Interannual Variability of Precipitation in an Ensemble of AMIP Climate Simulations Conducted with the CCC GCM2

XIAOLAN L. WANG* AND FRANCIS W. ZWIERS*

Canadian Centre for Climate Modelling and Analysis, Atmospheric Environment Service, Victoria, British Columbia, Canada

(Manuscript received 24 March 1997, in final form 27 March 1998)

ABSTRACT

In this paper log–linear analysis and analysis of variance methods were used to analyze the interannual variability and potential predictability of precipitation as simulated in an ensemble of six 10-yr Atmospheric Model Intercomparison Project climate simulations conducted with CCC GCM2, the second-generation general circulation model of the Canadian Centre for Climate Modelling and Analysis. Since observed 1979–88 sea surface temperatures (SSTs) and sea ice extent were prescribed as lower boundary conditions in all six simulations, it is possible to diagnose the extent to which the variability of the seasonal frequency, seasonal mean intensity, and seasonal total of precipitation is affected by the prescribed boundary conditions. The specified SST–sea ice forcing was found to significantly affect both the frequency and intensity of precipitation, particularly in the Tropics, but also in the temperate latitudes. Precipitation frequency appears to be more sensitive to the external forcing than precipitation intensity, especially over land areas. Potential predictability from internal sources such as land surface variations is generally small.

1. Introduction

The observed interannual variability of seasonal mean quantities arises from a number of sources. Daily weather variations, which are due primarily to the internal dynamics of the atmosphere, cannot be predicted in detail more than about two weeks in advance. These short timescale fluctuations induce unpredictable variability in seasonal mean atmospheric quantities, which is usually referred to as ''climate noise'' (or weather noise; Madden 1976). Slowly varying ''external'' boundary conditions, such as anomalies of sea surface temperature (SST), or sea ice extent can cause predictable variations in seasonal mean quantities. In addition, there is the possibility that the atmosphere's internal dynamics or slowly evolving surface properties (e.g., soil moisture and snow cover) may also generate potentially predictable interannual variability of seasonal mean quantities.

Unfortunately, the effects of these sources of variability are confounded in the real atmosphere. Consequently, it is difficult to diagnose their relative importance from the observational record alone. However, we can obtain estimates of the relative importance of some of these sources of potentially predictable variance in the climate simulated by a general circulation model (GCM) by conducting an ensemble of simulations in which some of the external boundary conditions are controlled. Several studies of this type have been reported, including Rowell (1998), Rowell and Zwiers (1997), Anderson and Stern (1996), Zwiers (1996, hereafter Z96), Kumar et al. (1996), Kumar and Hoerling (1995), Dix and Hunt (1995), and Stern and Miyakoda (1995).

The purpose of this paper is to diagnose the potential predictability of seasonal total precipitation, precipitation frequency, and precipitation intensity, using an ensemble of six 10-yr Atmospheric Model Intercomparison Project (AMIP) simulations (Gates 1992) conducted with the Canadian Climate Centre (CCC) GCM2 (McFarlane et al. 1992) in which the observed 1979–88 evolution of monthly mean SST and sea ice extent has been prescribed. Zwiers (1996) used the same GCM experiment to study the variability of seasonal mean 850-hPa temperature, 500-hPa height, and 300-hPa streamfunction. With this experimental design it is possible to partition the total interannual variability of a seasonal mean into (i) an external sources component that reflects the ''signal'' of the prescribed boundary conditions; (ii) an internal sources component that contains the effects of ''internal'' boundary conditions (CCC GCM2 has an interactive land surface), the atmosphere's internal dynamics on long timescales, and

* Additional affiliation: Department of Mathematics and Statistics, University of Victoria, Victoria, British Columbia, Canada.

Corresponding author address: Dr. Francis W. Zwiers, Canadian Centre for Climate Modelling and Analysis, University of Victoria, P.O. Box 1700, MS 339, Victoria, BC V8W 2Y2, Canada.
E-mail: Francis.Zwiers@ec.gc.ca

weather noise; and (iii) an intersimulation component that we dub the ''simulation effect.'' The latter accounts for small changes in the model configuration that occurred when CCC GCM2 was moved to a new computer (Z96 gives details) and for the effects of the choice of initial conditions for each simulation.

The main approach that has been used to study potential predictability is the analysis of variance (ANOVA), as is the case in the studies of Rowell and Zwiers (1997), Rowell (1998), Z96, Shea et al. (1995), Zwiers (1987), Madden (1976), among others. Kumar et al. (1996) and Kumar and Hoerling (1995) used a variant of the same technique while Dix and Hunt (1995) and Stern and Miyakoda (1995) used less formal techniques. ANOVA is a powerful approach for analyzing measurements on a continuous scale, such as those of precipitation intensity and seasonal total precipitation. However, reliability can be compromised when ANOVA models are used to approximate the behavior of discrete data such as the number of precipitation days in a season. This type of data is more appropriately analyzed with so-called log–linear models of the probability of the events that are counted. Thus we use different approaches to the analysis of variance depending upon the type of variable. Log-linear methods are frequently used in biomedical settings, but they have not often been applied to climatological problems.

An additional complication that arises when applying ANOVA techniques to precipitation data is that a different paradigm is needed to estimate the proportion of interannual variability that is induced by weather noise. While simple time series models (e.g., red noise as in Z96) can be applied to variables such as temperature and geopotential, a different algorithm is needed to represent the short timescale behavior of discontinuous (in time) variables such as precipitation. In this paper, we use a bootstrap algorithm (Zwiers 1990; Wilks 1997) for this purpose.

ANOVA is used in this study to assess the potential predictability of the intensity and seasonal total of precipitation, and the log-linear analysis (LLA) approach is used to analyze the predictability of precipitation frequency. In both cases, the short timescale behavior of precipitation is assessed using a bootstrap algorithm. A brief description of the Canadian Centre for Climate Modelling and Analysis (CCCMA) AMIP ensemble of simulations and the dataset is provided in section 2. The statistical methods used in this study are described in section 3. The results of the analysis are discussed in section 4, with a summary following in section 5.

2. Data and related subjects

The object of this study is six 10-yr daily precipitation time series for each model grid box obtained from an ensemble of six AMIP simulations conducted with the CCC GCM2. The CCC GCM2 is a spectral model with T32 resolution in the horizontal, 10 levels in the vertical, semi-implicit time stepping with a 20-min time step, and a full physics package. McFarlane et al. (1992) describes the model in detail. The experimental design is briefly described in the following paragraphs (Z96 gives full details).

The AMIP experimental protocol (Gates 1992) describes a 10-yr climate simulation that is forced with ''observed'' monthly mean SST and sea ice extent for the period from January 1979 to December 1988 as contained in the AMIP SST/sea ice dataset. The SST was varied on a daily basis by assigning the monthly mean SST to the midmonth day of each month and then interpolating linearly between midmonths. Sheng and Zwiers (1998) point out that this interpolating scheme slightly reduces the strength of the forcing experienced by the model during the extreme seasons compared to that experienced by the real atmosphere.

The ensemble consists of one AMIP simulation conducted on a CRAY XMP computer and five additional AMIP simulations conducted subsequently on an NEC SX/3 computer. All six simulations were forced with the same time evolving SST and sea ice extent boundary conditions. Minor coding changes were made when converting CCC GCM2 from one machine to the other, but no algorithmic changes were made. The surface topography in the CRAY simulation is different from that in the NEC simulations at three locations on the 2°N latitude (viz., 79°W, 101°E, and 120°E). The CRAY simulation was initiated from 1 January 1979 First Global Atmospheric Research Program Global Experiment (FGGE) III-B operational analyses (Bengtsson et al. 1982), while the NEC simulations were initialized from 1 January model states separated by 2-yr intervals in a long control simulation. Zwiers (1996) found some statistically significant differences between the climates simulated on the two computers. The cause of the differences was determined to be the source of the initialization data (FGGE III-B vs the previously simulated model state) rather than the computer used to perform the simulations.

The land surface was initialized from climatological conditions in all six simulations.[1] Thus large adjustments associated with the equilibration of the land surface with the rest of the climate system are not expected. Nonetheless, we discarded the first five months of each simulation to safeguard against the possibility of some initial adjustment resulting from the use of the time varying AMIP SSTs. We then carried out the analyses season by season, dividing a calendar year into four seasons in the usual way, namely, December–February (DJA), March–May (MAM), June–August (JJA), and September–November (SON). Thus we have available for analysis 10 JJA and SON seasons and 9 DJF and

[1] The soil moisture climatology from a previous simulation in which SSTs were prescribed from climatology.

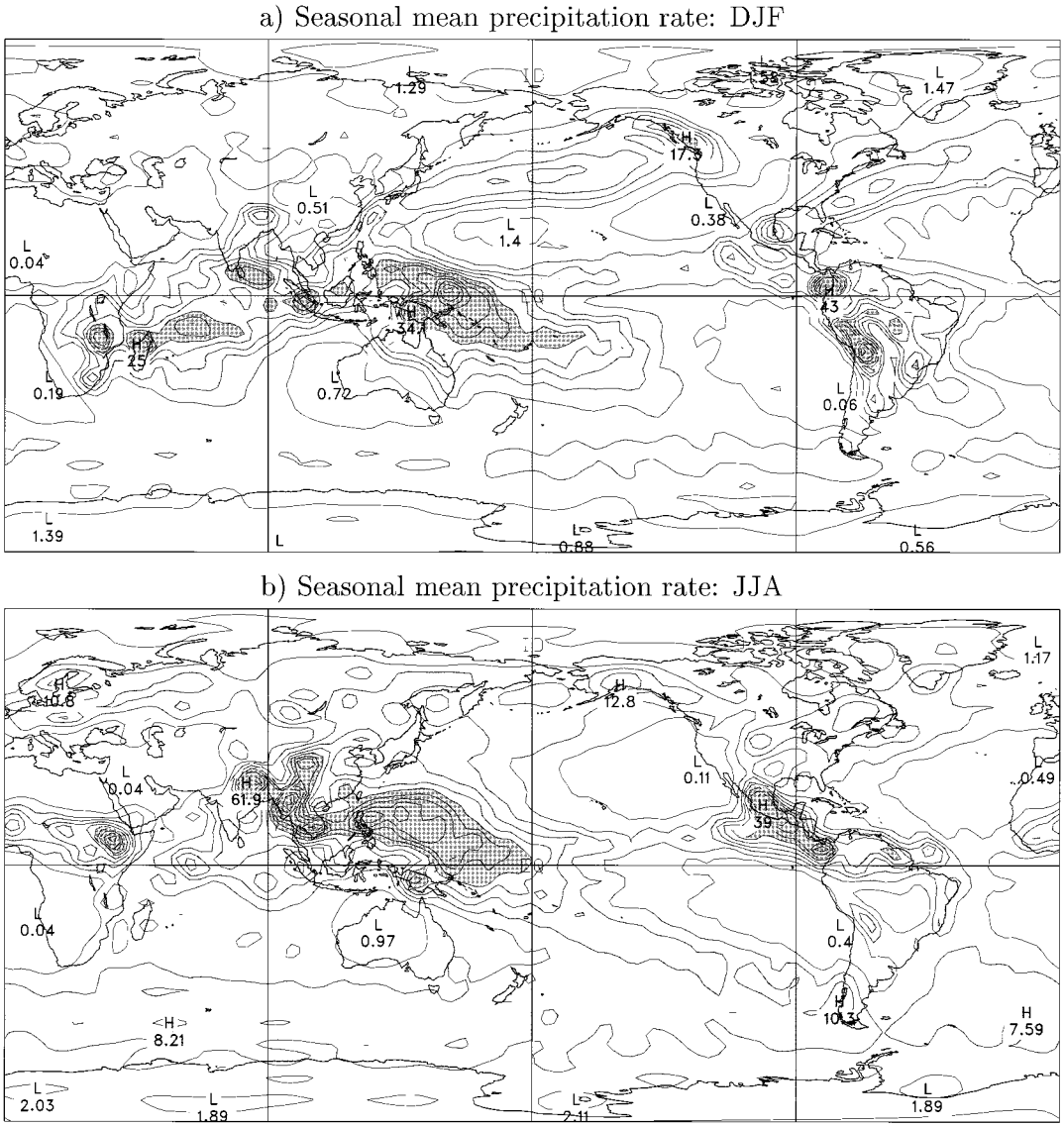


FIG. 1. Contour maps of seasonal mean precipitation simulated in the AMIP ensemble for (a) DJF and (b) JJA. Units are mm day^{-1}. The contour interval is 2.5 mm day^{-1}. Light and dark shadings indicate precipitation rates greater than 5.0 and 15.0 mm day^{-1}.

MAM seasons. Since CCC GCM2 does not take leap years into account, there are 92 days in MAM and JJA, 91 days in SON, and 90 days in DJF.

The DJF and JJA mean precipitation rates simulated in the AMIP ensemble are displayed in Fig. 1. The broad features of the mean simulated precipitation are similar to that which is observed (Jaeger 1976; Legates and Willmott 1990; Xie and Arkin 1996). However, in comparison with observations, the model produces too much precipitation in the western tropical Pacific Ocean. Orographically induced features of the Asian summer monsoon, such as the maximum over the Western Ghats of India, are not well represented in this modest resolution model. Also, precipitation is not correctly distributed in equatorial Africa in JJA.

At the opposite end of the time spectrum, the model appears to do a reasonable job of simulating extreme precipitation events (Zwiers and Kharin 1998). However it generates small precipitation events too frequently because it uses moist convective adjustment (McFarlane et al. 1992; Boer et al. 1984) and a simple scheme that deposits moisture on the surface when relative humidity exceeds 95%. We therefore used a threshold to classify days as wet or dry. We experimented with several thresholds ranging from 0.1 mm to values in excess of 2.0 mm. Small thresholds (less than 0.5 mm) appear to

give too many wet days. We therefore use a relatively large threshold, 1 mm, when classifying days as wet or dry, and use the exceedences above 1 mm when analyzing the intensity and seasonal accumulation of precipitation. Our analysis of the variability of the frequency, intensity, and seasonal accumulation of precipitation is not highly sensitive to this particular choice of threshold value.

3. Methodology

We analyze two distinctly different types of precipitation-related variables and therefore use two different analysis techniques. The seasonal mean intensity of precipitation and the seasonal accumulation of precipitation are measured on a continuous scale and, in the case of CCC GCM2, are approximately Gaussian (this is discussed further in section 3a). This allows us to use standard ANOVA methods to partition the interannual variability of these quantities into a number of variance components that can then be used to diagnose the relative importance of boundary forcing and other effects. Moreover, because the variance components are statistically independent, a number of hypotheses can be tested without having to refit the statistical model that is implicit in ANOVA. This model, the way in which the variance is partitioned, and the tests that can be performed are described in section 3a.

In contrast, the seasonal frequency of precipitation is a discrete random variable that can only take integer values. The seasonal frequency can be thought of as the sum of N Bernoulli (i.e., 0/1) random variables that take value 1 when it rains and 0 otherwise. Here, N is the length (in days) of a season. We assume that the probability of a wet day is constant within a given season, but the effects discussed in section 1 may well cause it to vary from year to year or simulation to simulation. The log–linear models discussed in section 3b are used to determine whether this is the case. These models are formally equivalent to those used in standard ANOVA, but inferences cannot be made as easily because there is not an equivalent partitioning of variability into variance components.

The effects of dependence between days must be taken into account when making inferences with both ANOVA and LLA. We do this by means of a moving blocks bootstrap. Details are given in section 3c.

a. The ANOVA approach

The ANOVA technique was originally defined by R. A. Fisher as the ''separation of the variance ascribable to one group of causes from the variance ascribable to other groups'' (Huitson 1966). This technique has been applied in atmospheric science in various guises in the past. Zwiers (1996) recently applied it using both one- and two-factor models to analyze the interannual variability and predictability of circulation variables, both in the CCCMA ensemble of AMIP climate simulations and in the relevant observed climate. Likewise, the two-factor mixed effects model is adopted in this study to analyze variabilities of the seasonal mean intensity and seasonal accumulation of precipitation in the ensemble.

A separate analysis of variance is performed in each season at every grid box. To simplify notation let X denote the variable of interest (i.e., seasonal mean precipitation intensity U or seasonal total precipitation T) at a given location and in a given season. The simulated values of X can be organized in a $Y \times S$ matrix with elements X_{ys}, which denote the value of X in year y and simulation s. Here Y is the number of years available for analysis (9 in DJF and MAM; 10 in JJA and SON) and S is the number of simulations (6).

We will represent the variation between the elements of the data matrix with a two-factor mixed effects model of the form

$$X_{ys} = \mu + \beta_y + \gamma_s + \delta_{ys} + \varepsilon_{ys}, \qquad (1)$$

where μ is the long-term mean value of X, β_y is the effect of boundary forcing, γ_s is the simulation effect, δ_{ys} is the internal sources (interaction) effect, and ε_{ys} represents weather noise. The β and γ terms are often referred to as *main effects* in the parlance of statisticians, and the δ term is referred to as the *two-factor interaction.* The phrase ''mixed effects'' alludes to the fact that we treat some elements of the model as being deterministic and others as being random. The elements of the model are described in detail in the following paragraphs.

The boundary-forced effects β_y, which are assumed to be deterministic, are deviations from the grand mean that are common to all simulations and are therefore presumed to arise from the prescribed SST–sea ice boundary conditions. We treat these effects as fixed because the boundary conditions are prescribed identically in each simulation. We assume that the simulation effects γ_s contain both a deterministic ''configuration effect,'' a_s, and a random effect, ξ_s, that is, $\gamma_s = a_s + \xi_s$. The configuration term represents systematic effects due to changes in the computer used to run the simulations and in the source of the initialization data. The random component ξ_s represents intersimulation variation caused by the choice of initial conditions. The latter are assumed to be independent, identically distributed (iid) zero mean Gaussian variables with variance σ_ξ^2.

The interaction effects δ_{ys} are assumed to be a purely random effect arising either from surface properties (e.g., soil moisture and snow cover) that are not constrained to evolve identically in each simulation or from the internal dynamics of the atmosphere (Z96). We assume that these terms are iid zero mean Gaussian random variables with variance σ_δ^2. The error terms ε_{ys}, which are also assumed to be iid Gaussian, represent interannual variability in the quantity of interest that is caused by daily ''weather'' variations. Finally, it is fur-

ther necessary to assume that the random elements of this model, ξ_s, δ_{ys}, and ε_{ys}, are mutually independent.

As with all models, this model is at best an approximation of reality that aids in the analysis of variability within a given framework. There are likely to be departures from the assumptions, such as the presence of weak dependence between the random elements, that will compromise the interpretation of results to some extent.

To apply the model it is necessary to assume that the observed random variable X_{ys} (i.e., U_{ys} or T_{ys}) is Gaussian. The central limit theorem suggests that this is reasonable since U_{ys} and T_{ys} are the result of many individual events in a season in most parts of the world. We also used an objective measure (the Kolmogorov–Smirnov statistic) to test the goodness-of-fit of the Gaussian distribution and found little evidence contrary to the Gaussian assumption.

The total sum of squares of variable X,

$$\mathrm{TSS}_X = \sum_s \sum_y (X_{ys} - X_{oo})^2,$$

can be partitioned into three statistically independent variance components:

$$\mathrm{TSS}_X = \mathrm{SSB}_X + \mathrm{SSS}_X + \mathrm{SSI}_X, \tag{2}$$

where

$$\mathrm{SSB}_X = S \sum_y (X_{yo} - X_{oo})^2, \tag{3}$$

$$\mathrm{SSS}_X = Y \sum_s (X_{os} - X_{oo})^2, \tag{4}$$

$$\mathrm{SSI}_X = \sum_s \sum_y (X_{ys} - X_{os} - X_{yo} + X_{oo})^2, \tag{5}$$

and a "$_o$" is used to replace a subscript when an arithmetic average is taken over that index (this notation is used throughout the paper). Substituting model (1) (note that $\gamma_s = a_s + \xi_s$) into (3), (4), and (5) and taking expectations, one obtains

$$E(\mathrm{SSB}_X) = S \sum_y \beta_y^2 + (Y - 1)(\sigma_\delta^2 + \sigma_\varepsilon^2), \tag{6}$$

$$E(\mathrm{SSS}_X) = Y \sum_s a_s^2 + Y(S - 1)\sigma_\xi^2 + (S - 1)(\sigma_\delta^2 + \sigma_\varepsilon^2), \tag{7}$$

$$E(\mathrm{SSI}_X) = (S - 1)(Y - 1)(\sigma_\delta^2 + \sigma_\varepsilon^2). \tag{8}$$

Comparison of (6) and (8) suggests that it may be possible to test the null hypothesis that the prescribed boundary conditions do not influence the simulated interannual variability (H_B: $\beta_y = 0$ for $y = 1, \ldots, Y$) with

$$F_{B,X} = \frac{\mathrm{SSB}_X}{(Y - 1)} \Bigg/ \frac{\mathrm{SSI}_X}{(S - 1)(Y - 1)}. \tag{9}$$

We would expect $F_{B,X}$ to take values near 1 when H_B is true. When H_B is false, $F_{B,X}$ will tend to be greater than 1 since the numerator in (9) then estimates a larger variance than the denominator. The assumptions we have made ensure that $F_{B,X}$ has the F distribution with $(Y - 1)$ and $(S - 1)(Y - 1)$ degrees of freedom when H_B is true. Thus H_B is tested by comparing $F_{B,X}$ with the critical value $F_{1-\alpha}[(Y - 1), (S - 1)(Y - 1)]$ where α denotes the selected significance level and the notation $F_{1-\alpha}[m, n]$ denotes the $(1 - \alpha)$-quantile of the F distribution with m and n degrees of freedom. An unbiased estimator of the proportion of the total variance in quantity X that is due to boundary-forced effects is (Rowell 1998)

$$P_{B,X} = \left(\mathrm{SSB}_X - \frac{\mathrm{SSI}_X}{S - 1}\right) \Bigg/ \mathrm{TSS}_X. \tag{10}$$

Analogously, the null hypothesis that the configuration changes (i.e., computer type and source of initial conditions) and the choice of initial conditions do not affect intersimulation variability (H_S: $a_s = 0$ for $s = 1, \ldots, S$ and $\sigma_\xi^2 = 0$) can be tested by comparing

$$F_{S,X} = \frac{\mathrm{SSS}_X}{(S - 1)} \Bigg/ \frac{\mathrm{SSI}_X}{(S - 1)(Y - 1)} \tag{11}$$

with $F_{1-\alpha}[(S - 1), (S - 1)(Y - 1)]$; and an unbiased estimator of the proportion of the total variance in quantity X arising from simulation effects is

$$P_{S,X} = \left(\mathrm{SSS}_X - \frac{\mathrm{SSI}_X}{Y - 1}\right) \Bigg/ \mathrm{TSS}_X. \tag{12}$$

If significant simulation effects are identified, it is possible to further decompose SSS_X into separate terms that can be ascribed to configuration effects and the initial conditions (see Z96).

In order to test the null hypothesis that variations from internal sources are not potentially predictable (H_I: $\sigma_\delta^2 = 0$), it is necessary to take the variability caused by the weather noise into account. Madden (1976), Zwiers (1987), Z96, and others describe methods that can be used when X is a quantity such as 500-hPa height that varies continuously in time. In these circumstances it is usually possible to obtain an independent estimate $\hat{\sigma}_{\varepsilon,X}^2$ of $\sigma_{\varepsilon,X}^2$ such that H_I can be tested by comparing

$$F_{I,X} = \frac{\mathrm{SSI}_X}{(S - 1)(Y - 1)} \Bigg/ \hat{\sigma}_{\varepsilon,X}^2 \tag{13}$$

against the critical values of a suitable F distribution. These estimates use time series models to take the short timescale variability of X into account.

We attempted to take a similar approach with precipitation by modeling its short-term stochastic properties with a conditional chain-dependent process (Katz and Parlange 1996). While this was generally successful at mid- and high latitudes, precipitation in CCC GCM2's climate exhibits more complex stochastic behavior at low latitudes than we were able to represent with these stochastic models.

We therefore used another approach for testing H_I. A bootstrapping procedure was used to estimate the 10%, 5%, and 1% significance-level critical values of SSI_X under H_I. The hypothesis was then tested by comparing the observed value of SSI_X with the estimated critical value $SSI_{X,1-\alpha}$ at the α significance level or, alternatively, by comparing

$$F_{I,X} = \frac{SSI_X}{SSI_{X,1-\alpha}} \tag{14}$$

with 1. Details of the bootstrapping procedure are given in section 3c.

b. The LLA approach

Let f_{ys} denote the seasonal total number of wet days (i.e., seasonal frequency of precipitation) for year y of simulation s ($y = 1, 2, \ldots, Y$ and $s = 1, 2, \ldots, S$). When days are independent the counts f_{ys} can be thought of as the number of successes in N Bernoulli trials where the probability of success in any trial may depend upon the year (reflecting a boundary effect), the simulation (the simulation effect), and perhaps the ''interaction'' between the two (i.e., internal variation in the probability of precipitation that arises from internal sources). A multiplicative model for the frequency of precipitation that takes these effects into account is (Kennedy 1983)

$$F_{ys} = GB_yS_sI_{ys}, \tag{15}$$

where F_{ys} is the expected seasonal frequency of precipitation and B_y, S_s, and I_{ys} are year, simulation, and interaction factors that modify the background precipitation frequency G. These factors are constrained so that

$$\prod_{y=1}^{Y} B_y = \prod_{s=1}^{S} S_s = \prod_{y=1}^{Y} I_{ys} = \prod_{s=1}^{S} I_{ys} = 1.$$

These constraints result in a model with exactly YS parameters that captures all of the variability in the observed precipitation frequencies f_{ys}; that is, $\hat{F}_{ys} = f_{ys}$, where $\hat{F}_{ys}$ denotes an estimate of F_{ys}. Hence the model is said to be ''saturated'' (Kennedy 1983).

The similarity between this model and the standard two-way ANOVA becomes more apparent when we take logs to obtain

$$L_{ys} = \ln(F_{ys}) = \mu + \beta_y + \gamma_s + \delta_{ys}, \tag{16}$$

where $\mu = \ln(G)$, $\beta_y = \ln(B_y)$, $\gamma_s = \ln(S_s)$, and $\delta_{ys} = \ln(I_{ys})$. This is a log-linear cell frequency model, a multiplicative model that becomes linear in its parameters after taking logs. The constraints in the multiplicative model become the familiar constraints that apply to a standard linear model, that is, $\beta_o = \gamma_o = \delta_{os} = \delta_{yo} = 0$. Corresponding maximum likelihood estimates of the parameters are (Everitt 1992; Kennedy 1983; Goodman 1978)

$$\hat{\mu} = l_{oo}, \qquad \hat{\beta}_y = l_{yo} - l_{oo}, \qquad \hat{\gamma}_s = l_{os} - l_{oo},$$
$$\hat{\delta}_{ys} = l_{ys} - l_{os} - l_{yo} + l_{oo}, \tag{17}$$

where $l_{ys} = \ln(f_{ys})$. As with the linear model, the main effects β_y and γ_s are the effects of boundary forcing and simulation, respectively, and the interaction between them, δ_{ys}, is presumed to represent variations in precipitation frequencies arising from internal sources.

As in ANOVA, our aim here is to test hypotheses about model terms β_y, γ_s, and δ_{ys}. The formal approach to developing tests is similar to that which is often used in regression analysis; a hierarchy of models is fitted and the importance of terms added between successive models in the hierarchy is judged by the change in the goodness-of-fit. We will first describe the appropriate tests under the assumption that there is no dependence between days and then go on to describe a method that copes with dependence.

The paradigm for constructing tests of statistical significance of terms β_y, γ_s, and δ_{ys} uses three hierarchies of log-linear models. The hierarchy

$$\ln(F_{ys}) = \mu + \gamma_s, \tag{18}$$
$$\ln(F_{ys}) = \mu + \beta_y + \gamma_s \tag{19}$$

is used to test the hypothesis that the prescribed boundary conditions do not affect precipitation frequency (i.e., H_B: $\beta_y = 0$ for $y = 1, 2, \ldots, Y$). The hierarchy

$$\ln(F_{ys}) = \mu + \beta_y, \tag{20}$$
$$\ln(F_{ys}) = \mu + \beta_y + \gamma_s \tag{21}$$

is used to test the hypothesis that precipitation frequency is not affected by the model configuration or the initial conditions (i.e., H_S: $\gamma_s = 0$ for $s = 1, 2, \ldots, S$). The hierarchy

$$\ln(F_{ys}) = \mu + \beta_y + \gamma_s \tag{22}$$
$$\ln(F_{ys}) = \mu + \beta_y + \gamma_s + \delta_{ys} \tag{23}$$

is used to test the hypothesis that internal variability in precipitation frequency is not potentially predictable (i.e., H_I: $\delta_{ys} = 0$ for $y = 1, 2, \ldots, Y$ and $s = 1, 2, \ldots, S$).

In each case, the hypothesis is tested by measuring the change in goodness-of-fit of the expected precipitation frequencies F_{ys} to the observed frequencies f_{ys} between two models. The goodness-of-fit can be assessed with either the likelihood-ratio chi-square statistic

$$G^2 = 2 \sum_y \sum_s (f_{ys}) \ln\left(\frac{f_{ys}}{F_{ys}}\right)$$
$$= 2 \sum_y \sum_s (f_{ys})[\ln(f_{ys}) - \ln(F_{ys})] \tag{24}$$

or with Pearson's nonparametric χ^2 statistic

$$\chi^2 = \sum_y \sum_s \frac{(f_{ys} - F_{ys})^2}{F_{ys}}.$$

The G^2 statistic is generally preferable since its ''additive property'' (Kennedy 1983) makes it more convenient to use. Both statistics are *approximately* χ^2 distributed with appropriate numbers of degrees of freedom when H is true.

By taking the difference between G^2 (24) for the two models in each hierarchical pair [(18, 19), (20, 21), and (22, 23)], one obtains the following likelihood-ratio statistics for testing H_B, H_S, and H_I:

$$G_B^2 = 2 \sum_y \sum_s (f_{ys}\hat{\beta}_y), \qquad G_S^2 = 2 \sum_y \sum_s (f_{ys}\hat{\gamma}_s),$$

$$G_I^2 = 2 \sum_y \sum_s (f_{ys}\hat{\delta}_{ys}).$$

These statistics measure the improvement in the fit of the more complex model in the hierarchical pair over the simpler model. They can be interpreted as measures of the strength of the boundary-forced, simulation, and internal sources effects, respectively. They are asymptotically χ^2 distributed with $\nu_B = (Y - 1)$, $\nu_S = (S - 1)$, and $\nu_I = (S - 1)(Y - 1)$ degrees of freedom when H_B, H_S, and H_I, respectively, are true. Hypotheses H_B, H_S, and H_I can therefore be tested by comparing

$$F_{B,F} = \frac{G_B^2}{\chi^2_{1-\alpha}[Y - 1]}, \tag{25}$$

$$F_{S,F} = \frac{G_S^2}{\chi^2_{1-\alpha}[S - 1]}, \tag{26}$$

$$F_{I,F} = \frac{G_I^2}{\chi^2_{1-\alpha}[(S - 1)(Y - 1)]} \tag{27}$$

with 1 where α is the selected significance level and $\chi^2_{1-\alpha}[\nu]$ is the critical value from the χ^2 distribution with ν degrees of freedom.

A practical problem associated with the calculation of likelihood-ratio chi-square statistics (e.g., G_B^2, G_S^2, and G_I^2) is that the observed frequency f_{ys} may occasionally be equal to zero. A widely accepted solution is to add a small quantity to all elementary cell frequencies to be analyzed (Everitt 1992; Kennedy 1983). This value was set to 0.1 in the current study.

So far, we have assumed that there is no dependence between days. When there is dependence the likelihood-ratio chi-square statistics G_B^2, G_S^2, and G_I^2 may have null distributions different from the χ^2 distributions described above. This contrasts with the standard linear model, where the statistics used to test for its main effects are not affected by dependence between days. Since the null distributions cannot be derived analytically, we used a Monte Carlo procedure to estimate the critical values of the likelihood-ratio chi squares under their respective null hypotheses. The procedure and the derived critical values are described in section 3c. If $G^2_{B,1-\alpha}$, $G^2_{S,1-\alpha}$, and $G^2_{I,1-\alpha}$ denote the critical values of G_B^2, G_S^2, and G_I^2, respectively, then H_B, H_S, and H_I are tested at the α significance level by comparing

$$F_{B,F} = \frac{G_B^2}{G^2_{B,1-\alpha}}, \tag{28}$$

$$F_{S,F} = \frac{G_S^2}{G^2_{S,1-\alpha}}, \tag{29}$$

$$F_{I,F} = \frac{G_I^2}{G^2_{I,1-\alpha}} \tag{30}$$

with 1, respectively.

c. The stochastic simulation of daily precipitation

In order to estimate the critical values of SSI_X, G_B^2, G_S^2, and G_I^2, one needs to simulate a random daily precipitation field series that has the same length and the same short-term dependence characteristics as the corresponding observed series but does not include boundary-forced, simulation, and interaction effects. This was accomplished by means of a moving block bootstrapping procedure (Wilks 1997) that ''resamples'' the daily precipitation fields. The procedure implicitly assumes that H_B, H_S, and H_I are all true. This results in conservative tests (i.e., they may operate at less than the nominal significance levels) because the critical values of SSI_X, G_B^2, G_S^2, and G_I^2 may be somewhat overestimated.

Our resampling procedure operates as follows. We suppose that the daily precipitation series has a memory of about N_d days. We chose $N_d = 15$, but the exact value of N_d used in the procedure below is not that important provided that the timescale of dependence is not underestimated. To preserve dependence in the resampled series on timescales shorter than the N_d days, we generated random series of daily precipitation fields by selecting spells of N_d continuous days from the original daily precipitation field series. Specifically, for a given N-day season such as DJF and an ensemble of S Y-yr GCM simulations, we let integer $K = (N \times Y \times S)/N_d$ and then

1) generated K uniform $[0, N - N_d]$ random integers $n_1, n_2, \ldots, n_K$, K uniform $[1, Y]$ integers $y_1, y_2, \ldots, y_K$, and K uniform $[1, S]$ integers $s_1, s_2, \ldots, s_K$;
2) constructed a daily precipitation sequence R_i^e, $i = 1, 2, \ldots, NYS$ by setting $R^e_{(k-1)N_d+j} = R_{n_k+j,y_k,s_k}$, $j = 1, 2, \ldots, N_d$, and $k = 1, 2, \ldots, K$, where R_{n_k+j,y_k,s_k} is the precipitation field simulated by the GCM on day $(n_k + j)$ of year y_k of simulation s_k;
3) grouped sequences R_i^e, $i = 1, 2, \ldots, NYS$ into chunks of N days representing a season, and then Y seasons representing a simulation; and
4) computed a realization of the SSI_X, G_B^2, G_S^2, and G_I^2 fields from the bootstrap sample $\{R^e_{nys}; n = 1, 2, \ldots, N, y = 1, 2, \ldots, Y, s = 1, 2, \ldots, S\}$.

We obtained empirical estimates of the distributions of

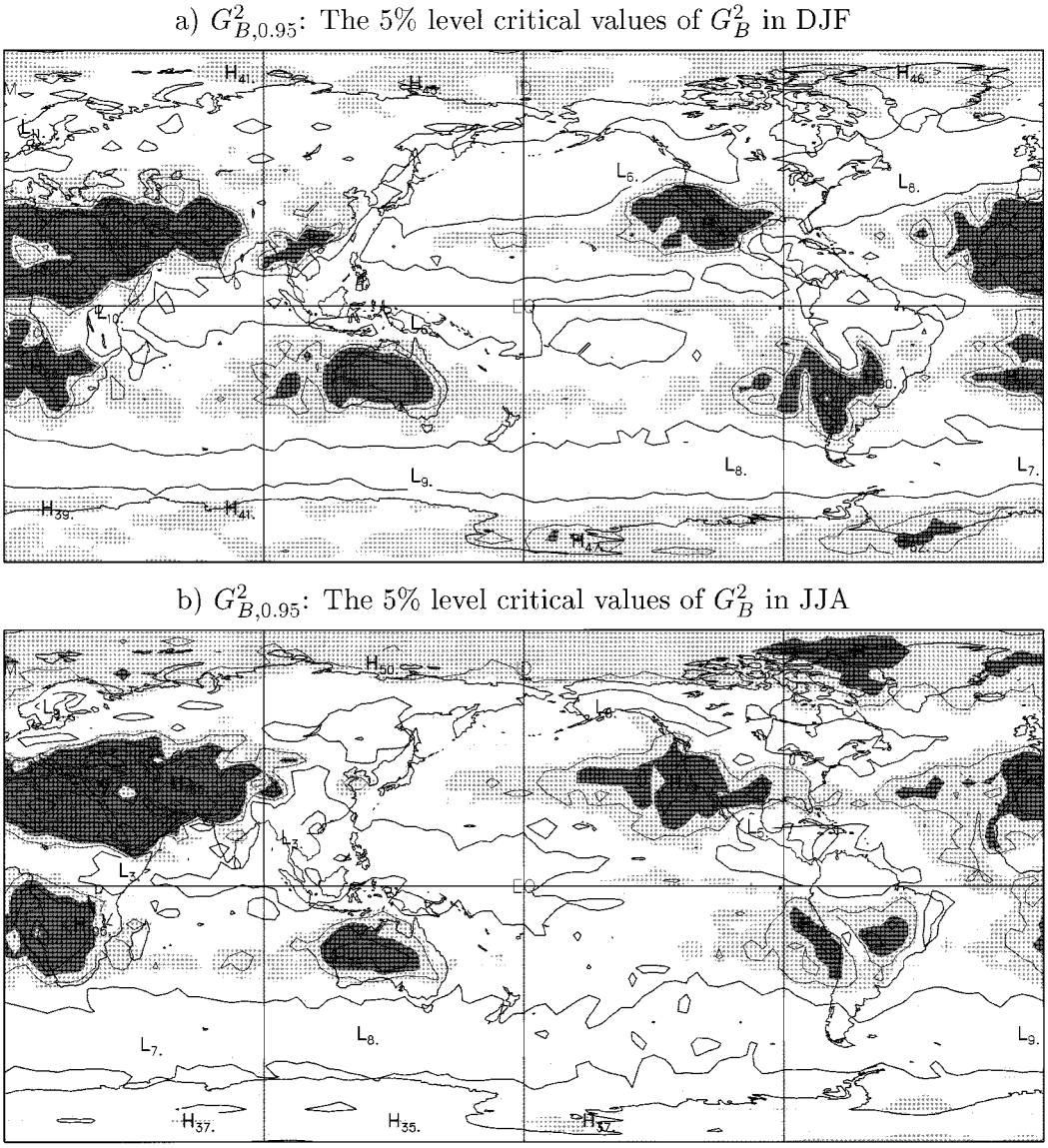


FIG. 2. Contour maps of the 5% level critical values of G^2_B in (a) DJF and (b) JJA, obtained from the 1000 Monte Carlo simulations described in section 3c. Light, medium, and dark shadings indicate locations where the critical values are greater than $\chi^2_{0.95}(Y-1)$, $2\chi^2_{0.95}(Y-1)$, and $3\chi^2_{0.95}(Y-1)$, respectively (see section 3c for more details).

SSI_X, G^2_B, G^2_S, and G^2_I under the assumption that H_B, H_S, and H_I all hold by repeating the steps above 1000 times. The 90th, 95th, and 99th percentiles of these distributions were taken as the critical values for the 10%, 5%, and 1% level significance tests described in sections 3a and 3b.

The effects of dependence between days can be seen by comparing the critical values of the statistics G^2_B, G^2_S, and G^2_I with those from the relevant χ^2 distributions. Figure 2 shows the estimates of $G^2_{B,0.95}$ obtained from the bootstrapping procedure described above, with light, medium, and dark shadings indicating locations where $G^2_{B,0.95} > \chi^2_{0.95}(Y-1)$, $G^2_{B,0.95} > 2 \times \chi^2_{0.95}(Y-1)$, and $G^2_{B,0.95} > 3 \times \chi^2_{0.95}(Y-1)$, respectively. The null distributions are narrower than $\chi^2(Y-1)$ in areas with frequent precipitation, and they are much wider than $\chi^2(Y-1)$ in the tropical convergence zones where precipitation is rare. These are both circumstances under which it can be anticipated that the χ^2 approximation to the null distribution of G^2_B will break down, even if dependence is not an issue. Elsewhere, the $\chi^2(Y-1)$ should approximate the null distribution of G^2_B quite well in the absence of dependence, but Fig. 2 shows that this is not the case.

The null distributions of other likelihood-ratio statistics, such as G^2_S and G^2_I, are likewise affected by the dependence. Contour maps of the critical values have structures similar to those shown in Fig. 2.

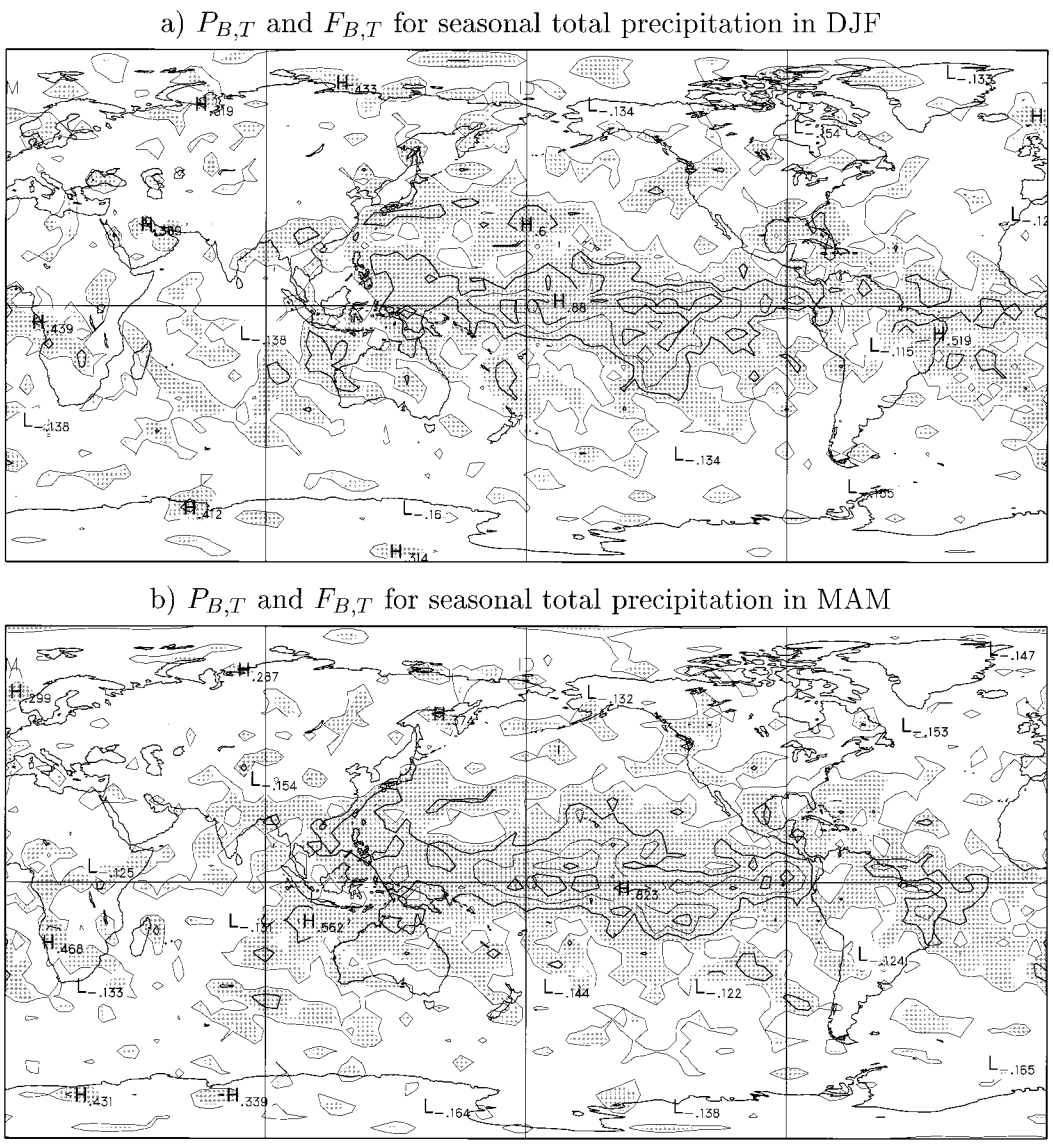


FIG. 3. Maps of $P_{B,T}$ (contoured) and $F_{B,T}$ (shaded) in (a) DJF, (b) MAM, (c) JJA, and (d) SON. Statistic $F_{B,T}$ is used to test for statistical significance of the boundary-forced effects on the interannual variations of seasonal total precipitation. The shading indicates locations where the $F_{B,T}$ (9) ratios, and hence the boundary-forced signals, are significant.

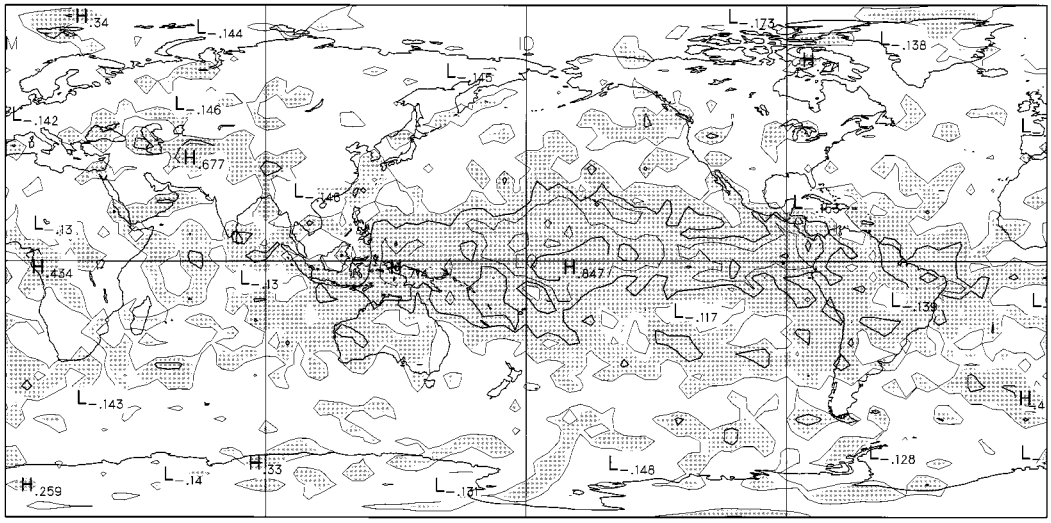
4. Interannual variability of precipitation

Using log-linear analysis, the statistical significance of the boundary-forced, simulation, and internal sources effects on precipitation frequency were assessed separately in each grid box. As a result, fields of the statistics $F_{B,F}$ (28), $F_{S,F}$ (29), and $F_{I,F}$ (30) were obtained for each season. Likewise, analysis of variance was performed on the ensembles of seasonal mean intensity and seasonal accumulation of precipitation, resulting in fields of the statistics $F_{B,X}$ (9), $F_{S,X}$ (11), and $F_{I,X}$ (14) (where X stands for the intensity U or the seasonal total T). The portion of the globe covered with significant statistics (at the 5% level) is given in Table 1 for all three

TABLE 1. Percentages of the globe covered with significant (at the 5% level) F ratios representing specific effects on the seasonal frequency, seasonal mean intensity, and seasonal totals of precipitation in different seasons of the year.

	F ratios	DJF (%)	MAM (%)	JJA (%)	SON (%)
Seasonal frequency	$F_{B,F}$	40.1	36.3	43.5	34.5
	$F_{I,F}$	10.2	8.9	9.6	6.2
	$F_{S,F}$	8.4	7.4	7.9	5.0
Seasonal mean intensity	$F_{B,U}$	23.7	26.0	28.4	26.3
	$F_{I,U}$	3.8	2.8	3.7	2.7
	$F_{S,U}$	4.5	5.1	5.6	5.0
Seasonal total precipitation	$F_{B,T}$	31.0	33.2	37.7	32.7
	$F_{I,T}$	6.1	3.6	4.7	3.8
	$F_{S,T}$	5.4	4.5	6.2	5.0

c) $P_{B,T}$ and $F_{B,T}$ for seasonal total precipitation in JJA



d) $P_{B,T}$ and $F_{B,T}$ for seasonal total precipitation in SON

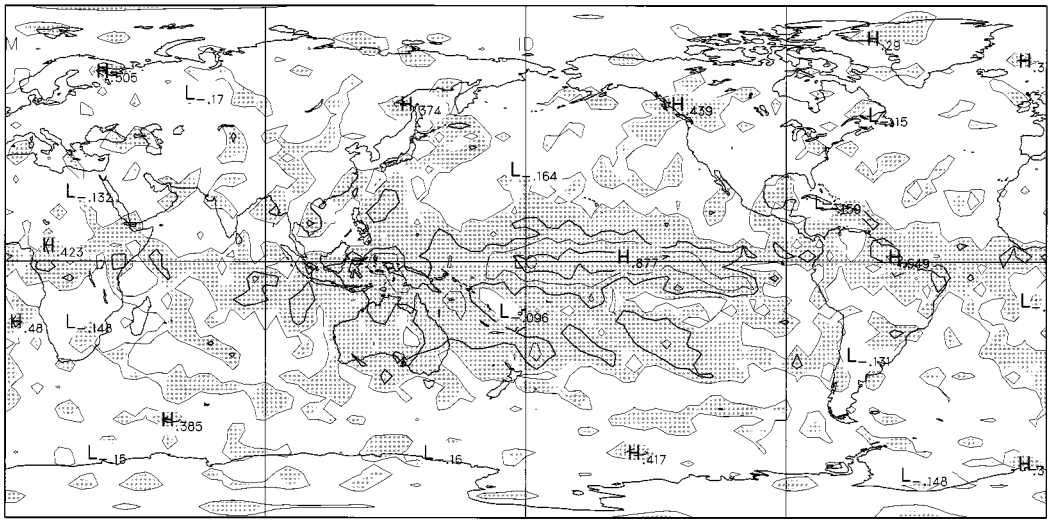


FIG. 3 (*Continued*) at the 5% level. Contour lines indicate the boundary-forced variance proportion $P_{B,T}$ (10). The contour interval and the lowest contour level are 0.2 and 0.1, respectively.

precipitation characteristics and all statistics described above. This table can be interpreted with Livezey and Chen (1983), who show that rejection rates of 7.8%, 6.9%, and 6.4% are field significant at the 5% level in fields with 250, 500, and 1000 equivalent degrees of freedom.

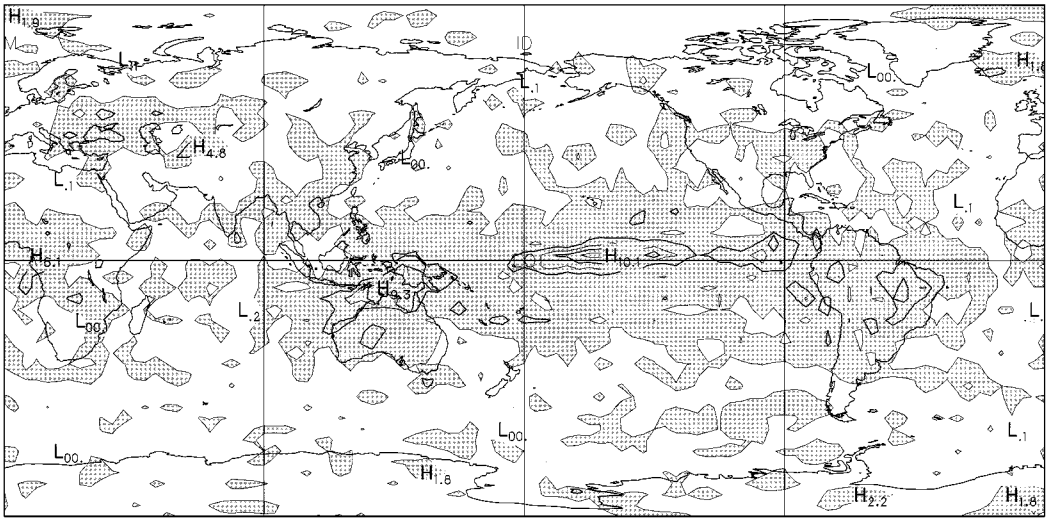
The field significance tests justify at least some interpretation of the spatial patterns of the test statistic. We attempt to do this below. However, interpretation is made difficult and uncertain because the statistics exhibit considerable spatial variability that is associated with the small-scale spatial correlation structure of precipitation in CCC GCM2. It is in fact very difficult to tell whether all but the largest regions are significant.

a. The boundary-forced interannual variability

The specified boundary conditions have significant effects on the frequency, intensity, and seasonal total precipitation over a substantial fraction of the globe. Depending upon the variable and season, rejection rates vary between 23.7% and 43.5%. That is to say, the boundary forced effects on all three aspects of precipitation are highly field significant.[2] We note that the CCC

[2] The smallest boundary effect rejection rate is 23.7%, which would be field significant in a field with more than about 15 spatial degrees of freedom (Livezey and Chen 1983).

a) $G_B^2/G_{B,0.95}^2$ for the frequency in JJA



b) $F_{B,U}/F_{0.95}[9, 45]$ for the intensity in JJA

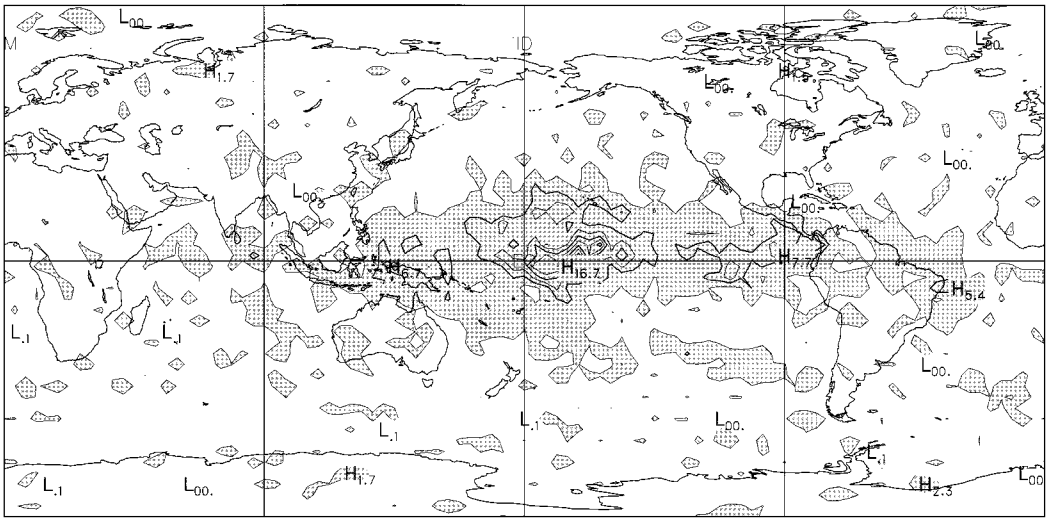


FIG. 4. Maps of (a) $F_{B,F} = G_B^2/G_{B,0.95}^2$ and (b) $F = F_{B,U}/F_{0.95}$ (9, 45) in JJA, which are used to test for statistical significance of the boundary-forced effects on the interannual variations of the frequency and intensity of precipitation, respectively. The shading indicates locations where $F_{B,F} > 1$ or $F > 1$; that is, the SST/sea ice forced signals are significant at the 5% level. The contour interval and the lowest contour level are 2.0 and 1.0, respectively.

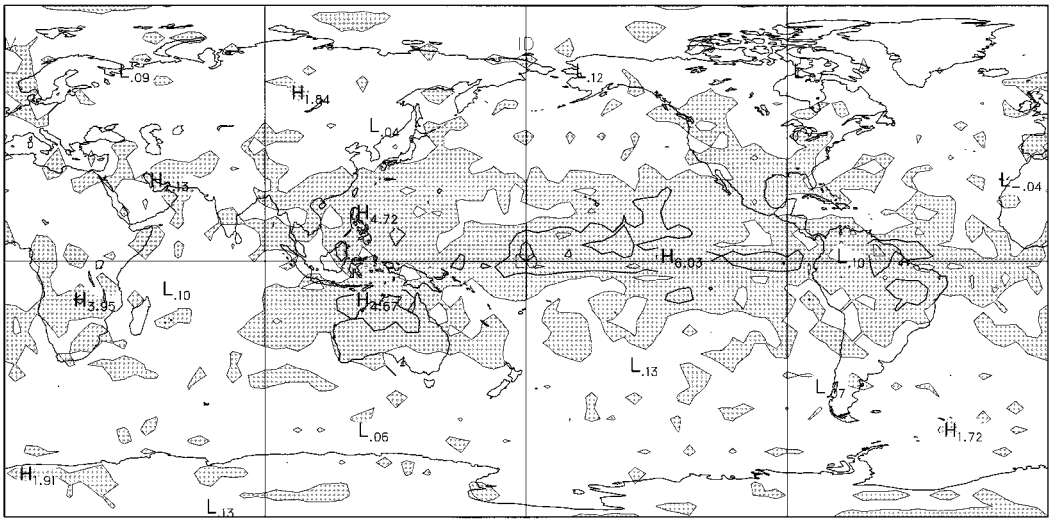
GCM2 simulated precipitation is slightly more sensitive to the prescribed boundary conditions in JJA than in other seasons (cf. Table 1).

Since the seasonal accumulation is a complex quantity involving both the frequency and the intensity, it will be used hereafter as a prototypical variable wherever a comparison between variabilities of frequency and intensity is not required.

The variance proportions $P_{B,T}$ (10) and the statistics $F_{B,T}$ (9), which are used to assess the statistical significance of the boundary-forced effects on seasonal total precipitation, are mapped in Fig. 3. The most salient feature in these diagrams is the region of large $P_{B,T}$ (contoured) and $F_{B,T}$ (shaded) values in the Tropics, especially in the tropical Pacific, where the boundary-forced effects are responsible for 40%–90% of the simulated interannual variability of the seasonal total precipitation. Significant signals were also diagnosed over equatorial South America, the equatorial Atlantic to western Africa, and southern Asia (Indonesian Archipelago). This is not surprising, because the variability of El Niño–Southern Oscillation (ENSO) dominates in the tropical Pacific, whereas natural variations from internal chaotic variability are weak over the Tropics.

The boundary-forced signal is much weaker and not strongly organized in the temperate latitudes. Never-

a) $G_B^2/G_{B,0.95}^2$ for the frequency in MAM



b) $F_{B,U}/F_{0.95}[8,40]$ for the intensity in MAM

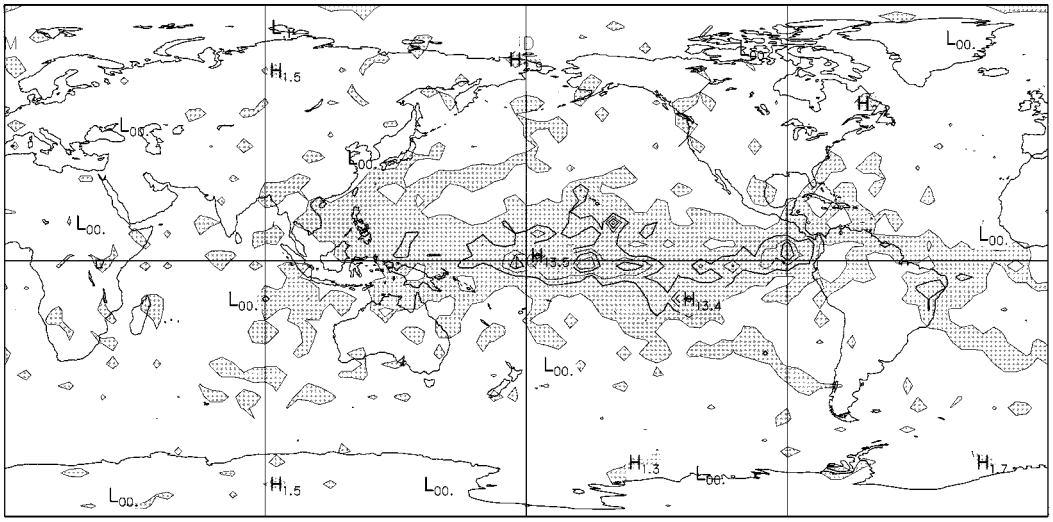


FIG. 5. The same as Fig. 4 but for MAM.

theless, apparently significant signals were diagnosed in the northwestern Pacific (over the North Pacific currents, east of Japan) and perhaps in the region from south China to India in DJF and MAM (see Fig. 3). There is a suggestion that boundary forcing may influence a small fraction of interannual precipitation variability in western and southern North America (in DJF and MAM), where the boundary forcing has also been diagnosed to have significant effects on the 500-hPa geopotential (Z96). Little signal was found over west-central Europe throughout the year. Over the Southern Hemisphere, boundary-forced effects were identified to be significant over the midlatitudes of the South Pacific and South America, the Atlantic and the Indian Oceans, and over most of Australia (especially in JJA and SON).

The patterns of significant boundary-forced signal diagnosed here bear some similarity to patterns of significant relationships between ENSO and observed precipitation over landmasses revealed by Xie et al. (1996), especially in DJF, and also to those of Ropelewski and Halpert (1989, 1987). Over the United States and Canada, the pattern of significant boundary-forced effects on DJF precipitation bears some resemblance to Shabbar and Barnston's (1996) and Barnston's (1994) geographical distributions of CCA skill in forecasting 3-month mean surface precipitation over Canada and the United States with double-weighted SST predictors for DJF/JFM; both indicate strong boundary-forced signals along southern Canada to the northern states (including the northern Great Lakes) and in the southeastern and south-central states (cf. Fig. 3a). In JJA, the CCC GCM2–simulated precipitation seems to contain much

less boundary-forced signal than does the observed as revealed by Xie et al. (1996), especially over North America and northern Eurasia.

Precipitation frequency appears to be more sensitive to the boundary forcing than does precipitation intensity in the sense that we are able to diagnose significant boundary-forced signals over a larger part of the globe. The total area affected by significant boundary-forced signals is generally much larger (about 10%–15% more of the globe; cf. Table 1) for precipitation frequency. In particular, over some land areas, the boundary conditions are diagnosed to have significant effects *only* on precipitation frequency. To illustrate this, Fig. 4 displays contour maps of the $F_{B,F}$ (28) and $F_{B,U}/F_{0.95}$ (9) ratios for JJA. Clearly, seasonal frequencies of precipitation over tropical western Africa, southeastern Europe, southeastern Asia, Australia, South America, and eastern North America are diagnosed to be significantly affected by the specified boundary conditions, whereas little variability of the corresponding intensity is ascribable to the boundary conditions. Similar features were also found over tropical Africa, Australia, India, and the Yangtze River basin of China in SON; and the contrast between frequency and intensity is striking over southern Africa, South China, Australia, and western and southern North America in MAM, as shown in Fig. 5.

b. The potentially predictable variations from internal sources

Evidence for potential predictability from internal sources is generally small in precipitation frequency. Significant $G^2_{I,F}$ values at the 5% level were identified over only 6.2%–10.2% of the globe. Most of the rejections occur over land areas. Zwiers (1996) obtained a similar result for 850-hPa temperature.

The internal sources effects on the seasonal mean intensity and seasonal total of precipitation were found to be statistically insignificant. The rejection rates range from 2.7% to 6.1% (cf. Table 1; this test and the previous one are probably conservative, i.e., the real significance level is likely less than 5%), which are not field significant. In other words, no potentially predictable interannual variability in the seasonal mean intensity and seasonal total of precipitation was identified to arise from the internal sources.

c. The simulation effects

Using the statistics $F_{S,F}$ (29), $F_{S,T}$, and $F_{S,U}$ (11) to test for the presence of the significant simulation effects, we obtained rejection rates ranging from 5.0% to 8.4% for the seasonal frequency and from 4.5% to 6.2% for the seasonal mean intensity and seasonal total of precipitation (cf. Table 1). The simulation effects on DJF, MAM, and JJA seasonal frequency of precipitation may be marginally field significant at the 5% significance level. However, no organized structure turns up in the resulting maps of test statistics (not shown) with the exception of a small region over northern South America, near 2°N, 79°W (and the downstream region), where the topography is different in the CRAY and NEC simulations. Simulation effects are not field significant when this area is discounted.

5. Summary and conclusions

Log-linear analysis and analysis of variance approaches were used to analyze the interannual variability and potential predictability of precipitation as it is simulated in the CCCMA ensemble of AMIP climate simulations.

The AMIP experiment produced six 10-yr realizations of a simulated climate that are physically consistent with the prescribed SST–sea ice forcing in the sense that the climate model computed the atmosphere's deterministic and stochastic response to the imposed forcing. This makes it possible for us to gain useful insights into the sources of interannual variability and potential predictability of precipitation, though the separation we have made here between externally and internally generated variability is one that cannot be made in the real atmosphere or in the climate of a coupled model (Z96).

Generally, the specified SST–sea ice forcing was found to have significant effects on both the frequency and the intensity of precipitation, particularly in the Tropics but also over some regions in the temperate latitudes. For example, the model suggests that 10%–40% of the total interannual variability of seasonal total precipitation over South China to India and over western and southern North America may be predictable from the lower boundary conditions, especially in DJF and MAM. We are generally able to detect the effect of the external forcing in precipitation frequency more easily than in the intensity, especially over land areas. For instance, over tropical western Africa, southeastern Europe, southeastern Asia, Australia, South America, and eastern North America, the interannual variability of JJA precipitation frequency appears to be significantly affected by the external forcing, whereas little variability of the corresponding intensity can be ascribed to the forcing. Patterns of significant boundary-forced effects on precipitation diagnosed here bear some similarity to the observed relationships between ENSO and global precipitation, especially in DJF, though the AMIP precipitation seems to contain less boundary-forced signal than does the observed precipitation in JJA, especially over North America.

Potentially predictable interannual variability from internal sources is generally small. The internal sources effects on the seasonal mean intensity and seasonal total of precipitation were diagnosed to be statistically insignificant. There is a hint that there may be a weak internal sources effect on the seasonal frequency of precipitation over some land areas. The simulation effects were found to be statistically insignificant.

Acknowledgments. We thank Greg Flato for his constructive comments on an earlier version of this manuscript. We also thank the referees for their helpful comments. Research by Xiaolan L. Wang was supported by the Canadian Institute for Climate Studies via a Collaborative Research Agreement (Collaborative Research Agreement 7 CICS-VARIABILITY 4-44300) with the University of Victoria.

REFERENCES

Anderson, J. L., and W. F. Stern, 1996: Evaluating the potential predictive utility of ensemble forecasts. *J. Climate,* **9,** 260–269.

Barnston, A. G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. *J. Climate,* **7,** 1513–1564.

Bengtsson, L., M. Kanamitsu, P. Kallberg, and S. Uppsala, 1982: FGGE 4-dimensional data assimilation at ECMWF. *Bull. Amer. Meteor. Soc.,* **63,** 29–43.

Boer, G. J., N. A. McFarlane, R. Laprise, J. D. Henderson, and J.-P. Blanchet, 1984: The Canadian Climate Centre spectral atmospheric general circulation model. *Atmos.–Ocean,* **22,** 397–429.

Dix, M. R., and B. G. Hunt, 1995: Chaotic influences and the problem of deterministic seasonal predictions. *Int. J. Climatol.,* **15,** 729–752.

Everitt, B. S., 1992: *The Analysis of Contingency Tables.* 2d ed. Chapman and Hall, 164 pp.

Gates, W. L., 1992: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc.,* **73,** 1962–1970.

Huitson, A., 1966: *The Analysis of Variance.* Charles Griffin, 83 pp.

Jaeger, L., 1976: Monatskarten des Niedersclags für die ganz Erde. Ber. Dtsch. Wetterdienstes, No. 139, 38 pp.

Katz, R. W., and M. B. Parlange, 1996: Mixtures of stochastic processes: Application to statistical downscaling. *Climate Res.,* **7** (2), 185–193.

Kennedy, J. J., 1983: *Analyzing Qualitative Data: Introductory Log-Linear Analysis for Behavioral Research.* Praeger, 262 pp.

Kumar, A., and M. P. Hoerling, 1995: Prospects and limitations of seasonal atmospheric GCM predictions. *Bull. Amer. Meteor. Soc.,* **76,** 335–345.

——, ——, M. Ji, A. Leetmaa, and P. Sardeshmukh, 1996: Assessing a GCM's suitability for making seasonal predictions. *J. Climate,* **9,** 115–129.

Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in gauge-corrected global precipitation. *Int. J. Climatol.,* **10,** 111–127.

Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.,* **111,** 46–59.

Madden, R. A., 1976: Estimates of the natural variability of time-averaged sea-level pressure. *Mon. Wea. Rev.,* **104,** 942–952.

McFarlane, N. A., G. J. Boer, and M. Lazare, 1992: The Canadian Climate Centre second-generation general circulation model and its equilibrium climate. *J. Climate,* **5,** 1013–1044.

Ropelewski, C. F., and M. S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.,* **115,** 1606–1626.

——, and ——, 1989: Precipitation patterns associated with the high index phase of the Southern Oscillation. *J. Climate,* **2,** 268–284.

Rowell, D. P., 1998: Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations. *J. Climate,* **11,** 109–120.

——, and F. W. Zwiers, 1997: Sources of atmospheric decadal variability. *Preprints, Seventh Conf. on Climate Variations,* Long Beach, CA, Amer. Meteor. Soc., 87–89.

Shabbar, A., and A. Barnston, 1996: Skill of seasonal climate forecasts in Canada using canonical correlation analysis. *Mon. Wea. Rev.,* **124,** 2370–2385.

Shea, D. J., N. A. Sontakke, R. A. Madden, and R. W. Katz, 1995: The potential for long-range prediction of precipitation over the southwest monsoon season: An analysis of variance approach. Preprints, *Sixth Int. Meeting on Statistical Climatology,* Galway, Ireland, 475–477.

Sheng, J., and F. W. Zwiers, 1998: A time-interpolation scheme for bottom boundary conditions in atmospheric general circulation models. *Climate Dyn.,* **14,** 609–613.

Stern, W., and K. Miyakoda, 1995: Feasibility of seasonal forecasts inferred from multiple GCM simulations. *J. Climate,* **8,** 1071–1085.

Wilks, D. S., 1997: Resampling hypothesis tests for autocorrelated fields. *J. Climate,* **10,** 65–82.

Xie, P., and P. A. Arkin, 1996: Analysis of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate,* **9,** 840–858.

——, B. Rudolf, U. Schneider, and P. A. Arkin, 1996: Gauge-based monthly analysis of global precipitation from 1971 to 1994. *J. Geophys. Res.,* **101** (D14), 19 023–19 034.

Zwiers, F. W., 1987: A potential predictability study conducted with an atmospheric general circulation model. *Mon. Wea. Rev.,* **115,** 2957–2974.

——, 1990: The effect of serial correlation on statistical inferences made with resampling procedures. *J. Climate,* **3,** 1452–1461.

——, 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dyn.,* **12,** 825–847.

——, and V. V. Kharin, 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO_2 doubling. *J. Climate,* **11,** 2200–2222.