A method for estimating and assessing modes of interannual variability in coupled climate models

S. Grainger\textsuperscript{1} \hspace{1cm} C. S. Frederiksen\textsuperscript{2} \hspace{1cm} X. Zheng\textsuperscript{3}

(Received 20 March 2015; revised 31 January 2016)

Abstract

The seasonal mean of a climate variable consists of: slow-external; slow-internal; and intraseasonal components. Using an analysis of variance-based method, the interannual variability of the seasonal mean from an ensemble of coupled atmosphere-ocean general circulation model (CGCM) realisations is separable into these three components. Eigenvalue decomposition is applied to the covariance matrices to obtain, for each component, the dominant modes of variability (eigenvectors) and their associated variance (eigenvalues) for the climate variable. Here, a method is described that assesses the modes of interannual variability in CGCMs against those obtained from reanalysis data based on observations. A metric is defined based on the pattern correlation

gives this article, © Austral. Mathematical Soc. 2016. Published February 17, 2016, as part of the Proceedings of the 17th Biennial Computational Techniques and Applications Conference. ISSN 1446-8735. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to this URL for this article.
between the observed and modelled modes of variability, and the ratio of their associated variances. This metric is applied to monthly mean southern hemisphere 500 hPa geopotential height from the second half of the 20th century. It is shown that CGCMs have clear differences in the slow-component of modes of interannual variability, related to external forcings and/or slowly-varying internal variability.

1 Introduction

The variability of the atmospheric circulation is controlled by many physical processes, which may act on time scales ranging from days to years. These processes, on their different timescales, influence the interannual variability of the seasonal mean of a climate variable [1]. Consequently, a seasonal mean climate anomaly is considered as a statistical random variable consisting of signal and noise components [2]. The signal is related to slowly varying (a season or longer) processes and is considered the slow component of interannual variability of the seasonal mean [3]. In a coupled atmosphere-ocean climate
Separation of variability

Given the conceptual model described in Section 1, the separation of the interannual variability of the seasonal mean into signal and noise components is possible given at least monthly data [1, 8]. Here, we assume that we have
2 Separation of variability

monthly mean data on a spatial grid for an ensemble of model realisations. In this case, after the annual cycle is removed, the monthly mean anomaly of a climate variable $x$ at any grid point is [8]

$$x_{\text{sym}} = \beta_y + \delta_{sy} + \varepsilon_{\text{sym}},$$  \hspace{1cm} (1)

where $s = 1, \ldots, S$ is the realisation number in an ensemble of size $S$, $y = 1, \ldots, Y$ is the year index in a sample of $Y$ years and $m = 1, 2, 3$ is the month index in a season. The slow-external component, independent of realisation, is $\beta_y$, $\delta_{sy}$ is the slow-internal component, taken to be constant over a season, and the intraseasonal component $\varepsilon_{\text{sym}}$ is the residual monthly departure of $x_{\text{sym}}$ from the slow components. Here, we are interested in the time series of the slow component, that is,

$$\mu_{sy} = \beta_y + \delta_{sy}.$$  \hspace{1cm} (2)

Zheng and Frederiksen [3] showed that it is possible to estimate covariance matrices for the slow and intraseasonal components of the interannual variability of the seasonal mean. The total seasonal mean covariance is estimated as the sum of the covariances of the external and internal components [8], that is,

$$\hat{V}(x_{\text{sym}}^1, x_{\text{sym}}^2) = \hat{V}(\beta^1_y, \beta^2_y) + \hat{V}(\delta^1_{sy}, \delta^2_{sy}, \varepsilon^1_{\text{sym}}, \varepsilon^2_{\text{sym}}),$$  \hspace{1cm} (3)

where $x_{\text{sym}}^1$ and $x_{\text{sym}}^2$ are the time series at any two grid points in the set $i = 1, \ldots, I$, the subscript $\circ$ denotes an average over an index $s, y$ or $m$, and $\hat{V}$ denotes an estimated covariance. Total internal covariance is estimated by

$$\hat{V}(\delta^1_{sy} + \varepsilon^1_{\text{sym}}, \delta^2_{sy} + \varepsilon^2_{\text{sym}}) = \frac{1}{Y(S-1)} \sum_{y=1}^{Y} \sum_{s=1}^{S} (x_{\text{sym}}^1 - x_{\text{symo}}) (x_{\text{sym}}^2 - x_{\text{symo}}),$$  \hspace{1cm} (4)

and the covariance of the slow-external component by

$$\hat{V}(\beta^1_y, \beta^2_y) = \hat{V}(x_{\text{symo}}^1, x_{\text{symo}}^2) - \frac{1}{S} \hat{V}(\delta^1_{sy} + \varepsilon^1_{\text{sym}}, \delta^2_{sy} + \varepsilon^2_{\text{sym}}).$$  \hspace{1cm} (5)
2 Separation of variability

where

\[ \hat{\mathcal{V}}(x_{oy_0}^1, x_{oy_0}^2) = \frac{1}{Y-1} \sum_{y=1}^{Y} (x_{oy_0}^1 - x_{oo}^1) (x_{oy_0}^2 - x_{oo}^2) \]  

(6)

is the ensemble mean seasonal mean covariance. In the special case of a single model realisation, that is \( S = 1 \), it is not possible to separately estimate the external and internal covariances; instead, the total seasonal mean covariance is estimated directly from equation (6).

For the intraseasonal component, Zheng and Frederiksen [3] showed that the covariance is able to be estimated as a function of monthly moments. In this case,

\[ \hat{\mathcal{V}}(\varepsilon_{sy_0}^1, \varepsilon_{sy_0}^2) = \frac{1}{9} \left[ \hat{\alpha}(3 + 4\hat{\phi}) \right] , \]  

(7)

where

\[ \hat{\alpha} = a/[2(1 - \hat{\phi})] \]  

(8)

is the covariance of the intraseasonal components within each month, and

\[ \hat{\phi} = (a + 2b)/[2(a + b)] , \quad 0 \leq \hat{\phi} \leq 0.1 , \]  

(9)

is the intermonthly correlation between consecutive months. The covariance and intermonth correlations of the intraseasonal components are assumed to be independent of months within a season. The two monthly moments are

\[ a = \frac{1}{2} \left[ \frac{1}{YS} \sum_{y=1}^{Y} \sum_{s=1}^{S} (x_{sy_1}^1 - x_{sy_2}^1) (x_{sy_1}^2 - x_{sy_2}^2) \right. \]

\[ + \left. \frac{1}{YS} \sum_{y=1}^{Y} \sum_{s=1}^{S} (x_{sy_2}^1 - x_{sy_3}^1) (x_{sy_2}^2 - x_{sy_3}^2) \right] , \]  

(10)

\[ b = \frac{1}{2} \left[ \frac{1}{YS} \sum_{y=1}^{Y} \sum_{s=1}^{S} (x_{sy_1}^1 - x_{sy_2}^1) (x_{sy_2}^2 - x_{sy_3}^2) \right. \]

\[ + \left. \frac{1}{YS} \sum_{y=1}^{Y} \sum_{s=1}^{S} (x_{sy_2}^1 - x_{sy_3}^1) (x_{sy_1}^2 - x_{sy_2}^2) \right] . \]  

(11)
The interannual covariance of the slow component is defined as the residual of the total seasonal mean covariance after the removal of the covariance of the intraseasonal component, that is,

\[ \hat{\mathbf{V}}(\mu_{sy}^1, \mu_{sy}^2) = \hat{\mathbf{V}}(x_{sy o}^1, x_{sy o}^2) - \hat{\mathbf{V}}(\varepsilon_{sy o}^1, \varepsilon_{sy o}^2). \]  

(12)

For all components, \((I \times I)\) covariance matrices are obtained by applying equations \((3), (7)\) and \((12)\) to all pairs of grid points. The modes of interannual variability of each component are defined as the empirical orthogonal functions (EOFs) obtained by eigenvalue decomposition of the corresponding covariance matrix, in descending order by variance explained [9]. The leading eigenvectors give the dominant modes of variability for each component, and the corresponding eigenvalues give the estimated variance associated with each mode.

3 Model assessment

The centred mean square difference between reference and model samples is [5]

\[ E' = \hat{\mathbf{V}} + \hat{\mathbf{V}}' - 2\sqrt{\hat{\mathbf{V}}' C}, \]  

(13)

where \(C\) is the correlation between the two samples, and \(\hat{\mathbf{V}}\) and \(\hat{\mathbf{V}}'\) are the estimated reference and model sample variances, respectively, for example the EOF associated variances defined in Section 2. Based on this, Grainger et al. [10] proposed a score for how well a CGCM reproduces the \(j\)th reference slow component mode of variability (slow-EOF):

\[ M_{\mu}^j = \frac{|R_j| \left(1 + R_{sst}^j\right)^2}{2 \left(\frac{\hat{\mathbf{V}}_{\mu}}{\hat{\mathbf{V}}_{\mu} + \hat{\mathbf{V}}_{\mu}'}/\right)}, \]  

(14)

where \(R_j\) is the pattern correlation between the model and reference slow-EOFs and \(R_{sst}^j\) is the pattern correlation between the model and reference
slow sea surface temperature (SST)–height covariance patterns. The method for calculating the slow SST–height covariance patterns is analogous to the covariance methodology in Section 2, and is detailed by Grainger et al. [11].

The application of equation (14) requires a one to one match between model and reference slow-EOFs. This is obtained by the following procedure.

1. For the leading \( J \) reference slow-EOFs, find the permutation of the leading \( J \) model slow-EOFs which maximises

\[
\sum_{j=1}^{J} |R_j^i| \left(1 + R_{sst}^i\right)^2.
\]

2. For each \( j = 1, \ldots, J \) reference mode, first check for higher order, that is, \( > J \) model modes with a higher score \( M_{\mu}^j > M_{\mu}^i \). Then check all model modes for ambiguous scores, defined as \( M_{\mu}^j \geq 0.75M_{\mu}^i \).

3. If higher order or ambiguous modes are identified, manually inspect the model slow-EOFs to resolve the one to one match.

### 4 Application

To illustrate the methodology, slow-EOFs of 500 hPa geopotential height for the southern hemisphere summer (December-January-February, DJF) and winter (June-July-August, JJA) are examined. CGCM data were obtained from the WCRP CMIP3 [6] and CMIP5 [7] multi-model datasets. Data from the Twentieth Century Reanalysis Project (20CR) [12] for the period 1951–2000 was used as the reference dataset. Observed SST data were obtained from the HADISST dataset [13]. All 500 hPa geopotential height data were mapped to a \( 2.5^\circ \times 2.5^\circ \) latitude/longitude grid, and SST data to a \( 2^\circ \times 2^\circ \) latitude/longitude grid.
Figure 1: (a) Leading three slow-EOFs (normalised to unit length) of 20CR southern hemisphere 500 hPa geopotential height for DJF for the period 1951–2000. (b) Slow SST–height covariance (mK) with HADISST SST for the slow-EOFs in (a). The estimated standard deviation (m) and variance explained (%) are given bottom left in (a) for each EOF.

The leading three slow-EOFs of 20CR southern hemisphere 500 hPa geopotential height in DJF and JJA are shown in Figures 1 and 2, respectively. In both seasons, the leading slow-EOF represents high latitude variability associated with the southern annular mode [4]. There is a protrusion into the South Pacific, particularly in JJA (Figure 2(a)). Slow-EOFs 2 and 3 in both seasons represent variability associated with ENSO [4], evident in the slow SST–height covariances (Figures 1(b) and 2(b)), which show a strong relationship with tropical Pacific SSTs.
Figure 2: (a) Leading three slow-EOFs (normalised to unit length) of 20CR southern hemisphere 500 hPa geopotential height for JJA for the period 1951–2000. (b) Slow SST–height covariance (mK) with HADISST SST for the slow-EOFs in (a. The estimated standard deviation (m) and variance explained (%) are given bottom left in (a) for each EOF.

The matching slow-EOFs were estimated for the same period using the ensemble over all realisations for each of the 23 CMIP3 and 45 CMIP5 models which were available. The model slow-EOFs were then evaluated against the 20CR slow-EOFs (see Section 3) using $J = 3$. For each season, an overall score is calculated by

$$M^{ss}_{\mu} = \frac{1}{3} \sum_{j=1}^{3} M^{j}_{\mu},$$

where $ss$ denotes a season, that is DJF or JJA. The overall scores thus
calculated are shown in Figure 3 for all models for both seasons. In order to rank the CGCMs, a weighting based on their relative spread within each season is used, that is,

$$M^{\text{tot}}_\mu = M^{\text{djf}}_\mu + 2.2M^{\text{jja}}_\mu.$$  \hspace{1cm} (16)

In DJF (Figure 3(a)), there is fairly consistent performance across most models, with many models having an overall score $> 0.4$. In contrast, the CGCMs reproduce the 20CR JJA slow-EOFs (Figure 2) less well, with only two CGCMs having an overall score $> 0.4$. Figure 3 also shows that the performance of CMIP5 CGCMs has improved relative to CMIP3, with generally higher overall scores in both seasons.

5 Conclusions

In this article, a method was formulated to assess the skill of climate models in reproducing the leading slowly varying modes of interannual variability of the seasonal mean. The method was applied to southern hemisphere 500 hPa geopotential height. Coherent spatial patterns, represented by the slow-EOFs, of interannual variability for summer (DJF) and winter (JJA) were estimated for the CMIP5 and CMIP3 datasets for the period 1951–2000. These were assessed against reference slow-EOFs from 20CR data for the same period.

The 20CR slow-EOFs are best reproduced in DJF. The slow-EOFs are less well reproduced in JJA. The spread of results in both seasons enables the definition of a metric ranking model overall performance. There are clear improvements in the CMIP5 dataset over CMIP3 during both summer and winter. The largest individual improvements (not shown) in CMIP5 CGCMs are in the spatial structures of the slow-EOFs related to ENSO variability and their slow SST–height covariances.

The method developed in this article is generally applicable to any climate variable for which the interannual variability of the components are separable.
Figure 3: Model overall score (equation (15)) for (a) DJF and (b) JJA for all CMIP3 (blue) and CMIP5 (red) CGCMS. Models are shown in order of the total overall score (equation (16)).
The slow-EOFs of northern hemisphere 500 hPa geopotential height in CMIP5 models will be assessed. The method is able to track improvements in the interannual variability of future multi-model datasets as they become available.

Acknowledgements We acknowledge the modelling groups, the Program for Climate Model Diagnosis and Intercomparison and the WCRP’s Working Group on Coupled Modeling for their roles in making available the CMIP3 and CMIP5 multi-model datasets. This work is partially supported by the Australian Climate Change Science Program of the Australian Department of the Environment.

References


References


Author addresses

1. **S. Grainger**, Environment and Research Division, Bureau of Meteorology, Melbourne, Victoria 3001, Australia.

   mailto:c.frederiksen@bom.gov.au

3. **X. Zheng**, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China.