Vol. 24: 95–102, 2003 | CLIMATE RESEARCH Clim Res | Published July 28

Seasonal and spatial variations of cross-correlation matrices used by stochastic weather generators

Justin T. Schoof*, Scott M. Robeson

Atmospheric Science Program, Department of Geography, Indiana University, Bloomington, Indiana 47405, USA

ABSTRACT: We examine seasonal and spatial variations of stochastic-weather-generator (SWG) parameters and their impact on simulated weather sequences. Using daily weather observations from 29 stations across the contiguous United States, we estimate monthly station-specific parameters that are compared with the constant parameters that frequently are used in SWG applications. A WGEN-type SWG is then used to generate a 100 yr record of daily maximum and minimum air temperature and daily total solar radiation at each station. These sequences are compared to sequences generated with constant parameters. While the means and standard deviations of the generated sequences are in agreement, the SWG with station-specific parameters preserves relationships between variables. This is evident in both the lag-0 and lag-1 cross-correlations between generated variables and derived variables, such as diurnal temperature range. These results suggest that literature-based SWG parameters may be appropriate for applications where monthly values of the means and standard deviations of generated variables are of interest. For applications that require proper simulation of relationships between variables, station-specific parameterizations are recommended.

KEY WORDS: Stochastic weather generator · Climate simulation · Climate variability · Autoregressive parameters

1. INTRODUCTION

For many applications, the historical climate record is inadequate due to short or incomplete data records, or lack of appropriate spatial coverage. As a result, models of observed daily weather sequences, or stochastic weather generators (SWGs), are often used to supplement the historical record or to provide data for locations where weather data are not routinely collected (Johnson et al. 1996, Wilks & Wilby 1999). As time-series models with several interconnected components, SWGs simulate sequences for a number of variables, which typically include daily maximum and minimum air temperature (T_{max} and T_{min}) and total daily solar radiation (R), using a multivariate autoregressive process.

The generated sequences are designed to have the desired cross-correlations between T_{max}, T_{min}, and R using 2 matrices, **A** and **B**, which are estimated using lag-0 and lag-1 cross-correlations (see Section 3.2). In many SWG implementations, **A** and **B** are treated as constant with respect to location, time of year, and wet/dry status. Hayhoe (2000) examined bi-monthly variations in SWG parameters for 3 stations in Canada and found spatial and seasonal variability in observed cross-correlations. Several authors (e.g. Wilks & Wilby 1999) have suggested using location- and time-specific parameters in an effort to account for spatial and seasonal variability.

In this study, the magnitude of the spatial and seasonal variability of these stochastic model parameterizations is investigated over a larger number of stations and wider range of climates than have been studied in the past. Using daily data from the contiguous USA, we examine the spatial and seasonal differences in the values of the lag-0 and lag-1 cross-correlations, and hence **A** and **B**. We also examine differences between simulated weather series when **A** and **B** are held con-

*Email: jschoof@indiana.edu

Fig. 1. The contiguous United States and the 29 stations used in this study

stant and when they are allowed to vary by location and time of year. Rather than focusing on traditional statistics used to evaluate SWGs (monthly means and variances, monthly maximum and minimum values, freeze-free periods, etc.), we primarily examine (1) the correlation structure of the generated sequences and (2) variables that are closely related to the correlation structure, such as diurnal temperature range.

2. DATA

To estimate the impacts of seasonally and spatially varying autoregressive parameters, daily data were extracted for 29 climatically diverse locations across the contiguous USA. (Fig. 1), essentially the same stations used in the landmark study of Richardson & Wright (1984). Hourly air temperature, solar radiation, and precipitation values for 1961 to 1990 were available at these 29 locations through the Solar and Meteorological Surface Observation Network dataset (SAMSON; NCDC/NREL 1993) and updates, available from the National Climatic Data Center, Asheville, NC. To allow comparison and combination with cooperative climatic data, daily T_{max} and T_{min} (°C) were calculated using the now-common 07:00 h observation time (Janis 2002). Daily total solar radiation values (MJ m^{-2} d^{-1}) were produced by numerical integration of hourly solar radiation observations.

3. DESCRIPTION OF WEATHER GENERATOR

The SWG used in this research is based on the well-known and commonly used WGEN model (Richardson & Wright 1984). Using a number of parameters estimated from observed data, the model traditionally generates daily values of precipitation occurrence, precipitation amount, T_{max}, T_{min}, and R. Our primary objective was to evaluate the impact of varying the parameterizations of **A** and **B**; therefore, precipitation amount was not simulated. The individual components of the model are described below.

3.1. Precipitation-occurrence component

Precipitation occurrence is simulated by a 2-state, first-order Markov chain. The occurrence of precipitation depends on 2 parameters: p_{01}, the probability of a wet day following a dry day, and p_{11}, the probability of a wet day following a wet day. Depending on the precipitation occurrence simulation for the previous day, a uniform [0,1] random number is compared to the appropriate transition probability. If the random number is less than the transition probability, a wet day is simulated. Otherwise, a dry day is simulated. Previous work has shown that changes in precipitation parameters can effect the moments of conditioned variables generated by the model. Therefore, if a SWG is to be used in a climate-change context, additional adjustments to model parameters may be needed (Katz 1996).

3.2. Temperature and radiation component

Daily values of T_{max}, T_{min}, and R are simulated by a first-order multivariate stochastic process, as described by Matalas (1967). To produce stationary time series, harmonic analysis first is used to construct annual cycles of daily means and standard deviations of the input variables. Using a Fourier transform, annual cycles are fit to daily means and standard deviations using the first 3 harmonics. These annual-cycle harmonics are fit separately for wet and dry days. In some cases, a given day of the year may have few wet or dry occurrences; therefore, a 15 d moving window was used to construct the daily means and standard deviations. In particularly dry areas, SWG users may consider using longer moving windows. The time series are then reduced to standardized residual elements by subtracting the daily means and dividing by the standard deviations, as defined by the harmonics.

The SWG simulates daily residuals of T_{max}, T_{min}, and R for Day i with

$$\mathbf{X}_i = \mathbf{A}\mathbf{X}_{i-1} + \mathbf{B}\boldsymbol{\varepsilon}_i \tag{1}$$

where $\mathbf{X}_i$ is a (3 × 1) matrix containing the current day's standardized values of T_{max}, T_{min}, and R; $\mathbf{X}_{i-1}$ is a (3 × 1) matrix containing the previous day's standardized values of T_{max}, T_{min}, and R; $\boldsymbol{\varepsilon}_i$ is a (3 × 1) vector of independent values from a standard normal distribution; and **A** and **B** are (3 × 3) matrices given by

$$\mathbf{A} = \mathbf{M}_1 \mathbf{M}_0^{-1} \tag{2}$$

$$\mathbf{BB}^{\mathrm{T}} = \mathbf{M}_0 - \mathbf{M}_1 \mathbf{M}_0^{-1} \mathbf{M}_1^{\mathrm{T}} \tag{3}$$

where $\mathbf{M}_0$ is the (3×3) matrix of lag-0 cross correlations and $\mathbf{M}_1$ is the (3×3) matrix of lag-1 cross correlations. For example, $\mathbf{M}_0(1,2)$ is the correlation between $T_{\max}$ and $T_{\min}$ and $\mathbf{M}_1(1,2)$ is the correlation between $T_{\max}$ and $T_{\min}$ lagged by 1 d. While $\mathbf{A}$ can be estimated directly, $\mathbf{B}$ is estimated by defining a new matrix, $\mathbf{Z} = \mathbf{BB}^{\mathrm{T}}$ (see Greene 2000). Then by spectral decomposition, $\mathbf{Z} = \mathbf{CLC}^{\mathrm{T}}$, where $\mathbf{C}$ is the matrix of eigenvectors of $\mathbf{BB}^{\mathrm{T}}$ and $\mathbf{L}$ has the eigenvalues of $\mathbf{BB}^{\mathrm{T}}$ on the diagonal and zeros elsewhere. Since $\mathbf{BB}^{\mathrm{T}} = \mathbf{Z}^{1/2}\mathbf{Z}^{1/2\mathrm{T}} = \mathbf{Z}$, $\mathbf{B} = \mathbf{Z}^{1/2}$. Then by Greene's Theorem 2.10, estimates of $\mathbf{B}$ can then be computed as $\mathbf{B} = \mathbf{CL}^{1/2}\mathbf{C}^{\mathrm{T}}$. After generation of the residual series with Eq. (1), the daily harmonics described above are used to produce dimensional values of $T_{\max}$, $T_{\min}$, and R, based on wet/dry status.

The use of a standard normal distribution for all 3 elements of $\boldsymbol{\varepsilon}_i$ may not be appropriate in many situations, particularly for solar radiation. For this reason, some SWGs (e.g. LARS-WG; Semenov & Barrow 1997) have used more complex distributions for R. Other SWGs (e.g., CLIGEN; Nicks & Gander 1993, 1994) have addressed this issue by constraining the generated R data between a maximum value based on station location and sun angle and a minimum value of 5% of the maximum value. In an examination of 15 US climate stations, Harmel et al. (2002) found that even $T_{\max}$ and $T_{\min}$ were not generally normally distributed in each month, results that have wide implications for further SWG research. Nonetheless, the focus of this research is not on $\boldsymbol{\varepsilon}_i$, and the vast majority of SWGs are still based on assumptions of normality. SWG users must decide how these assumptions impact their particular application.

4. OBSERVED RELATIONSHIPS

In many implementations, WGEN-type models use fixed values of $\mathbf{M}_0$ and $\mathbf{M}_1$—and therefore $\mathbf{A}$ and $\mathbf{B}$—irrespective of location and time of year. Richardson (1982) provides the following values:

$$\mathbf{M}_0 = \begin{pmatrix} 1.000 & 0.633 & 0.186 \\ 0.633 & 1.000 & -0.193 \\ 0.186 & -0.193 & 1.000 \end{pmatrix} \quad \mathbf{M}_1 = \begin{pmatrix} 0.621 & 0.445 & 0.087 \\ 0.563 & 0.674 & -0.100 \\ 0.015 & -0.091 & 0.250 \end{pmatrix}$$
$$\mathbf{A} = \begin{pmatrix} 0.567 & 0.086 & -0.002 \\ 0.253 & 0.504 & -0.050 \\ -0.006 & -0.039 & 0.244 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 0.781 & 0.000 & 0.000 \\ 0.328 & 0.637 & 0.000 \\ 0.238 & -0.341 & 0.873 \end{pmatrix} \tag{4}$$

While the literature-based correlation matrices may be appropriate during some seasons at some locations, the observed and literature-based (constant) correlation matrices can be very different when location and the entire calendar year are considered. The differences in the correlation matrices ultimately dictate the variability in the estimated elements of the $\mathbf{A}$ and $\mathbf{B}$ matrices. In the following sections, observed values of $\mathbf{M}_0$ and $\mathbf{M}_1$, and hence the estimates of $\mathbf{A}$ and $\mathbf{B}$, are examined.

4.1. Seasonal/spatial variability of lag-0 correlation coefficients ($\mathbf{M}_0$)

The literature-based values of $\mathbf{M}_0$ agree with observations at some locations during some months; however, examination of station-specific monthly correlations suggests that the literature values may not be appropriate for all locations year round (Figs. 2 & 3).

During the late summer months, the literature-based value of $\mathbf{M}_0(1,2)$—0.633—is similar to values observed at many stations. However, during the winter months, the literature-based correlation is lower than the observed correlation at most stations (Figs. 2a & 3a). The correlations between temperature and radiation [$\mathbf{M}_0(1,3)$ and $\mathbf{M}_0(2,3)$] are more seasonally and spatially variable than $\mathbf{M}_0(1,2)$ (Figs. 2b,c & 3b,c). For these elements, the literature-based correlations are appropriate at some locations during the transition seasons, but they are generally too strong during the winter months and too weak during the summer months.

4.2. Seasonal/spatial variability of lag-1 correlation coefficients ($\mathbf{M}_1$)

The data used in this study suggest that the elements of $\mathbf{M}_1$, the lag-1 correlation matrix, are also seasonally and spatially variable within the contiguous USA. As with the elements of $\mathbf{M}_0$, the literature-based values are appropriate at some locations and some times of year [e.g. $\mathbf{M}_1(2,2)$ in November; Fig. 4e], but fail to accommodate the range of values observed over the study area in most months (Fig. 4). For some elements of $\mathbf{M}_1$ the literature-based value is entirely outside the range of observations during particular months [e.g. $\mathbf{M}_1(2,3)$ in November, December, January, and February; Fig. 4h]. (Note that in all of the boxplots shown, spatial variability can be inferred from the amount of variation in any given box-and-whiskers, although some of the variation also is due to sampling variability.)

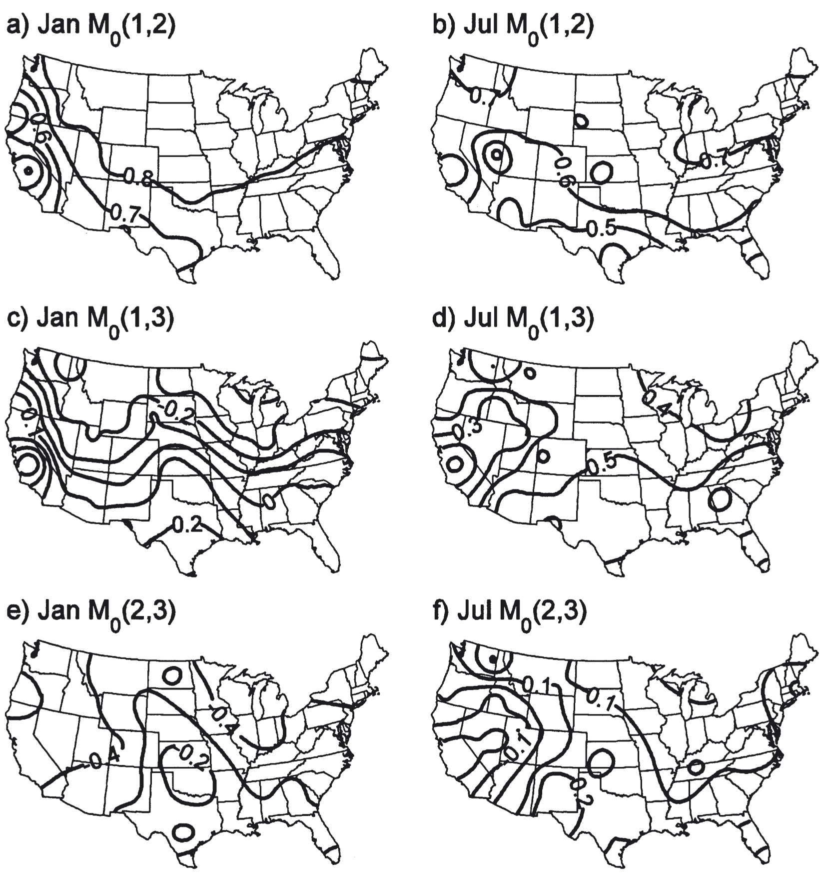

Fig. 2. Contour plots of lag-0 correlations ($\mathbf{M}_0$). (a) January correlation between daily $T_{\max}$ and $T_{\min}$: $\mathbf{M}_0(1,2)$; (b) July correlation between daily $T_{\max}$ and $T_{\min}$: $\mathbf{M}_0(1,2)$; (c) January correlation between $T_{\max}$ and R: $\mathbf{M}_0(1,3)$; (d) July correlation between $T_{\max}$ and R: $\mathbf{M}_0(1,3)$; (e) January correlation between daily $T_{\min}$ and R: $\mathbf{M}_0(2,3)$; and (f) July correlation between daily $T_{\min}$ and R: $\mathbf{M}_0(2,3)$. Literature-based values of $\mathbf{M}_0$ are given in Eq. (4)

4.3. Seasonal/spatial variability of A and B

Although seasonal variations in the elements of **A** are generally small, several elements of **A** exhibit substantial spatial variability in individual months (Fig. 5). The observed values of each element are generally in poor agreement with literature-based values, with most stations consistently having values that differ from the literature-based values. In general, the individual elements of **B** exhibit more seasonal variability and less spatial variability than the elements of **A** (Fig. 6). Observed values of $\mathbf{B}(1,2)$ and $\mathbf{B}(3,2)$ are different from the literature-based values at all stations during all months. For other elements of **B**, especially those exhibiting seasonal variability, the literature-based values are appropriate only at particular times and locations. In general, the literature-based values of **B** do not agree with those computed with spatially and seasonally variable values of $\mathbf{M}_0$ and $\mathbf{M}_1$ (Fig. 6). Although the estimated **A** and **B** elements vary over relatively large spatial scales, they do not have obvious relationships with physiographic characteristics, such as latitude, longitude, and elevation. This result has implications for interpolation of SWG parameters (e.g. Semenov & Brooks 1999).

5. WEATHER-GENERATOR IMPLEMENTATION

Variability in the elements of the **A** and **B** matrices suggests that using station-specific parameters may have important impacts on data generated with an SWG. To investigate the effects of these spatially and seasonally varying parameterizations, the SWG described in Section 3 was used to produce 100 yr sequences of daily $T_{\max}$, $T_{\min}$, and R for each station in our analysis (Fig. 1; as described in Section 2). The SWG was run in 2 modes. First, the elements of **A** and **B** were held constant according to the literature-based values (i.e. values given by Richardson 1982), producing generated data that is hereafter referred to as LGEN. In the second mode, the monthly values of **A** and **B** estimated from historical data for each individual station were used, producing data that is hereafter referred to as ABGEN.

6. EVALUATION OF GENERATED DATA

6.1. Means and standard deviations

Means and standard deviations of the generated variables (not shown) are in general agreement with observations for both versions of the generator (LGEN and ABGEN). These means and standard deviations are largely dependent on the harmonics used to depict the means and standard deviations—and not on the correlation structure of the variables. Since these harmonics do not differ between the ABGEN and LGEN simulations, the small differences in these values are not unexpected. For the eastern half of the contiguous USA, differences between observed and generated $T_{\max}$ are less than 1°C in all months. For stations in the western USA, differences are less than 1°C during the summer months and 1 to 2°C during the winter months. Differences between observed and generated

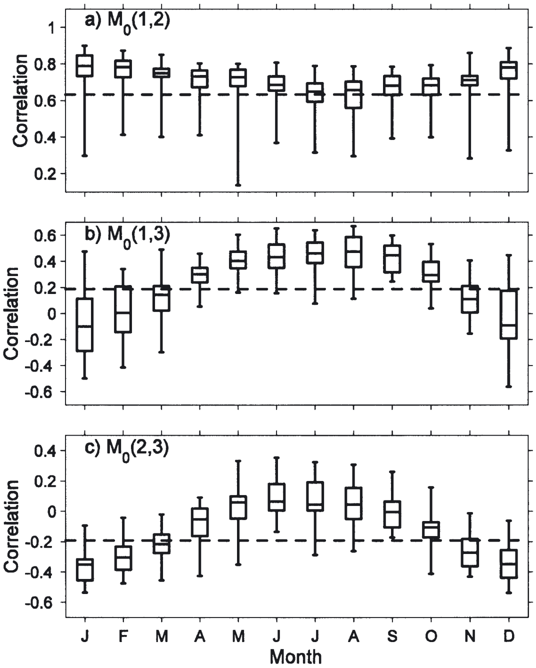

Fig. 3. Distributions (boxplots) of lag-0 correlations ($\mathbf{M}_0$). (a) Correlation between daily T_{max} and T_{min}: $\mathbf{M}_0(1,2)$; (b) correlation between T_{max} and R: $\mathbf{M}_0(1,3)$; and (c) correlations between daily T_{min} and R: $\mathbf{M}_0(2,3)$. Each box shows the distribution of correlations across the 29-station network and depicts the maximum and minimum values, as well as the inter-quartile range and median. The dashed line represents the literature-based value

T_{min} are less than 1°C at all locations during the spring and summer months, with differences of 1 to 2°C in the east during the winter. During the late autumn and winter months (October–February), differences between observed and generated R are less than 1 MJ m^{-2} d^{-1} in the western half of the US. During the spring, these differences are as large as 3 MJ m^{-2} d^{-1}. The errors in R are not unexpected and result from poor agreement in the lower tail of the distribution. While observed values are physically bounded at zero, modeled values of R can become negative. In the 100 yr simulations conducted here, negative R values were not generated at most stations. When this fundamental simulation error did occur, R was set to zero. Some alternatives, such as constraining the generated R values in physically plausible ways, are available (see Section 3.2).

ABGEN and LGEN produce means and standard deviations with nearly identical spatial and temporal variability (and therefore are not shown). In terms of these means and standard deviations, the station-specific ABGEN does not provide any improvement over the constant-parameter LGEN.

6.2. Correlations between generated variables

While correlations between simulated variables are not routinely used to evaluate SWGs, preservation of the correlation structure between the variables is critical for impacts modeling in agriculture and hydrology, where multiple input series of daily weather variables are routinely employed. Since the correlation structure between the generated variables is fundamentally dependent on the values of **A** and **B**, the station-specific generator should better replicate the observed correlations between variables. Lag-0 and lag-1 cross-correlations confirm that station-specific, monthly parameterization of **A** and **B** produces a better match between simulated and observed correlations (Figs. 2 & 7). Data simulated with the constant, literature-based values for **A** and **B** resulted in larger differences between observed and generated correlations (Figs. 2 & 7). (Note that the correlations depicted in Fig. 7 are

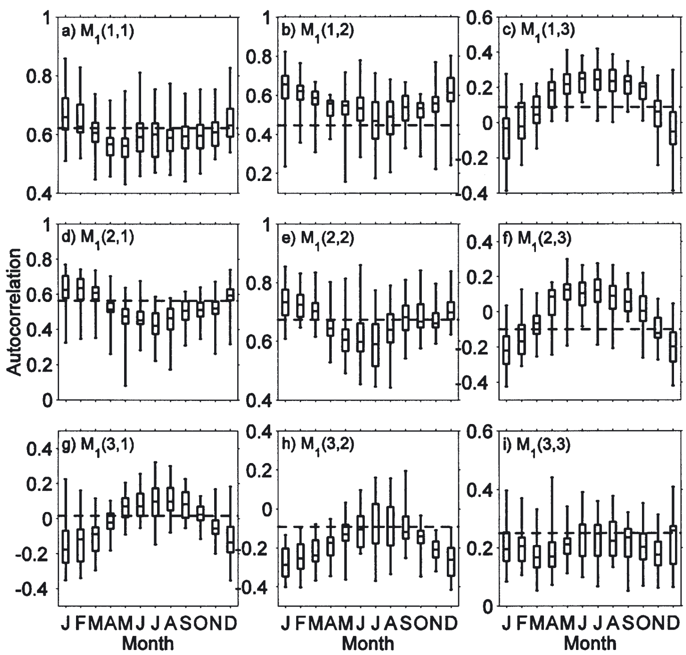

Fig. 4. Boxplots of lag-1 correlations ($\mathbf{M}_1$). (a) Lag-1 correlation between T_{max} and T_{max}; (b) lag-1 correlation between T_{max} and T_{min}; (c) lag-1 correlation between T_{max} and R; (d) lag-1 correlation between T_{min} and T_{max}; (e) lag-1 correlation between T_{min} and T_{min}; (f) lag-1 correlation between T_{min} and R; (g) lag-1 correlation between R and T_{max}; (h) lag-1 correlation between R and T_{min}; and (i) lag-1 correlation between R and R. Each box shows the distribution of correlations across the 29-station network and depicts the maximum and minimum values, as well as the inter-quartile range and median. The dashed line represents the literature-based value

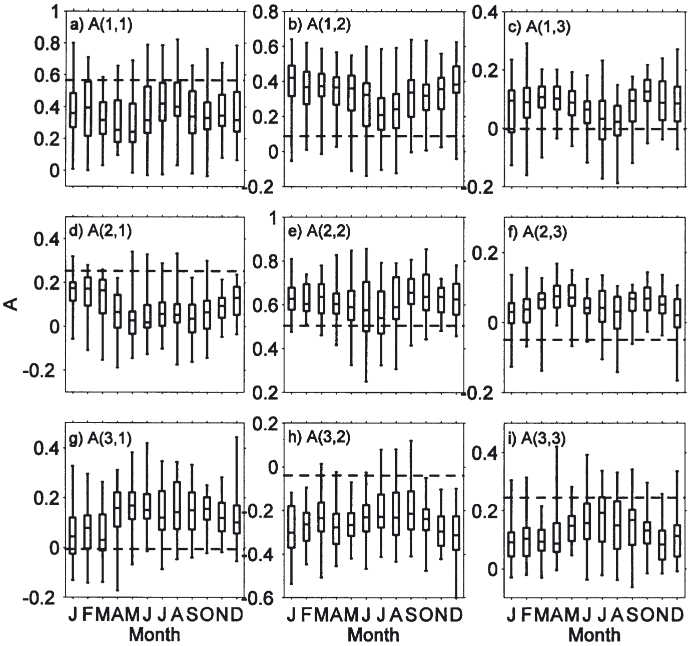

Fig. 5. Boxplots of elements of **A**. (a) **A**(1,1); (b) **A**(1,2); (c) **A**(1,3); (d) **A**(2,1); (e) **A**(2,2); (f) **A**(2,3); (g) **A**(3,1); (h) **A**(3,2); and (i) **A**(3,3). Each box shows the distribution of coefficients across the 29-station network and depicts the maximum and minimum values, as well as the inter-quartile range and median. The dashed line represents the literature-based value

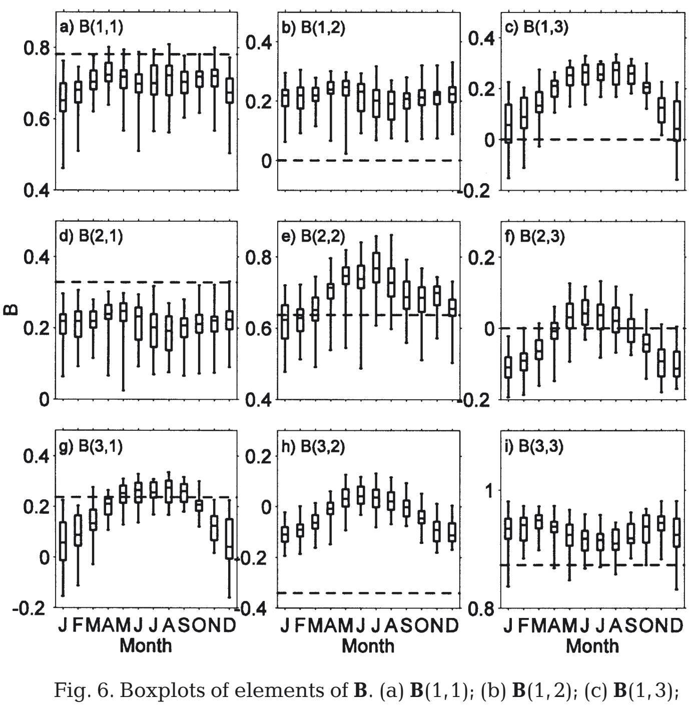

Fig. 6. Boxplots of elements of **B**. (a) **B**(1,1); (b) **B**(1,2); (c) **B**(1,3); (d) **B**(2,1); (e) **B**(2,2); (f) **B**(2,3); (g) **B**(3,1); (h) **B**(3,2); and (i) **B**(3,3). Each box shows the distribution of coefficients across the 29-station network and depicts the maximum and minimum values, as well as the inter-quartile range and median. The dashed line represents the literature-based value

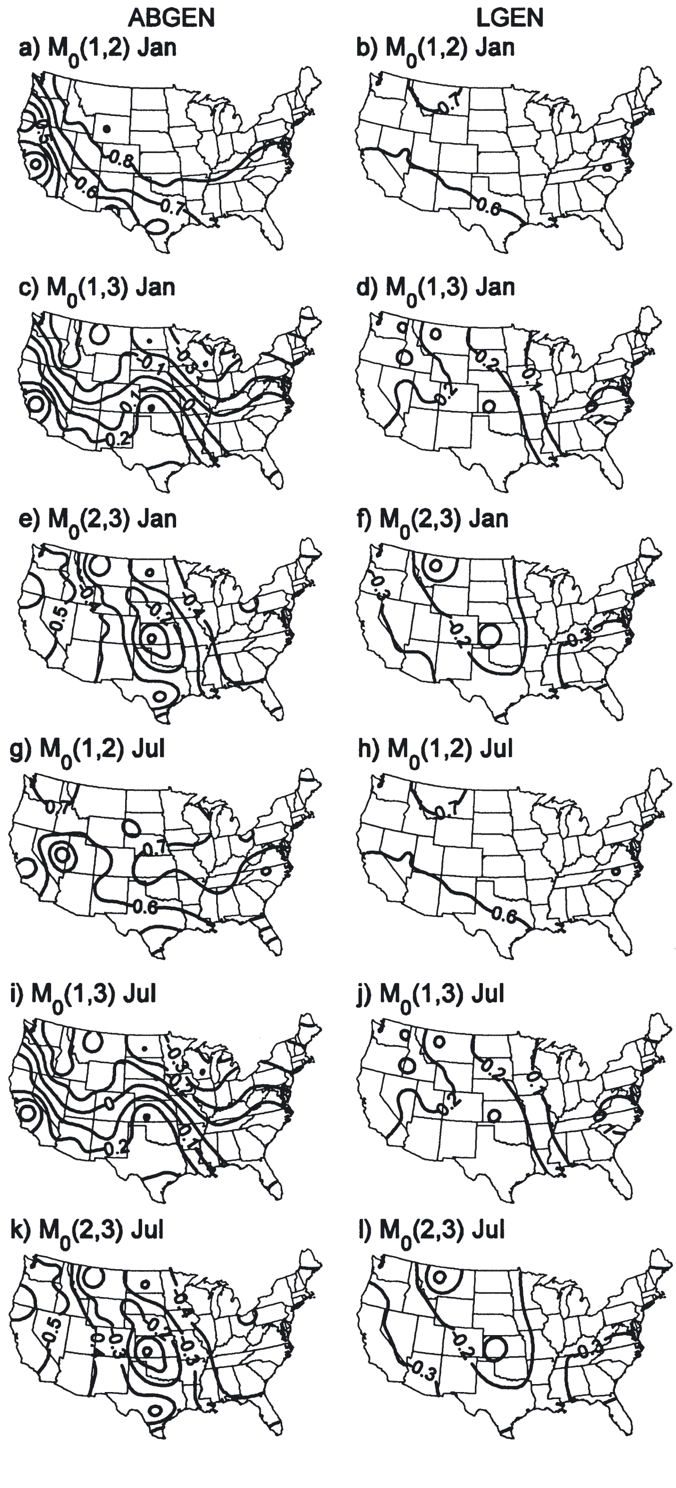

Fig. 7. Contour plots of lag-0 correlations for ABGEN (left) and LGEN (right) generated data. (a) January $\mathbf{M}_0(1,2)$ ABGEN; (b) January $\mathbf{M}_0(1,2)$ LGEN; (c) January $\mathbf{M}_0(1,3)$ ABGEN; (d) January $\mathbf{M}_0(1,3)$ LGEN; (e) January $\mathbf{M}_0(2,3)$ ABGEN; (f) January $\mathbf{M}_0(2,3)$ LGEN; (g) July $\mathbf{M}_0(1,2)$ ABGEN; (h) July $\mathbf{M}_0(1,2)$ LGEN; (i) July $\mathbf{M}_0(1,3)$ ABGEN; (j) July $\mathbf{M}_0(1,3)$ LGEN; (k) July $\mathbf{M}_0(2,3)$ ABGEN; and (l) July $\mathbf{M}_0(2,3)$ LGEN. Literature-based values are given in Eq. (4)

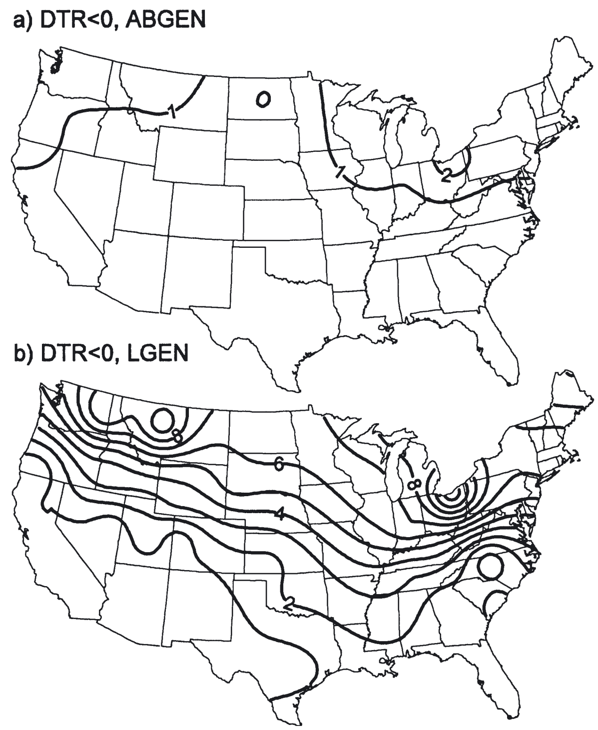

Fig. 8. Contour plots of the number of times per year (in a 100 yr simulation) the SWG simulates negative diurnal temperature range for (a) ABGEN and (b) LGEN

computed using the generated data and include the annual cycle harmonics; therefore, some of the variability in these maps results from differences in the harmonics.)

For each element of $\mathbf{M}_0$, absolute differences between observed and generated values are larger for LGEN than ABGEN. The maximum absolute differences between observed and generated $\mathbf{M}_0(1,2)$, the correlation between T_{max} and T_{min}, are 0.49 for LGEN and 0.27 for ABGEN. For $\mathbf{M}_0(1,3)$, the correlation between T_{max} and R, the maximum absolute difference is 0.68 for LGEN, compared with 0.23 for ABGEN. Maximum absolute differences between observed and generated $\mathbf{M}_0(2,3)$, the correlation between T_{min} and R, are 0.46 for LGEN and 0.22 for ABGEN.

The lag-1 cross-correlations computed with the generated data are also different from the observed lag-1 cross-correlations ($\mathbf{M}_1$). For ABGEN, the absolute differences between generated and observed elements of $\mathbf{M}_1$ are typically small (<0.1 for all temperature–temperature correlations and <0.2 for all temperature–radiation correlations). For LGEN, the differences can be quite large. For example, the maximum absolute difference between generated and observed $\mathbf{M}_1(2,1)$ for the month of May is 0.48 for LGEN, compared with only 0.10 for ABGEN.

6.3. Diurnal temperature range

Diurnal temperature range (DTR ≡ $T_{max} - T_{min}$) also can be an effective evaluation tool for SWGs. Given that DTR is a function of both T_{max} and T_{min}, its accurate simulation requires that the relationships between these 2 variables be preserved. Using the 100 yr simulations described above, LGEN (with constant **A** and **B**) produces many more days with negative DTR (i.e. a fundamental simulation error) than ABGEN at most stations used in this study (Fig. 8). While both SWGs (LGEN and ABGEN) simulated monthly means and standard deviations of generated variables well, the frequency distribution of DTR is not simulated as well by the literature-based generator. Although monthly mean DTR is similar in both models, the station-based generator produces much better agreement between observed and simulated standard deviation of DTR (Fig. 9), especially during the winter months. The cause for these simulation errors in LGEN can be

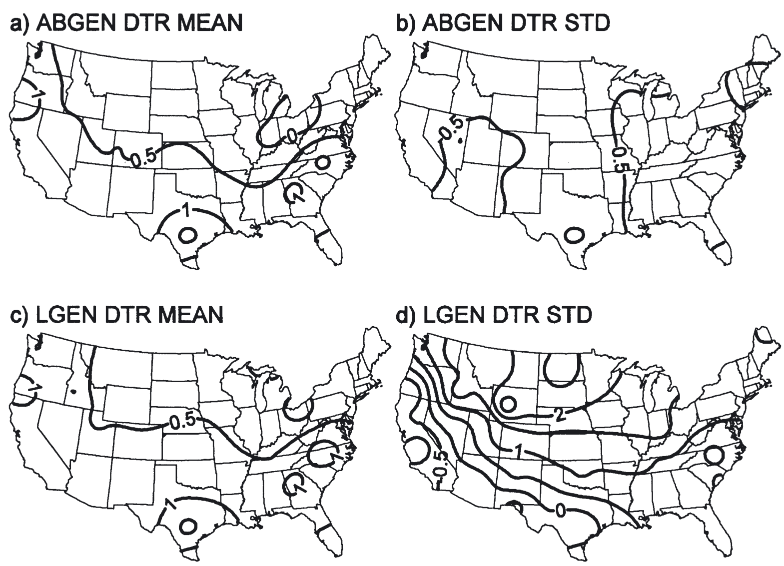

Fig. 9. Contour plots of the errors in simulating the mean (left) and standard deviation (right) of diurnal temperature range (DTR) during January. (a) ABGEN mean DTR minus observed mean DTR and (b) ABGEN standard deviation of DTR minus observed standard deviation of DTR. (c) LGEN mean DTR minus observed mean DTR and (d) LGEN standard deviation of DTR minus observed standard deviation of DTR

traced to $\mathbf{M}_0(1,2)$, the element of the correlation matrix that relates the current day's T_{max} and T_{min}. Errors in the LGEN standard deviation of DTR are highly correlated with differences in literature-based and observed values of $\mathbf{M}_0(1,2)$ (monthly correlations range from –0.92 to –0.70, significant at the 99% level).

The station-specific generator also reproduces the relationships between temperature and radiation more accurately. Because DTR is closely linked to cloud cover and precipitation (Leathers et al. 1998), and radiation is a reasonable surrogate for cloud cover, allowing the relationships between temperature and radiation to vary by location and time of year helps to improve the simulation of temporal variability in DTR.

7. DISCUSSION AND CONCLUSIONS

In this study, the effects of SWG parameterizations have been investigated. Using historical data from 29 stations in the US, we examined the spatial and seasonal differences in the lag-0 and lag-1 cross-correlations between T_{max}, T_{min}, and R. These correlations ultimately determine the nature of the **A** and **B** matrices used in the SWG, and they were found to have profound spatial and seasonal variations.

To investigate the impacts of the seasonal and spatial variability in the elements of these matrices, 100 yr simulations for 29 stations were undertaken with (1) **A** and **B** assumed constant (values from Richardson 1982) and (2) **A** and **B** computed for each individual station on a monthly basis.

The simulations were compared to observed data using statistical and graphical methods. The results suggest that monthly means and standard deviations of each simulated variable agree with observed values for both simulations; however, the literature-based generator failed to preserve relationships between variables. This shortcoming is evident in both the simulated diurnal temperature range (DTR) and in the correlations between simulated variables.

Our findings suggest that literature-based values may be appropriate for applications where monthly values of the means and standard deviations of generated variables are of interest. For applications that require proper simulation of relationships between variables, station-specific parameterizations are more appropriate. In addition, because SWGs are now being used in climate-change studies (e.g. GCM downscaling research; Semenov & Barrow 1997, Wilks 1999), additional caution is warranted. While SWG parameters will certainly change as climate changes, the magnitude of changes will vary seasonally and spatially.

Acknowledgements. This manuscript was substantially improved by the comments of Robert E. Davis and 2 anonymous reviewers.

LITERATURE CITED

Greene WH (2000) Econometric analysis. Prentice Hall, Upper Saddle River, NJ

Harmel RD, Richardson CW, Hanson CL, Johnson GL (2002) Evaluating the adequacy of simulating maximum and minimum daily air temperature with the Normal distribution. J Appl Meteorol 41:744–753

Hayhoe HN (2000) Improvements in stochastic SWG for diverse climates. Clim Res 14:75–87

Janis MJ (2002) Observation-time-dependent biases and departures for daily minimum and maximum air temperatures. J Appl Meteorol 41:588–603

Johnson GL, Hanson CL, Hardegree SP, Ballard EB (1996) Stochastic weather simulation: overview and analysis of two commonly used models. J Appl Meteorol 35: 1878–1896

Katz (1996) Use of conditional stochastic models to generate climate change scenarios. Clim Change 32:237–255

Leathers DJ, Palecki MA, Yarnal B (1998) Climatology of the daily temperature range annual cycle in the United States. Clim Res 9:197–211

Matalas NC (1967) Mathematical assessment of synthetic hydrology. Water Resour Res 3:937–945

NCDC/NREL (1993) Solar and Meteorological Surface Observation Network (SAMSON, available from NCDC). Available at http://www.ncdc.noaa.gov

Nicks AD, Gander GA (1993) Using CLIGEN to stochastically generate climate data inputs to WEPP and other water resource models. USGS Water-Resources Investigations Report 93-4018, USGS, Fort Collins, CO

Nicks AD, Gander GA (1994) CLIGEN: a SWG for climate inputs to water resource and other models. In: Watson DG, Zarzueta FS, Harrison TV (eds). Proceedings of the 5th International Conference on Computers in Agriculture, Orlando, FL. American Society of Agricultural Engineers, St. Joseph, MI, p 903–909

Richardson CW (1982) Dependence structure of daily temperature and solar radiation. Trans Am Soc Agric Eng 25: 735–739

Richardson CW, Wright DA (1984) WGEN: A model for generating daily weather variables. US Dept. of Agriculture, Agricultural Research Service, Publ. No. ARS–8, Washington, DC

Semenov MA, Barrow EM (1997) Use of a stochastic SWG in the development of climate change scenarios. Clim Change 35:397–414

Semenov MA, Brooks RJ (1999) Spatial interpolation of the LARS-WG SWG in Great Britain. Clim Res 11:137–148

Wilks DS (1999) Multisite downscaling of daily precipitation with a SWG. Clim Res 11:125–136

Wilks DS, Wilby RL (1999) The SWG game: a review of stochastic weather models. Prog Phys Geog 23:329–357

Editorial responsibility: Robert Davis, Charlottesville, Virginia, USA

Submitted: February 4, 2003; Accepted: May 28, 2003
Proofs received from author(s): July 7, 2003