Interannual Variability and Predictability in an Ensemble of Climate Simulations with the MRI-JMA AGCM

Xiaogu ZHENG

National Institute of Water and Atmospheric Research, Wellington, New Zealand

Masato SUGI

Meteorological Research Institute, Tsukuba, Japan

and

Carsten S. FREDERIKSEN

Bureau of Meteorology Research Centre, Melbourne, Australia

(Manuscript received 1 June 2001, in final form 11 September 2003)

Abstract

An analysis of variance approach, for systematically studying and evaluating the interannual variability and predictability of seasonal mean fields, is demonstrated using an ensemble of six 50-year simulations of the 500 hPa geopotential height field from the Meteorological Research Institute-Japan Meteorological Agency (MRI-JMA) global atmosphere model forced by observed sea surface temperatures (SSTs). The model performance is analysed, for the seasons June-July-August (JJA) and December-January-February (DJF), and compared with NCEP reanalysis.

The magnitude and geographical distribution of the unpredictable weather noise variability (interannual variability of seasonal mean series arising from intraseasonal variability) are simulated well by the model. The spatial distribution of the potentially predictable variability (the interannual variability of seasonal means from which sampling error due to weather noise variability is removed) is simulated fairly well for DJF, but relatively poorly for JJA.

Despite this, however, several extratropical wintertime patterns, identified as the EOFs of the seasonal mean fields, are well simulated by the model. These include the two patterns forced by the El Niño-Southern Oscillation (ENSO): the Tropical Northern Hemisphere (for DJF), and the Tropical Southern Hemisphere (for JJA) patterns. In addition, the three patterns forced by the tropical Indian Ocean: Western Pacific (for DJF), Meridional Wavetrain (for JJA), and Southern Indian Ocean (for JJA) patterns, are also well simulated.

1. Introduction

Recently, ensembles of multi-decadal simulations using atmospheric general circulation models (AGCMs) forced by observed SSTs have been applied extensively to understand climate variability and the predictability of seasonal forecasts (e.g., Kumar et al. 1996; Zwiers 1996; Sugi et al. 1997; Rowell 1998; Zheng and Frederiksen 1999). For such studies, it is important to assess how well the interannual variability arising from different sources is simulated by the AGCMs. This paper is devoted to an Analy-
sis of Variance study (e.g., Zwiers 1996; Zheng and Frederiksen 1999) of an ensemble of six 50-year simulations of the 500-hPa geopotential heights produced with the Meteorological Research Institute-Japan Meteorological Agency (MRI-JMA) AGCM.

A common assumption (Madden 1976) is that a seasonal mean of meteorological values is composed of two parts, named the potentially predictable component and the weather noise component. The former is the seasonal “population” mean (i.e., derived from the set of all possible outcomes) which is unchanged during the season. Therefore, it is more likely the response to slowly varying external forcings (mainly SSTs, but also other factors, for example, ice coverage, vegetation coverage, greenhouse gas concentration, and sulphate aerosol loading), and the response to slowly varying internal dynamics (e.g., the quasi-biennial oscillation in the stratosphere). This component is potentially predictable at the long-range (at least three months in advance) because the slowly varying forcings, and the slowly varying internal dynamics, may be predictable at the long-range. In contrast, the weather noise component is the departure of the seasonal “sample” mean (i.e., derived from a subset of all possible outcomes) of daily values from the seasonal “population” mean. Thus, it is the response to rapidly varying day-to-day weather events and is potentially long-range unpredictable because daily weather is unpredictable beyond a couple of weeks (Lorenz 1970; Leith 1973). This component should average to zero if sampled over a very large number of realizations, but not necessarily if averaged over a single realization of a single season. The seasonal mean of the weather noise component will have its largest contribution from intra-seasonal variability, since the shorter period weather events will effectively be filtered out by the seasonal mean operator. In mid-latitudes, the weather noise component can explain about 50% of the variance of seasonal mean time series of temperature, geopotential height and wind (Zheng et al. 2000). Although the two components cannot be numerically separated from each other, their variability can be estimated using daily data (Madden 1976) or using monthly data (Zheng et al. 2000). A comparison of the simulated and observed spatial distributions and magnitudes of these two components may provide useful diagnostics to identify deficiencies in the AGCM simulations.

It is convenient to consider the potentially predictable component further as composed of two parts: the SST-forced component and the internal source component (Zwiers 1996; Zheng and Frederiksen 1999; Frederiksen and Zheng 2000). The former is potentially predictable by AGCMs given perfect SST forecasts. The latter is unpredictable by AGCMs, unless one has both a perfect SST forecast and correct initial conditions for every season (Lin 2001). For the multi-decadal AGCM simulations studied in this paper, the initial states for all seasons are significantly different from the truth, so the internal source component is unpredictable by the AGCM. Only the SST-forced component could possibly be predictable, depending on the response of the model to this lower boundary forcing.

In assessing a model’s predictive skill, the correlation between the observed seasonal means, and the simulated ensemble seasonal means, is typically used as an index of the estimated real-world forecast skill for an ensemble of finite size. Recently, Zheng and Frederiksen (1999), using daily data, and Zheng et al. (2000), using monthly data, have provided estimates of two new correlation indices that could provide additional useful information about a model’s predictive skill. The first, the correlation between the observed seasonal mean and the simulated SST-forced component, is an index of the estimated real-world forecast skill for an infinite ensemble of simulations, can be used to assess whether a larger ensemble would give more predictability. The second, the correlation between the simulated SST-forced components and the observed potentially predictable components, provides a lower bound on the model’s skill in simulating SST-forced variability. Together, these indices provide a means of understanding the prediction errors and possibly improving the predictions. Here, the indices are applied to the global monthly data from the model, and the National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR) reanalyses (Kalnay et al. 1996). We shall use the reanalysis data as our proxy for the observed 500 hPa geopotential height.
Another approach to assessing a model’s predictive skill is to consider the principal components of the dominant modes of variability. That is, rather than applying the above analysis of variance (ANOVA) and correlation analysis at each grid-point, the analysis could be done on a reduced set of time series, which represent the dominant global patterns of interannual variance. Venzke et al. (1999) proposed a method for determining patterns of the SST-forced component in ensembles of model simulations. However, such patterns may be subject to errors arising from model deficiencies, and therefore are not used in this paper. Here, we propose to evaluate the model by comparing the principal component time series that are derived as the projection of the model and observation data onto the dominant modes of the observation. These modes may suffer from more sampling error than modes derived from models, for there is only one realization available. However, they are not subject to model error, for they are derived directly from observations. Applying the above ANOVA and skill indices to the resulting time series provides an indication of the skill of the model in reproducing these modes. In this paper, we use NCEP data as the observation, because observations are assimilated into the NCEP model.

The plan of this paper is as follows. In section 2, we derive the formulae for estimating the interannual variability and covariability. The description of the data is in section 3. The application of these formulae to a study of interannual variability and predictability, both for global fields and for the dominant patterns of interannual variability, are in section 4. Our conclusions are in section 5.

2. Formula for variability

In the discussion to follow, it is assumed that the mean annual cycle has been removed by a harmonic equivalent analysis (e.g., Epstein 1991). Any uncertainties introduced by the estimation of the mean annual cycle are ignored.

2.1 Statistical model for monthly mean

We represent a monthly mean of a meteorological variable, within a particular season obtained over a number of years from an ensemble of multi-decadal AGCM simulations starting from different initial states, as (Zwiers 1996)

\[ x_{\text{sym}}(r) = \beta_y(r) + \delta_y(r) + e_{\text{sym}}(r), \] (1)

where, \( s = 1, \ldots, S \) denotes the simulation member, \( y = 1, \ldots, Y \) the year, \( m = 1, 2, \) and 3 the month within a given 3-month season and \( r = 1, \ldots, R \) the location. The terms in the statistical model have the following interpretation: \( \beta_y(r) \) is the externally forced component (e.g., SST-forcing), \( \delta_y(r) \) the internal source component and \( e_{\text{sym}}(r) \) the weather noise component. It should be noted that in this statistical model, SST-forced intraseasonal variability is included in the weather noise component.

For convenience, we introduce the vector notation \( \delta_y = \{ \delta_y(r), r = 1, \ldots, R \} \) for each of the components. Then \( \delta_y \) is assumed to be purely random and statistically independent with respect to \( s \) and \( y \). Also, \( e_{\text{sym}} \) is assumed to comprise a stationary stochastic process with mean zero and is statistically independent and identically distributed with respect to \( s \) and \( y \).

Similarly, an observed monthly mean is represented as (Zheng et al. 2000)

\[ x'_{\text{sym}}(r) = \beta'_y(r) + \delta'_y(r) + e'_{\text{sym}}(r), \] (2)

where \( \beta'_y(r), \delta'_y(r) \) and \( e'_{\text{sym}}(r) \) are analogous to the definitions of \( \beta_y(r), \delta_y(r) \) and \( e_{\text{sym}}(r) \), respectively. Furthermore, \( \delta_y, e_{\text{sym}}, \delta'_y \) and \( e'_{\text{sym}} \) are assumed to be statistically independent of each other (Zwiers 1996).

Following the notation of Zwiers (1996), a “circle” represents a subscript when an average is taken over that index. Thus, for example, the prediction of a seasonal mean (i.e., a simulated ensemble seasonal mean for that season) can be expressed as

\[ x_{\text{oyo}}(r) = \beta_y(r) + \delta_y(r) + e_{\text{oyo}}(r). \] (3)

Similarly, an observed seasonal mean can be expressed as

\[ x'_{\text{oyo}}(r) = \beta'_y(r) + \delta'_y(r) + e'_{\text{oyo}}(r). \] (4)

Following the notation of Zheng and Frederiksen (1999), the symbol \( V \) denotes the true variance of a variable, or the true covariance of two variables, and the symbol \( C \) denotes the true correlation coefficient between two variables. \( V \) and \( C \) represent the corresponding estimated values.
The following assumptions about monthly means of a simulated or observed climate variable (i.e., \( \epsilon_{xym} \) and \( \epsilon'_{xym} \)) are necessary for estimating their variance (Zheng et al. 2000). Since the daily time series of a climate variable within a season is in general assumed to be stationary, so are the monthly statistics. In particular, the variance is assumed to be independent of month; i.e.,

\[
V(\epsilon_{xym}) = V(\epsilon_{xym}) = V(\epsilon_{xym}).
\]

The same is also assumed to be true for the inter-monthly covariance, i.e.,

\[
V(\epsilon_{xym}, \epsilon_{xym}) = V(\epsilon_{xym}, \epsilon_{xym}).
\]

Strictly speaking, these assumptions may not apply well in transition seasons, but they are probably reasonable for winter or summer seasons. Recognizing that day-to-day weather events are unpredictable beyond a week or two, we further assume that the intraseasonal components are uncorrelated if they are a month or more apart, i.e.,

\[
V(\epsilon_{xym}, \epsilon_{xym}) = 0.
\]

2.2 Variance of weather noise component

Under the assumptions contained in Eqs. (5)–(7),

\[
E \left( \begin{array}{c} \epsilon_{xym} \\ \epsilon_{xym} \\ \epsilon_{xym} \end{array} \right) = \sigma^2 \left( \begin{array}{ccc} 1 & \phi & 0 \\ \phi & 1 & \phi \\ 0 & \phi & 1 \end{array} \right),
\]

where \( E \) denotes the expectation value based on all years and simulations,

\[
\sigma^2 = V(\epsilon_{xym}), \quad m = 1, 2, 3, (9)
\]

denotes the variance of \( \epsilon_{xym} \) and

\[
\phi = \frac{1}{\sigma^2} V(\epsilon_{xym}, \epsilon_{xym}) = \frac{1}{\sigma^2} V(\epsilon_{xym}, \epsilon_{xym}),
\]

is the inter-monthly correlation. In addition, using Eqs. (9) and (10), it follows that,

\[
E \left( \begin{array}{c} \epsilon_{xym} - \epsilon_{xym} \\ \epsilon_{xym} - \epsilon_{xym} \\ \epsilon_{xym} - \epsilon_{xym} \end{array} \right) = \sigma^2 \left( \begin{array}{ccc} 2 - 2\phi & 2\phi - 1 \\ 2\phi - 1 & 2 - 2\phi \end{array} \right).
\]

Since, from Eq. (1), \( \epsilon_{xym} - \epsilon_{xym} = x_{xym} - x_{xym} \) and \( \epsilon_{xym} - \epsilon_{xym} = x_{xym} - x_{xym} \) are known, the moment estimation can be applied to derive

\[
\hat{\sigma}^2(2 - 2\phi) = a \quad \text{and} \quad \hat{\sigma}^2(2\phi - 1) = b,
\]

where

\[
a = \frac{1}{2} \frac{1}{\sum_{i=1}^{S}} \sum_{y=1}^{Y} (x_{xym} - x_{xym})^2 \\
+ (x_{xym} - x_{xym})^2,
\]

\[
b = \frac{1}{\sum_{i=1}^{S}} \sum_{y=1}^{Y} ((x_{xym} - x_{xym}) (x_{xym} - x_{xym})).
\]

Solving Eq. (12) for \( \phi \) and \( \sigma^2 \) we can derive the following estimates

\[
\phi = \frac{a + 2b}{2(a + b)},
\]

and

\[
\hat{\sigma}^2 = \frac{a}{2(1 - \phi)}.
\]

These parameters, estimated here using a moment estimation approach, are the same as those derived by Zheng et al. (2000) using maximum likelihood estimation (MLE). Our approach is more general in that it does not assume that the weather noise is normally distributed, which is a requirement of MLE. However, if the weather noise distribution is skewed or heavily tailed, the estimation error using moment estimation will be larger.

In order to reduce the estimation error, we constrain the estimated \( \phi \) to lie within the interval [0, 0.1] based on the assumption that the daily variable is a red noise process with first order autocorrelation coefficient generally between 0 and 0.9 (see the Appendix of Zheng et al. 2000 for a detailed proof). From Eqs (7), (9) and (10),

\[
V(\epsilon_{xym}) = \frac{1}{9} \sum_{m=1}^{3} V(\epsilon_{xym}, \epsilon_{xym}) = \hat{\sigma}^2 (3 + 4\phi)/9.
\]

Since \( \epsilon_{xym} \) is statistically independent of index \( s \), then

\[
V(\epsilon_{xym}) = \frac{1}{S} V(\epsilon_{xym}).
\]

The variance for the weather noise component of the observations, \( V(\epsilon'_{xym}) \), can be estimated by setting \( S = 1 \) and \( \epsilon_{ym} = \epsilon_{ym} \).
2.3 Variance of SST and internal source components

Since $\delta_{xy} + \epsilon_{xoyo}$ is assumed to be statistically independent with respect to index $s$, then

$$V(\delta_{xy} + \epsilon_{xoyo}) = \frac{1}{S} V(\delta_{xy} + \epsilon_{xoyo}).$$

Equations (19) and (20) imply that

$$V(\delta_{xy} + \epsilon_{xoyo}) = \frac{S}{S-1} V(\delta_{xy} + \epsilon_{xoyo} - \delta_{xy} - \epsilon_{xoyo}).$$

From Eq. (1), $\delta_{xy} + \epsilon_{xoyo} - \delta_{xy} - \epsilon_{xoyo} = x_{xoyo} - x_{xoyo}$. Thus, its variance can be estimated as

$$V(\delta_{xy} + \epsilon_{xoyo} - \delta_{xy} - \epsilon_{xoyo}) = \frac{1}{YS} \sum_{y=1}^{Y} \sum_{s=1}^{S} (x_{xoyo} - x_{xoyo})^2.$$  

It follows further from Eqs (21) and (22) that

$$V(\delta_{xy} + \epsilon_{xoyo}) = \frac{1}{YS} \sum_{y=1}^{Y} \sum_{s=1}^{S} (x_{xoyo} - x_{xoyo})^2.$$  

Then, using Eqs (19) and (23) we have

$$V(\delta_{xy} + \epsilon_{xoyo}) = \frac{1}{YS} \sum_{y=1}^{Y} \sum_{s=1}^{S} (x_{xoyo} - x_{xoyo})^2.$$  

From Eq. (1), $x_{xoyo} = \beta_{y} + \delta_{xy} + \epsilon_{xoyo}$, and its variance can be estimated as

$$V(x_{xoyo}) = \frac{1}{Y-1} \sum_{y=1}^{Y} (x_{xoyo} - x_{xoyo})^2.$$  

Assuming $\beta_{y}$ is statistically independent of $\delta_{xy} + \epsilon_{xoyo}$ (Venzke et al. 1999; Zwiers 1996; Zheng and Frederiksen 1999), it follows that

$$\hat{V}(\beta_{y}) = \hat{V}(x_{xoyo}) - \hat{V}(\delta_{xy} + \epsilon_{xoyo}),$$

and

$$\hat{V}(x_{xoyo}) = \hat{V}(\beta_{y}) + \hat{V}(\delta_{xy} + \epsilon_{xoyo}).$$

Furthermore, assuming $\delta_{xy}$ and $\epsilon_{xoyo}$ are statistically independent of each other (Zwiers 1996; Zheng and Frederiksen 1999), then it follows that

$$V(\delta_{xy}) = \hat{V}(\delta_{xy} + \epsilon_{xoyo}) - \hat{V}(\epsilon_{xoyo}),$$

and

$$V(\beta_{y} + \delta_{xy}) = \hat{V}(x_{xoyo}) - \hat{V}(\epsilon_{xoyo}).$$

Since $S = 1$ for observations, Eq. (23) cannot be applied to estimate $\hat{V}(\delta_{xy} + \epsilon_{xoyo})$. Nevertheless, the variance of $x_{xoyo}$ can be directly estimated as

$$\hat{V}(x_{xoyo}) = \frac{1}{Y-1} \sum_{y=1}^{Y} (x'_{xoyo} - x''_{xoyo})^2,$$

and the variance of $(\beta'_{y} + \delta'_{y})$ can then be estimated as

$$\hat{V}(\beta'_{y} + \delta'_{y}) = \hat{V}(x'_{xoyo}) - \hat{V}(\epsilon_{xoyo}).$$

2.4 Correlations

Since the boundary forcing is the only possible source of the covariability between the AGCM simulated and observed seasonal mean field, that is,

$$V(x_{xoyo}, x'_{xoyo}) = V(\beta_{y}, \beta'_{y}),$$

then it can be assumed that,

$$V(x_{xoyo}, x'_{xoyo}) = V(\beta_{y}, x'_{xoyo}) = V(\beta_{y}, \beta'_{y} + \delta'_{y}) = V(\beta_{y}, \beta'_{y}).$$

It follows from the definition of the correlation between two fields that

$$\hat{C}(x_{xoyo}, x'_{xoyo}) = \frac{\hat{V}(x_{xoyo}, x'_{xoyo})}{\sqrt{\hat{V}(x_{xoyo}) \hat{V}(x'_{xoyo})}},$$

$$\hat{C}(\beta_{y}, x'_{xoyo}) = \frac{\hat{V}(\beta_{y}, x'_{xoyo})}{\sqrt{\hat{V}(\beta_{y}) \hat{V}(x'_{xoyo})}},$$

$$\hat{C}(\beta_{y}, \beta'_{y} + \delta'_{y}) = \frac{\hat{V}(\beta_{y}, \beta'_{y} + \delta'_{y})}{\sqrt{\hat{V}(\beta_{y}) \hat{V}(\beta'_{y} + \delta'_{y})}}.$$  

Zheng and Frederiksen (1999) show that $\hat{C}(x_{xoyo}, x'_{xoyo})$ represents the estimated real world forecast skill for a finite number of simulations. As the number of simulation increases towards infinity, $x_{xoyo}$ tends towards $\beta_{y}$ by the law of large number. Therefore, $\hat{C}(\beta_{y}, x'_{xoyo})$ is an estimate of the real-world forecast skill for infinite simulations. Since $\hat{V}(x_{xoyo})$ is greater than $\hat{V}(\beta_{y})$, $\hat{C}(x_{xoyo}, x'_{xoyo})$ is less than $\hat{C}(\beta_{y}, x'_{xoyo})$. A large difference between $\hat{C}(x_{xoyo}, x'_{xoyo})$ and $\hat{C}(\beta_{y}, x'_{xoyo})$ would therefore imply a greater potential to improve the forecast skill by increasing the ensemble size.
An interesting question is how many simulations \((S)\) are required for the estimated real-world forecast skill \(C(x_{\text{oyo}}, x'_{\text{yo}})\) to reach a value \(p < C(\beta_y, x'_{\text{yo}})\) \(\) from Eqs (34)–(35),

\[
p = C(x_{\text{oyo}}, x'_{\text{yo}}) = C(\beta_y, x'_{\text{yo}}) + \sqrt{V(\beta_y)/V(x_{\text{oyo}})}
\]

\[
= C(\beta_y, x'_{\text{yo}}) + \sqrt{V(\beta_y)/(V(\beta_y) + (V(\delta_x) + V(\epsilon_{\text{oyo}}))/S)}.
\]  

(37)

If the number of simulation years \((Y)\) is large enough, then all variances and correlations in Eq. (37) would be close to the true values. Solving Eq. (37) for \(S\) and substituting the true values by the estimated values we obtain

\[
S = \frac{V(\delta_x + \epsilon_{\text{oyo}})p^2}{V(\beta_y)(C^2(\beta_y, x'_{\text{yo}}) - p^2)}.
\]  

(38)

In particular, the estimated \(S\) is infinite if \(p = C(\beta_y, x'_{\text{yo}})\).

Finally, \(\hat{C}(\beta_y, \beta'_y)\), which can be defined as

\[
\hat{C}(\beta_y, \beta'_y) = \hat{V}(x_{\text{oyo}}, x'_{\text{yo}})/\sqrt{V(\beta_y)\hat{V}(\beta'_y)}.
\]  

(39)

represents the skill in simulating the externally forced variability. A value of \(\hat{C}(\beta_y, \beta'_y)\) close to unity would indicate there is little error in simulating the externally forced variability. Unfortunately, \(\hat{C}(\beta_y, \beta'_y)\) is difficult to estimate because \(\hat{V}(\beta'_y)\) cannot be estimated from a single realization. However, since \(\hat{V}(\beta'_y + \delta'_y)\) is greater than \(\hat{V}(\beta'_y)\), then \(\hat{C}(\beta_y, \beta'_y + \delta'_y)\) is less than \(\hat{C}(\beta_y, \beta'_y)\). Therefore, \(\hat{C}(\beta_y, \beta'_y + \delta'_y)\) does provide a lower bound on \(\hat{C}(\beta_y, \beta'_y)\).

2.5 Caveats

When using the formulae for the variance of each component and the covariance between two components, the following caveats should be considered.

Firstly, for the formulae related to the externally forced components \(\beta_y\) and \(\beta'_y\), \(Y\) should be large in order to limit the sample error. For the formulae not related to the externally forced component, such as \(\hat{V}(\epsilon_{\text{oyo}})\) and \(\hat{V}(\delta_x + \epsilon_{\text{oyo}})\), the sampling error can be limited by increasing the number of ensemble members \(S\). The decreasing rate of sampling error is of the order of \(1/S\).

Secondly, the smaller the potential predictability (i.e., \(\hat{V}(\beta_y)/\hat{V}(x_{\text{oyo}})\)), the more likely the correlation \(\hat{C}(\beta_y, x'_{\text{yo}})\) will be distorted. This can be understood as follows. Suppose \(\xi\) is the estimation error of \(\hat{V}(\beta_y)/\hat{V}(x_{\text{oyo}})\). Then, from Eqs (34) and (35),

\[
\hat{C}(\beta_y, x'_{\text{yo}}) = \hat{C}(x_{\text{oyo}}, x'_{\text{yo}})/\sqrt{V(\beta_y)/V(x_{\text{oyo}})} + \xi.
\]  

(40)

Since the gradient of the function \(1/\sqrt{x}\) is sharper near zero, \(\xi\) has a more significant impact on \(\hat{C}(\beta_y, x'_{\text{yo}})\) for smaller \(V(\beta_y)/V(x_{\text{oyo}})\). From Eqs (34) and (36),

\[
\hat{C}(\beta_y, \beta'_y + \delta'_y) = \hat{C}(x_{\text{oyo}}, x'_{\text{yo}})/\sqrt{V(\beta_y)/V(x_{\text{oyo}})\hat{V}(\beta'_y + \delta'_y)/\hat{V}(x'_{\text{yo}})}.
\]  

Similarly if \(\hat{V}(\beta_y)/\hat{V}(x_{\text{oyo}})\) or \(\hat{V}(\beta'_y + \delta'_y)/\hat{V}(x'_{\text{yo}})\) is small, \(\hat{C}(\beta_y, \beta'_y + \delta'_y)\) is likely to be a larger error.

Finally, some assumptions, such as the decomposition of the data implied in Eqs (1) and (2), and the independence between the different components, may not exactly fit the meteorological data. This could introduce statistical model errors into the estimates.

3. Data

In this study, simulated and observed monthly mean fields of 500-hPa height are used. The simulated fields are taken from an ensemble of 50-year long simulations with a T42 version of the MRI-JMA98 atmospheric model (Shibata et al. 1999). The model was developed at MRI based on JMA-GSM9603, Japan Meteorological Agency global model (JMA 1997). The model has 30 unevenly spaced levels, the top of which is at 0.4-hPa. The model includes comprehensive physical processes: radiation, cumulus convection, large-scale condensation, planetary boundary layer processes, simplified biosphere, and orographic gravity wave drag.

The model has been integrated for a 50-year period from 1 January 1949 to 31 December 1998. For the lower boundary condition over the ocean, the observed SST analyzed at Hadley Centre, UK Meteorological Office (HadISST, Rayner et al. 2000) was used, and is the main source of external forcing. The 50-year runs
were repeated six times using the same SSTs, but with different initial conditions. The initial conditions of the atmosphere, and the land surface of the six runs, are obtained from the computed states for January 1 of the different years in the preliminary ten-year integration of the same model with climatological SST.

The monthly-mean global fields of the 500-hPa height from the NCEP/NCAR reanalyses (Kalnay et al. 1996), over the period January 1949 to December 1998, are taken as our proxy observed data. The data were sub-sampled on a 5° × 5° latitude/longitude grid. Artificial climate “jumps” are minimized in the reanalyses by the use of a fixed assimilation/modeling system, but trends related to changes in observing systems may still be present (e.g., Kidson 1999).

The SSTs used in the correlation analysis are the tropical SSTs (20°S–20°N) from the Hadley Centre global ice and sea surface temperature data set (HADISST1.1) sub-sampled on a 4° × 4° latitude/longitude grid, over the period June 1958 to February 1996. The extra-tropical SSTs have been excluded from our analysis because they are generally forced by the extratropical atmospheric circulation and may not be independent of the tropical SSTs (see, for example, Lau and Nath 1994, 1996). Including them in the correlation analysis may overestimate the SST forced variability.

For convenience, the tropical SSTs are partitioned into three regions: Pacific (for the Pacific Ocean), Atlantic (for the Atlantic Ocean) and Indian (for the Indian Ocean and the sea surrounding Indonesia). Principal component analysis is applied to the seasonal mean SST field for each region. Three dominant SST EOFs and PCs are derived for the Pacific and Indian regions and two dominant EOFs and PCs for the Atlantic region in seasons December-January-February (DJF) and June-July-August (JJA). They are shown in Fig. 1.

4. Results

4.1 Global interannual variability

Here, we examine how well the spatial distribution and magnitude of interannual variability are simulated. The estimated weather noise variability $V(\varepsilon_{\text{sys}})$, the potentially predictable variability $V(\beta_y + \delta_y)$ and the potential predictability $V(\beta_y + \delta_y)/V(\varepsilon_{\text{sys}})$ are shown in Fig. 2 for the NCEP/NCAR 500-hPa heights for austral and boreal winter. The corresponding terms $[V(\varepsilon_{\text{sys}}), V(\beta_y + \delta_y)$ and $V(\beta_y + \delta_y)/V(\varepsilon_{\text{sys}})]$ for the simulations are shown in Fig. 3. The area-weighted spatial correlation coefficients between the observed and the simulated variability, and the area-weighted ratios $\sum_{r=1}^{R} w(r) \sqrt{V(\varepsilon_{\text{sys}}(r))}/\sum_{r=1}^{R} w(r)$, $\sqrt{V(\beta_y(\varepsilon_{\text{sys}}(r)) + \delta_y(r))}/\sum_{r=1}^{R} w(r) \sqrt{V(\varepsilon_{\text{sys}}(r))}$, are listed in Table 1.

The weather noise variability is well simulated. The spatial distribution of weather noise variability estimated from the reanalysis data (leftmost panels of Fig. 2) is very similar to that estimated from the simulated data (leftmost of Fig. 3), and this is indicated by spatial area weighted correlation coefficients that are very close to unity (Table 1). The magnitudes of the weather noise variability estimated from reanalysis data (leftmost panels of Fig. 2) are also similar to those estimated from simulated data (leftmost panels of Fig. 3), and this is indicated by the area-weighted ratios $\sum_{r=1}^{R} w(r) \sqrt{V(\varepsilon_{\text{sys}}(r))}/\sum_{r=1}^{R} w(r) \sqrt{V(\varepsilon_{\text{sys}}(r))}$ that are also close to unity (Table 1).

The potentially predictable variability is simulated fairly well in DJF, but not so well in JJA. In DJF, the two spatial distributions (bottom central panels of Figs. 2 and 3) are more similar than those in JJA (top central panels of Figs. 2 and 3). In fact, the spatial area weighted correlation between the observed and simulated in DJF still indicates reasonable skill with values around 0.6 and 0.7 (Table 1), while the skill in JJA is relatively lower with values of about 0.4 which is still statistically significant. The magnitudes of potentially predictable variability for both seasons are underestimated, especially for JJA. This is indicated by the area weighted ratios of the observed over the simulated potentially predictable variability, which are about 1.20 in DJF, but are around 1.4 and 1.6 in JJA.

The rightmost panel of Fig. 2 shows the potential predictability of the NCEP data. There is relatively large potential predictability over the southern part of the Eurasian Continent, in the vicinity of the Tibetan Plateau in both seasons, but particularly in DJF, with potential
Fig. 1. Dominant EOFs of interannual variation (1958–1995) in DJF and JJA tropical (20°S–20°N) SSTs in the Indian (“ind”), Pacific (“pac”) and Atlantic (“atl”) sectors. The amount of variance explained appears in parenthesis.
predictability as high as 0.9. There is also a region of very high potential predictability in the southern Indian Ocean. These two regions form a somewhat symmetric pattern about the equator and could suggest some form of Indian Ocean SST-forcing. However, this pattern of potential predictability is not seen in the model results (rightmost panel of Fig. 3), except slightly during DJF. Furthermore, this is not seen at higher levels. For example, the potential predictability of the 200 hPa NCEP-NCAR geopotential height field (not shown) is much more zonal in these regions. Whether, over the Eurasian region, this is an artefact of the NCEP-NCAR reanalysis in the region of very high topography, or whether it is due to Indian Ocean SST-forcing associated with the Asian monsoon, is something the authors hope to

<table>
<thead>
<tr>
<th></th>
<th>JJA SH</th>
<th>JJA NH</th>
<th>DJF SH</th>
<th>DJF NH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial area-weighted correlation between $\sqrt{V(y_{x,y})}$ and $\sqrt{V(x_{y,y})}$</td>
<td>0.97</td>
<td>0.96</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>$\sum_{r=1}^{R} w(r) \sqrt{V(y_{x,y}(r))} / \sum_{r=1}^{R} w(r) \sqrt{V(x_{y,x}(r))}$</td>
<td>0.96</td>
<td>1.04</td>
<td>1.00</td>
<td>1.08</td>
</tr>
<tr>
<td>Spatial area-weighted correlation between $\sqrt{V(\beta_{y} + \delta_{y})}$ and $\sqrt{V(\beta_{x} + \delta_{x})}$</td>
<td>0.40</td>
<td>0.41</td>
<td>0.59</td>
<td>0.68</td>
</tr>
<tr>
<td>$\sum_{r=1}^{R} w(r) \sqrt{V(\beta_{y}^{x}(r) + \delta_{y}^{x}(r))} / \sum_{r=1}^{R} w(r) \sqrt{V(\beta_{y} + \delta_{x}(r))}$</td>
<td>1.54</td>
<td>1.40</td>
<td>1.22</td>
<td>1.24</td>
</tr>
</tbody>
</table>

Fig. 2. Variance decomposition for NCEP/NCAR 500-hPa geopotential heights. The top row applies to northern summer (JJA) and the bottom row to winter (DJF). The leftmost column shows the standard deviation of $x_{y,y}$. The central column shows the standard deviation of $\beta_{y} + \delta_{y}$ and the rightmost column shows the potential predictability $V(\beta_{y}^{x} + \delta_{y}^{x})/V(x_{y,y})$. Units for the left and central columns are meters. The 5 hPa contour is dashed. Solid contours are drawn from 10 to 50 m at every 5 m. Values in the right column are non-dimensional and are from 0.1 to 0.9 at every 0.1, where 0.3 is for 10% significance, 0.4 is for 3% significance and 0.5 is for more than 0.5% significance.
pursue in a future study. Here, our interest is mainly to evaluate the skill of the MRI-JMA model.

The simulated internal dynamics variability \( V(\delta_y) \) is shown in Fig. 4 (leftmost panel), and the proportion of the simulated externally forced variability in the simulated total potentially predictable variability \( V(\beta_x + \delta_y) \) is also shown in Fig. 4 (middle column). It is generally significant in the tropics and insignificant in the extratropics. It appears that a significant part of the simulated potentially predictable variability outside the 30S to 30N latitude-band is due to low frequency internal dynamics variability. As will be shown in subsection 4.3, this variability in the internal dynamics is largely due to the North Atlantic Oscillation (NAO) and the Scandinavian (SCA) teleconnection pattern in the NH and high latitude mode (HLM) in the SH. The importance of internal dynamics is further borne out when one considers the simulated SST-forced potential predictability \( V(\beta_x)/V(x_{sys}) \) (rightmost panel of Fig. 4). \( V(\beta_x)/V(x_{sys}) \) is also generally significant in the tropics and insignificant in the extratropics, except for DJF in the Southern Hemisphere.

4.2 Global predictability

The grid-pointwise estimates of the real-world forecast skill for the six simulations \( C(x_{sys},x_{sys}) \), the real-world forecast skill for an infinite number of simulations \( C(\beta_x,x_{sys}) \), and the lower bound of the skill in simulating the externally forced variability \( C(\beta_x,\beta_x + \delta_y) \) are shown in Fig. 5. The estimated real-world forecast skill for the six simulations \( C(x_{sys},x_{sys}) \) is only highly significant \( [C(x_{sys},x_{sys}) > 0.6] \) in the tropics, and is more significant for DJF. The externally forced variability is simulated well in the tropics \( [C(\beta_x,\beta_x + \delta_y) > 0.6] \), and also in some extratropical areas, such as in the Australian and New Zealand region. However, further diagnosis is required to confirm whether the significant correlation \( C(\beta_x,\beta_x + \delta_y) \) is physically robust, or due to statistical errors.

An important diagnostic for the significance of the correlation is to examine the observed potential predictability \( V(\beta_x + \delta_y)/V(x_{sys}) \) (rightmost panels of Fig. 2), and the simulated SST-
Fig. 4. The top row applies to northern summer (JJA) and the bottom row to winter (DJF). The leftmost column shows the simulated internal source estimates (the standard deviation of $\delta_y$). The central column shows the proportion of SST-forced variability in potentially predictable variability, i.e., $\frac{V(\beta_y)}{V(\beta_y + \delta_y)}$. The rightmost column shows the simulated externally forced potential predictability, i.e., $\frac{V(\beta_y)}{V(x_{Yo})}$. Unit for the left is hPa. The solid contours are drawn at 10, 20, 30 and 40, and the dashed contours are drawn at 5 and 2. Values in the central and right columns are non-dimensional. The solid contours are drawn at 0.5, 0.7 and 0.9 and the dashed contours are drawn at 0.4, 0.3 and 0.2.

Fig. 5. The leftmost column shows the estimated real-world forecast skill of six simulations given perfect SST prediction, i.e., $C(x_{Yo}, x_{Yo})$. The central column shows the real-world forecast skill for infinite simulations, i.e., $C(\beta_y, x_{Yo})$. The rightmost column shows the model's performance in simulating the forced component, i.e., $C(\beta_y, \beta_y + \delta_y)$. Top row is for JJA, bottom row for DJF. Units are non-dimensional. The solid contours are drawn at 0.6, 0.8 and 0.9. The dashed contours are drawn at 0.4, 0.3 and 0.2.
forced predictability $\hat{V}(\beta_y)\tilde{V}(x_{yc0})$ (rightmost panels of Fig. 3). It has been shown in subsection 2.3 that the closer to unity the indexes are, the more reliable the estimated correlation $C(\beta_y, \beta'_x + \delta'_y)$. Otherwise, the correlation may be distorted (Zheng and Frederiksen 1999). For this reason, correlation coefficients $C(\beta_y, x_{yc0})$ and $C(\beta_y, \beta'_x + \delta'_y)$ are set to zero if either of the two ratios is less than 0.1.

4.3 Extra-tropical winter circulation patterns

Extra-tropical winter circulations are more extensively studied than summer patterns. Northern Hemisphere winter teleconnection patterns, and modes of interannual variation of geopotential height, have been identified and studied over many decades (for example, Wallace and Gutzler 1981; Barnston and Livezey 1987). There is strong evidence for the existence of a number of important teleconnections. These include the Pacific North American (PNA) pattern, the North Atlantic Oscillation/Arctic Oscillation (NAO/AO), the Western Pacific (WP) pattern, the East Atlantic (EA) pattern, the East Atlantic/Western Russia (EA/WR) pattern, the Eurasian (EU) pattern, the Scandinavian (SCA) pattern and the Tropical Northern Hemisphere (TNH) pattern. Previous observational studies of Southern Hemisphere winter teleconnections in the height field (for example, Kidson 1999; Karoly et al. 1989; Kiladis and Mo 1998) have commonly found four dominant patterns of inter-annual variation. These are the Southern Hemisphere high latitude mode (HLM) (or Antarctic Oscillation AAO), the Tropical Southern Hemisphere (TSH) pattern (or ENSO pattern), the Southern Hemisphere meridional wavetrain (MW) pattern (see, for example, Figs. 5(a)–(c) of Karoly et al. 1989) and the Pacific/South American (PSA) pattern.

In the current study, the circulation patterns are derived by varimax rotation of all significant EOFs of the seasonal mean field for the NH extratropics (20–90°N) and the SH extratropics (90–20°S) derived from 1958–1995 (matching our SST data) NCEP/NCAR 500-hPa height. Here, Kaiser’s rule of thumb (Wilks 1995) is used to determine how many significant EOFs should be rotated. The top eight significant EOFs are rotated for both Hemispheres. The rotated patterns are shown in Fig. 6 for NH and Fig. 7 for SH. Their explained variance is also shown in the figures. For NH, the patterns 1–8 resemble the PNA, NAO, WP, EU, TNH, EA/WR, EA and SCA patterns. For SH, the patterns 1, 2, 3 and 5 are close to the HLM, MW, TSH and PSA patterns. We will refer to pattern 4 as the Southern Indian Ocean (SIO) pattern. The reason will be explained later in this section.

To identify the relation with SST for each pattern, the correlation between the principal component (PC) series of each circulation pattern, and the PC series of each SST pattern during 1958–1995 (Fig. 1), are listed in Table 2 for boreal winter (DJF) and Table 3 for austral winter (JJA). Since our focus is on interannual variability, linear trends are removed from the seasonal mean PC time series.

For the Northern Hemisphere, the TNH is correlated with Pacific Ocean ENSO SST pattern 1 (with correlation $-0.59$) and the WP with Indian Ocean SST pattern 1 (with correlation $+0.47$). Furthermore, the EA and EA/WR patterns are also associated with Pacific Ocean SST pattern 2 and Indian Ocean SST pattern 1 (with correlation $0.43$). For the Southern hemisphere, the TSH pattern is associated with ENSO with a correlation of $0.57$ with Pacific Ocean SST pattern 1. The MW pattern is correlated with Indian Ocean SST pattern 2 (with correlation $-0.42$).

Height pattern 4, in the Southern Hemisphere, is particularly interesting in that it has a significant trend in its PC time series. The PC time series for Indian Ocean SST pattern 1, during JJA, also displays a significant trend, and appears to be related to a warming of the equatorial Indian Ocean during the latter half of the twentieth century. Both PC time series are significantly (at 1% level) negatively correlated (with correlation $-0.55$). At an opposite phase to that shown in Fig. 6, height pattern 4, with an anomalous high/low pair in the mid/high latitudes of the southern Indian Ocean, is consistent with an implied low pressure anomaly over the equatorial Indian Ocean, associated with anomalous SST warming (see Fig. 1). This suggests a tropical/extra-tropical teleconnection associated with warming tropical Indian Ocean SSTs. For this reason, we shall refer to height pattern 4 as the Southern Indian Ocean (SIO) pattern. If the linear trend
Fig. 6. Eight Rotated EOFs for the NH. Percentage variance explained is in brackets.
Fig. 7. Eight Rotated EOFs for the SH. Percentage variance explained is in brackets.
is removed from both PC time series, there is no significant correlation. In that case, height pattern 4 is only weakly correlated with Indian Ocean SST pattern 2 (with correlation +0.32, significant at 5% level).

Some important patterns, such as the NAO, HLM and SCA, are not significantly correlated with any tropical SST pattern. Therefore, the potential predictability of NAO, HLM and SCA patterns (see the third rows of Tables 4 and 5) is likely largely due to low frequency internal dynamics.

The modeled principal component (PC) time series are defined here as the projections of the modeled data onto the EOFs of winter NCEP-NCAR 500 hPa geopotential height for NH and SH. Again, since our focus is on interannual variability, linear trends are removed from the seasonal mean PC time series. Comparing these time series with the NCEP PCs will provide us with some indication of how well these dominant modes of observed variability are simulated, given a perfect SST forecast, and consequently a deeper understanding of the predictability inherent in the MRI-JMA AGCM. We do this by using the analysis techniques of the previous two subsections, applied to the two PC time series.

The correlation $\hat{C}(x_{yo}, x'_{yo})$ for the corresponding PCs of each pattern are shown in Tables 4–5 for NH and SH, respectively. The correlation for WP, TNH, MW and SIO are significant at 1% level (0.35) and TSH is also significant at a 5% level (0.28). The observed and simulated time series of these patterns are shown in Fig. 8. For other patterns, the correlations are not significant at a 5% level.

The potential predictability of the simulated principal components $\hat{C}(x_{yo}, x'_{yo})$ is also listed in Tables 4 and 5. They are greater than 0.45 for all patterns with significant correlation $\hat{C}(x_{yo}, x'_{yo})$.

We have shown by correlation analysis with tropical SSTS that the WP, MW and SIO patterns are related to variability of Indian Ocean SSTs. This suggests that the temporal variability of the WP and SIO patterns are simulated well by JMA-MRI model. It is also well-known that the TNH and TSH patterns are

<table>
<thead>
<tr>
<th>Patterns</th>
<th>Most significant Correlation with SST PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1 (PNA)</td>
<td>$+0.43$ with ind.3</td>
</tr>
<tr>
<td>PC2 (NAO)</td>
<td>$-0.24$ with pac.2</td>
</tr>
<tr>
<td>PC3 (WP)</td>
<td>$+0.47$ with ind.1</td>
</tr>
<tr>
<td>PC4 (EU)</td>
<td>$-0.34$ with pac.2</td>
</tr>
<tr>
<td>PC5 (TNH)</td>
<td>$-0.59$ with pac.1</td>
</tr>
<tr>
<td>PC6 (EA/WR)</td>
<td>$+0.43$ with ind.1</td>
</tr>
<tr>
<td>PC7 (EA)</td>
<td>$+0.43$ with pac.2</td>
</tr>
<tr>
<td>PC8 (SCA)</td>
<td>$+0.22$ with ind.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patterns</th>
<th>Most significant Correlation with SST PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1 (HLM)</td>
<td>$-0.12$ with atl.2</td>
</tr>
<tr>
<td>PC2 (MW)</td>
<td>$-0.42$ with ind.2</td>
</tr>
<tr>
<td>PC3 (TSH)</td>
<td>$+0.57$ with pac.1</td>
</tr>
<tr>
<td>PC4 (SIO)</td>
<td>$+0.32$ with ind.2</td>
</tr>
<tr>
<td>PC5 (PSA)</td>
<td>$+0.31$ with pac.1</td>
</tr>
<tr>
<td>PC6</td>
<td>$-0.14$ with pac.3</td>
</tr>
<tr>
<td>PC7</td>
<td>$-0.34$ with ind.3</td>
</tr>
<tr>
<td>PC8</td>
<td>$-0.34$ with ind.2</td>
</tr>
</tbody>
</table>
The JMA-MRI model also simulates their temporal variability well. The corresponding correlation coefficients between the simulated externally forced component and the observed seasonal means \( \mathcal{C}(\beta_y, x_{yo}) \) is about a quarter more than the \( \mathcal{C}(\gamma_{yo}, x_{yo}) \) value. This suggests that a larger ensemble would improve the skill of the model in simulating these modes. For example, the real world forecast skill of the six simulations \( \mathcal{C}(\gamma_{yo}, x_{yo}) \) is 0.29 for the TSH pattern. Using Eq. (36), we estimate that it could be 0.35 if the simulation number \( S \) increases to 9.

### 5. Summary and conclusions

Interannual variability and predictability of a seasonal mean field have been studied using an ensemble of six 50-year 500 hPa geopotential height fields simulated with the Meteorological Research Institute-Japan Meteorological Agency global atmosphere model. Both the geographical distribution and the magnitude of the unpredictable weather noise variability are well simulated by the model. The spatial distribution of the potentially predictable variability in DJF is simulated fairly well.
well, but relatively poorly in JJA. The magnitude of the potentially predictable variability is underestimated, especially in JJA.

The potential predictability is generally high in the tropics (30°S to 30°N), and most of the potentially predictable variance in the tropics is externally (SST) forced variability. Therefore, it might be expected that the interannual variability in the tropics should be well simulated by an atmospheric general circulation model forced with observed SSTs. The results show that this is indeed the case. On the other hand, the potential predictability is generally low in the extratropics, and a considerable part of that variability is low frequency (interannual/supra-annual) internal source variability. As a result, the interannual variability in the extratropics is generally not well simulated by the model. However, it should be noted that the model does have high skill in simulating the potentially predictable variability in some areas in the extratropics, such as the Australian–New Zealand region. This suggests that some regional patterns of variability may be more predictable than other patterns in some seasons.

A comparison of observed and modeled principal component time series showed that the model skillfully reproduced the five observed patterns of wintertime interannual variability that are strongly forced by tropical SST. That is, the Western Pacific and Tropical Northern Hemisphere patterns in the NH during DJF, and the Tropical Southern Hemisphere, the South Indian Ocean and the Southern Meridional Wavetrain patterns in the SH during JJA. For these modes the externally forced component is well simulated. The temporal variability of the Western Pacific and South Indian Ocean patterns is well simulated in the MRI-JMA model. The success of the model in simulating the Western Pacific, Meridional wavetrain and Tropical Southern Hemisphere patterns indicates the potential of the MRI-JMA AGCM in seasonal forecasting for Japan, Australia and New Zealand.

**Acknowledgments**

This study was initiated during a visit by one of the authors (XZ) to the Institute of Statistical Mathematics (ISM) in Tokyo under the support of a Visiting Fellowship offered by the Japanese Ministry of Education, Science, Sports and Culture. XZ thanks Prof. T. Ozaki of ISM for his hospitality. The study has been continued under the support of the Foundation for Research Science and Technology of New Zealand under contract CO1X0202. The reanalysis data we used were obtained from the NCAR Data Support Section. We also thank the two anonymous reviewers and Drs Brett Mullan and James Ranwick for helpful comments.
References


Epstein, E.S., 1991: Determining the optimum number of harmonics to represent normals based on multiyear data. *J. Climate*, 4, 1047–1051.


JMA, 1997: Outline of operational numerical weather prediction at Japan Meteorological Agency. *NPD/JMA*.


