Acknowledgments. During the course of this work one of us (Y. Chao) was supported by a grant from the NOAA program EPOCS, and the other (S. G. H. Philander) by a NOAA TOGA Grant NA90AA-D-AC404. Computations were performed on CYBER-205 at NOAA’s Geophysical Fluid Dynamics Laboratory. Programming support from Tony Rosati, Bill Hurlin, Ron Pacanowski, and Mike Cox is gratefully acknowledged. One of us (Y. Chao) also acknowledges the support of a grant by NASA RTP 578-21-13 to Jet Propulsion Laboratory (JPL) and by JPL Contract 958658 to the University of California at Los Angeles.

REFERENCES

Simulation of the Asian Summer Monsoon with the CCC GCM-1

FRANCIS W. ZWIERS

Canadian Climate Centre, Downsview, Ontario, Canada

(Manuscript received 1 November 1991, in final form 20 May 1992)

ABSTRACT

The climate literature contains a considerable amount of indirect evidence that there is a connection between the size of the spring Tibetan snowpack and the strength of the subsequent Asian summer monsoon. This paper reports on a study that was conducted to search for evidence of a direct snow–monsoon interaction in a simulated climatology derived from two long integrations of the Canadian Climate Centre’s GCM version 1. Statistical methods based on a combination of empirical orthogonal function analysis and canonical correlation analysis were the primary investigative tools. Only a weak signal was found. It is therefore concluded that either the simulated variability of the snow on Tibet is too small, the model does not react appropriately to the simulated variability, or the true natural snow–monsoon mechanism is weak and any snow–monsoon connection relies upon a third factor. The first possibility is considered to be remote: the model simulates substantial interannual variability of Tibetan snow. The second and third possibilities are more likely. In particular, the physical mechanism that is thought to connect Tibetan snow with the Asian monsoon may not be properly simulated in the model.

1. Introduction

The purpose of this study is to look for evidence of statistical relationships between Tibetan or Eurasian surface conditions prior to the Asian summer monsoon and the subsequent strength of some of its components in a long climate simulation.

The Asian summer monsoon is characterized by a sharp increase in precipitation amount that occurs between the middle of May and the middle of June, the climatological date of onset being approximately 31 May (Das 1986). The variability of precipitation rates within the monsoon season, which extends from June to September, is large. Periods of high and low precipitation that occur during the monsoon season are referred to as “active” and “break” periods, respectively, in the classical monsoon literature. The weakening of monsoon precipitation and the collapse of the associated large-scale circulation structures, which occurs during September, is referred to as the “withdrawal” phase.

The onset of the monsoon is symptomatic of a number of large-scale circulation changes that occur in the atmosphere. Low-level anticyclonic flow develops over the Indian Ocean. This circulation advects moist air from the Arabian Sea over the Indian subcontinent and eastward across the Bay of Bengal and the Indochina Peninsula. Direct heating of the atmosphere by the Tibetan Plateau, together with the effects of increasing latent heat release, cause upper-level anticyclonic circulation to develop around the Tibetan Plateau (Murakami 1987; Krishnamurti and Sugi 1987) and low-level cyclonic circulation to develop to the east and west of the Tibetan massif. About a week prior to onset there is a poleward shift of the Northern Hemisphere westerly jet across the Himalayas and the establishment of an easterly jet to the south of the Himalayas [Murakami 1987; Tao and Chen 1987 (after Zhao et al. 1984); Mohanty et al. 1985]. This flow regime is well developed by June and is maintained through July and August before collapsing during September.

A substantial number of observational studies have examined the relationship between Tibetan or Eurasian snow cover and the strength of the subsequent Indian monsoon. Blandford (1884) was apparently the first worker to correlate increased winter snow at observing posts in the Himalayas with reduced rainfall in India. Walker (1910) extended this work by applying statistical techniques to data gathered between 1876 and 1908. He found a negative correlation between point source observations of spring Himalayan snow cover and the subsequent Indian summer monsoon rainfall. Much later, Hahn and Shukla (1976) reported on a correlation study based on 1967–1973 satellite-derived Eurasian winter snow cover extents and corresponding surface observations of June–September Indian rainfall. The results of their study support Walker’s evidence. Dey and Bhanu Kumar (1982, 1983) and Dickson (1984) report on similar studies that also sup-

© 1993 American Meteorological Society

Many previous studies have also considered the relationship between surface and lower-atmospheric heating prior to the Indian monsoon and the subsequent Indian monsoon rainfall. For example, Jagannathan and Khandekar (1962) found that geopotential heights between 700 and 400 mb over northern India during March were significantly correlated with subsequent summer monsoon rainfall over peninsular India. Mooley et al. (1986) describe significant relationships between the April 500-mb ridge location and the subsequent Indian monsoon rainfall. Shukla and Mooley (1987) use January–April Darwin pressure change and the April 500-mb ridge location as Indian monsoon rainfall predictors.

There have been a small number of studies of snow–monsoon relations in the climates simulated by general circulation models (GCMs). Published work includes that of Barnett et al. (1989, hereafter referred to as BDSR) and Yasunari et al. (1991, hereafter referred to as YKT). Barnett et al. describe an experiment that was designed to test the response of the Asian summer monsoon to changes in Asian surface properties and snowfall rates. They found that prescribed albedo changes did not have a significant impact on the simulated Asian summer monsoon; however, prescribed snow accumulation rate changes did have significant effects, characterized by weaker (stronger) monsoon precipitation during heavy (light) snow simulations. Yasunari et al. describe an experiment designed to test the response of the simulated climate to prescribed changes in the size of the March Eurasian snowpack. They found that the prescribed changes had a substantial effect on the strength of the hydrological cycle at midlatitudes during the subsequent summer. They also report a somewhat weaker Asian summer monsoon in response to an enhanced snowpack and suggest that the monsoon response is weakened by evaporation/convective feedback.

The internal mechanism that is thought to connect Tibetan (or perhaps Eurasian) spring snow mass with the subsequent Asian summer monsoon is indirect (BDSR). The Asian summer monsoon has been described as a large “sea breeze,” in which surface sensible heating of the Tibetan Plateau induces low-level convergence of moisture-laden air over southern Asia. Condensation of the moisture results in heavy precipitation and the release of large amounts of latent heat. The heat release promotes rising motion that leads to low- (high-) level convergence (divergence), thus strengthening the monsoon circulation.

A large Tibetan snow mass is thought to weaken or delay the onset of the monsoon by keeping the Tibetan surface cool and inhibiting surface sensible heating. Snow–albedo feedback first acts to delay the melting of a large snowpack. Melting, when it subsequently occurs, inhibits surface sensible heating through the absorption of the latent heat of fusion. Some of the resulting meltwater is stored in the form of soil moisture, and this water continues to inhibit Tibetan surface sensible heating through the absorption of the latent heat of vaporization. Thus it is hypothesized that Tibetan snow mass influences the strength of the Asian summer monsoon by delaying the time at which significant surface sensible heating can occur and by constraining the amount of heating that can take place.

In this study we examine such monsoon mechanisms by applying statistical methods to a 58-year GCM-simulated climatology. The model used to generate the climatology is the T20 version of the Canadian Climate Centre (CCC) GCM (GCM-1, Boer et al. 1984a,b). Such a sample is sufficiently large that it is possible to conduct statistical investigations of many naturally occurring climate mechanisms in simulations without having to perform GCM experiments with perturbed surface conditions or hydrological cycles. The 58-year dataset is longer than any observed dataset of monthly mean analyses, has complete global coverage, and is not plagued by the effects of time-dependent changes in observing and analysis systems. It thus constitutes a useful simulated climate analysis tool.

In contrast to the observational work cited above, the focus of this study, as that of BDSR, is primarily on the Southeast Asian summer monsoon rather than the Indian summer monsoon. The reason is that modest resolution GCMs are not well adapted to simulating regional variations in monsoon structure.

The paper is organized as follows. GCM-1 is briefly described in section 2. The simulated and observed data used in this study are described in section 3. Section 4 contains a description of the simulated monsoon and comparisons with the observed monsoon. The statistical techniques are discussed in section 5, the results are presented in section 6, and section 7 contains a summary and some conclusions.

2. The model

GCM-1 is fully described by Boer et al. (1984a,b) and will only be briefly described here. It has been used for a variety of purposes in a number of previous studies including Boer (1989), Boer and Laze (1988), Lambert (1987), Zwiers (1987), Zwiers and Boer (1987), and Zwiers and Hamilton (1986).

GCM-1 is a spectral model that is run at T20 truncation. There are ten unevenly spaced levels in the ver-
tical. Long- and shortwave radiative heating is explicitly calculated with both diurnal and annual cycles. Calculation of the surface fluxes of heat, moisture, and momentum, and the horizontal and vertical subgrid-scale diffusion of these quantities are incorporated in the model as is a parameterization of gravity wave drag. The physics of the model include the calculation of a surface moisture budget that incorporates snow on the surface and groundwater in liquid and frozen forms. Evapotranspiration from the surface depends upon water availability (in the form of snow and both liquid and frozen soil moisture), the proportion of the soil moisture “bucket” that is full, and the proportion of the surface that is covered with snow. Snow–albedo feedback is included in the calculation. Grid squares completely covered with snow have an albedo of 0.7. Sea surface temperatures are specified from climatology. The surface temperature and heat budget over land and sea is calculated from energy budget considerations. Zonally averaged cloudiness and ozone are specified.

3. The data

a. Simulations

The simulated climate data used in this paper was extracted from 50- and 20-year simulations conducted with GCM-1. The details of the initialization of the two simulations are sufficiently different that the two model integrations represent two independent realizations of the same climate. A 38-year pooled monthly climatology was derived from the 50-year simulation. Tape handling difficulties caused data losses that account for the discrepancy between the lengths of the simulation and the derived climatology. It is this monthly climatology that is used to describe the simulated Asian summer monsoon in section 4. The simulated climate is essentially the same as that described by Boer et al. (1984b), Zwiers and Storch (1989) have previously described other results obtained from this simulation.

A seasonal, rather than monthly, climatology was produced from the 20-year simulation. Data extracted from this run are included in our analysis of monsoon mechanisms. Previously described results obtained from this simulation include Boer (1990), Lambert (1988a), Zwiers (1987), and Zwiers and Boer (1987).

Time series of “daily” data were extracted from both simulations for a selection of prognostic and diagnostic variables at the surface and in the free atmosphere. These time series were used in this study to derive samples containing a total of 58 monthly and seasonal means. We used simulated winds, rain and snow fall amounts, ground temperature, and available groundwater to examine the model’s Asian summer monsoon and to study monsoon mechanisms.

b. Observations

We compared the model’s precipitation climatology with that of Legates and Willmott (1990). The model’s snowfall climatology was compared with those of Schutz and Bregman (1988) and Kopanev and Lipovaskaya (1978). The simulated winds were compared with a 10-year climatology derived from NMC analyses for the period 1979–1988. While not perfect (Lambert 1988b) this is one of the few available sources of global upper-air wind data.

4. The simulated monsoon

We will show below that the large-scale circulation features of the monsoon are generally well simulated by GCM-1. We will also show that the model’s precipitation field shows monsoon variations; however, the model is not able to simulate all of the precipitation changes associated with the observed monsoon because these arise from complex interactions between the large-scale flow and smaller-scale orographic features that are not well represented in our modest-resolution climate model.

a. Precipitation

Simulated eastern hemisphere monthly mean precipitation fields for April and July are displayed in Fig. 1 together with corresponding mean precipitation fields obtained from Legates and Willmott (1990). While the model simulates the large-scale features of the precipitation field with some fidelity, several prominent nonclimatological features can be identified. The largest such feature is the local maximum of precipitation that is evident in April on the slopes of the Mongolian Plateau to the east of the Tibetan Plateau. As the monsoon season develops, this feature moves from 36°N, 113°E to about 29°N, 100°E in July, when it is in fact the dominant feature of the model’s Asian summer monsoon. It remains at about the same location in August, but with reduced intensity. The feature then begins to move to the northeast again and regains its April position in October, although precipitation rates in October are much less than those in April. A second nonclimatological feature is the large local precipitation maximum centered on the northern coast of New Guinea and a third feature is the precipitation maximum centered over the Seychelles Islands in the Indian Ocean.

Despite the difficulties with GCM-1’s precipitation field, it does exhibit several monsoon features. Precipitation over India increases from less than 2 mm day$^{-1}$ in April to 4–6 mm day$^{-1}$ in June, July, August, and September (JJAS) and then decreases to about 2 mm day$^{-1}$ in October. This compares reasonably well with mean JJAS all India rainfall rates of about 7 mm day$^{-1}$ derived from 1871 to 1978 station data compiled by Mooley and Parthasarathy (1984). Precipitation over
the Malay and Indochina peninsulas increases from 4–8 mm day$^{-1}$ in April to 8–12 mm day$^{-1}$ in May and June. This compares reasonably well with Legates and Willmott although the observed dataset shows small-scale structures with large precipitation rates that the model is not able to simulate. In July, however, simulated precipitation over Indochina decreases relative to June while observed precipitation rates in the area increase. The decrease is related to movement to the southwest of the large precipitation area over China. The intensity and geographical extent of this simulated feature roughly matches that observed in July over the southeastern flanks of the Tibetan Plateau. Simulated precipitation rates remain low over the Indochina Peninsula in August and September (not shown), and the precipitation maximum over the eastern slopes of the Tibetan Plateau again moves to the northeast and regains a distinct nonclimatological character.

To the extent that it is possible to judge, snowfall is reasonably well simulated by the model. Comparison of October, December, and February mean maps (not shown) with those of Kopanev and Lipovaskaya (1978; Figs. 1–3) and Schutz and Bregman (1988) reveals that the snow depth simulated in early winter (October and December) is less than that observed. February snow depth compares reasonably well with climatology except over the Tibetan Plateau. April snow depth (Fig. 2) is generally greater than observed. This is also the case in May. The simulated snow depth is derived from simulated snow mass using the snow mass–snow density relationship described by McFarlane et al. (1991).

We speculate that many of the discrepancies between the observed and simulated snow depth relate to the model's land-surface processes package. In early winter snowfall does not accumulate on the ground until the

---

**Fig. 1.** Monthly mean precipitation rate (mm day$^{-1}$). Cross hatching indicates regions with precipitation rates greater than 6 mm day$^{-1}$. (a) April GCM-1. (b) April observed (Legates and Willmott 1990). (c) As in (a) except for July. (d) As in (b) except for July.

**Fig. 2.** April mean snow depth (cm). Cross hatching indicates snow depth greater than 40 cm. (a) GCM-1. (b) Observed (Shutz and Bregman 1988).
underlying soil is frozen. This takes considerable time because the soil is represented as a single layer. In late winter snowmelt is delayed over much of Eurasia because the surface processes model does not correctly simulate the effect of vegetation masking in the region. Snow accumulation is also sustained later into the season than observed because the entire soil layer must melt before the 0°C ground-temperature line can migrate north.

b. Low-level circulation

The characteristics of the low-level circulation of the model and NMC analyses in the monsoon region are displayed in Fig. 3 in terms of the mean July 850-mb winds. The model has a well-developed, if somewhat weak, monsoon circulation in July. The easterly trades across the south Indian Ocean and the anticyclonic circulation around the Arabian Sea is well represented in the model, but mean wind speeds are less than observed. The maximum simulated wind speed over the Arabian Sea is slightly less than 10 m s⁻¹ whereas observed speeds exceed 20 m s⁻¹.

The seasonal development of the monsoon circulation in the model is similar to that observed. Characteristic features of the monsoon circulation, such as the anticyclonic flow over the Arabian Sea and the low-level westerlies that blow across the Indian and Indochina peninsulas first appear in the May mean circula-

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 4. Mean July zonal wind averaged zonally between 60°E and 90°E (m s⁻¹). Positive values indicate westerly winds. Cross hatching indicates wind speed greater than 20 m s⁻¹. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 4. Mean July zonal wind averaged zonally between 60°E and 90°E (m s⁻¹). Positive values indicate westerly winds. Cross hatching indicates wind speed greater than 20 m s⁻¹. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)

![Fig. 3. Mean July 850-mb winds (m s⁻¹). Contours are drawn at 5 m s⁻¹ intervals. (a) GCM-1. (b) Observed (NMC).](image)
parable to that observed but it is not properly closed. The easterly jet is well positioned in the model at about 125-mb but is about 5 m s$^{-1}$ weaker than observed.

The development in time of the upper-level circulation is somewhat different than observed. This can be seen from Fig. 5 which shows the mean evolution with time of the zonal wind at 200 mb averaged between 60°E and 90°E. The NH westerly jet shifts northward at about the same time and rate in both the simulated and observed climates: the mean 20 m s$^{-1}$ contour crosses 25°N on 23 May in the simulated climate and does so on 24 May in the observed climate; however, the observed climate generally has easterlies on the equator for all of May and June while such easterlies often do not appear in the simulated climate until June. The mean date of onset, as determined by

![Fig. 5. Evolution of zonal wind (m s$^{-1}$) at 200 mb between 1 May and 30 June. The wind is averaged zonally between 60° and 90°E. Positive values indicate westerlies. Narrow cross hatching indicates westerlies stronger than 30 m s$^{-1}$. Wide cross hatching indicates easterlies stronger than 10 m s$^{-1}$. (a) GCM-L (b) Observed (NMC).](image)

the date at which mean 200-mb zonal wind becomes easterly at 15°N, is approximately two weeks later in the simulated climate (7 June) than in the observed climate (23 May).

The model does a good job of simulating the observed variability of the upper-level monsoon circulation. Observed and simulated interannual variability of the 200-mb zonal wind (not shown) are very similar. Variation in the date of onset is also similar. To estimate this variability, an alternate subjective determination of the date of onset was made. Individual Hovmöller diagrams of the evolution of the zonal wind for each year were examined to determine the date at which “sustained” easterly 200-mb zonal winds appeared at 15°N. According to this analysis, the mean date of onset in the observed climate is 19 May and that in the simulated climate is 31 May. In both cases, the standard deviation of the date of onset is 8.7 days.

d. Available water

The monsoon control mechanism hypothesized by BDSR and others involves snow–albedo feedback and the subsequent evaporation of meltwater carried over as soil moisture. Evidence of this type of control mechanism will be found only in the simulated climate if there is significant interannual variation of the total available water, defined as the sum of accumulated snow plus liquid and frozen groundwater. Figure 6 illustrates the mean and standard deviation of the total available water in April in the simulated climate. In April (Fig. 6a) there are substantial amounts of avail-
able water on the highest parts of the Tibetan region and most of the northern half of Eurasia. Corresponding variability (Fig. 6b) is large in the Tibetan region and over the Central Siberian Plateau. In June (not shown) there is less than 10 cm of available water over most of the Tibetan region. Northwestern China is an exception because it is affected by monsoon precipitation in the simulated climate.

By July (not shown), the carry over throughout the Tibetan region not affected by monsoon precipitation is reduced, on average, to less than 5 cm of water. At this depth moisture evaporates from the soil moisture “bucket” incorporated into GCM-1 at no more than 50% of full potential. The variability of soil moisture is correspondingly small. Consequently, we might speculate that the indirect monsoon control mechanism as postulated will not impose strong control over the simulated monsoon because of the limited potential for large interannual variations of evaporative cooling of the Tibetan surface concurrent with the monsoon.

5. Analysis strategy

The relationships that we examine are related to the monsoon control mechanism discussed in the Introduction. Recall that the postulated control mechanism is indirect: it involves the accumulated Tibetan (or Eurasian) snowpack, the subsequent snowmelt and carry over of melt water as soil moisture, the reduction of surface heating induced by evaporative cooling of the surface, and a consequent weakening of the monsoon.

The most direct part of the postulated control process is the latter component that involves surface heating. The connection between surface heating and monsoon precipitation should be more or less concurrent. If the mechanism operates in GCM-1 one might expect to see a strong statistical relationship between quantities such as June, July, and August (JJA) mean surface heating over Tibet (or Eurasia) and JJA mean Asian monsoon precipitation. Surface heating of the atmosphere is determined by surface heat flux, latent heat release near the surface, and direct radiative heating of the atmosphere near the surface. Surface heat flux (HFS) is by far the largest of these components. We thus looked for evidence of the most direct part of the monsoon control process by looking for statistical connections between JJA Tibetan (or Eurasian) HFS and concurrent JJA Asian monsoon precipitation.

The presence of a statistical relationship between soil moisture in Tibet (or Eurasia) and the subsequent strength of the monsoon would be an indication that a substantial part of the postulated indirect control mechanism operates in the climate of GCM-1. On a monthly time scale one might expect to see the strongest connection between May Tibetan (or Eurasia) soil moisture and June Asian monsoon precipitation because variability of soil moisture in May may still be affected by the size of the spring snowpack. We therefore also looked for evidence of a statistical connection between these fields. Note that it is not possible to search for a concurrent relationship between available water and monsoon precipitation because the two variables are strongly connected in those parts of the Tibetan (or Eurasian) region that are affected by monsoon precipitation.

Finally, the presence of a statistical relationship between Tibetan (or Eurasian) snow mass in late spring and summer Asian monsoon precipitation would be a strong indication that the entire indirect control mechanism operates in the climate of GCM-1. April has the greatest extent and interannual variability of monthly mean Eurasian snow accumulation in the climate of GCM-1. Thus one might anticipate that the strongest statistical relationship between snow and precipitation on a monthly time scale exists between April Tibetan (or Eurasian) snow and June Asian monsoon precipitation. We therefore look for a statistical connection between these fields.

The main quantities and regions considered in our analysis of the mechanisms previously discussed are:

- Precipitation in the part of Southeast Asia that contains most of the simulated monsoon activity. This region, designated A in Fig. 7, extends from 90° to 124°E and from 8° to 41°N. The Indian region is excluded because the Indian summer monsoon is not well simulated by GCM-1 and because the simulated interannual variability of precipitation is this region is small. Observational studies, such as that of Lau and Li (1984), suggest that the East Asian monsoon may be governed by something other than a simple snow–monsoon connection; however, the response obtained by BDSR in their experiments with a similar GCM is largely confined to the region considered in this study.
- Surface heat flux over Tibet and Eurasia. Tibet (region B in Fig. 7) is defined to extend from 62° to 113°E and from 25° to 53°N. Eurasia (region C in

![Fig. 7. The FGGE surface topography truncated to 120 resolution. The units are m x 10^-2. Regions A, B, and C outline areas used in the CCA of simulated monsoon relationships described in the text.](image)
Fig. 8) is defined as the region extending from 17° to 141°E and from 25° to 64°N.
- Available soil moisture and snow mass in Tibet and Eurasia.
- A vertical cross section of zonal winds between the equator and 45°N which is averaged between 60° and 90°E.

The principal statistical tool used in our analysis is canonical correlation analysis (CCA, see appendix A). CCA is used to search for linear statistical relationships between pairs of climatic fields such as Tibetan surface heat flux and Southeast Asian precipitation.

The complete data analysis process included a number of steps, however. First, a pair of fields were selected for study as previously discussed. Second, the dimensionality of the chosen fields was drastically reduced using empirical orthogonal functions (EOFs). The first five EOFs were retained in all cases. It was felt that this dimension reduction would retain information on a regional scale that might be important relative to the monsoon but would filter out small-scale information that could cloud any regional relationships that might be present. The dimension reduction is also necessary to facilitate the canonical correlation analysis.

The spectra obtained from the EOF analysis are summarized in Table 1. The first five EOFs represent more than 50% of total variance in every instance except JJA Eurasian HFS and May Eurasian available water. In the latter cases the first five EOFs represent about 35% and 40% of total variance, respectively. The proportion of variance explained by the EOFs depends upon both size of the region and spatial scale of the variability. Given two fields defined in the same region, the spectrum for the noisier field will be flatter and a fixed number of EOFs will explain less of the total variability.

Third, the five-dimensional fields resulting from the EOF analysis were analyzed using CCA. The CCA produces pairs of five-dimensional canonical patterns that represent modes of variation of the EOF coefficients that are optimally correlated. The canonical patterns can be interpreted as sets of EOF coefficients. These coefficients are used to recombine the EOFs to obtain pairs of spatial patterns that can then be interpreted as modes of variation of the analyzed fields that are optimally correlated. These patterns, linear combinations of the EOFs, are referred to as spatial canonical patterns (SCPs) in this paper. Corresponding to each pair of canonical patterns is a pair of coefficient time series that represent the variation of the amplitudes of the canonical patterns with time. These coefficients are referred to as canonical variables (CVs).

6. Results

a. JJA mean surface heat flux and precipitation

Relationships between HFS and Southeast Asian precipitation were analyzed by correlating Tibetan and Eurasian JJA mean HFS with Southeast Asian JJA mean precipitation. The mean and interannual standard deviation of JJA Eurasian HFS is displayed in Fig. 8. Note the large interannual variability in Southeast Asian coastal regions, the southern part of the Tibetan region, and the Kazakh Uplands (Fig. 8b).

There are strong statistical connections between concurrent HFS and precipitation. The largest canonical correlation of Tibetan (Eurasian) HFS with precipitation is 0.76 (0.80) when bias corrected. This correlation is significantly different from zero. The 95% confidence interval for the first canonical correlation of Tibetan HFS with precipitation extends from 0.62 to 0.85. The two CV pairs (Tibetan or Eurasian) extract essentially the same information about the relationship between HFS and Southeast Asian precipitation. The Tibetan spatial canonical pattern (SCP, see Fig. 9a) is very similar to the Eurasian SCP (not shown). The same is true of the corresponding precipitation SCPs generated by the two analyses. Apparently information about the Tibetan surface state accurately reflects this component of the monsoon control mechanism. We will therefore discuss only the Tibetan analysis in the remainder of this section.

The SCP for precipitation (Fig. 9b) explains approximately 16% of the total JJA precipitation variation. The pattern shown in Fig. 9b is similar to the second EOF of the Southeast Asian precipitation field.
This finding is also supported by composite analyses conducted using both the HFS and precipitation CVs as selection keys. For example, Fig. 12a (12b) illustrates JJA mean precipitation for the 6 years with the highest (lowest) values of the HFS CV. High values of the CV correspond to high heating rates in Tibet and vice versa. Monsoon precipitation is organized as a single large feature over southern China and the Indochina peninsula during high heating rate years (Fig. 12a) and is reduced over much of this area during low heating rate years (Fig. 12b). The difference between the two composites (Fig. 12c) is significant in a region over southern China, where precipitation increases during high-HFS years, and in a region on the eastern flanks of the Tibetan massif, where precipitation decreases during high-HFS years. The difference between the corresponding HFS composites (not shown) shows the general structure of the HFS SCP.

Some features of the moisture budget and the monsoon circulation also change in response to changes in the HFS CV.

(which explains 15% of total variation). Thus the mode of variation illustrated in Fig. 9b is one of the dominant modes of precipitation variation.

The HFS SCP (Fig. 9a) has a positive center of action situated over central Tibet and a negative center situated over central China. The latter is an indication that the relationship described by the 0.76 canonical correlation is partly confounded with the direct local cooling effect of precipitation. That is, increased surface heat flux over Tibet corresponds to reduced surface heat flux and increased precipitation in the monsoon region. The corresponding precipitation SCP (Fig. 9b) is characterized by a large center of action situated over southwestern China flanked by a pair of smaller centers of opposite sign.

The scatter diagram shown in Fig. 10 illustrates the statistical relationship between the CVs that corresponds to the SCPs of Fig. 9. It clearly indicates that in the simulated climate, positive surface heating anomalies in the Tibetan region are associated with enhanced monsoon precipitation in southern China where the primary feature of the model's rendition of the Asian summer monsoon is situated.

It can also be shown that the precipitation mode of variation depicted in Fig. 9b captures at least part of the monsoon variability simulated by the model. This is illustrated in Fig. 11 which shows the annual cycle of the mean and interannual standard deviation of the pattern coefficient. The mean precipitation coefficient (Fig. 11a) is elevated above the background level during the monsoon season and large variability is sustained in this mode during this period (Fig. 11b). Together, these results support the hypothesis that monsoon variability is at least partially controlled by Tibetan surface conditions.
The vertically integrated moisture budget satisfies
\[
\frac{\partial w}{\partial t} + \nabla \cdot Q = E - P, \tag{1}
\]
where the precipitable water \( w \) is given by
\[
w = \int_{p_0}^{p} \frac{\beta d p}{g}. \tag{2}
\]
Following Boer (1982), the quantity \( p_0 \) in (2) has a constant value that is always greater than the surface pressure \( p_s \), and the quantity \( \beta \) is a step function defined by
\[
\beta = \begin{cases} 
1, & \text{if } p \leq p_s \\
0, & \text{if } p > p_s.
\end{cases} \tag{3}
\]
Assuming that the mean time rate of change for precipitable water is negligible, (2) reduces to
\[
\nabla \cdot \bar{Q} = \bar{E} - \bar{P}. \tag{4}
\]
That is, time-averaged moisture convergence is approximately balanced by the time-averaged difference between evaporation and precipitation.

We computed the monthly mean moisture convergence potential function and flux vectors from the simulated difference between evaporation and precipitation using the method outlined by Boer and Sargent (1985). The mean JJA moisture flux and potential function for the six high heating rate years is displayed in Fig. 13a. This diagram, which is similar in structure to that of the 58-year mean, shows the net moisture sources and sinks in the “tropical” half of the eastern hemisphere. It is reasonable to expect the structure of the simulated moisture flux field to be similar to that of the observed mean energy flux because moisture and energy fluxes are closely related in the tropics (moisture flux in the tropics is approximately proportional to the inverse of energy flux). This is indeed the case as can be seen by comparing Fig. 13a with the mean divergent component of the July energy budget derived from FGGE data by Boer and Sargent (1985, Fig. 6b).

---

**Fig. 10.** The canonical pattern pair coefficients for the canonical correlation between Tibetan JJA HFS and Southeast Asian JJA precipitation.

**Fig. 11.** The annual cycle of the mean and interannual standard deviation of the precipitation canonical-pattern coefficient derived from the correlation of Tibetan JJA HFS with Southeast Asian JJA precipitation. The coefficient is normalized so that its JJA mean has zero mean and unit variance. (a) Mean. (b) Standard deviation.
The difference between the divergent part of the simulated moisture flux for the six greatest and least surface heating years is displayed in Fig. 13b. We see that southern China is a relative moisture sink during years with high rates of Tibetan surface heating. These results also support the hypothesis that greater Southeast Asian summer precipitation is associated with greater Tibetan surface heating of the atmosphere.

The mean zonal wind, averaged between 60° and 90°E, for the six years with greatest surface heating, is displayed in Fig. 14a and the difference between the six greatest and least surface heating years is illustrated in Fig. 14b. The modest westerly and easterly jets in the difference field, located at 30° and 15°N, respectively, indicate that there is some strengthening of the upper-level monsoon circulation in the simulated climate during years in which there is increased Tibetan surface heating. This is accompanied by a strengthening of the westerly flow across the Indian subcontinent and surrounding water near the surface.

Fig. 12. Composite analysis of Southeast Asian JJA precipitation using the Tibetan JJA HFS canonical pattern coefficient as the selection key. (a) Composite mean of the six simulated years with the highest coefficient values (mm day⁻¹). Precipitation rates in excess of 10 mm day⁻¹ are shaded. (b) As (a) but for the six simulated years with the lowest coefficient values. (c) Difference between (a) and (b) (mm day⁻¹). Differences significant at the 5% level according to a two-sided t test are shaded.
In summary, the first Tibetan HFS–Southeast Asian precipitation canonical correlation contains strong evidence that there is direct, natural control of at least some portion of the simulated Asian summer monsoon by surface heating on the Tibetan Plateau. Virtually identical information was extracted in correlations that used Tibetan and Eurasian HFS suggesting that it is regional changes on the Tibetan Plateau that affect the monsoon rather than broader-scale changes in surface heating from the Eurasian landmass. The evidence for the connection was strengthened by the fact that the Tibetan HFS signal was also apparent in other monsoon related quantities such as moisture convergence and the upper-air winds.

The CCA of Tibetan (Eurasian) HFS and Southeast Asian precipitation produced a second canonical correlation that was nominally significantly different from zero. This correlation is not discussed because we have substantially less confidence in it than the first correlation.

b. May available water and June precipitation

The CCA was repeated with Eurasian (and Tibetan) May available water and Southeast Asian June precipitation. Available water is generally large in May on the western half of the Tibetan massif (where the effects of the accumulated snowpack are felt) and on its eastern flank (which is affected by the precipitation maximum that is perennially situated over China). Only a small portion of the Tibetan massif enjoys substantial available water variability. Thus only this region has the potential to control the Asian summer monsoon.

There is some weak statistical evidence of a connection between Eurasian (primarily Tibetan) May available water and Southeast Asian June precipitation. The largest canonical correlation, 0.38 (bias corrected), may be significantly different from zero. Its nominal 95% confidence interval is (0.13, 0.58). The SCPs that correspond to the first CV pair are illustrated in Figs. 15a,b.

The main features of the available water SCP (Fig. 15a) are a positive center of action on the western flank of the Tibetan massif and a negative center over its central region. This localized mode of variation rep-
resents 7% of the total variation of May available water in region C. The corresponding precipitation SCP (Fig. 15b) contains positive centers of action over central China and the northern Malaysian Peninsula. It represents 11% of the total June precipitation variability in region A. Figures of the mean and standard deviation of the coefficient of the precipitation SCP (not shown) show that the mode of variability illustrated in Fig. 15b captures some of the features of the model's monsoon climate over Southeast Asia.

There is other evidence that suggests that the connection between available water and precipitation is real. For example, similar SCPs are obtained when Tibetan May available water is correlated with June precipitation. Also, the derived SCPs are evident in composite analyses of the simulated available water and precipitation data; however, the composite differences were not found to be significant.

In summary, the statistical evidence tentatively supports the hypothesis that indirect processes that constrain Tibetan surface heating during the monsoon season act to reduce the strength of the monsoon. It is evident from both the canonical correlation analysis and the related composite analyses that increased available water in the central Tibetan region in May is related to reduced monsoon precipitation over China in June. We conjecture that the signal is weak (i) because May available water does not completely determine available water in subsequent months, and (ii) because there is only limited potential for the control of the summer monsoon via evaporative cooling of the Tibetan region due to the low variability of available water in the region.

c. April snow mass and June precipitation

As expected from our discussion of available water variability, the evidence for a connection between April Asian snow mass and subsequent June monsoon precipitation is weak in the climate of GCM-1. None of the canonical correlation interval estimates in our analysis of Tibetan April snow mass and Southeast Asian June precipitation excludes zero, indicating that no correlation is significantly different from zero. The largest canonical correlation in this case is 0.18 (bias corrected). Similar results were obtained when we correlated Eurasian April snow mass with June precipitation.

Some observational studies, such as that of Khanka- dekar (1991), obtain larger correlations between December–February mean Eurasian snow cover and the subsequent June–September Indian monsoon precipitation than they do between shorter lag correlations such as that between March–May mean Eurasian snow cover and the subsequent Indian monsoon. We repeated our analysis using simulated December–February mean snow mass and JJA precipitation, but did not find a significant correlation.

7. Summary and conclusions

GCM-1 does a credible job of simulating the large-scale features of the observed Asian summer monsoon given its modest resolution. The model's performance is less satisfactory in its simulation of precipitation. The observed Asian summer monsoon is characterized by increased precipitation over India, all of Southeast Asia, and Indonesia. In contrast, the primary feature of the simulated monsoon is a large precipitation maximum located over China that intensifies and moves to the Southeast as the monsoon season develops. Our diagnostics of monsoon control mechanisms focused on the region that contains this feature of the simulated precipitation field. A recently improved version of the model with higher resolution, interactive clouds, and an interactive slab ocean ameliorates many of the problems associated with the simulated monsoon region precipitation.

There is relatively strong evidence that concurrent Tibetan surface heating does affect Southeast Asian
monsoon precipitation. The dominant CV pair lends support to the idea that enhanced surface heating of the atmosphere over the Tibetan Plateau results in enhanced Southeast Asian monsoon precipitation. The same signal was seen in composite analyses of precipitation, moisture flux, and zonal wind.

There is only weak evidence to support the hypothesis that the Asian summer monsoon is controlled by internal mechanisms involving Tibetan spring snow mass in the climate simulated by GCM-1. A weak, nominally statistically significant correlation was found between May available water and June precipitation, suggesting that at least part of the hypothesized indirect monsoon control mechanism operates in the simulated climate; however, a significant relationship was not found between the simulated April snowpack and subsequent June precipitation.

It was argued that the simulated snowpack itself could not greatly influence the monsoon because (i) the part of the Tibetan region affected by large interannual snowpack variability is small; and hence (ii) snowpack variability does not result in large variability of available water carryover into the monsoon season. The lack of water carryover implies that the spring snowpack cannot strongly influence the monsoon through the cooling of the Tibetan surface in summer by the evaporation of moisture carried over from the melting of the snowpack. Apparently noise masks whatever signal may be present.

The results of this study contrast sharply with those of BDSR and YKT. Both of these experiments showed evidence of a snow–monsoon connection. The former report a particularly strong signal, probably because the forcing imposed in their experiment was stronger than that used in the YKT experiment.

Soil moisture carryover seems to be the key difference. Yasunari et al. documented substantial changes in the heat and water balance over Eurasia, which take place in two stages in the climate of the MRI-GCM. In the early stage (April), the albedo effects of enhanced snow cover dominate and result in a reduction of the total heating of the atmospheric column and a snow storage effect. In the latter stage (August), there is relative cooling of the Tibetan surface due to evaporation of soil moisture. Apparently, increased precipitation maintains ground wetness during the summer months after the spring snowpack has been increased. Barnett et al. also note that substantial amounts of soil moisture are carried over into summer when the spring snowpack is enhanced, and that this (and the “protective” albedo effect of the snowpack) is fundamental to the modulation of the monsoon. They do not discuss the mechanism that sustains the soil moisture.

In contrast to BDSR and YKT, there is little evidence that soil moisture is effectively carried over within the natural (i.e., unperturbed) workings of GCM-1. This is probably due to the model’s lack of interactive clouds and deficiencies in its surface processes module. Albedo feedbacks that result from interactive clouds likely play an important part in the maintenance of ground wetness. The surface processes module, which contains a simple one-layer soil model and treats snow as part of the soil layer, does not store meltwater effectively. Consequently, the effect of snowpack variations may be reduced to that of the corresponding albedo variations alone. Barnett et al. showed that albedo alone did not have a profound effect in their experiments.

Acknowledgments. I would like to thank George Boer for suggesting this study and for useful discussion of this work. I would like to thank M. L. Khedkar and H. von Storch for their useful comments on the manuscript.

APPENDIX A

Canonical Correlation Analysis

Canonical correlation analysis (CCA), first developed by Hotelling (1936), is employed to uncover linear relationships between pairs of multidimensional random variables, such as wind and precipitation fields, which are each defined at a number of grid points and possibly a number of vertical levels. For example, in the case of zonal wind and precipitation fields, the purpose of CCA is to find the linear wind index (i.e., combination of grid points) and linear precipitation index that are most strongly correlated with each other.

From an algebraic point of view, wind field observations can be thought of as realizations of a random vector $\mathbf{U}$ in an $l \times m$ dimensional vector space (where $l$ is the number of levels and $m$ is the number of grid points at each level). Similarly, precipitation field observations can be regarded as realizations of a random vector $\mathbf{P}$ in a $k$ dimensional vector space (where $k$ is the number of precipitation grid points). The purpose of canonical correlation analysis is to find the wind pattern $a_i$ and the precipitation pattern $b_i$ for which the corresponding coefficients $a_i^T \mathbf{U}$ and $b_i^T \mathbf{P}$ are most strongly correlated. Having found patterns $a_i$ and $b_i$, one proceeds to obtain a second set of patterns $a_2$ and $b_2$ such that coefficients $a_2^T \mathbf{U}$ and $b_2^T \mathbf{P}$ are most strongly correlated subject to the constraint that they are not correlated with coefficients $a_i^T \mathbf{U}$ and $b_i^T \mathbf{P}$. These coefficients are referred to as canonical variables (CVs).

CCA is similar to empirical orthogonal function (EOF) analysis. In EOF analysis the object is to decompose a random field into a collection of uncorrelated coefficients (often called the principal components) that optimally capture the variability of the field. Corresponding to each coefficient is a field (the EOF) that describes the mode of variation represented by that coefficient. In CCA the object is to decompose a pair of random fields into pairs of random coefficients.
(the CVs) that optimally capture the (linear) covariability of the two fields. Corresponding to each CV pair is a pair of fields (which we will call canonical patterns) that describe the modes of variation in the random fields under study that are most closely related to each other. A brief technical description of CCA follows. More complete descriptions are given by Press (1982) and Kendal et al. (1983). Examples of climatological applications of CCA include Karl et al. (1989) and others cited in that paper.

a. Technical description of CCA

For the sake of brevity, let us suppose that we have observed \( n \) pairs of realizations of an \( l \)-dimensional random vector \( \mathbf{U} \) and a \( k \)-dimensional vector \( \mathbf{P} \). Without loss of generality, assume that \( l < k \). We will denote the observations by the \((l + k)\)-dimensional vectors

\[
\begin{pmatrix} \mathbf{U}_i \\ \mathbf{P}_i \end{pmatrix}, \quad i = 1, \ldots, n. \tag{A.5}
\]

Assume that the \((l + k)\)-dimensional random vector

\[
\begin{pmatrix} \mathbf{U} \\ \mathbf{P} \end{pmatrix}
\]

has a variance–covariance matrix \( \Sigma \) given by

\[
\begin{pmatrix} \Sigma_{UU} & \Sigma_{UP} \\ \Sigma_{PU} & \Sigma_{PP} \end{pmatrix}. \tag{A.6}
\]

The problem at hand is to find the linear combinations

\[
\mathbf{U}_i = a'_i \mathbf{U} \quad \text{and} \quad \mathbf{P}_i = b'_i \mathbf{P} \tag{A.8}
\]

that are most strongly correlated. That is, we wish to maximize the correlation \( \rho_1 \) given by

\[
\rho_1 = \rho(U_1, P_1) = a'_1 \Sigma_{UU} a_1 / \left[ (a'_1 \Sigma_{UU} a_1)(b'_1 \Sigma_{PP} b_1) \right]^{1/2}. \tag{A.9}
\]

The correlation between \( \mathbf{U}_1 \) and \( \mathbf{P}_1 \) is independent of their scaling. It is therefore necessary to constrain the linear combinations so that

\[
\text{var}(U_1) = a'_1 \Sigma_{UU} a_1 = 1
\]

and

\[
\text{var}(P_1) = b'_1 \Sigma_{PP} b_1 = 1. \tag{A.10}
\]

Thus, the problem is to maximize (A.9) subject to (A.10). Press (1982) shows that this reduces to the eigenproblems

\[
\begin{align*}
(\Sigma_{UP} \Sigma_{PP}^{-1} \Sigma_{PU} - \lambda^2 \Sigma_{UU}) a &= 1, \\
(\Sigma_{PU} \Sigma_{UU}^{-1} \Sigma_{UP} - \lambda^2 \Sigma_{PP}) b &= 1, \tag{A.11}
\end{align*}
\]

which have \( l \) identical nonzero solutions for \( \lambda^2 \). The square root of the largest eigenvalue \( \lambda_1 \) maximizes (A.9) and the corresponding eigenvectors \( a_1 \) and \( b_1 \) define the required CVs.

It quickly follows that the \( i \)th pair of CVs, defined as the \( \mathbf{U} \) and \( \mathbf{P} \) coefficients, respectively, which are most strongly correlated subject to the constraints that

\[
\rho(U_i, U_j) = 0, \quad \text{for} \quad j < i
\]

\[
\rho(P_i, P_j) = 0, \quad \text{for} \quad j < i
\]

\[
\rho(U_j, P_m) = 0, \quad \text{for} \quad j < i, m < i, j \neq m \tag{A.12}
\]

have correlation \( \lambda_{ij} \), where \( \lambda_i^2 \) is the \( i \)th largest eigenvalue of (A.11). These CVs are given by

\[
U_i = a'_i \mathbf{U} \quad \text{and} \quad P_i = b'_i \mathbf{P}, \tag{A.13}
\]

where \( a_i \) and \( b_i \) are the corresponding eigenvectors.

In practice one does not know the true joint variance–covariance matrix \( \Sigma \). This is overcome by substituting the corresponding maximum likelihood estimate

\[
\hat{\Sigma} = \begin{pmatrix} \hat{\Sigma}_{UU} & \hat{\Sigma}_{UP} \\ \hat{\Sigma}_{PU} & \hat{\Sigma}_{PP} \end{pmatrix}, \tag{A.14}
\]

where \( \hat{\Sigma}_{UU} \) is given by

\[
\hat{\Sigma}_{UU} = \frac{1}{n} \sum_{i=1}^{n} (U_i - \bar{U})(U_i - \bar{U})' / n.
\]

The overbar indicates averaging; \( \hat{\Sigma}_{UP}, \hat{\Sigma}_{PU}, \) and \( \hat{\Sigma}_{PP} \) are computed analogously. Then, if \( \mathbf{U} \) and \( \mathbf{P} \) are jointly Gaussian, the substitution of (A.14) into (A.11) will produce maximum likelihood estimates of the canonical correlations and the corresponding canonical patterns.

b. Inference about canonical correlation estimates

While several methods exist to make inferences about the canonical correlations, the literature seems to be devoid of results concerning the stability of the corresponding canonical patterns. One possible approach to the question of pattern stability, which we do not pursue in this paper, is to use bootstrapping techniques [see Zwiers (1990) for a review] to make a rough estimate of the sampling variability of the canonical patterns.

A frequently used method of inference is to test a sequence of null hypotheses that the last \( l - m + 1 \) canonical correlations are zero; that is, tests of

\[
H_0: \rho_m = \rho_{m+1} = \cdots = \rho_l = 0 \tag{A.15}
\]

are conducted sequentially for \( m = 1, 2, \ldots, l \) until \( H_0 \) is accepted. The test used was proposed by Bartlett (1947). It is based on the statistic
\[-\left\{ n - 1 - m - \frac{1}{2} (l + k - 1) \right\} + \sum_{i=1}^{l} r_i^{-2} \ln \left[ \prod_{j=m+1}^{l} \left( 1 - r_j^2 \right) \right], \quad (A.16)\]

where $r_1, \ldots, r_l$ are the estimated canonical correlations. This statistic is approximately distributed as a Chi-squared random variable with $(l - m) \times (k - m)$ degrees of freedom under the null hypothesis. Note that the combined significance level of an inference based on a sequence of such tests will be less than the significance level of inferences made with individual tests. Note also that this method of inference is similar to the asymptotic principal component selection rules discussed in Priesendorfer et al. (1981).

Muirhead and Watermann (1980) have shown that statistics like (A.16) are not particularly robust against departures from the assumption that $U$ and $P$ are jointly distributed as multivariate Gaussian random vectors. This should not be a serious consideration in the present study because the precipitation data are obtained from regions in which there is frequent precipitation and because all the observations consist of monthly means. Thus, we can appeal to the law of large numbers for both precipitation and nonprecipitation data to be assured that they have Gaussian properties. Also, the data are obtained from a simulated climate and are thus not subject to the coding and data-handling errors that might compromise a nonrobust statistic computed from real observations.

An alternative method of inference used in this paper is to construct interval estimates of the canonical correlations that take sampling variability into account. Glynn and Muirhead (1978) show that the estimator

$$\hat{\theta}_i = Z_i - \frac{1}{2n_i} \left[ l + k - 2 + r_i^2 \right] + 2(1 - r_i^2) \sum_{j \neq i}^{n} \frac{r_j^2}{(r_j^2 - r_i^2)} \quad (A.17)$$

is unbiased for $\theta_i$ up to $O(n^{-2})$, where $\theta_i = \tanh^{-1}(\hat{\theta}_i)$ and $Z_i = \tanh^{-1}(r_i)$. They also show that $\text{var} (\hat{\theta}_i) = 1/n$ up to $O(n^{-2})$. We therefore construct an approximate bias-corrected 95% confidence interval for the $i$th canonical correlation by computing the interval \[\left[ \tanh(\hat{\theta}_i - 2/\sqrt{n}), \tanh(\hat{\theta}_i + 2/\sqrt{n}) \right].\] The results of some preliminary Monte Carlo experiments we have conducted with this statistic suggest that while improved, substantial bias can remain even after adjustment when samples are small.

REFERENCES


---, and ---, 1984b: The climatology of the Canadian Climate Centre General Circulation Model as obtained from a five-year simulation. *Atmos.–Ocean*, 22, 430–473.


