mathbf{w}}^T \mathbf{z}'^{j} = \bar{x}^j + \sum_{k=1}^{K} \tilde{w}_k z'^{j}_k \qquad \text{(C5)} \\ &= \bar{x}^j + \sum_{k=1}^{K} \alpha \tilde{w}_k x'^{j}_k \end{aligned}$$

where the row vector $\tilde{\mathbf{w}}^T$ is equal to the term in square brackets. Note that the sum of the K elements of $\tilde{\mathbf{w}}$ satisfies $\sum_{k=1}^{K} \tilde{w}_k = 1$. Hence, if we define $\mathbf{z}^j = \bar{\mathbf{x}}^j + \mathbf{z}'^{j} = \bar{\mathbf{x}}^j + \alpha \mathbf{x}'^{j}$ to be the adjusted ensemble,

$$\mu_e^j = \tilde{\mathbf{w}}^T \bar{\mathbf{x}}^j + \tilde{\mathbf{w}}^T \mathbf{z}'^{j} = \tilde{\mathbf{w}}^T \left(\bar{\mathbf{x}}^j + \mathbf{z}'^{j} \right) = \tilde{\mathbf{w}}^T \mathbf{z}^j. \qquad \text{(C6)}$$

Since each distinct α value defines a unique weight vector $\tilde{\mathbf{w}}^T$ together with a unique adjusted ensemble $\mathbf{z}^j$, (C6) describes the complete set of adjusted ensembles whose weighted mean gives the same minimum error variance estimate as (A8). To obtain an ensemble transformation that only involves positive weights, we choose $\alpha = 1$ if all of the preliminary weights $w_k \ge 0$. Otherwise, we choose α so that the smallest weight is zero: $\min(\tilde{w}_k) = 0$. This is achieved by setting $\alpha = 1 - K \min(w_k)$ where $\min(w_k)$ is the lowest of the (negative) preliminary weights. Having chosen α in this way, we can then ensure that the variance constraints (C1) and (C2) are satisfied by letting

$$\tilde{x}_k^j = \mu_e^j + \beta\left(\bar{x}^j + \alpha x'^{j}_k - \mu_e^j \right) \qquad \text{(C7)}$$

where

$$\beta = \sqrt{\frac{s_e^2}{\dfrac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K} \tilde{w}_k \left[\left(z_k^j - \mu_e^j \right) \right]^2}} \qquad \text{(C8)}$$

since we want

$$\frac{1}{J}\sum_{j=1}^{J}\sigma_e^{2j} = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{w}_k\left[\beta\left(z_k^j - \mu_e^j\right)\right]^2 = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{w}_k\left[\beta\left(\bar{x}_k^j + \alpha x'^j_k - \mu_e^j\right)\right]^2 = s_e^2 \quad \text{(C9)}$$

as required by (C2).

With these relationships in hand, we can now prove that the transformed ensemble given by (C7) has a weighted mean equal to μ_e and a time averaged weighted variance equal to s_e^2. To prove that its mean equals μ_e, use (C7) to show that

$$\begin{aligned}
\sum_{k=1}^{K}\tilde{w}_k\tilde{x}_k &= \frac{1}{K}\sum_{k=1}^{K}\tilde{w}_k\left[\mu_e + \beta\left(\bar{x} + \alpha x'_k - \mu_e\right)\right] \\
&= \mu_e - \beta\left(\mu_e\right) + \beta\sum_{k=1}^{K}\tilde{w}_k\left(\bar{x} + \alpha x'_k\right) \\
&= \mu_e + \beta\left(\mu_e - \mu_e\right), \quad \text{because} \quad \left[\bar{x} + \sum_{k=1}^{K}\tilde{w}_k\left(\alpha x'_k\right)\right] = \mu_e \quad \text{from (C5)} \\
&= \mu_e
\end{aligned} \qquad \text{(C10)}$$

as was required. Furthermore,

$$\begin{aligned}
\frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{w}_k\left(\tilde{x}_k^j - \mu_e^j\right)^2 &= \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{w}_k\left[\beta\left(\bar{x}^j + \alpha x'^j_k - \mu_e^j\right)\right]^2 \\
&= \beta^2\frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{w}_k\left[\left(z_k^j - \mu_e\right)\right]^2, \quad \text{because} \quad \mathbf{z}^j = \bar{\mathbf{x}}^j + \alpha\mathbf{x}'^j \\
&= \left\{\frac{s_e^2}{\frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{w}_k\left[\left(z_k^j - \mu_e^j\right)\right]^2}\right\}\frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{w}_k\left[\left(z_k^j - \mu_e\right)\right]^2 \\
&= s_e^2
\end{aligned} \qquad \text{(C11)}$$

as was required. Note that the correlation of each model in the perturbed ensemble given by (C7) with the original model is equal to one.