

Statistical Procedures for Making Inferences about Precipitation Changes Simulated by an Atmospheric General Circulation Model

RICHARD W. KATZ

Department of Atmospheric Sciences, Oregon State University, Corvallis, 97331, and Environmental and Societal Impacts Group, National Center for Atmospheric Research,¹ Boulder, CO 80307

(Manuscript received 7 February 1983, in final form 16 May 1983)

ABSTRACT

A statistical methodology is presented for making inferences about changes in mean daily precipitation from the results of general circulation model (GCM) climate experiments. A specialized approach is required because precipitation is inherently a discontinuous process. The proposed procedure is based upon a probabilistic model that simultaneously represents both occurrence and intensity components of the precipitation process, with the occurrence process allowed to be correlated in time and the intensities allowed to have a non-Gaussian distribution. In addition to establishing whether the difference between experiment and control daily means is statistically significant, the procedure provides confidence intervals for the ratio of experiment to control median daily precipitation intensities and for the difference between experiment and control probabilities of daily precipitation occurrence. The technique is applied to the comparison of winter and summer precipitation data generated in a control integration of the Oregon State University atmospheric GCM.

1. Introduction

A statistical methodology will be presented for making inferences about changes in mean daily precipitation from the results of general circulation model (GCM) climate experiments. Unlike many atmospheric variables such as temperature that can be viewed as continuous processes, precipitation is inherently a discontinuous process, either occurring or not occurring. The precipitation process can be regarded as consisting of two components, one being the sequence of occurrences (and non-occurrences) of precipitation and the other being the amount (or intensity) of precipitation for time periods during which precipitation is occurring. Because of these two components, a specialized approach is required to make statistical inferences about precipitation parameters.

Although the existing literature on the statistical analysis of GCM climate experiments (e.g., Chervin and Schneider, 1976; Laurmann and Gates, 1977) has not dealt with this problem, the statistical procedure to be proposed can be viewed as an extension of a procedure applied in previous work to temperature time series generated by a GCM (Katz, 1982). This technique will be based upon a probabilistic model that simultaneously represents both the occurrence and intensity components of the precipita-

tion process. In particular, the precipitation occurrence process is allowed to be correlated in time, and the precipitation intensities are allowed to have a non-Gaussian distribution. The specific probabilistic model for precipitation that will be employed is a special case of a chain-dependent process (Katz, 1977a), a model that has been applied to real daily precipitation measurements by Katz (1977b). The model is similar to one employed by Todorovic and Woolhiser (1975). Closely related procedures have been used by Klugman and Klugman (1981) to test for climatic change with actual precipitation measurements, and by Crow (1978) to assess the statistical significance of weather modification experiments to augment natural precipitation.

The appropriate statistical methodology for making inferences about changes in mean daily precipitation is introduced in Section 2, and the application of this methodology to GCM climate experiments is discussed in Section 3. Section 4 presents the results of a test application of the statistical procedure to precipitation data generated in a control integration of the Oregon State University (OSU) atmospheric GCM. Finally, Section 5 consists of some concluding remarks and suggestions for future research.

2. Statistical methodology

a. Probabilistic model

The daily precipitation process consists of a bivariate sequence of random variables $\{(J_t, X_t): t = 1, 2, \dots\}$.

¹ The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Here J_t represents the occurrence (or nonoccurrence) of precipitation on the t th day; i.e.,

$$J_t = \begin{cases} 1, & \text{if precipitation occurs on } t\text{th day} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

We call a day for which $J_t = 1$ a *wet day*, and a day for which $J_t = 0$ a *dry day*. The total number of wet days that occur in a n -day time period is denoted by N . We note that N is a random variable with $0 \leq N \leq n$ and that it can be expressed as

$$N = \sum_{t=1}^n J_t. \quad (2)$$

It is assumed that the precipitation occurrence process $\{J_t; t = 1, 2, \dots\}$ constitutes a two-state first-order Markov chain with parameters

$$p = \Pr[J_t = 1],$$

$$P_{ij} = \Pr[J_{t+1} = j | J_t = i], \quad i, j = 0, 1. \quad (3)$$

Here p is the unconditional probability of a wet day and P_{ij} is a transition probability; for example, P_{11} is the conditional probability of a wet day given the previous day is wet and P_{01} is the conditional probability of a wet day given the previous day is dry. Such a model commonly has been employed to allow for day-to-day dependence in the precipitation occurrence process (e.g., Gabriel and Neumann, 1962).

The second component $\{X_t; t = 1, 2, \dots\}$ of the precipitation process represents the amount of precipitation that occurs on the t th day. It is convenient to consider a related process $\{Y_k; k = 1, 2, \dots\}$, where Y_k denotes the amount of precipitation (or intensity) that occurs on the k th wet day. It is assumed that the Y_k 's are independent and identically distributed random variables with common population mean $\mu = E(Y_k)$ and common population variance $\sigma^2 = \text{Var}(Y_k)$.

This probabilistic model for daily precipitation could be called a *chain-dependent process*, since it is a special case of such a process (Katz, 1977a, 1977b). It is a special case because the simplifying assumption has been made that the Y_k 's are identically distributed. This assumption is supported by the recent empirical work of Chin and Miller (1980). The probabilistic model proposed by Todorovic and Woolhiser (1975) imposes an additional requirement, namely, that the Y_k 's are exponentially distributed. In this paper, no specific form of parametric distribution for daily precipitation intensity is assumed. We will, however, apply the logarithmic transformation to adjust for the positive skewness of the intensity distribution (see Section 3a).

b. Distribution of total and mean daily precipitation

The total precipitation in n days (S_n) is given by

$$S_n = \sum_{t=1}^n X_t = \sum_{k=1}^N Y_k. \quad (4)$$

From (4), S_n can be viewed as a "random sum"; i.e., a sum of a random number of independent and identically distributed random variables. Like an ordinary or nonrandom sum of independent and identically distributed random variables, a random sum has an approximately Gaussian distribution as the sample size n tends to infinity (e.g., Billingsley, 1968, 143–150). The fact that the distribution of total precipitation S_n is asymptotically Gaussian also follows from the central limit theorem for chain-dependent processes (Katz, 1977a).

To specify the asymptotic distribution of S_n completely, it only remains to determine appropriate normalizing constants. One way to derive these normalizing constants is by conditioning on the random variable N , since given a value of N , S_n is an ordinary nonrandom sum. General expressions for the mean and variance of one random variable in terms of its conditional mean and variance given another random variable are (e.g., Lindgren, 1968, 113 and 119)

$$E(S_n) = E[E(S_n|N)], \quad (5)$$

$$\text{Var}(S_n) = E[\text{Var}(S_n|N)] + \text{Var}[E(S_n|N)]. \quad (6)$$

Now, for a random sum, (5) and (6) reduce to (e.g., Feller, 1968, 301)

$$E(S_n) = E(N)\mu, \quad (7)$$

$$\text{Var}(S_n) = E(N)\sigma^2 + \text{Var}(N)\mu^2. \quad (8)$$

To obtain expressions for the mean and variance of N , we invoke the representation (2) of N as a sum of random variables assumed to constitute a realization of a two-state first-order Markov chain. For sums of such a stochastic process, it has been shown (e.g., Gabriel, 1959) that

$$E(N) = np, \quad (9)$$

$$\text{Var}(N) \approx np(1-p) \frac{1+d}{1-d}. \quad (10)$$

Here $d = P_{11} - P_{01}$ and is the first-order autocorrelation (or "persistence parameter") for the Markov chain $\{J_t; t = 1, 2, \dots\}$. If the persistence parameter d equals zero, then the J_t 's are independent random variables; otherwise, the J_t 's are dependent. For daily precipitation occurrences, d is generally positive.

Substituting (9) and (10) into (7) and (8) yields

$$E(S_n) = np\mu, \quad (11)$$

$$\text{Var}(S_n) \approx nV^2, \quad (12)$$

where

$$V^2 = p\sigma^2 + p(1-p) \frac{1+d}{1-d} \mu^2. \quad (13)$$

The first term on the rhs of (8) can be regarded as the variance of a nonrandom sum of $\approx np$ indepen-

dent random variables [i.e., the expected number of wet days (9)]. The second term on the rhs of (8) arises because total precipitation actually is a random sum, and this term would be present even if daily precipitation occurrences were independent. Nevertheless, the effect of the Markovian dependence is, for persistence parameter $d > 0$, to inflate this second term by the factor $(1 + d)/(1 - d)$. This sort of variance inflation is analogous to that which occurs in the case of a first-order autoregressive process (e.g., Leith, 1973).

Applying (11) and (12), the test statistic

$$\frac{S_n - np\mu}{n^{1/2}V} \tag{14}$$

tends to a standard Gaussian distribution (i.e., having zero mean and unit variance) as the sample size n tends to infinity. The expression (13) for the approximate variance of the total precipitation in n days is a special case of that given in Katz (1977b) for a more general chain-dependent process for daily precipitation. Klugman and Klugman (1981) have used the more general version of (13), in which the order of the Markov chain is allowed to be higher than first and the precipitation intensities are not identically distributed, to estimate the variance of actual precipitation totals.

In making comparisons between different precipitation time series, it is convenient to deal with mean daily precipitation

$$\bar{X} = \frac{1}{n} S_n, \tag{15}$$

rather than total precipitation. In particular, making comparisons in terms of precipitation totals is inappropriate if the different time series vary in length. Using (11) and (12), the population mean and approximate variance of the sample mean \bar{X} are given by

$$E(\bar{X}) = p\mu, \quad \text{Var}(\bar{X}) \approx \frac{1}{n} V^2, \tag{16}$$

where V^2 is given by (13). Since the quantity V^2 is a constant, it is evident from (16) that the variance of mean daily precipitation tends to zero as the sample size n tends to infinity. The distribution of mean daily precipitation \bar{X} is approximately Gaussian because the test statistic (14) also can be expressed as

$$\frac{\bar{X} - p\mu}{V/n^{1/2}} \tag{17}$$

c. Estimation of parameters

To apply the probabilistic model for daily precipitation described in Section 2a, the parameters of the model need to be estimated from the data at hand. We assume that a sample consisting of n consecutive days of data for the precipitation process $\{(J_t, X_t)$:

$t = 1, 2, \dots, n\}$ is available. The parameters p, P_{11} , and P_{01} , of the two-state first-order Markov chain model for the daily occurrence of precipitation are estimated using the following relative frequencies (e.g., Billingsley, 1961):

$$\hat{p} = \frac{N}{n}, \quad \hat{P}_{11} = \frac{n_{11}}{n_{10} + n_{11}}, \quad \hat{P}_{01} = \frac{n_{01}}{n_{00} + n_{01}}. \tag{18}$$

In (18), n_{ij} denotes the number of times that $J_t = i$ and $J_{t+1} = j$, for $t = 1, 2, \dots, n - 1, i = 0, 1$, and $j = 0, 1$. This statistic n_{ij} is called a *transition count* for the sample $\{J_t; t = 1, 2, \dots, n\}$. Then the persistence parameter d is estimated simply by

$$\hat{d} = \hat{P}_{11} - \hat{P}_{01}. \tag{19}$$

The parameters μ and σ^2 of the precipitation intensity process $\{Y_k; k = 1, 2, \dots, N\}$ are estimated using the ordinary sample mean and variance for the Y_k 's; i.e.,

$$\hat{\mu} = \frac{1}{N} \sum_{k=1}^N Y_k, \tag{20}$$

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{k=1}^N (Y_k - \hat{\mu})^2. \tag{21}$$

Finally, the variance of the mean daily precipitation \bar{X} is estimated by substituting the parameter estimates (18)–(21) into (13). Letting $\widehat{\text{Var}}(\bar{X})$ denote the estimator of $\text{Var}(\bar{X})$, we obtain

$$\widehat{\text{Var}}(\bar{X}) = \frac{1}{n} \hat{V}^2, \tag{22}$$

where

$$\hat{V}^2 = \hat{p}\hat{\sigma}^2 + \hat{p}(1 - \hat{p}) \frac{1 + \hat{d}}{1 - \hat{d}} \hat{\mu}^2. \tag{23}$$

Because the test statistic (17) depends on the unknown variance of \bar{X} , we replace $\text{Var}(\bar{X})$ with its estimator $\widehat{\text{Var}}(\bar{X})$, given by (22). This new test statistic

$$\frac{\bar{X} - p\mu}{\hat{V}/n^{1/2}} \tag{24}$$

also has a limiting distribution that is standard Gaussian. The variance of the total precipitation S_n can be estimated in the same manner, simply substituting the quantity \hat{V}^2 , given by (23), in place of V^2 in (12).

3. Application to GCM climate experiments

We are given a time series of simulated precipitation data of length n_c , $\{[J(c), X(c)]; t = 1, 2, \dots, n_c\}$, generated by a GCM control run, and another time series of simulated precipitation data of length n_e , $\{[J(e), X(e)]; t = 1, 2, \dots, n_e\}$, generated by a GCM climate experiment run. It is assumed that the experiment time series $[J(e), X(e)]$ is independent of the control time series $[J(c), X(c)]$. The problem of concern is to test whether the observed difference between the

control and experiment mean daily precipitation is statistically significant. To apply the results of Section 2, the control and experiment precipitation time series must satisfy the requirements stated in that section. For now, we assume that these requirements are indeed satisfied.

a. Tests of significance

To perform the tests of significance, separate probabilistic models of the type described in Section 2a first are fit to the control and experiment precipitation time series, using the methods of parameter estimation described in Section 2c. Following the notation of Section 2a, $J_t(c)$ denotes the occurrence of precipitation on the t th day of the GCM control run [with parameters $p_c, P_{11}(c), P_{01}(c)$, and d_c], N_c denotes the number of wet days in the n_c -day control run, $X_t(c)$ denotes the amount of precipitation on the t th day of the control run, and $Y_k(c)$ denotes the precipitation intensity on the k th wet day of the control run [with parameters μ_c and σ_c^2]. Similarly, $J_t(e)$ denotes the occurrence of precipitation on the t th day of the GCM climate experiment run [with parameters $p_e, P_{11}(e), P_{01}(e)$, and d_e], N_e denotes the number of wet days in the n_e -day experiment run, $X_t(e)$ denotes the amount of precipitation on the t th day of the experiment run, and $Y_k(e)$ denotes the precipitation intensity on the k th wet day of the experiment run [with parameters μ_e and σ_e^2].

Because the probability distribution of real daily precipitation intensities is known to be positively skewed (e.g., Katz, 1977b; Todorovic and Woolhiser, 1975), the Gaussian approximation (14) or (24) for the distribution of total or mean daily precipitation is not very accurate unless the sample size n is relatively large. For instance, in the application of a chain-dependent process to a set of actual precipitation data by Katz (1977b), the Gaussian approximation differs substantially from the exact distribution of total precipitation when $n = 20$. We work with transformed precipitation intensities

$$Y_k^*(c) = \ln[Y_k(c)], \tag{25}$$

$$Y_k^*(e) = \ln[Y_k(e)] \tag{26}$$

to obtain new data, $\{Y_k^*(c); k = 1, 2, \dots, N_c\}$ and $\{Y_k^*(e); k = 1, 2, \dots, N_e\}$, that have an approximately Gaussian distribution. When the asymptotic distribution theory introduced in Section 2 is applied to the logarithmically transformed data, rather than to the original data, the distribution of the test statistic (14) or (24) is more closely approximated by the standard Gaussian distribution. The use of the logarithmic transformation is consistent with the assumption sometimes made that daily precipitation intensity has a lognormal distribution (e.g., Crow, 1978). However, even if the precipitation intensity has only an approximately lognormal distribution, the test statistic

(14) or (24) will be more closely approximated by the Gaussian distribution when the logarithmically transformed data, rather than the original data, are employed.

For the logarithmically transformed control and experiment precipitation intensities, population means are denoted by

$$\mu_c^* = E[Y_k^*(c)], \quad \mu_e^* = E[Y_k^*(e)], \tag{27}$$

and population variances by

$$(\sigma_c^*)^2 = \text{Var}[Y_k^*(c)], \quad (\sigma_e^*)^2 = \text{Var}[Y_k^*(e)]. \tag{28}$$

The analogues to total and mean daily precipitation are

$$S_{n_c}^*(c) = \sum_{k=1}^{N_c} Y_k^*(c), \quad S_{n_e}^*(e) = \sum_{k=1}^{N_e} Y_k^*(e), \tag{29}$$

$$\bar{Y}_c^* = \frac{1}{n_c} S_{n_c}^*(c), \quad \bar{Y}_e^* = \frac{1}{n_e} S_{n_e}^*(e). \tag{30}$$

The expressions derived in Section 2b hold as well when dealing with the analogues (29) and (30). The only change is that the precipitation intensity parameters now are given by (27) and (28).

A test statistic can be constructed on the basis of (24). The statistic

$$Z = \frac{\bar{Y}_e^* - \bar{Y}_c^*}{(n_c^{-1} \hat{V}_c^2 + n_e^{-1} \hat{V}_e^2)^{1/2}} \tag{31}$$

has an approximately standard Gaussian distribution under the null hypothesis that $E(\bar{Y}_c^*) = E(\bar{Y}_e^*)$. Here

$$\hat{V}_c^2 = \hat{p}_c (\hat{\sigma}_c^*)^2 + \hat{p}_c (1 - \hat{p}_c) \frac{1 + \hat{d}_c}{1 - \hat{d}_c} (\hat{\mu}_c^*)^2, \tag{32}$$

$$\hat{V}_e^2 = \hat{p}_e (\hat{\sigma}_e^*)^2 + \hat{p}_e (1 - \hat{p}_e) \frac{1 + \hat{d}_e}{1 - \hat{d}_e} (\hat{\mu}_e^*)^2. \tag{33}$$

In this case, the parameter estimates, $\hat{\mu}_c^*$, $(\hat{\sigma}_c^*)^2$, $\hat{\mu}_e^*$, and $(\hat{\sigma}_e^*)^2$, are based on substituting the logarithmically transformed data $Y_k^*(c)$ and $Y_k^*(e)$ into (20) and (21). That is,

$$\hat{\mu}_c^* = \frac{1}{N_c} \sum_{k=1}^{N_c} Y_k^*(c), \tag{34}$$

$$(\hat{\sigma}_c^*)^2 = \frac{1}{N_c - 1} \sum_{k=1}^{N_c} [Y_k^*(c) - \hat{\mu}_c^*]^2, \tag{35}$$

$$\hat{\mu}_e^* = \frac{1}{N_e} \sum_{k=1}^{N_e} Y_k^*(e), \tag{36}$$

$$(\hat{\sigma}_e^*)^2 = \frac{1}{N_e - 1} \sum_{k=1}^{N_e} [Y_k^*(e) - \hat{\mu}_e^*]^2. \tag{37}$$

The parameter estimates, $\hat{p}_c, \hat{d}_c, \hat{p}_e$, and \hat{d}_e , for the Markov chains are computed by means of (18) and (19).

b. Confidence intervals

It is difficult to convert the results of the test of significance (described in Section 3a) into a confidence interval for the change in mean daily precipitation. This difficulty arises both because the test of significance is based on the logarithmically transformed data, rather than on the original precipitation data, and because of the complexity introduced through total precipitation being a random sum. Under the assumption that precipitation occurrences are independent, rather than Markovian dependent, Crow (1978) has obtained an approximate confidence interval for the ratio of daily precipitation means. An alternative approach that we will take is to consider separately the precipitation occurrence process and the precipitation intensity process. Given that mean daily precipitation has changed, this approach has the advantage that it provides more detailed information about the nature of how the precipitation process has changed. For instance, how much of the change in total precipitation is due to a change in frequency of occurrence and how much is due to a change in intensity can be identified.

First, we discuss the problem of obtaining a confidence interval for the difference $\mu_e^* - \mu_c^*$ between the experiment and control precipitation intensity parameters. Because the daily precipitation intensities are assumed to be independent (see Section 2a), an approximate $[100(1 - \alpha)]\%$ confidence interval for $\mu_e^* - \mu_c^*$ is given by

$$\hat{\mu}_e^* - \hat{\mu}_c^* \pm z_{\alpha/2} \left[\frac{1}{N_c} (\hat{\sigma}_c^*)^2 + \frac{1}{N_e} (\hat{\sigma}_e^*)^2 \right]^{1/2}, \quad (38)$$

where $z_{\alpha/2}$ satisfies

$$\Pr\{Z > z_{\alpha/2}\} = \alpha/2, \quad (39)$$

for Z a standard Gaussian random variable. A test of whether precipitation intensity has changed (i.e., a null hypothesis of $\mu_e^* = \mu_c^*$) also can be based on (38).

Since $\mu_c^* = E\{\ln[Y_k(c)]\}$ and $\mu_e^* = E\{\ln[Y_k(e)]\}$, the confidence interval (38) is in terms of parameters for the logarithmically transformed precipitation intensities. Applying the exponential transformation to (38), a $[100(1 - \alpha)]\%$ confidence interval for the ratio of the median experiment precipitation intensity to the median control precipitation intensity is given by

$$\begin{aligned} & \exp\left\{(\hat{\mu}_e^* - \hat{\mu}_c^*) - z_{\alpha/2} \left[\frac{1}{N_c} (\hat{\sigma}_c^*)^2 + \frac{1}{N_e} (\hat{\sigma}_e^*)^2 \right]^{1/2}\right\} \\ & < \exp(\mu_e^* - \mu_c^*) \\ & < \exp\left\{(\hat{\mu}_e^* - \hat{\mu}_c^*) + z_{\alpha/2} \left[\frac{1}{N_c} (\hat{\sigma}_c^*)^2 + \frac{1}{N_e} (\hat{\sigma}_e^*)^2 \right]^{1/2}\right\}. \end{aligned} \quad (40)$$

The median is more appropriate than the mean as a measure of central tendency, because the distribution of precipitation intensity is asymmetric.

Second, we discuss the problem of obtaining a confidence interval for the difference between the experiment and control probabilities of precipitation occurrence. Using (10), an approximate $[100(1 - \alpha)]\%$ confidence interval for $p_e - p_c$ is

$$\hat{p}_e - \hat{p}_c \pm z_{\alpha/2} \left[\frac{1}{n_c} \hat{p}_c(1 - \hat{p}_c) \frac{1 + \hat{d}_c}{1 - \hat{d}_c} + \frac{1}{n_e} \hat{p}_e(1 - \hat{p}_e) \frac{1 + \hat{d}_e}{1 - \hat{d}_e} \right]. \quad (41)$$

Here $z_{\alpha/2}$ is defined by (39). A test of whether the probability of precipitation occurrence has changed (i.e., a null hypothesis of $p_e = p_c$) also can be based on (41).

c. GCM data

1) POOLED DATA

The statistical calculations for estimating the variance of mean daily precipitation are slightly more complicated if the data consist of GCM runs for more than one year, such as L runs each of length n' for a total sample size of $n = n'L$. In this situation, the only change required is a revision of the definition of the transition count n_{ij} , for $i = 0, 1$, and $j = 0, 1$. Now

$$n_{ij} = \sum_{l=1}^L n_{ij}(l), \quad (42)$$

where $n_{ij}(l)$ denotes the number of times $J_l = i$ and $J_{l+1} = j$, for $t = (l - 1)n' + 1, (l - 1)n' + 2, \dots, ln' - 1$ (see Anderson and Goodman, 1957).

2) STATIONARITY

To minimize the effects of seasonal cycles, a winter season or a summer season of at most three months should be employed (e.g., December–February or June–August).

3) SAMPLE SIZE

The test procedure (31) is based on a large sample approximation, requiring that the expected frequency of wet days be relatively large. Even if the overall sample size is large, the expected frequency of wet days will be small if the probability of occurrence of precipitation is small enough. Thus, care must be taken in applying the procedure to dry regions.

4. Preliminary test calculations

To test the operational feasibility of the proposed procedure for making statistical inferences about

changes in mean daily precipitation, a data sample from a three-year control integration of the OSU atmospheric GCM is used. Three consecutive winter season simulations (1 December–28 February) and three consecutive summer season simulations (1 June–31 August) were analyzed. Daily total precipitation data for nine grid points at scattered locations in the United States at both 34 and 46°N were examined. The three-year winter season data sets each have a sample size of 270 ($=3 \times 90$), whereas the three-year summer season data sets each have a sample size of 276 ($=3 \times 92$).

a. Model validation

The probabilistic model for daily precipitation described in Section 2a requires several assumptions about the nature of the precipitation process. The winter and summer GCM precipitation data sets were employed to verify some of these assumptions. In particular, it is assumed that the precipitation occurrence process $\{J_t; t = 1, 2, \dots\}$ constitutes a first-order Markov chain. When dealing with real daily precipitation measurements, there is some evidence indicating that higher than first-order Markov chains should be fitted (e.g., Chin, 1977). Higher than first-order, as well as first-order, Markov chains were fit to the winter and summer GCM precipitation occurrence data, and an automatic model selection criterion was employed to determine the appropriate order (see Katz, 1981). In most cases, a first-order chain was selected, although a higher-order chain was occasionally selected.

The methodology presented in Section 2 could be generalized to the case of a higher than first-order Markov chain, by applying some basic results for general chain-dependent processes (Katz, 1977a). However, the resulting variance formulas would be quite complicated, requiring the inversion of a matrix. Klugman and Klugman (1981) give an example of actual daily precipitation measurements in which estimates of the variance of total precipitation are obtained both under the assumption of a first-order Markov chain and of a second-order chain. For this example, the estimated variance is greater when a second-order chain is fit. Nevertheless, it is difficult to draw any general conclusion about the effects on the estimated variance of total precipitation of the assumption that the Markov chain is first order.

Another assumption concerns the use of the logarithmic transformation, (25) and (26), to remove the possible skewness from the distribution of precipitation intensities (i.e., the Y_k 's). Histograms for the winter and summer GCM data sets indicate that the probability distributions of precipitation intensities are positively skewed. For example, Fig. 1 shows the histograms of precipitation intensities for winter and summer at one particular GCM grid point (100°W, 34°N). To remove this skewness, both the square root

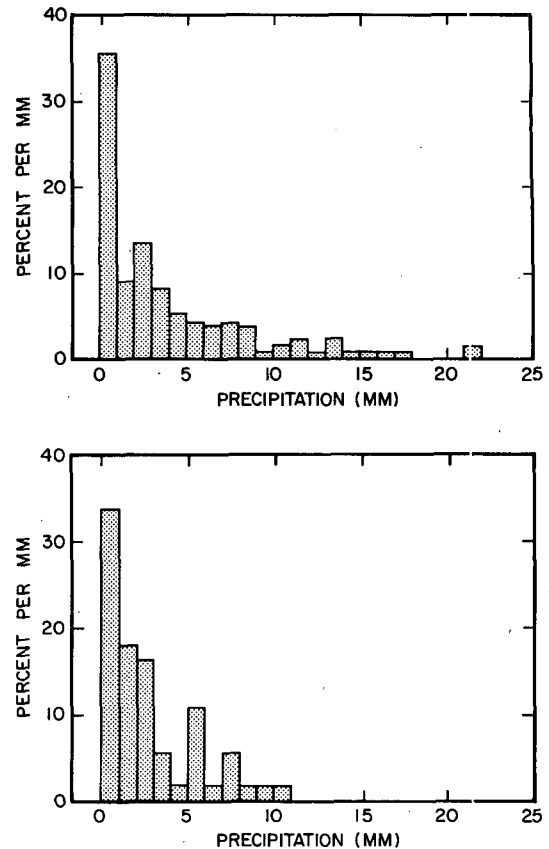


FIG. 1. Histogram of daily precipitation intensities at one OSU GCM grid point (100°W, 34°N): (a) during winter (December–February) and (b) during summer (June–August).

and logarithmic transformations were tried. Since the logarithm is a limiting case of a power transformation, the square root can be regarded as intermediate between no transformation and the logarithmic transformation. Histograms of the transformed precipitation intensities suggest that the logarithmically transformed data have distributions that are more nearly Gaussian than are the distributions of the square-root transformed data. In general, the square root transformation does not remove enough of the positive skewness in the precipitation intensity data. The logarithmic transformation, on the other hand, appears to remove nearly all of the positive skewness in most cases. A power transformation between the square root and logarithm might be more appropriate, as sometimes the logarithmically transformed data are slightly negatively skewed. Fig. 2 shows the histograms of the logarithmically transformed precipitation intensities for winter and summer at the same grid point as Fig. 1.

The precipitation intensities also are required to be independent. Real daily precipitation intensities exhibit little, if any, dependence (e.g., Katz, 1977b). To check for dependence, the summer and winter

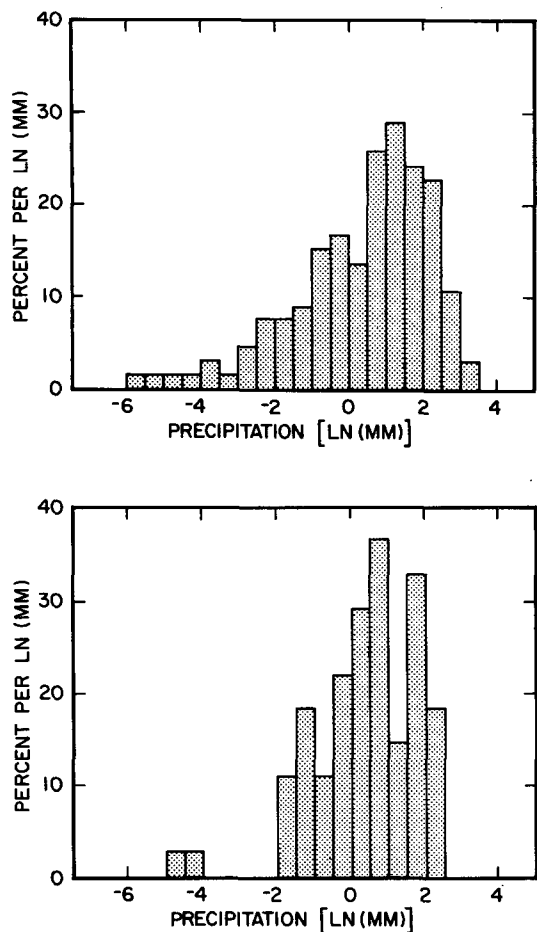


FIG. 2. Histogram of logarithmically transformed daily precipitation intensities at one OSU GCM grid point (100°W, 34°N): (a) during winter (December–February) and (b) during summer (June–August).

GCM precipitation intensity data were modeled by fitting low-order autoregressive processes to the logarithmically transformed data (similar to the approach employed by Katz, 1982). In many cases, a zero-order autoregressive process (i.e., an independent process) was selected as the best model. On the other hand, for several data sets a first-order or higher-order autoregressive process (i.e., a dependent process) was selected as the best model, with the suggestion of a small degree of positive autocorrelation.

The methodology in Section 2 could be generalized to the case of amounts of precipitation on consecutive wet days being dependent, again resulting in more complicated variance formulas. Estimates of the variance of total precipitation have been obtained under the assumption that daily precipitation intensities are independent and, alternatively, that they are generated by a first-order autoregressive process. Because the sample first-order autocorrelation coefficient is more frequently positive than negative, the estimated

variance of total precipitation tends to be slightly greater when a first-order autoregressive process is fit to daily precipitation intensities. If GCM simulated precipitation intensities actually do exhibit dependence, then comparisons with the corresponding real precipitation measurements might constitute a useful diagnostic check.

b. Tests of significance

To demonstrate how tests for significant differences between mean daily precipitation from GCM control and experiment time series may be conducted, the statistical procedure outlined in Section 3a was applied to the winter and summer GCM control runs. Here we are testing whether the difference between mean daily precipitation, averaged over all summer days (both dry and wet), and mean daily precipitation, averaged over all winter days, is statistically significant. The results are summarized in Table 1, including the sample means (34) and (36) of the logarithmically transformed data, the estimated standard deviations of these means [based on (32) and (33)], the test statistic *Z* (31), and the associated probability value (or *P*-value) denoted by *P*. The differences are statistically significant at the 5% level (i.e., *P* values smaller than 0.05) for six of the nine locations. Comparing the results by geographical location, the degree

TABLE 1. Tests of significance for winter versus summer mean daily precipitation (based on logarithmically transformed intensities).

| Grid point number | Location | Mean (ln mm) [Standard deviation of mean (ln mm)] | | Test statistic <i>Z</i> | <i>P</i> value |
|-------------------|-------------|------------------------------------------------------|--------------------|-------------------------|--------------------|
| | | Winter | Summer | | |
| 1 | 115°W, 34°N | 0.989 (0.1106) | 0.365 (0.0820) | -4.53 | <10 ⁻⁵ |
| 2 | 100°W, 34°N | 0.180 (0.0790) | 0.068 (0.0427) | -1.24 | 0.215 |
| 3 | 90°W, 34°N | 0.324 (0.0842) | -0.059 (0.0515) | -3.89 | <10 ⁻³ |
| 4 | 80°W, 34°N | 0.545 (0.0765) | 0.152 (0.0580) | -4.09 | <10 ⁻⁴ |
| 5 | 120°W, 46°N | 0.866 (0.0902) | 0.792 (0.0630) | -0.67 | 0.503 |
| 6 | 100°W, 46°N | 0.048 (0.0811) | 0.540 (0.0825) | 4.26 | <10 ⁻⁴ |
| 7 | 90°W, 46°N | -0.020 (0.0847) | 0.140 (0.0496) | 1.63 | 0.103 |
| 8 | 80°W, 46°N | 0.015 (0.0798) | 0.782 (0.0846) | 6.60 | <10 ⁻¹⁰ |
| 9 | 65°W, 46°N | 0.269 (0.0770) | 0.616 (0.0803) | 3.13 | 0.017 |

of spatial coherency is reasonably high. The observed total precipitation is less in summer than in winter for each of the four locations having common latitude 34°N , with three of these differences being statistically significant. Excluding the grid point located near the Pacific Coast (grid point number 5), the observed total precipitation is greater in the summer than in the winter for each of the four remaining locations having common latitude 46°N , with three of these differences being statistically significant.

c. Confidence intervals

The procedures for determining confidence intervals for changes in precipitation occurrence and intensity (described in Section 3b) also were applied to winter and summer GCM control runs. Table 2 gives the confidence intervals, based on (40), for the ratio (summer to winter) of median precipitation intensities. If the confidence interval contains the value one, then the change is not statistically significant at the 5% level. Hence, it is evident that the changes are statistically significant for seven of the nine locations. The estimated median daily precipitation intensity is less in summer than in winter for each of the four locations having common latitude 34°N , with three of these observed differences being statistically significant. Again excluding grid point number 5, the estimated median daily precipitation intensity is significantly greater in summer than in winter for each of the four remaining locations having common latitude 46°N .

Table 3 gives the confidence intervals, based on (41), for the difference between summer and winter probabilities of precipitation occurrence. In this case, if the confidence interval contains the value zero, then the difference is not statistically significant at the 5% level. We note that the changes are statistically significant for four of the nine locations. The observed relative frequency of daily precipitation occurrence is lower in summer than in winter for each of the four

TABLE 2. Confidence intervals for ratio (summer to winter) of median daily precipitation intensities.

| Grid point number | Estimated ratio of medians | 95% confidence interval for ratio of medians | |
|-------------------|----------------------------|----------------------------------------------|-------------|
| | | Lower limit | Upper limit |
| 1 | 0.554 | 0.401 | 0.766* |
| 2 | 0.974 | 0.594 | 1.597 |
| 3 | 0.394 | 0.255 | 0.690* |
| 4 | 0.565 | 0.386 | 0.828* |
| 5 | 0.796 | 0.630 | 1.005 |
| 6 | 2.275 | 1.667 | 3.106* |
| 7 | 1.451 | 1.021 | 2.061* |
| 8 | 3.378 | 2.528 | 4.514* |
| 9 | 1.733 | 1.240 | 2.422* |

* Significant at 5% level.

TABLE 3. Confidence intervals for difference between summer and winter probabilities of daily precipitation occurrence.

| Grid point number | Probability of precipitation occurrence | | 95% confidence interval for difference in probabilities (summer - winter) |
|-------------------|-----------------------------------------|--------|---------------------------------------------------------------------------|
| | Winter | Summer | |
| 1 | 0.682 | 0.424 | -0.258 ± 0.262 |
| 2 | 0.489 | 0.199 | $-0.290 \pm 0.242^*$ |
| 3 | 0.496 | 0.214 | $-0.282 \pm 0.202^*$ |
| 4 | 0.548 | 0.359 | $-0.189 \pm 0.179^*$ |
| 5 | 0.744 | 0.848 | 0.104 ± 0.124 |
| 6 | 0.678 | 0.605 | -0.073 ± 0.175 |
| 7 | 0.611 | 0.399 | $-0.212 \pm 0.202^*$ |
| 8 | 0.633 | 0.630 | -0.003 ± 0.170 |
| 9 | 0.544 | 0.591 | 0.047 ± 0.174 |

* Significant at 5% level.

locations having common latitude 34°N , with all of these differences being statistically significant at the 5% level or just above. Only one of the observed differences in relative frequency of daily precipitation occurrence for the five locations having common latitude 46°N is statistically significant.

It is interesting to compare the outcomes of the three different tests of significance: the overall test (Table 1), the test for change in median daily precipitation intensity (Table 2), and the test for change in probability of daily precipitation occurrence (Table 3). For grid point numbers 3 and 4, all three tests result in statistical significance at the 5% level. Further, the direction of the changes is consistent among the three tests; i.e., lower mean precipitation, lower median precipitation intensity, and lower probability of precipitation occurrence in the summer than in the winter. For grid point number 5, all three tests result in a lack of statistical significance. The other six grid points do not have complete agreement among the three tests. Grid point number 7 is an instructive case, the median intensity being significantly higher in summer, whereas the probability of occurrence is significantly higher in winter. These two contrasting effects tend to cancel out, so that the overall test is not statistically significant.

What magnitude of climatic change could be detected is a question of practical importance in the design of GCM experiments. The estimated standard deviations on which the confidence intervals presented in Tables 2 and 3 are based can be converted into approximate changes in precipitation intensity and probability of occurrence that could be detected. Depending on the geographical location, the smallest relative increase (decrease) in median daily precipitation intensity that would be identified as statistically significant (i.e., P value smaller than 0.05) ranges from $\sim 32\text{--}57\%$ (24–36%) in winter and from $\sim 19\text{--}92\%$ (16–48%) in summer. The smallest change in probability of daily precipitation occurrence that could be detected ranges

from ~ 0.13 – 0.17 in winter and from ~ 0.11 – 0.33 in summer. These approximate changes rely on the assumption that GCM control and experiment runs, each three years in length, are available. Of course, climatic changes that are smaller in magnitude could be detected if longer GCM control and experiment runs were available.

5. Concluding remarks

A procedure for making statistical inferences about changes in mean daily precipitation from the results of GCM climate experiments has been presented. The application of this technique has been demonstrated through the use of GCM control data to compare winter and summer precipitation. Besides establishing whether the difference between seasonal means is statistically significant, the procedure provides confidence intervals for the ratio of seasonal median precipitation intensities and for the difference between seasonal probabilities of precipitation occurrence.

We have concentrated on the analysis of simulated precipitation time series from GCM climate experiments. The same methodology could be applied to real precipitation measurements. As part of a comprehensive GCM diagnostic study, comparisons also could be made with the results of fitting the probabilistic model in a similar manner to corresponding observed precipitation data.

The technique proposed here is based on the choice of a particular form of probabilistic model to represent the precipitation process. This probabilistic model could be generalized in several respects. In particular, higher than first-order Markov chains could be fitted to the daily precipitation occurrence data. The model also could be generalized by allowing the amounts of precipitation on consecutive wet days to be correlated. Finally, transformations other than the logarithm could be applied to precipitation intensities in an attempt to obtain data having a more nearly Gaussian distribution.

Acknowledgments. The author thanks W. Lawrence Gates for suggesting this topic, William McKie for providing programming assistance, and Allan Murphy for supplying comments on this work. An

earlier version of this paper, entitled "Procedures for Determining the Statistical Significance of Precipitation Changes Simulated by an Atmospheric General Circulation Model," appeared as Climatic Research Institute Report No. 33, Oregon State University (March 1982). This research was supported by National Science Foundation Grant ATM-8001702.

REFERENCES

- Anderson, T. W., and L. A. Goodman, 1957: Statistical inference about Markov chains. *Ann. Math. Statist.*, **28**, 89–110.
- Billingsley, P., 1961: Statistical methods in Markov chains. *Ann. Math. Stat.*, **32**, 12–40.
- , 1968: *Convergence of Probability Measures*. Wiley, 253 pp.
- Chervin, R. M., and S. H. Schneider, 1976: On determining the statistical significance of climate experiments with general circulation models. *J. Atmos. Sci.*, **33**, 405–412.
- Chin, E. H., 1977: Modeling daily precipitation occurrence process with Markov chain. *Water Resour. Res.*, **13**, 949–956.
- , and J. F. Miller, 1980: On the conditional distribution of daily precipitation amounts. *Mon. Wea. Rev.*, **108**, 1462–1464.
- Crow, E. L., 1978: Confidence limits for seeding effect in single-area weather modification experiments. *J. Appl. Meteor.*, **17**, 1652–1660.
- Feller, W., 1968: *An Introduction to Probability Theory and its Applications*, Vol. I, 3rd ed. Wiley, 509 pp.
- Gabriel, K. R., 1959: The distribution of the number of successes in a sequence of dependent trials. *Biometrika*, **46**, 454–460.
- , and J. Neumann, 1962: A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quart. J. Roy. Meteor. Soc.*, **88**, 90–95.
- Katz, R. W., 1977a: An application of chain-dependent processes to meteorology. *J. Appl. Prob.*, **14**, 598–603.
- , 1977b: Precipitation as a chain-dependent process. *J. Appl. Meteor.*, **16**, 671–676.
- , 1981: On some criteria for estimating the order of a Markov chain. *Technometrics*, **23**, 243–249.
- , 1982: Statistical evaluation of climate experiments with general circulation models: A parametric time series modeling approach. *J. Atmos. Sci.*, **39**, 1446–1455.
- Klugman, M. R., and S. A. Klugman, 1981: A method for determining change in precipitation data. *J. Appl. Meteor.*, **20**, 1506–1509.
- Laurmann, J. A., and W. L. Gates, 1977: Statistical considerations in the evaluation of climate experiments with atmospheric general circulation models. *J. Atmos. Sci.*, **34**, 1187–1199.
- Leith, C. E., 1973: The standard error of time-average estimates of climate means. *J. Appl. Meteor.*, **12**, 1066–1069.
- Lindgren, B. W., 1968: *Statistical Theory*, 2nd ed. Macmillan, 521 pp.
- Todorovic, P., and D. A. Woolhiser, 1975: A stochastic model of n -day precipitation. *J. Appl. Meteor.*, **14**, 17–24.