A method for evaluating the modes of variability in general circulation models

S. Grainger[1]   C. S. Frederiksen[2]   X. Zheng[3]

Abstract

The seasonal mean of an atmospheric climate variable is considered to be a statistical random variable with two components: a slow component related to slowly varying forcing from external and internal atmospheric sources (time scale of a season or more), and an intraseasonal component related to forcing from weather variability with time scale less than a season. Here, a method is proposed to compare the modes of variability obtained from eigenvalue decomposition of the slow and intraseasonal covariance matrices estimated from reanalysis data with modes of variability estimated from a set of coupled general circulation models. As an example, the method is applied to the Southern Hemisphere summer 500hPa geopotential height for the period 1951–2000. The method is applicable to many other atmospheric climate variables and datasets.

Contents

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. How these physical processes, or modes, are affected under climate change scenarios is able to be investigated by General Circulation Models (GCMs). However, in order to better assess the reliability of any changes in variability, it is necessary to compare the modes of variability estimated from GCMs with those estimated from a reference sample.

The seasonal mean of a climate variable is considered to be a statistical random variable consisting of signal and noise components [3]. The signal is related to slowly varying (a season or more) boundary, or external, forcings or internal atmospheric variability and is considered as the 'slow' component of interannual variability of the seasonal mean [10]. The noise is related to internal atmospheric variability with time scales of about two weeks to a season and is considered as the 'intraseasonal' component of variability [10].

Frederiksen and Zheng [3] reviewed methods to estimate the intraseasonal

variance. Recently, a method has been formulated to estimate the intraseasonal covariance from monthly mean data [10]. The method has been applied to reanalysis data of the $500\,\text{hPa}$ geopotential height field, in both hemispheres, for summer and winter [2, 4]. However, it is also possible to apply the method to individual realisations from GCMs. Here, by interpolating all GCM realisations onto the same grid as the reanalysis data, the intraseasonal and slow modes of variability are compared.

2 Methodology

2.1 Statistical model

Given the statistical model described in the Introduction, the monthly mean anomaly of a climate variable x at a geographical location r ($r = 1, \ldots, R$) is represented as [10]

$$x_{ym}(r) = \mu_y(r) + \epsilon_{ym}(r), \qquad (1)$$

where m ($m = 1, 2, 3$) is the month index in the season, y ($y = 1, \ldots, Y$) is the year index, Y is the total number of years, $\mu_y(r)$ is the seasonal population mean in year y taken from an infinite population, and $\epsilon_{ym}(r)$ is the residual monthly departure of $x_{ym}(r)$ from $\mu_y(r)$. Here, we assume that the vector $\{\epsilon_{y1}(r), \epsilon_{y2}(r), \epsilon_{y3}(r)\}$ is a stationary and independently distributed random vector with respect to year y. In practice, neither $\mu_y(r)$ or $\epsilon_{ym}(r)$ are able to be directly calculated. However, the covariance V between x at two locations r_1 and r_2 is

$$\begin{aligned} V(x_{yo}(r_1), x_{yo}(r_2)) \;=\; & V(\epsilon_{yo}(r_1), \epsilon_{yo}(r_2)) + V(\mu_y(r_1), \mu_y(r_2)) \\ & + V(\mu_y(r_1), \epsilon_{yo}(r_2)) + V(\mu_y(r_2), \epsilon_{yo}(r_1)), \quad (2) \end{aligned}$$

where a subscript "o" represents the average over an index (m or y).

Zheng and Frederiksen [10] applied stationary assumptions to the statistics of monthly data. The intraseasonal covariance $V(\epsilon_{yo}(r_1), \epsilon_{yo}(r_2))$ is then estimated as

$$\hat{V}(\epsilon_{yo}(r_1), \epsilon_{yo}(r_2)) = \frac{1}{9} \left[\hat{\sigma}_2^2 (3 + 4\hat{\phi}) \right] , \tag{3}$$

where

$$\hat{\sigma}^2 = \frac{a}{2(1-\hat{\phi})} \approx V(\epsilon_{ym}(r_1), \epsilon_{ym}(r_2)), \quad m = 1, 2, 3, \tag{4}$$

is the covariance of the intraseasonal component in each month,

$$\hat{\phi} = \frac{a + 2b}{2(a+b)}, \quad 0 \le \hat{\phi} \le 0.1, \tag{5}$$

is the intermonthly correlation between consecutive months, and

$$
\begin{aligned}
a &= \frac{1}{2} \left\{ \frac{1}{Y} \sum_{y=1}^{Y} [x_{y1}(r_1) - x_{y2}(r_1)] [x_{y1}(r_2) - x_{y2}(r_2)] \right. \\
&\quad \left. + \frac{1}{Y} \sum_{y=1}^{Y} [x_{y2}(r_1) - x_{y3}(r_1)] [x_{y2}(r_2) - x_{y3}(r_2)] \right\} , \tag{6} \\
b &= \frac{1}{2} \left\{ \frac{1}{Y} \sum_{y=1}^{Y} [x_{y1}(r_1) - x_{y2}(r_1)] [x_{y2}(r_2) - x_{y3}(r_2)] \right. \\
&\quad \left. + \frac{1}{Y} \sum_{y=1}^{Y} [x_{y2}(r_1) - x_{y3}(r_1)] [x_{y1}(r_2) - x_{y2}(r_2)] \right\} . \tag{7}
\end{aligned}
$$

The variables a and b in equations (6) and (7) represent monthly moments of $x_{ym}(r)$ and are directly sampled. The total interannual covariance is estimated from the sample seasonal mean as

$$\hat{V}(x_{yo}(r_1), x_{yo}(r_2)) = \frac{1}{Y-1} \sum_{y=1}^{Y} [x_{yo}(r_1) - x_{oo}(r_1)] [x_{yo}(r_2) - x_{oo}(r_2)] . \tag{8}$$

The slow covarince is then estimated as the residual

$$\hat{V}(\mu_y(r_1), \mu_y(r_2)) = \hat{V}(x_{yo}(r_1), x_{yo}(r_2)) - \hat{V}(\epsilon_{yo}(r_1), \epsilon_{yo}(r_2)). \qquad (9)$$

In order to reduce sampling error when estimating the intraseasonal covariance, the truncation method described by Zheng and Frederiksen [10] is used. To ensure that the estimated intraseasonal and slow covariance matrices are positive semi-definite, the nearest positive semi-definite matrix is estimated using the method of Higham [5]. Finally, the untruncated intraseasonal and slow components of variability at each location r are estimated using the method of Zheng et al. [11]. This method also assumes stationarity of the monthly statistics. Since stationarity implies that inter-monthly covariances are equal, the equations used at location r are obtained by substituting $r = r_1 = r_2$ into equations (3)–(9) above. The intraseasonal and slow covariance matrices are then scaled such that their diagonal is equal to the respective grid point variability. The total, intraseasonal and slow modes of variability are then obtained from the eigenvalue decomposition [8] of the repsective estimated covariance matrices.

2.2 Comparison between samples

The relationship between the centred root-mean square difference E' between two samples x and x' is

$$E'^2 = \sigma_x^2 + \sigma_{x'}^2 - 2\sigma_x\sigma_{x'}C, \qquad (10)$$

where σ is the sample standard deviation and C is the correlation between x and x'. Taylor [9] noted that equation (10) resembles the law of cosines in trigonometry. This enables graphical representation of the statistical properties of comparisons between patterns from different samples relative to a reference sample. By interpolating fields from all samples onto the same grid as the reference sample, then for the eigenvalue decomposition of the covariance matrices we have the following.

1. Correlation of the eigenvectors, or Empirical Orthogonal Functions (EOFs), gives the pattern correlation between the reference and other sample modes of variability.

2. If the EOFs are normalised to unit length, then the eigenvalues are the variances for each mode of variability. The square root of the eigenvalue is the standard deviation of the mode of variability.

3. The estimates of the standard deviation of the reference and other sample modes of variability will be comparable, since the sum of the eigenvalues is equal to the trace of the matrix.

3 Example

To illustrate the methodology, a comparison was done between reanalysis and GCM data for the 500 hPa geopotential height for the Southern Hemisphere summer (December-January-February) for the second half of the 20th Century. The reanalysis data was obtained from the National Centers for Environmental Prediction (NCEP) National Center for Atmospheric Research dataset [6] for the period 1951–2000. Data from the supplied $2.5° \times 2.5°$ latitude/longitude grid was sub-sampled to $5° \times 5°$, with thinning towards the poles as described by Zheng and Frederiksen [10]. The leading four EOFs for the intraseasonal and slow covariance are shown in Figures 1 and 2. The EOFs explain, collectively, 55% and 83% respectively of the intraseasonal and slow covariance. An analysis of the EOFs, using data from 1949–2002, was given by Frederiksen and Zheng [4]. These EOFs are the reference EOFs for this comparion.

GCM data were obtained from the World Climate Research Programme (WCRP) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset [7]. Data for 68 realisations from 22 GCMs were obtained for the CMIP3 experiment using 20th Century forcings. The CMIP3 realisations
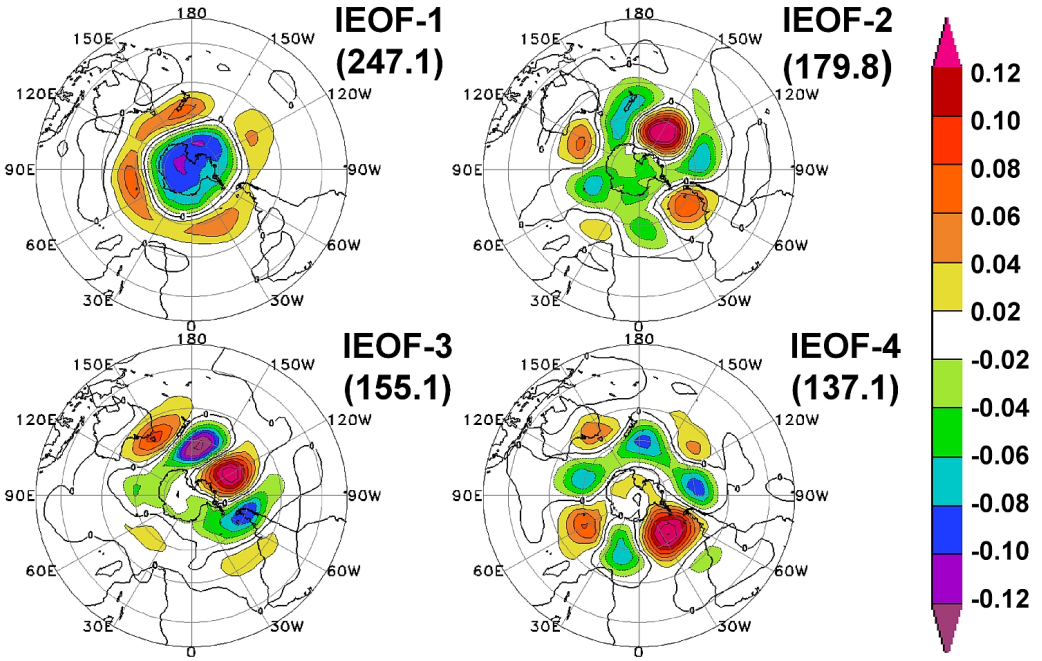
FIGURE 1: Leading four EOFs of the intraseasonal covariance from the NCEP reanalysis for 1951–2000 Southern Hemisphere summer 500 hPa geopotential height. The EOFs are normalised to unit length. The square root of the eigenvalue (units m) is given next to each EOF.
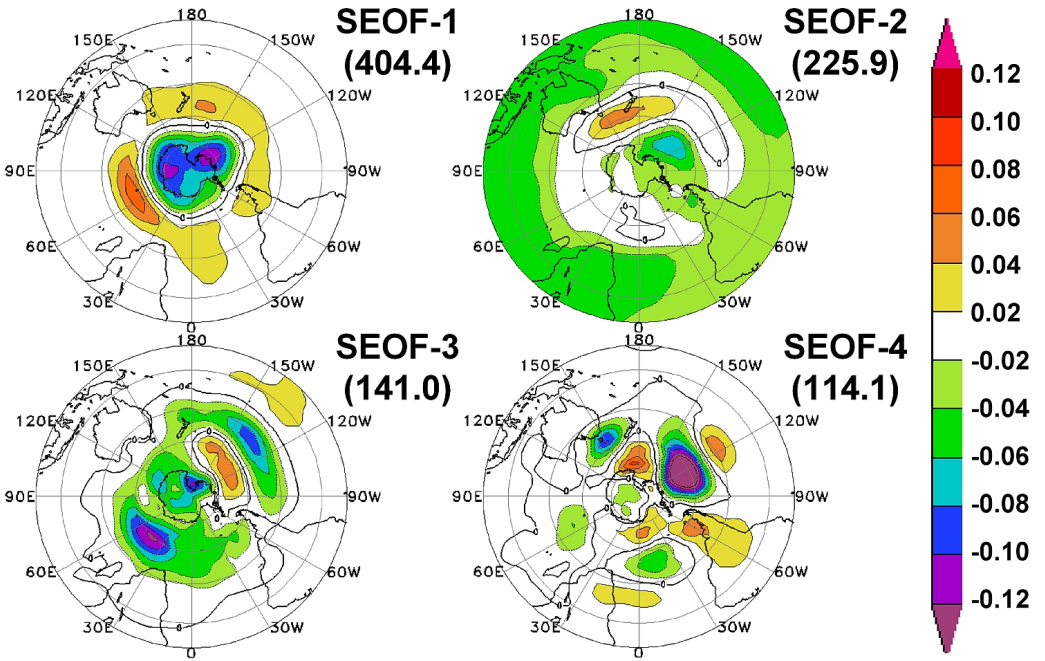
FIGURE 2: Leading four EOFs of the slow covariance from the NCEP reanalysis for 1951–2000 Southern Hemisphere summer 500 hPa geopotential height. The EOFs are normalised to unit length. The square root of the eigenvalue (units m) is given next to each EOF.
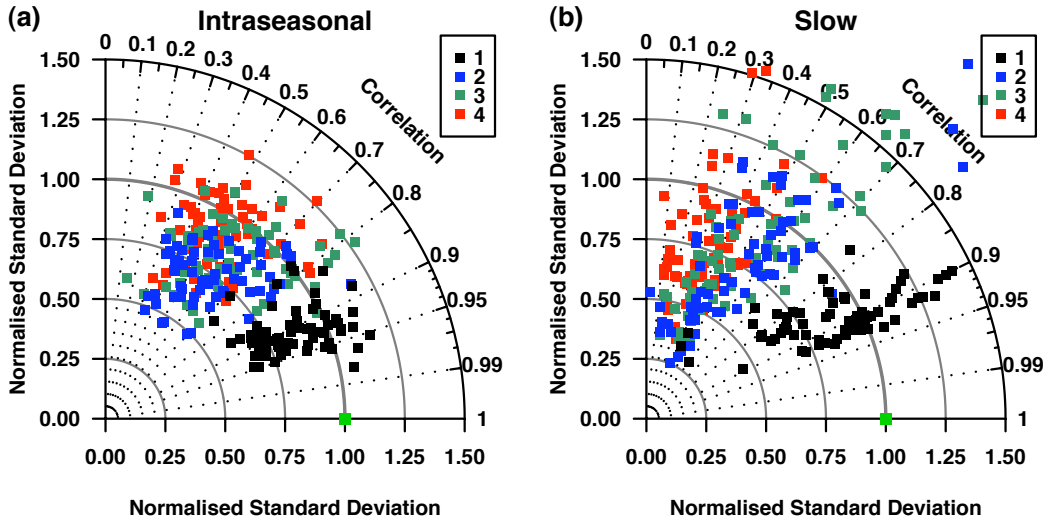
FIGURE 3: Pattern correlation and normalised standard deviation between CMIP3 realisations and reference data for each reference EOF for (a) intraseasonal, and (b) slow covariance. Each point represents the best match from a CMIP3 realisation to the given reference EOF.

were interpolated onto the NCEP reanalysis grid using distance based weights with a finite radius [1]. For each CMIP3 realisation, the highest pattern correlation from the leading six EOFs with the reference modes of variability was obtained. Where two or more EOFs from a CMIP3 realisation had similar pattern correlations with the reference mode, the best match was re-checked by manual inspection of the EOFs.

The results for all 68 CMIP3 realisations for the intraseasonal and slow covariance components are shown in Figure 3 as Taylor diagrams [9]. For each mode of variability, the standard deviations have been normalised by the reference standard deviation. This does not change the pattern correlation, and enables the results from different modes of variability to be shown on the same diagram. The radial distance gives the normalised standard deviation

and the azimuth gives the pattern correlation. For both the intraseasonal and slow covariance, the leading mode (the Southern Annular Mode [4]) is generally well reproduced. For the intraseasonal covariance (Figure 3a), modes 2, 3 and 4 are less well reproduced, although they are present in most CMIP3 realisations in some order. However, for the slow covariance (Figure 3b), modes 2, 3 and 4 are often poorly reproduced, reflected in the low pattern correlation and/or normalised standard deviation.

Figure 3 shows that the CMIP3 realisations generally underestimate the variability associated with all the intraseasonal and slow modes, as indicated by normalised standard deviations of less than 1.0. To check if this is due to the CMIP3 realisations preferring higher modes of variability, the sum of the eigenvalues was calculated for each component. Results (Figure 4) indicate that the eigenvalue sums for both the intraseasonal and slow covariances are generally underestimated in the CMIP3 realisations with respect to the reference data. This appears to extend from the total covariance eigenvalue sum (not shown) also being generally underestimated.

4   Conclusions

We have described how an existing method for estimating the intraseasonal and slow modes of variability can be extended to comparing data from GCM realisations with a reference dataset. By interpolating the GCM data onto the same grid as the reference data, consistent comparisons of the variability of each mode and the pattern correlation are obtained. For realisations from the 20th Century climate scenario of the CMIP3 dataset, the leading mode of the intraseasonal and slow covariance for 500 hPa geopotential height in the Southern Hemisphere summer is generally well reproduced. Modes 2, 3 and 4 are less well reproduced, although the intraseasonal modes are better reproduced than the slow modes. The variability of all modes in the CMIP3 realisations is generally underestimated. This appears to extend from the

(a)

Intraseasonal Covariance Eigenvalue Sum

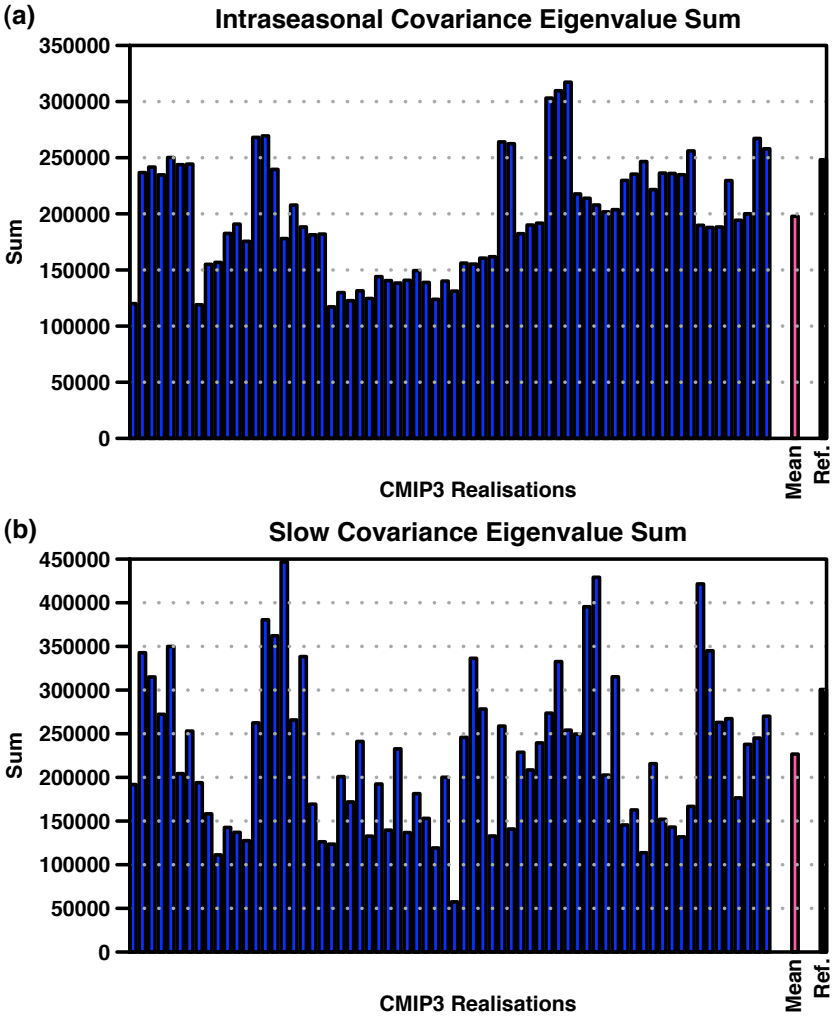**(b)**

Slow Covariance Eigenvalue Sum

FIGURE 4: Eigenvalue sum from the EOF decomposition of the (a) intraseasonal, and (b) slow covariance matrices. Unlabelled bars are individual CMIP3 realisations from the 20th Century experiment. The 'mean' bar denotes the sum averaged over all 68 CMIP3 realisations. The 'Ref.' bar denotes the NCEP reanalysis.

CMIP3 realisations underestimating the reference total covariance eigenvalue sum.

References

[1] G. P. Cressman. An operational objective analysis system. *Mon. Wea. Rev.*, 87:367–374, doi:10.1175/1520-0493(1959)087¡0367:AOOAS¿2.0.CO;2, 1959. C407

[2] C. S. Frederiksen and X. Zheng. Variability of seasonal-mean fields arising from intraseasonal variability. part 2, application to nh winter circulations. *Climate Dynamics*, 23:193–206, doi:10.1007/s00382-004-0429-6, 2004. C401

[3] C. S. Frederiksen and X. Zheng. Coherent Structures of Interannual Variability of the Atmospheric Circulation: The Role of Intraseasonal Variability. Frontiers in Turbulence and Coherent Structures, *World Scientific Lecture Notes in Complex Systems*, Vol. 6, Eds Jim Denier and Jorgen Frederiksen, World Scientific Publications, 87–120, 2007. C400

[4] C. S. Frederiksen and X. Zheng. Variability of seasonal-mean fields arising from intraseasonal variability: Part 3: Application to SH winter

and summer circulations. *Climate Dynamics*, 28:849–866, doi:10.1007/s00382-006-0214-9, 2007. C401, C404, C408

[5] N. Higham. Computing the nearest correlation matrix—a problem from finance. *I.M.A. J. Numerical Analysis*, 22:329–343, doi:10.1093/imanum/22.3.329, 2002. C403

[6] E. Kalnay et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, 77:437–471, doi:10.1175/1520-0477(1996)077¡0437:TNYRP¿2.0.CO;2, 1996. C404

[7] G. A. Meehl et al. The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.*, 88:1383–1394, doi:10.1175/BAMS-88-9-1383, 2007. C404

[8] H. von Storch and F. W. Zwiers. Statistical Analysis in Climate Research. *Cambridge University Press*, 484pp, 1999. C403

[9] K. E. Taylor. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 106:7183–7192, http://www.agu.org/pubs/crossref/2001/2000JD900719.shtml, 2001. C403, C407

[10] X. Zheng and C. S. Frederiksen. Variability of seasonal-mean fields arising from intraseasonal variability. Part 1, methodology. *Climate Dynamics*, 23:177–191, doi:10.1007/s00382-004-0428-7, 2004. C400, C401, C402, C403, C404

[11] X. Zheng, M. Sugi and C. S. Frederiksen. Interannual variability and predictability in an ensemble of climate simulations with the MRI-JMA AGCM. *J. Met. Soc. Jap.*, 82:1–18, doi:10.2151/jmsj.82.1, 2004. C403

Author addresses

1. **S. Grainger**, Centre for Australian Weather and Atmospheric Research, Bureau of Meteorology, GPO Box 1289, Melbourne, Victoria, 3001, AUSTRALIA.
   mailto:s.grainger@bom.gov.au

2. **C. S. Frederiksen**, Centre for Australian Weather and Atmospheric Research, Bureau of Meteorology, GPO Box 1289, Melbourne, Victoria, 3001, AUSTRALIA.

3. **X. Zheng**, National Institute of Water and Atmospheric Research, Wellington, NEW ZEALAND.