A weather generator for hydrological, ecological, and agricultural applications

Valeriy Y. Ivanov,[1,2,3] Rafael L. Bras,[1] and David C. Curtis[4]

Received 2 August 2006; revised 19 May 2007; accepted 14 June 2007; published 4 October 2007.

[1] This paper presents a weather generator that allows simulation of hydrometeorological variables representative of a given geographic location: precipitation, total cloud cover, incoming shortwave radiation, air temperature, humidity, and wind speed. The approach captures the essential relationships among the quantities of interest, while modeling the diurnal variation of weather conditions at the hourly scale. Precipitation is considered to be the key driver of simulated hydrometeorological conditions, which leads to a consistent covariation of the weather variables. The generator was calibrated and validated with data from three meteorological stations located in New Mexico, Arizona, and Oklahoma. The set of variables reproduced by the weather generator can serve as input to a number of models of environmental systems, involving hydrological, ecological, water resources, and agricultural applications. The model is also suitable for creating scenarios of climate regimes (e.g., dry versus wet climates) useful in sensitivity studies. The source codes of the weather generator, manual, and test applications are publicly available.

Citation: Ivanov, V. Y., R. L. Bras, and D. C. Curtis (2007), A weather generator for hydrological, ecological, and agricultural applications, *Water Resour. Res.*, *43*, W10406, doi:10.1029/2006WR005364.

1. Introduction

[2] The forcing of weather and the features of its temporal variability need to be considered in most hydrological, ecological, water resources, and agricultural applications. Using series of observed meteorological data is always the best option to account for the local weather and climate characteristics, however, this may lead to underrepresentation of extreme events. Using only observed meteorological data also makes it impossible to explore the effects of changes in climate. Weather generators have been commonly used to overcome these issues [*Wilks and Wilby*, 1999]. While new models have been recently constructed [e.g., *Parlange and Katz*, 2000; *Schnorbus and Alila*, 2004; *Schoof et al.*, 2005], it can be noted that they generally (1) do not simulate all the required variables in a single framework, suggesting implicit simplifying assumptions, (2) use daily time step in contrast to the hourly scale important for capturing diurnal behavior, and (3) have intensive data requirements due to the common Fourier series representation of parameter seasonality. The climate simulator of *Curtis and Eagleson* [1982] overcomes many of the above problems and the weather generator presented here builds on their ideas.

2. Formulation of Weather Generator

[3] The weather generator simulates precipitation, total cloud cover, shortwave radiation, air temperature, atmospheric humidity, and wind speed. With these variables, it is possible to drive detailed models of the water and energy balances. The diurnal cycle is important for all these quantities, so hourly simulations are required. In a consistent climate most of these variables are physically and thermodynamically related. A sound weather generator must preserve most of such interactions to maintain consistency and realism.

2.1. Precipitation

[4] Precipitation is the most crucial meteorological variable for many applications. Its presence or absence also affects statistics of many other hydrometeorological variables. A large variety of stochastic models of precipitation has been developed over the past years [e.g., *Bonta*, 2004; *Buishand*, 1978; *Chin*, 1977; *Gabriel and Neuman*, 1962; *Marien and Vandewiele*, 1986; *Morissey and Krajewski*, 1993; *Smith and Schreiber*, 1973; *Woolhiser*, 1992; *Woolhiser and Osborn*, 1985]. The full bibliography of rainfall stochastic models is quite extensive, and the reader is referred to the review papers by *Georgakakos and Kavvas* [1987] and *Foufoula-Georgiou and Krajewski* [1995] for a concise summary. In the following, the discussion only deals with studies relevant to the described weather generator.

[1]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
[2]Center for the Environment, Harvard University, Cambridge, Massachusetts, USA.
[3]Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan, USA.
[4]Carlton Engineering, Inc., Shingle Springs, California, USA.

0043-1397/07/2006WR005364

[5] Among the widely used precipitation models [*Wilks and Wilby*, 1999] are the seasonally varying first-order two-state Markov models [e.g., *Richardson*, 1981] and the "renewal" models, also referred to as the spell length models. An apparent advantage of the latter type of models, which simulate precipitation by generating random numbers from the fitted dry and wet spell-length distributions, is in the knowledge of the time limits of a current hydrometeorological period (dry or wet spell) at any time. As will be shown later, this facilitates an explicit coupling of the precipitation model and other components of the weather generator, such as cloud cover and air and dew point temperatures. The renewal model approach is adopted in the discussed weather generator. While recognizing the existence of more advanced options, a simple Poisson arrival model that implies the exponential distribution for dry spell (interstorm) periods [*Todorovic*, 1968; *Todorovic and Yevjevich*, 1969] is implemented in the discussed framework. This model does not represent any diurnal effects on wet or dry spells. Nevertheless, as will become evident in the following discussion, any hourly simulator of rainfall arrival (or observed data) could be used to drive the generation of the rest of the hydrometeorological variables. The influence of the rainfall regime will be consistently accounted for by the coupling mechanisms.

[6] *Grayman and Eagleson* [1969] showed that storm durations and interstorm times could be treated as independent events. Thus the precipitation model can be expressed by successive sampling from the fitted probability density functions. Time between two successive storms t_b [hour] follows the exponential distribution with parameter μ_b [hour], which is the mean time between storms. The storm duration t_r [hour] is also simulated using the exponential distribution, with μ_r [hour] as the mean storm duration. *Grayman and Eagleson* [1969] showed that storm depths were highly dependent on storm durations. Storm depths h [mm] are thus assumed here to follow a gamma distribution conditioned by the ratio t_r/μ_r, with the parameter μ_d [mm], which is the mean storm depth.

[7] The following procedure is used to simulate rainfall. At some initial time t_0, an interstorm duration t_b is generated from the fitted exponential distribution. The period $[t_0, t_0 + t_b]$ is considered dry. When the time reaches $[t_0 + t_b]$, the storm duration t_r is generated. Using the value set for t_r, a storm depth h is generated from the gamma distribution. The period $[t_0 + t_b, t_0 + t_b + t_r]$ is then considered wet. When time reaches $[t_0 + t_b + t_r]$ the process is repeated to determine the next storm-interstorm sequence. The adopted method assumes the model of rectangular pulses that considers a uniform rainfall intensity throughout the entire period t_r. Such a simplification is not a serious constraint for a variety of applications that do not exhibit high sensitivity to rainfall rates. The model of rectangular pulses is still commonly used and appears in the current work. Nevertheless, many applications do call for more sophisticates models that allow for diurnal and intrastorm variability. As will become clear, the weather generator presented here is fairly independent of the choice of rainfall driver. More advanced schemes (e.g., that include intrastorm and diurnal rate variability) can be implemented [e.g., *Bonta*, 2004] in the context of the model presented in this paper.

2.2. Cloudiness

[8] Simulating cloudiness is necessary when one needs to explicitly consider the components of the energy balance, their temporal dynamics, and dependence on occurrence of precipitation events. The cloud cover, however, is not explicitly modeled by most of the existing weather generators. They commonly circumvent the problem by simulating the net solar radiation thus only implicitly accounting for daily cloudiness (e.g., as in commonly used models WGEN, CLIGEN, and USCLIMATE (later GEM)) [*Richardson*, 1981; *Nicks and Gander*, 1993, 1994; *Hanson et al.*, 1994, 2002]. Where studies have been performed [e.g., *Falls*, 1974; *Chia and Hutchinson*, 1991; *Aguiar and Collares-Pereira*, 1992], the cloud cover is treated independently of other hydrometeorological variables at the daily timescale, which is theoretically incorrect. *Curtis and Eagleson* [1982] provide a framework for hourly cloud cover simulation using information on the occurrence of intrastorm and interstorm periods.

[9] Cloud cover is defined here as the fraction of the celestial dome covered by clouds. The cloudiness process, $N(t)$ [dimensionless], is therefore bounded by "0" (clear sky) and "1" (overcast). Intermediate values can define a variety of hydrometeorological conditions, e.g., "0.2", scattered, ..., "0.7", broken, etc., *Curtis and Eagleson* [1982] consider $N(t)$ as a random nonstationary process composed of intrastorm and interstorm periods. During the intrastorm period, the expected value of the mean of the process is close to 1.0, while if $N(t)$ is examined near the middle of a sufficiently long interstorm period, the expected value is usually quite different from 1.0. The central assumption made in the model of *Curtis and Eagleson* [1982] is that there is a loosely centered subregion, R_0, around the midpoint of the interstorm period in which the process $N(t)$ can be assumed stationary. By examining the first and second moment properties of the process, they conclude that the "fair weather" cloud cover process in this subregion is unaffected by approaching or receding precipitation systems: $E(N(t) \mid t_b)_{t \in R_0} = E(N(t))_{t \in R_0} = M_0$, $\mathrm{Var}(N(t) \mid t_b)_{t \in R_0} = \mathrm{Var}(N(t))_{t \in R_0} = \sigma_m^2$, M_0 [dimensionless] is the mean and σ_m^2 [dimensionless] is the variance of the fair weather $N(t)$. Another major assumption is that there is a smooth transition of mean and variance from the boundaries of fair weather period:

$$N(t) = M_0 + (1 - M_0)(1 - J(t)) + m(t)J(t), \tag{1}$$

where $J(t)$ is an assumed transition function and $m(t)$ is the stationary sequence of correlated deviations with $E(m(t)) = 0$ and $\mathrm{Var}(m(t)) = \sigma_m^2$, and autocorrelation function $\rho_N(\tau)$, where τ [hour] is the lag. The time varying conditional expectation and variance of cloud cover under this assumption are obtained as [*Curtis and Eagleson*, 1982]

$$E(N(t) \mid t_b) = M_0 + (1 - M_0)(1 - J(t)), \tag{2}$$

$$\mathrm{Var}(N(t) \mid t_b) = \sigma_m^2 J^2(t), \tag{3}$$

where $t \in t_b$. The autocorrelation structure of the cloud cover process (1) is not affected by $J(t)$ and is identical to the autocorrelation function of the process $m(t)$: $\rho_N(\tau) = \rho_m(\tau)$. The transition function $J(t)$ is assumed to be of an exponential form

$$J(t) = \left(1 - e^{-\varsigma(t-t_0)}\right)\left(1 - e^{-\gamma(t_0+t_b-t)}\right), \quad (4)$$

where ς and γ [h^{-1}] are decay coefficients controlling the transition rates from the boundaries (end/beginning of precipitation events) to/from the region R_0. These rates are assumed to be equal in the current model implementation. As follows from (4), for $t \in R_0$, $\lim_{t_b \to \infty} J(t) = 1$. $J(t)$ reaches a value close to 1.0 for all reasonable values of the decay coefficients and therefore the fair weather cloudiness is essentially simulated as $\lim_{t_b \to \infty} N(t) = M_0 + m(t)$, where $m(t)$ is taken to be a first-order Markov process:

$$m(t) = \rho_m(1)m(t-1) + \varepsilon(t)\sigma_m\sqrt{1 - \rho_m^2(1)}, \quad (5)$$

where $\varepsilon(t)$ is a random deviate. In the model of *Curtis and Eagleson* [1982], $\varepsilon(t)$ is assumed to be a normally distributed variable, $N(0, 1^2)$. However, a modification is introduced here and $\varepsilon(t)$ is simulated as a random deviate from the Beta probability distribution. The lower and upper bounds (y_1, y_2) of this distribution are found from (1) and (5) at every simulation step, i.e., the knowledge of $m(t-1)$ and $J(t)$ at every time t allows one to derive y_1 and y_2 by imposing the requirement for the cloud cover $N(t)$ to be in the range [0, 1]. The Beta distribution shape parameters a and b are estimated based on conditioning by the cloudiness value at $(t-1)$ (see section 4.2).

2.3. Shortwave Radiation

[10] The spatial and temporal distribution of surface solar irradiance exerts one of the fundamental controls on the land-surface energy and water dynamics. While hydrological estimation is typically limited to using only global shortwave radiation, a number of applications require radiation partition into type as well as wave band. For instance, the spectral band [0.38 μm ÷ 0.71 μm] corresponds to photosynthetically active radiation (PAR), which constitutes one of the principal determinants of biomass production. Therefore an ecological or ecohydrological simulation must consider PAR as one of the inputs. The methodology presented here attempts to satisfy such needs, permitting the parameterization of the model for essentially any geographic location and a wide range of atmospheric conditions.

[11] For all practical purposes, the Sun radiates its energy at a constant rate. Outside the atmosphere, at the mean solar distance, the beam irradiance, i.e., the solar constant, S_o, is 1366.1 [W m^{-2}], as adopted by the American Society for Testing and Materials [*Darula et al.*, 2005]. To allow for the varying solar distance (due to eccentricity of the Earth's orbit), a ratio of the actual Earth-Sun to the mean Earth-Sun distance, r [dimensionless], is introduced so that $S'_o = S_o/r^2$, where $r = 1.0 + 0.017\cos[(2\pi/365)(186 - \mathrm{JDay})$, where JDay is the Julian day. Several other variables defining the Sun's position with respect to a location on Earth are used in the following. Their definitions are provided in auxiliary material Text S1, section A.[1]

[12] The model considers two wide bands of solar spectrum: the ultraviolet (UV)/visible (VIS) band, $B_{\Lambda 1}$ [0.29 μm ÷ 0.70 μm], where ozone absorption and molecular scattering are concentrated, and the infrared in near- and short-wavelength range (NIR), $B_{\Lambda 2}$, [0.70 μm ÷ 4.0 μm], where water and mixed gases absorptions are concentrated. The spectrum separation into two bands facilitates the transmittance modeling of beam and diffuse clear sky irradiances because overlap between scattering and selective absorption is limited [*Gueymard*, 1989]. In the presence of clouds, the chosen spectral limits are also convenient due to the above differences in absorption properties by water droplets. Moreover, separate treatment of these two bands allows one to explicitly compute PAR, which is used in the process of leaf photosynthesis. According to *Slingo* [1989], the energy contained in the two bands, $B_{\Lambda 1}$ and $B_{\Lambda 2}$, is 46.6% and 53.4% of S'_o, respectively. These fractions are applied to obtain the extraterrestrial flux densities in the two bands: $S_{o\Lambda 1}$ and $S_{o\Lambda 2}$ [W m^{-2}].

2.3.1. Clear Sky

2.3.1.1. Direct Beam Irradiance

[13] It is assumed that direct rays entering the atmosphere encounter extinction processes, which are limited to the effects of [*Gueymard*, 1989] ozone, uniformly mixed gases, water vapor, and aerosol absorption, and Rayleigh and aerosol scattering. The transmittances T_X due to an extinction process X are defined in Text S1, section B. Separate extinction layers are considered, so that the total direct beam flux at the ground at normal incidence is $S_b^\mu = S_{o\Lambda 1}\prod_X T_{X,1} + S_{o\Lambda 2}\prod_X T_{X,2} = S_{b\Lambda 1}^\mu + S_{b\Lambda 2}^\mu$.

2.3.1.2. Diffuse Irradiance

[14] The diffuse irradiance at the ground level is modeled as a combination of three components I_{dX} corresponding to (the X index) the two scattering layers (molecules and aerosols) and to a backscattering process between ground and sky (Text S1, section B2). The total diffuse flux at the ground is therefore $S_d = \sum_X I_{dX,1} + \sum_X I_{dX,2} = S_{d\Lambda 1} + S_{d\Lambda 2}$.

2.3.2. Cloudy Sky

[15] Clouds alter transmission and reflection properties of the atmosphere. Therefore the effects of cloud cover need to be carefully accounted for. Cloud parameterizations provided in hydrological literature typically use empirical functions of the fractional cloud cover [e.g., *Becker*, 2001] or specify fixed bulk properties. However, cloud radiative characteristics strongly depend on cloud type, structure, and density [*Thomas and Stamnes*, 1999]. Also, different spectral intervals exhibit distinct absorption and scattering characteristics [*Slingo and Schrecker*, 1982].

[16] The framework used here relies on parameterizations developed by *Stephens* [1978] and *Slingo* [1989]. On the basis of both observational and theoretical evidence, these studies infer that the cloud radiative properties are mainly determined by the cloud total vertical liquid water path, LWP [g m^{-2}]. *Stephens* [1978] showed that the broadband optical thickness is essentially the same for clouds of different types that have the same LWP. *Slingo* [1989] introduced an additional independent relationship between

[1] Auxiliary materials are available in the HTML. doi:10.1029/2006WR005364.

cloud radiative properties and the effective radius of drop size distribution. The advantage of using LWP is that this quantity can be obtained by satellite microwave radiometry [e.g., *Han et al.*, 1994] and allows the introduction of seasonality effects in the cloud properties.

[17] The approach developed by *Stephens* [1978] is used here to derive the effective radius of drop size distribution based on LWP. The four-band model of *Slingo* [1989] is then used to derive the cloud transmittances and reflectances for the incident clear sky direct $S^{\mu}_{b\Lambda i}$ and diffuse $S^{\mu}_{d\Lambda i}$ fluxes (assumed to be incident on top of clouds).

2.3.2.1. Direct Beam Irradiance

[18] The model uses the simulated cloud cover, N [0.0 ÷ 1.0], to differentiate between the cloudless, $(1 - N)$, and cloudy, N, sky fractions. The direct beam flux from the clear sky fraction is assumed to reach the ground surface unaltered. The cloudy sky fraction is assumed to contain a homogeneous layer of clouds characterized by LWP [g m^{-2}]. The total direct beam normal irradiance $S^{\mu}_{B'\Lambda j}$ in each of the bands j, j = 1,. . .,4 of *Slingo*'s [1989] parameterization (Text S1, section C) is estimated as a linear combination of the fluxes from the clear sky and cloudy fractions of the sky dome:

$$S^{\mu}_{B'\Lambda j} = S^{\mu}_{b\Lambda i}\left[(1 - N) + T_{DBj}\,N\right]\frac{k_j}{K}, \tag{6}$$

where T_{DBj} [dimensionless] is the cloud transmissivity for the direct beam flux in band j, k_j is the respective fraction of solar irradiance at the top of the atmosphere in each band [0.466, 0.320, 0.181, 0.033] (Text S1, section C1), and K = 0.466 if i = 1, K = 0.534 if i = 2. The fluxes in the two bands $B_{\Lambda 1}$ and $B_{\Lambda 2}$ are then obtained as $S^{\mu}_{B\Lambda 1} = S^{\mu}_{B'\Lambda 1}$ and $S^{\mu}_{B\Lambda 2} = \sum_{j=2}^{4} S^{\mu}_{B'\Lambda j}$.

2.3.2.2. Diffuse Irradiance

[19] As above, diffuse flux from the clear sky fraction is assumed to reach the ground surface unaltered. The diffuse radiative flux at the cloud bottom may result from both the diffuse and beam radiation incident at the cloud top. The total diffuse irradiance $S_{D'\Lambda j}$ in each of the bands j = 1,. . .,4 (Text S1, section C2) is estimated as a linear combination of the fluxes from the clear sky and cloudy fractions:

$$S_{D'\Lambda j} = (1 - N)S_{d\Lambda i} + N\left[T_{DIRj}\,S^{\mu}_{b\Lambda i} + T_{DIFj}\,S_{d\Lambda i}\right]\frac{k_j}{K}, \tag{7}$$

where T_{DIRj} [dimensionless] is the diffuse transmissivity for direct incident radiation and T_{DIFj} [dimensionless] is the diffuse transmissivity for diffuse incident radiation. The fluxes in the two bands $B_{\Lambda 1}$ and $B_{\Lambda 2}$ are then obtained as: $S_{D\Lambda 1} = S_{D'\Lambda 1}$ and $S_{D\Lambda 2} = \sum_{j=2}^{4} S_{D'\Lambda j}$.

2.3.3. Terrain Effects

[20] The spatial distribution of solar radiation over the surface is strongly affected by small-scale topographic features such as slope angle, aspect, and screening or reflection effects from the surrounding terrain. Any application that takes into account the geometry of irradiated surface needs to explicitly consider the corresponding effects of shading or exposure. The implemented approach follows *Olseth et al.* [1995]. Details can be found in Text S1, section D.

2.4. Air Temperature

[21] A number of stochastic weather generators that include the capability for modeling air temperature have been proposed for agricultural simulations and climate studies [e.g., *Richardson and Wright*, 1984; *Hanson et al.*, 1994, 2002; *Nicks and Gander*, 1993, 1994]. These models typically simulate daily maximum and minimum temperature. The majority of these models are based on the multivariate stationary process that permits autocorrelation in the individual time series and cross correlations between the time series of air temperature and precipitation, e.g., conditioning by the wet or dry day occurrence (e.g., WGEN and USCLIMATE [*Richardson and Wright*, 1984; *Hanson et al.*, 1994]). The common problem, however, is that the temperature diurnal variation is neglected. Moreover, these approaches are seriously handicapped because they do not consider the effects of other variables (e.g., cloud cover) on a continuous basis. The model of *Curtis and Eagleson* [1982] deals with both problems. The model assumes that the hourly air temperature, $T(t)$ [°C], is a sum of two variables: a deterministic air temperature component $\tilde{T}(t)$ and a random variate $\delta T(t)$:

$$T(t) = \tilde{T}(t) + \delta T(t). \tag{8}$$

2.4.1. Definition of the Deterministic Component $\tilde{T}(t)$

[22] The deterministic component is built on an empirical method of *Bryan* [1964] that attributes temporal variation of the air temperature to the divergence of radiative heat flux and the divergence of eddy heat flux. In essence, an assumption is made that hourly temperature increments can be regressed on several hydrometeorological variables:

$$\frac{d\tilde{T}(t)}{dt} = b_0 - b_1\tilde{T}(t) + b_2K(t)s(t) + b_3K(t)r(t) + b_4q(t), \tag{9}$$

where b_i-s (i = 0, 1,. . ., 4) are the regression coefficients, $s(t)$ and $r(t)$ are the variables of the Sun position and geographic location, $K(t) = 1 - 0.65N^2(t)$ is the radiation attenuation factor, and $q(t)$ is the estimate of incoming longwave radiation. Expression (9) excludes several terms of the original model of *Curtis and Eagleson* [1982]: the ground temperature (not a standard measurement variable and not readily available), and wind speed and direction (assumed to have generally minimal contribution in modifying the air temperature). Equation (9) relates the change in temperature to a number of factors that operate throughout the daily and seasonal cycle. For example, higher values of cloud cover result in lower amplitude of the daily temperature because of the terms containing $s(t)$ and $r(t)$. The term containing $q(t)$ is nonzero throughout the entire day and should explain some of the differences in cooling observed on clear nights as opposed to cloudy nights. Absolute magnitudes of $s(t)$ and $r(t)$ vary for different seasons and geographic locations (Text S1, section E).

[23] The first-order differential equation (9) can be solved if the initial temperature is provided. *Curtis and Eagleson* [1982] provide a solution method summarized in Text S1, section E.

2.4.2. Definition of the Random Deviate Component $\delta T(t)$

[24] The deterministic component in the model of *Curtis and Eagleson* [1982] essentially represents the expected temperature value. It cannot explain all of the temperature variability and therefore the random deviate $\delta T(t)$ [°C] is introduced, defined as: $\delta T_o(t) = T_o(t) - \tilde{T}(t)$, where $T_o(t)$ [°C] is the observed air temperature and $\tilde{T}(t)$ [°C] is the deviation component. The deviations are assumed to be approximated by a first-order Markov process:

$$\delta T(t) = \overline{\delta T} + \rho_{\delta T}(1)\left(\delta T(t-1) - \overline{\delta T}\right) + \varepsilon_T(t)\sigma_{\delta T}\sqrt{1-\rho^2_{\delta T}(1)}, \tag{10}$$

where $\overline{\delta T}$, $\sigma^2_{\delta T}$, and $\rho_{\delta T}(1)$ are the mean, variance, and lag-1 value of autocorrelation of random temperature deviates, respectively, and $\varepsilon_T(t)$ is the standard normal deviate.

2.5. Air Humidity

[25] Some weather generators do include simulation of quantities that can be translated into air humidity. *Ahmed* [1974] generated air humidity within the framework of a multivariate model using mean weekly relative humidities. *Hanson et al.* [2002] simulated the daily dew point assuming that it is normally distributed. *Parlange and Katz* [2000] extended the model of *Richardson* [1981] to include the dew point temperature as a component of multivariate stochastic process. *Curtis and Eagleson* [1982] offer a multiple linear regression model for cases when the cross correlation of dew point with other hydrometeorological variables is significant. The model uses 11 parameters. Commonly, however, dew point temperatures are found to stay relatively constant during the day [e.g., *Glassy and Running*, 1994], which suggests the possibility of a simpler model with fewer parameters.

[26] Observations have shown that nightly minimum temperatures, T_{mind} [°C], tend to come into equilibrium with daily dew point temperatures, T_{dew} [°C] [*Running et al.*, 1987]. Therefore T_{mind} is often used as an indirect measure of T_{dew}. However, *Kimball et al.* [1997] showed that substantial differences may exist between these two quantities, especially in arid and semiarid climates. On the basis of long-term records for a number of climatic regions, *Kimball et al.* [1997] proposed an empirical model that allows for the adjustment of daily dew point temperature with respect to minimum air temperature using information on daily potential evaporation and a degree of region aridity:

$$T_{\text{dew}} = T_{\text{mind}}\left[-0.127 + 1.121\left(1.003 - 1.444\text{EF} + 12.312\text{EF}^2 - 32.766\text{EF}^3\right) + 0.0006\Delta T_d\right] - 273.15, \tag{11}$$

where ΔT_d [°C] is the amplitude of daily air temperature and EF [dimensionless] is the evaporative factor, $0 \le \text{EF} \le 1$ that *Kimball et al.* [1997] define as

$$\text{EF} = \frac{1}{P_{\text{ann}}}\frac{E_p}{\rho_w}D_{LH}, \tag{12}$$

where Ep [kg m^{-2} s^{-1}] is the daily potential evapotranspiration, ρ_w [kg m^{-3}] is the water density, D_{LH} [s] is the day length (Text S1, section A), and P_{ann} [m] is the annual precipitation. When estimating E_p following *Priestley and Taylor* [1972], *Kimball et al.* [1997] assume that the ground surface has a seasonally constant albedo, daily net radiation can be approximated using air temperature, and that ground heat flux is 10% of net radiation. *Kimball et al.* [1997] show that the model (11) improves estimates of T_{dew} based on T_{mind}.

[27] The major difficulty in using (11) is to compute the daily values of E_p and ΔT_d, since both quantities have to be available at the beginning of each day when T_{dew} is estimated. Both E_p and ΔT_d, however, are determined by a number of variables simulated at the hourly scale that cannot be readily predicted for the entire day. For instance, cloudiness affects temperature estimation and since both models use random deviates, there is no exact method to predict daily air temperature amplitude in advance. It is assumed that adjusted values of E_p and ΔT_d from the previous day, $(d-1)$, can be used for estimation of T_{dew} on the current day d. Since cloud cover affects net radiation and therefore energy available for evapotranspiration, an adjustment factor based on the radiation attenuation factor, $K(t)$, is introduced:

$$\tilde{A}(d) = \frac{\tilde{K}(d)}{\overline{K}(d-1)} = \frac{1 - 0.65^2\tilde{N}(d)}{1 - 0.65^2\overline{N}(d-1)}, \tag{13}$$

where $\overline{K}(d-1)$ is the average factor value for the previous day and $\tilde{K}(d)$ is the mean expected value for the day of estimation. The use of "renewal" model of rainfall arrival allows for the estimation of $\tilde{K}(d)$: at any time between successive storms, both the end time of the previous and beginning time of the next storm are known. The cloud cover model (1) is thus used to estimate an expected value of cloudiness for the following day, from which the factor $\tilde{A}(d)$ is then estimated. Once $\tilde{A}(d)$ is known, the expected values of E_p and ΔT_d that appear in equations (11)–(12) are approximated as $\tilde{E}_p(d) = \tilde{A}(d)\overline{E}_p(d-1)$, $\Delta\tilde{T}_d(d) = \tilde{A}(d)\Delta T_d(d-1)$.

[28] Two other variables are required for the estimation of T_{dew}: T_{mind} and P_{ann}. The problem of computing T_{mind} is similar to the one described above. Air temperature at the hour preceding sunrise can be usually associated with T_{mind}. Hence the air temperature from (8) at the hour preceding sunrise is taken as T_{mind} and used in the estimation of T_{dew}. *Kimball et al.* [1997] used the mean annual precipitation P_{ann} for all days throughout the year. Here, it is assumed that a monthly basis is more appropriate since different months/seasons have different degree of dryness. Therefore P_{ann} is considered as a precipitation parameter for each month, P^*_{ann}.

2.6. Wind Speed

[29] Typically, the cross-correlation coefficients between wind speed and other hydrometeorological variables are small. For example, using hourly weather data for Massachusetts and Kansas, *Curtis and Eagleson* [1982] estimated cross-correlation with maximum values of 0.35, usually around 0.1 (for air and dew point temperature and cloud cover). *Parlange and Katz* [2000] used daily data from

Table 1. Parameters Used in the Weather Generator

Parameter	Definition	Estimation Period
μ_b [hour]	mean time between storms	month-season
μ_r [hour]	mean storm duration	month-season
μ_d [mm]	mean storm depth	month-season
M_0 [dimensionless]	mean fair weather cloudiness	month
σ_m^2 [dimensionless]	variance of fair weather cloudiness	month
$\rho_m(1)$ [dimensionless]	lag-1 autocorrelation of fair weather cloudiness	month
$\varsigma = \gamma$ [h^{-1}]	cloudiness decay rates from the end of precipitation event/"fair weather" region to "fair weather" region/beginning of precipitation event	month
a and b [dimensionless]	shape parameters of the Beta distribution (conditioned by cloudiness at the previous hour)	month
S_o [W m^{-2}]	solar constant, S_o = 1366.1	constant
ΔGMT [hour]	time difference between the local time zone and GMT	constant
u_o [cm]	ozone amount in a vertical column	month-season
β [dimensionless]	Angström turbidity parameter	constant/season
ω_{Ai} [dimensionless]	aerosol single-scattering albedos (VIS and NIR bands)	constant/season
ρ_g [dimensionless]	spatial average regional albedo	constant/season
LWP$_o$ [g m^{-2}]	cloud total liquid water path for overcast conditions	month
a_i − s and b_i − s	(i = 0, 1,. . ., 4) regression coefficients in the equation for "deterministic" hourly temperature increment	month
$\overline{\delta T}$ [°C]	mean of random temperature deviates	month
$\sigma_{\delta T}^2$ [°C]	variance of random temperature deviates	month
$\rho_{\delta T}(1)$ [dimensionless]	lag-1 autocorrelation of random temperature deviates	month
$\tilde{A}(d)$ [dimensionless]	parameter used for adjustment of $\overline{E}_p(d)$ and $\Delta T_d(d-1)$	constant
P_{ann}* [mm]	precipitation index in dew temperature model	month
$\overline{W}$ [m s^{-1}]	mean wind speed	month-season
σ_s^2 [m^2 s^{-2}]	variance of wind speed	month-season
$\rho_s(1)$ [dimensionless]	lag-1 autocorrelation of wind speed	month-season
$\gamma_s(t)$ [dimensionless]	skewness of wind speed	month-season

Oregon and reported cross correlation not exceeding 0.2 (for daily maximum and minimum air temperatures and dew point). Therefore wind speed has been conventionally simulated as an independent variable [*Curtis and Eagleson*, 1982; *Nicks et al.*, 1990; *Hanson et al.*, 2002].

[30] The distribution of both hourly and daily wind speed is positively skewed [*Carlin and Haslett*, 1982; *Hennessey*, 1977]. To generate skewed hourly wind speed data, while preserving the first two moments of its distribution, *Curtis and Eagleson* [1982] use the *AR*(1) model, where the random term forces skewness on the results of the autoregressive model, leading to an approximately Gamma distribution of wind speed [*Maass et al.*, 1962]:

$$W_s(t) = \overline{W}_s + \rho_s(1)\left(W_s(t-1) - \overline{W}_s\right) + \varepsilon_t \sigma_s \sqrt{1-\rho_s^2(1)}, \quad (14)$$

where $\overline{W}_s$ [m s^{-1}], σ_s^2 [m^2 s^{-2}], and $\rho_s(1)$ [dimensionless] are the mean, variance, and lag-1 value of autocorrelation of wind speed, respectively. The term $\varepsilon_t(t)$ is [*Wilson and Hillerty*, 1931] $\varepsilon_t = \frac{2}{\gamma_\varepsilon}\left[1 + \frac{\gamma_\varepsilon \psi_t}{6} - \frac{\gamma_\varepsilon^2}{36}\right]^3 - \frac{2}{\gamma_\varepsilon}$, where ψ_t is a standard normal deviate and $\gamma_\varepsilon(t)$ is the skewness of $\varepsilon_t(t)$: $\gamma_\varepsilon = (1-\rho_s^3)\gamma_s/(1-\rho_s^2)^{1.5}$, where $\gamma_s(t)$ is the skewness determined from the wind speed data. While *Curtis and Eagleson* [1982] considered variation of $\overline{W}_s$ and σ_s^2 with time of the day, this implementation assumes that these parameters are time-invariant.

3. Data

[31] The model was tested using data from three meteorological stations: Albuquerque International Airport (New Mexico), Tucson International Airport, (Arizona), and Tulsa International Airport, (Oklahoma). This paper only shows the representative results for the location of Albuquerque (*Ivanov* [2006] and Text S1, section F, provides the other test results). The Albuquerque International Airport (New Mexico) is located at 35.05°N, 106.617°W. The data were

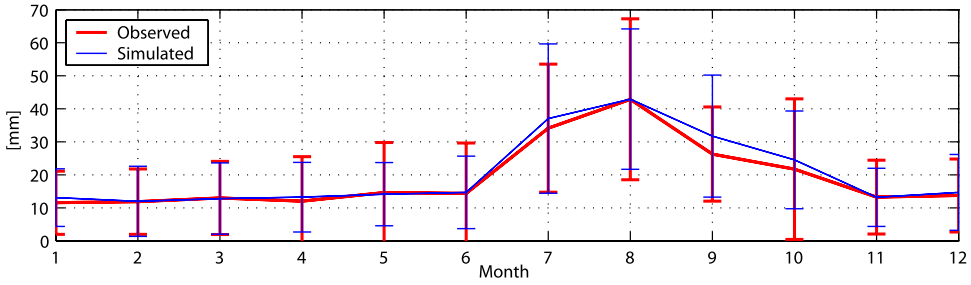

Figure 1. Observed and simulated mean monthly rainfall (Albuquerque, New Mexico). The vertical bars denote the estimated standard deviation of the monthly value.

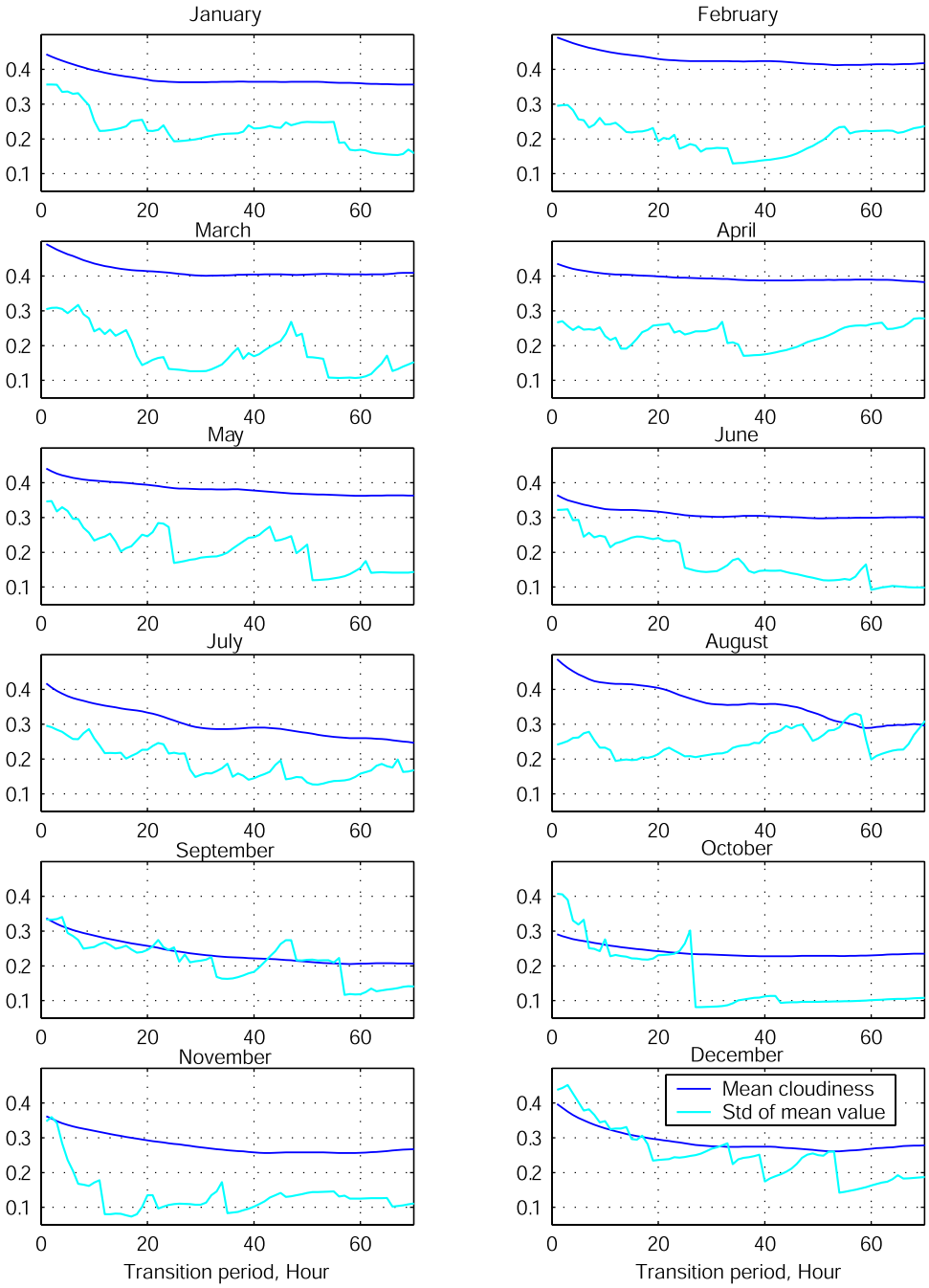

Figure 2. Estimated mean cloud cover value and standard deviation of the estimate as a function of the length of transition period (Albuquerque, New Mexico).

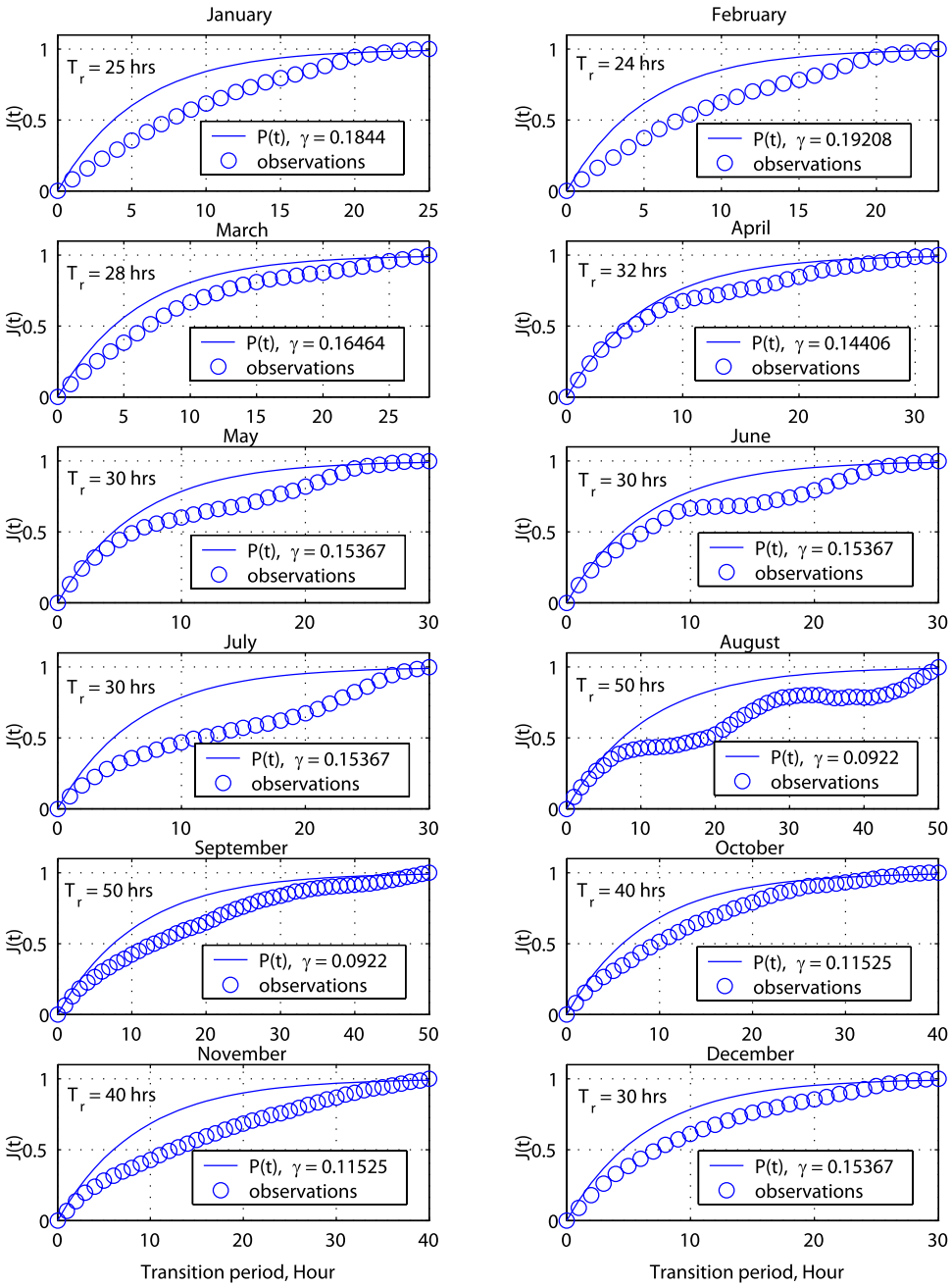

Figure 3. Analytical and observed transition functions $J(t)$ corresponding to the estimated transition period lengths (Albuquerque, New Mexico).

Table 2. Parameter Values of M_0, σ_m^2, and $\rho_m(1)$ of the Cloud Cover Model for the Location of Albuquerque, New Mexico

	ALB		
Month	M_0	σ_m^2	$\rho_m(1)$
January	0.363	0.407	0.916
February	0.424	0.410	0.909
March	0.402	0.400	0.893
April	0.391	0.389	0.896
May	0.381	0.376	0.896
June	0.302	0.349	0.908
July	0.288	0.321	0.879
August	0.323	0.354	0.853
September	0.210	0.329	0.918
October	0.228	0.345	0.922
November	0.257	0.357	0.904
December	0.275	0.384	0.911

available for the period of January 1961 to December 1995. The climate of Albuquerque is arid continental with a wide seasonal range of temperatures. More than 75% of the daylight hours have sunshine. The air is normally dry and humid days are rare. Nearly half of the annual precipitation in Albuquerque results from afternoon and evening thunderstorms during the summer. Thunderstorm frequency increases rapidly at the beginning of July, peaks during August, then diminishes by the end of September. Thunderstorms are usually brief and sometimes heavy. Small amounts of precipitation fall in the winter. Temperatures in Albuquerque are those characteristic of a dry, high-altitude, continental climate. The average daily range of temperature is relatively high. High temperatures during the winter are near 10°C with only a few days on which the temperature is below zero. In summer, daytime maxima are about 32°C. Sustained winds of 5.4 [m s^{-1}] or less occur approximately 80% of the time.

4. Parameter Estimation and Model Verification

[32] Parameter estimation is discussed and statistics derived from the observed hydrometeorological data are compared with those obtained from a 100-year simulation period. The generated series correspond to complete weather simulations, which involve synthetic modeling of the entire set of climate variables, accounting for the discussed linkages among them. A complete list of model parameters is provided in Table 1.

4.1. Precipitation

[33] The required parameters for the precipitation simulator include the mean values of time between storms μ_b, storm duration μ_r, and storm depth μ_d. The major challenge in estimating these parameters is separating point precipitation records into statistically independent storms [*Restrepo-Posada and Eagleson*, 1982]. Multiple mesoscale precipitation events are embedded in a single synoptic-scale disturbance and each event may produce intervals of rainfall followed by periods without rainfall before the next event arrives. A statistical method is generally required when operating with series of point precipitation data.

[34] While identifying independent interarrival times consistent with Poisson arrivals process, *Restrepo-Posada and Eagleson* [1982] found $t_{b\min}$ (the minimum separation interval between independent storms) ranging between 8 and 76 hours at various locations. Dry climates had higher values, while more humid climates were found to have lower values. *Restrepo-Posada and Eagleson* [1982] thus concluded that for precipitation models that assume rectangular pulse of storm rate, a strict requirement of independence is operationally impractical. If such long separation intervals (e.g., 8–76 hours) are imposed, long storm durations would result and these intervals would contain many periods without precipitation. It was therefore suggested that for problems where the dynamic response of a hydrologic system to precipitation inputs is of major interest, the rainfall model parameters should be estimated from raw storm data. *Grace and Eagleson* [1967] and *Sariahmed and Kisiel* [1968] used autocorrelation of successive storm depths as a surrogate indicator for storm independence. The time lag at which autocorrelation was not significantly different from zero was used as the $t_{b\min}$. They found $t_{b\min}$ to be between 2 and 3 hours (for convective storms). *Curtis and Eagleson* [1982] used the value of 2 hours. The same methodology is used here and for Albuquerque (New Mexico) the minimum lull duration is taken to be 3 hours.

[35] For climates with pronounced precipitation seasonality, the parameters μ_b, μ_r, and μ_d have to account for the intraannual variability. Precipitation seasons are identified by analyzing the mean annual cycles of precipitation. For Albuquerque, four seasons were identified: months November–June, July–August, September, and October. The corresponding mean values of μ_d/μ_r [mm h^{-1}], μ_r [hour], and μ_b [hour] are [0.778/3.61/152.7], [1.661/2.44/70.3], [1.161/3.197/85.9], and [1.048/3.96/122.7] for each season, respectively. Since the parameters are sampled from the assumed analytical distributions, their statistical properties are inherently preserved in the simulated data. Figure 1 illustrates the annual cycle of the rainfall process for Albuquerque.

4.2. Cloud Cover

[36] The parameters of the cloud cover model are M_0, σ_m^2, $\rho_m(1)$, γ, a, and b. The procedure of parameter estimation follows that of *Curtis and Eagleson* [1982] with some modifications.

[37] The existence of stationary interstorm "fair weather" cloud cover process is the central assumption of the model. Therefore identification of sequences of these periods in series of meteorological data is essential. The methodology introduced by *Curtis and Eagleson* [1982] employs an iterative approach that estimates the mean cloud cover for some subregion Δt within an interstorm period between successive precipitation events. Each interstorm period of length T_{is} [hour] is considered to be constrained by the last hour of the first rainfall event and the first hour of the following one. By successively eliminating 1 hour from both ends ($\Delta\tau_1 = 1$ hour, $\Delta\tau_2 = 2$ hours, etc.), a number of subregions, not exceeding in total ($T_{is}/2 - 1$), can be defined for each period. For any given subregion, Δt_k, corresponding to k eliminated hours from each end, a mean value of the cloud cover is estimated over all interstorm periods in the precipitation record exceeding $2k$ hours. Since $k \in [0, T_{\mathrm{is\,max}}/2 - 1]$, where $T_{\mathrm{is\,max}}$ is the maximum duration of an interstorm period in the record, a vector of the mean values of cloud cover of length ($T_{\mathrm{is\ max}}/2 - 1$) is obtained.

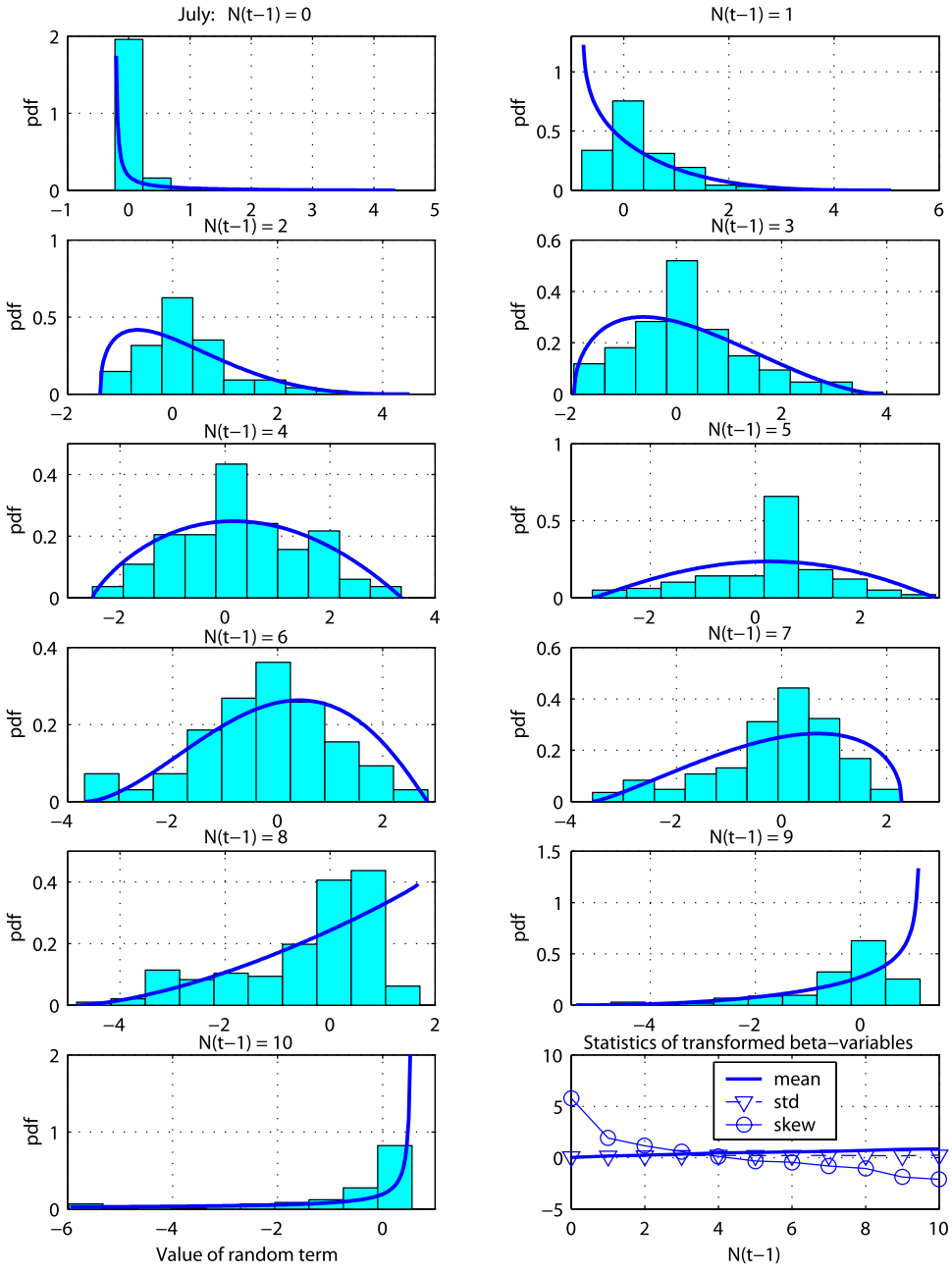

Figure 4. Histograms of deviates $\varepsilon(t)$ in the cloud cover model and the corresponding probability density function (solid line) approximated with the Beta distribution. The cloud cover $N(t-1)$ for the month of July is given on a [0, 10] basis (Albuquerque, New Mexico).

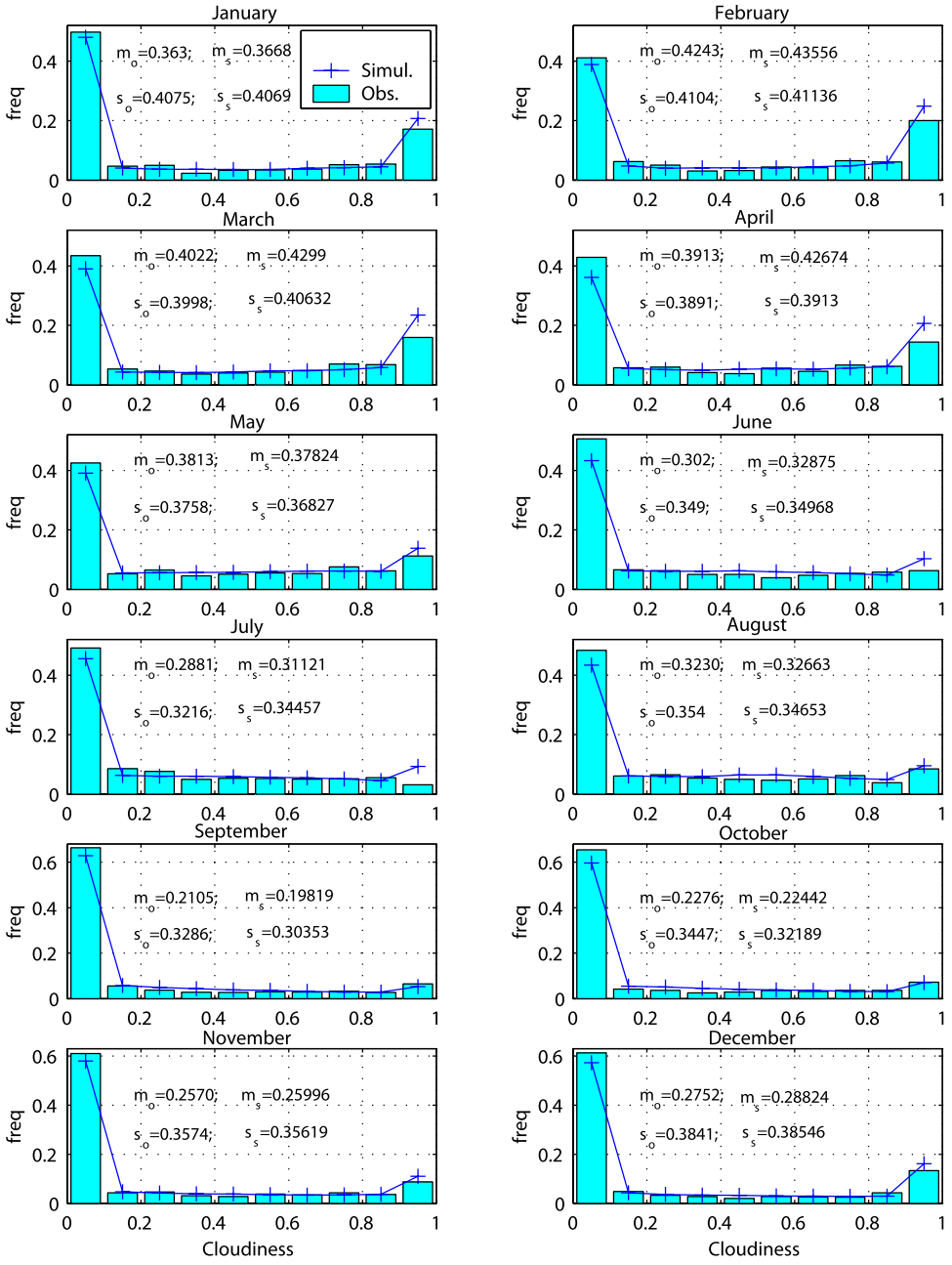

Figure 5. Observed and simulated cloud cover distribution (Albuquerque, New Mexico). Symbols *m* and *s* are the mean and standard deviation values, correspondingly, for the observed (subindex "o") and simulated (subindex "s") data.

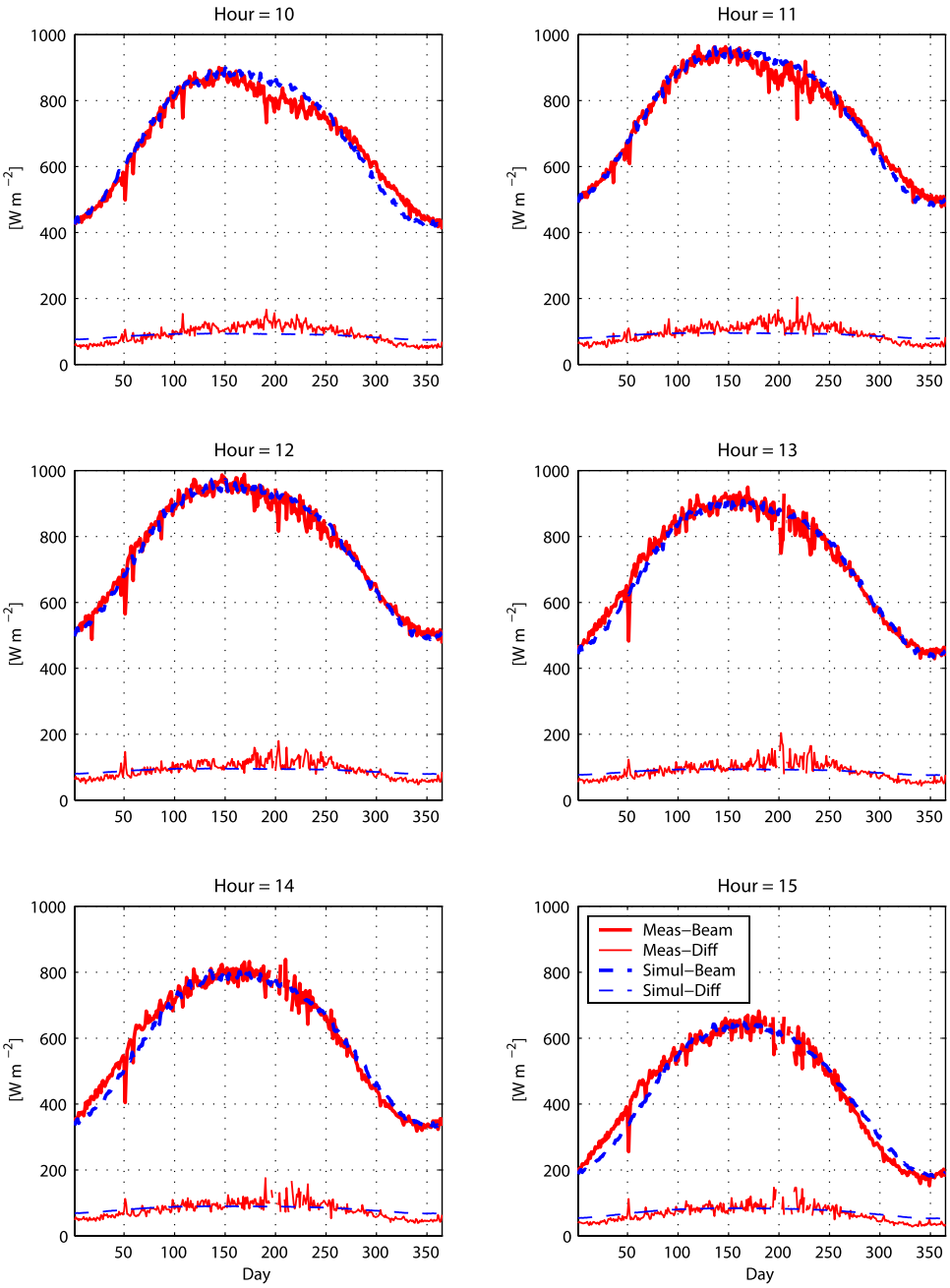

Figure 6. Seasonal cycles of the observed and simulated direct beam and diffuse flux on a horizontal surface in clear sky conditions for hours 10–15 (on a 24-hour basis, Albuquerque, New Mexico).

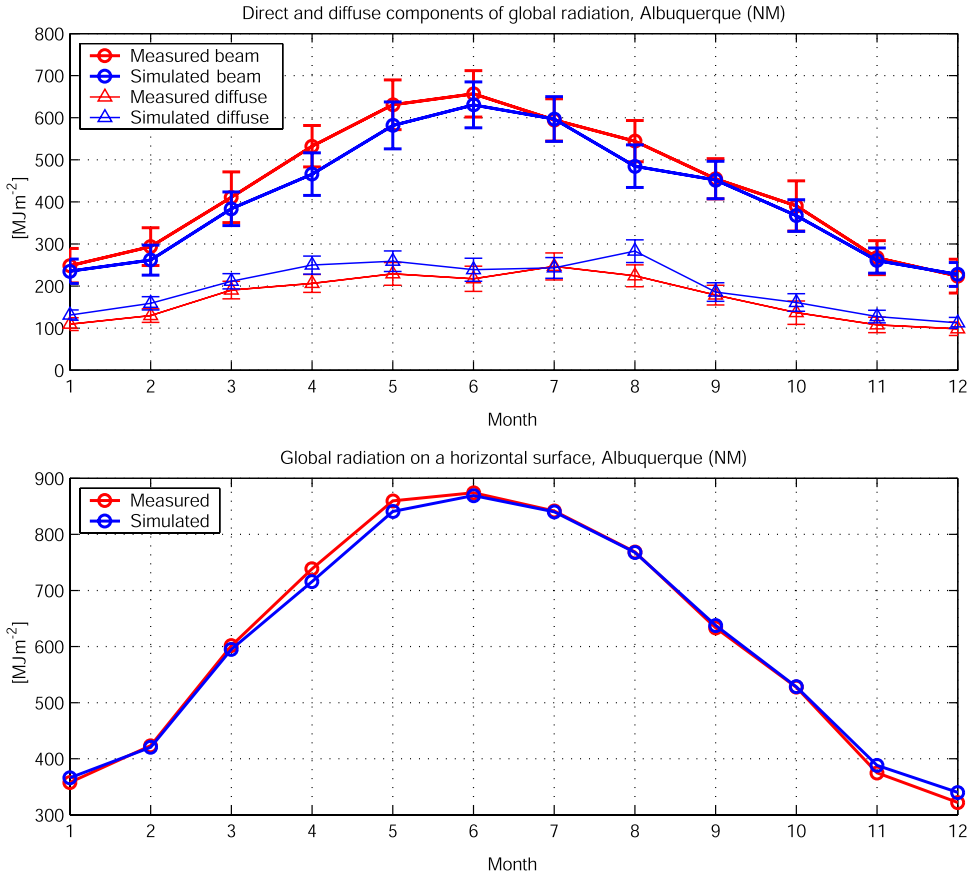

Figure 7. Seasonal cycles of mean monthly observed and simulated direct beam flux, diffuse flux, and global radiation on a horizontal surface for all sky conditions (Albuquerque, New Mexico).

[38] *Curtis and Eagleson* [1982] argue that with the increasingly larger number of eliminated hours, the estimated mean value of cloud cover stabilizes, reaching the fair weather mean value, M_0. The number of hours, T_r, eliminated from both ends of all interstorm periods after which there is no significant change in the mean cloudiness value, is considered to be the length of the transition period. Consequently, a necessary condition for an interstorm period to contain a fair weather cloud cover sequence is to be of duration $T_{is} > 2T_r$ [hour].

[39] A note has to be made regarding a particular case of interstorm periods for which the described approach is not appropriate. Sometimes, passing atmospheric precipitation systems do not necessarily result in rainfall at a given location. However, the cloud cover process is obviously nonstationary during such periods and the estimated mean value can be significantly affected. The approach used here cannot identify such a situation, which would, perhaps, require auxiliary information about cloud vertical structure and spatial characteristics of the precipitation process. Nevertheless, the procedure is efficient for most of the interstorm periods and results in reasonable estimates of the transition period as long as the above situation does not occur often. Still, caution has to be taken when interpreting the results of the described method.

[40] Figure 2 illustrates the outlined procedure for the Albuquerque station. In addition to the mean values, the standard deviation of the mean estimate is plotted. For the selected values of the transition period T_r, both the analytical and observed transition functions $J(t)$ are plotted in Figure 3. As one can see, the exponential form of $J(t)$ fits the observed cloud cover transition sufficiently well some of the months. In several cases, the performance is not as satisfactory, which may question the validity of the assumed exponential form for the location of Albuquerque (compare with results in Text S1, section F, for other locations). There can be also observed a particularly considerable bias for the months of July and August, which is likely attributed to the situation described in the note above (these months are associated with the monsoon period for Albuquerque).

[41] Once T_r is established, the fair weather sequences contained in the interstorm periods of length $T_{is} > 2T_r$ are combined in a new series containing only fair weather cloud cover values. For these series, created for each month or the entire period of analysis, the parameters M_0, σ_m^2, and $\rho_m(1)$ are determined by conventional methods. Estimated values of the parameters for Albuquerque are provided for reference in Table 2. As can be seen, the fair weather cloudiness varies throughout the year sometimes changing in magnitude by a factor of two between September and February. Cloudiness exhibits high autocorrelation at lag 1 hour ($\sim$0.9), which is relatively unchanged throughout the year. The parameter γ, with $\gamma = \varsigma$, is computed as $\gamma = 4.61/T_r$ [*Curtis and Eagleson*, 1982]. The parameters a and b are estimated by analyzing random deviates $\varepsilon(t)$, which are calculated from the observed cloud cover series by inverting

Table 3. Regression Parameters b_i of the Air Temperature Model for the Location of Albuquerque, New Mexico

Month	b_0	b_1	b_2	b_3	b_4
January	−5.52	0.139	3.70	18.0	14.5
February	−5.32	0.145	3.56	17.3	14.2
March	−4.76	0.147	3.16	16.4	13.2
April	−4.03	0.147	2.74	16.0	12.0
May	−2.87	0.156	2.74	14.9	10.4
June	−0.776	0.156	2.99	13.3	6.73
July	2.60	0.168	3.36	10.2	0.763
August	2.31	0.183	3.44	10.3	1.79
September	−0.403	0.156	3.15	12.8	5.62
October	−2.66	0.147	3.20	17.1	9.29
November	−4.30	0.149	3.40	18.0	12.3
December	−5.07	0.145	3.54	19.9	13.4

(1) and (5). The estimation of $\varepsilon(t)$ is conditioned by the cloud cover at time $(t-1)$. Therefore 11 vectors of deviates are composed from the cloud cover records for every month, each vector corresponds to one of the values of $N(t-1)$: 0.0, 0.1, …, 1.0. For each $N(t-1)$, the corresponding distribution of deviates is approximated by the Beta distribution with parameters a and b estimated from these deviates. Figure 4 illustrates the procedure for the month of July for Albuquerque. The last plot shows the skewness of the deviates as well as their first two moments scaled to be in the range [0, 1]. The behavior of these variables with respect to $N(t-1)$ is similar for all months: while the mean and standard deviation are essentially constant throughout the entire range of $N(t-1)$ values, the skewness of the deviates varies significantly, changing its sign from positive to negative. The probability density functions of the Beta distribution, corresponding to the same $N(t-1)$, can significantly vary among different months. Since the variability is quite substantial for most months, a and b are estimated on a monthly basis.

[42] When the parameter values have been estimated, the performance of the model is evaluated based on the first two moments and frequency histogram of the cloud cover simulated for fair weather periods. In general, the frequency distributions of fair weather cloudiness tend to be U-shaped with spikes at zero and one, as Figure 5 illustrates. As can be seen, there is a generally good agreement between the observed and simulated data. A note has to be made concerning the cloud cover during consecutive months that have different cloudiness statistics (e.g., August and September). The cloud cover model considers monthly values of the parameters. Cloudiness of an interstorm period that overlaps 2 months is therefore simulated using parameters for both months. The procedure that identifies the fair weather cloud cover sequences considers an interstorm period starting in 1 month and ending in a subsequent month as belonging to 1 month only, depending on the relative duration of the dry spell within each month. Consequently, the cloud cover statistics derived for any given month can be affected by the presence of interstorm periods during which cloudiness is simulated using two parameter sets.

4.3. Shortwave Radiation

[43] The model includes the following parameters (Text S1, sections B and C): the ozone amount in a vertical column u_o [cm], the Angström turbidity parameter β (same for both VIS and NIR bands), the aerosol single-scattering albedos ω_{Ai} [dimensionless], and the spatial average regional albedo ρ_g [dimensionless]. The liquid water path, LWP [g m^{-2}], is a measurable state of cloud thickness [*Han et al.*, 1994], however, this quantity may not be readily available. Furthermore, when the radiative transfer model produces forcing in conjunction with the stochastically generated variables, it is reasonable to link LWP to features of regional climatology.

[44] The parameters u_o, β, ω_{Ai}, and ρ_g are derived for clear sky atmospheric conditions. The parameter u_o is assumed to be seasonally constant (Text S1, section B) and $u_o = 0.34$ is used here. The parameter β is obtained through calibration, which is done by comparing the observed and simulated total direct beam flux for clear sky conditions. For example, Figure 6 illustrates the annual cycle of the mean observed and simulated direct beam radiation on a horizontal surface for a variety of daylight hours for different Julian days. The seasonally invariant calibrated value of the parameter is $\beta = 0.017$. Similarly, the albedos ω_{Ai} and ρ_g are calibrated to obtain the proper cycle of the clear sky diffuse radiation. In this example, the values are $\rho_g = 0.1$, $\omega_{A1} = 0.92$, and $\omega_{A2} = 0.833$. The assumed seasonal invariance of these parameters can partly contribute to the smaller than observed variability of the simulated clear sky diffuse radiation throughout the year (Figure 6).

[45] Once clear sky radiative fluxes are reproduced at a satisfactory level, LWP is considered as the model parameter to account for overcast and partially cloudy sky conditions. Note that for proper estimation, the parameters of both the rainfall and cloudiness model need to be already obtained. Seasonally varying monthly values of LWP_o for overcast conditions are assumed here to account for the intra-annual differences in cloud structure and origin. Furthermore, it is assumed that LWP for any sky condition is simply $\mathrm{LWP}_N = e^{N\ln(\mathrm{LWP}_o)} - 1$, where N is bounded by "0" (clear sky) and "1" (overcast, $\mathrm{LWP}_N = \mathrm{LWP}_o$). The monthly values of LWP_o are calibrated based on comparison between the simulated and observed direct beam and diffuse fluxes for all sky conditions. Figure 7 illustrates the annual cycles of the radiative fluxes for Albuquerque (New Mexico) with the calibrated values of LWP_o = 110 for January, 80 for February, 80 for March, 60 for April, 65 for May, 75 for June, 95 for July, 60 for August, 100 for September, 100

Table 4. Parameters $\overline{\delta T}$, $\sigma^2_{\delta T}$, and $\rho_{\delta T}(1)$ of the Air Temperature Model for the Location of Albuquerque, New Mexico

	ALB		
Month	$\overline{\delta T}$	$\sigma^2_{\delta T}$	$\rho_{\delta T}(1)$
January	0.0059	2.947	0.930
February	0.0161	3.032	0.928
March	0.00048	3.188	0.926
April	−0.0064	3.184	0.924
May	−0.0208	3.018	0.918
June	−0.0455	2.901	0.913
July	−0.0063	2.638	0.903
August	−0.0185	2.366	0.891
September	0.0013	2.7752	0.923
October	0.0018	3.125	0.925
November	0.0246	3.133	0.925
December	0.0154	2.891	0.924

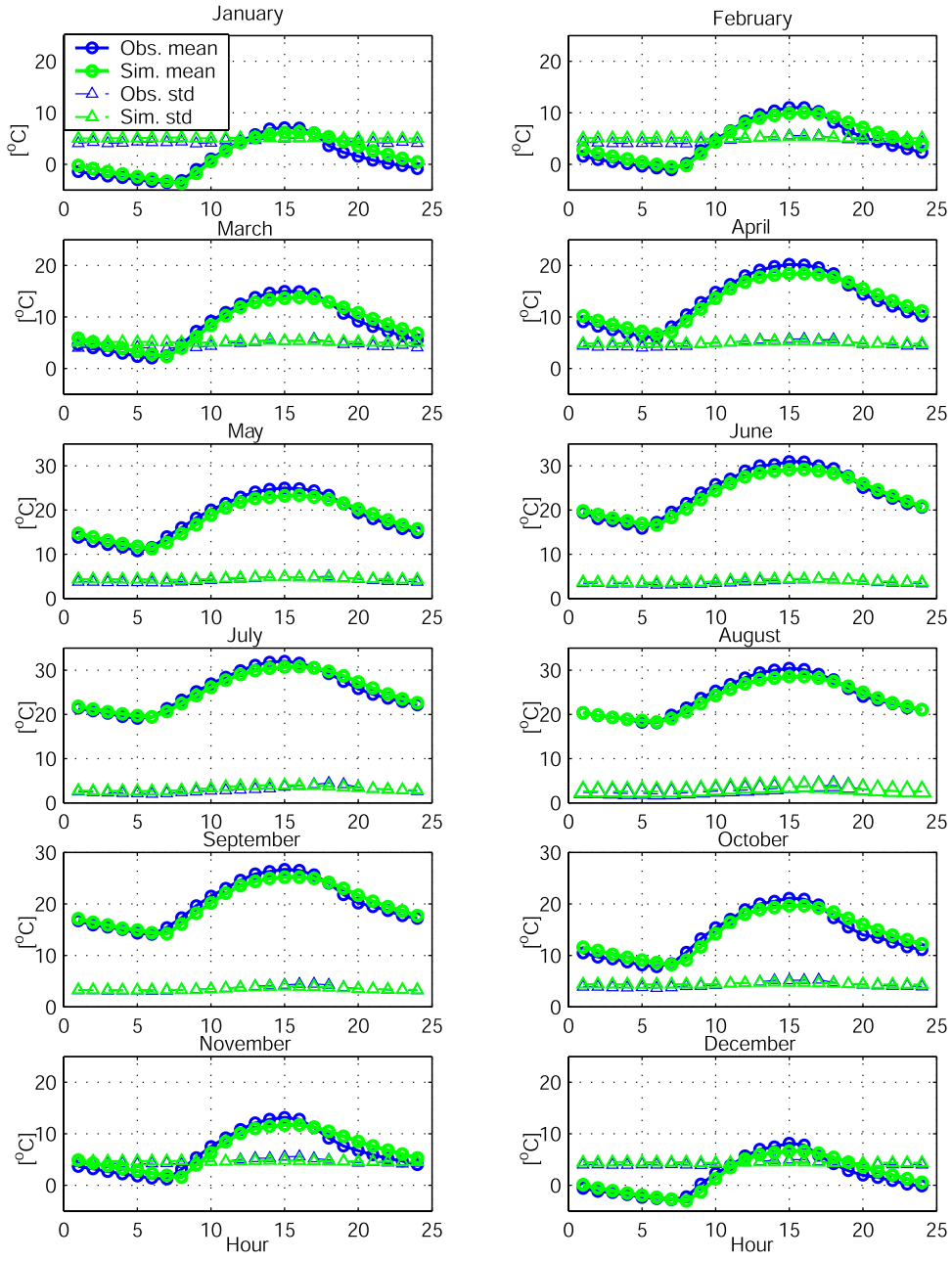

Figure 8. Observed and simulated daily cycles of air temperature and its standard deviation (Albuquerque, New Mexico).

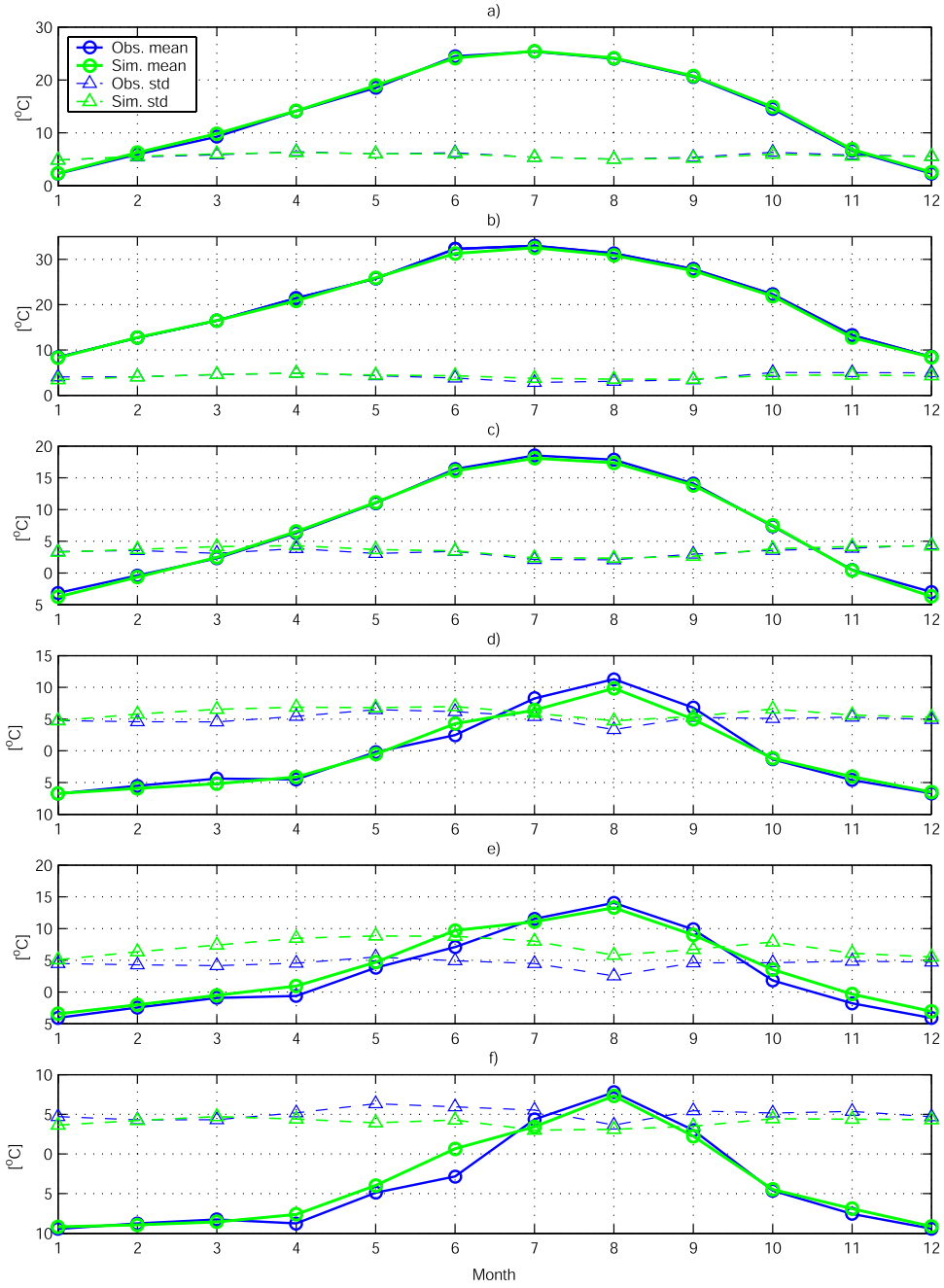

Figure 9. Observed and simulated mean values of air and dew point temperature and their standard deviations (Albuquerque, New Mexico). Mean monthly values and daily standard deviation of (a) air temperature; (b) maximum air temperature; (c) minimum air temperature; (d) dew point temperature; (e) maximum dew point temperature; and (f) minimum dew point temperature.

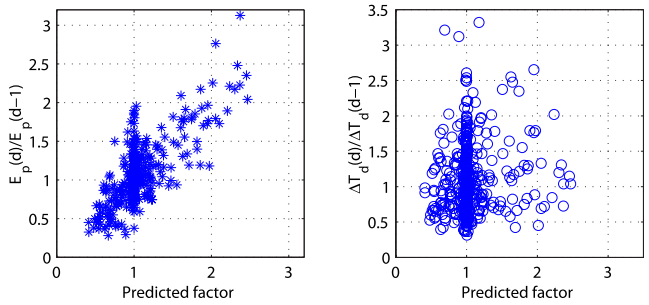

Figure 10. Comparison of $\tilde{A}(d)$, the predicted factor, as defined in (13) and the factor obtained from the simulated data when both $\overline{E}_p(d)$ and ΔT_d become known (Albuquerque, New Mexico).

for October, 130 for November, 150 for December. An alternate procedure would be to obtain LWP seasonality from other sources, such as World Climate Research Programme International Satellite Cloud Climatology Project (ISCCP) cloud data products (*Rossow and Schiffer*, 1991] that have near-global coverage [*Han et al.*, 1994].

4.4. Air Temperature

[46] The parameters of the air temperature model that need to be estimated involve the regression coefficients b_i-s ($i = 0, 1, \ldots, 4$), $\overline{\delta T}$, $\sigma^2_{\delta T}$, and $\rho_{\delta T}(1)$ (Text S1, sections E1 and E2). The b_i-s are estimated first on a monthly or seasonal basis, or as representative values for the entire period of interest. Table 3 provides monthly estimates of the regression parameters for air temperature data in Albuquerque (New Mexico). Once the regression parameters have been estimated, the air temperature model can be used to simulate the deterministic component of the hourly temperature change, starting at midnight of each day. The initial value is taken as the deterministic temperature component estimated for 2300 LT of the previous day. As discussed in section 2.4, the difference between the observed and estimated deterministic temperature components defines the temperature random deviate. Consequently, series of deviates can be estimated for each period of interest, e.g., for each month or season. The parameters $\overline{\delta T}$, $\sigma^2_{\delta T}$, and $\rho_{\delta T}(1)$ (equation (10)) are then obtained using conventional estimation techniques (Table 4).

[47] When simulating hourly temperature, it is important to reproduce both its mean daily cycle and features of its variability. Figure 8 illustrates the mean daily cycles of air temperature and its standard deviation for each month. The model (8) mimics the daily air temperature fluctuations quite well. On a monthly scale, the air temperature statistics are also well reproduced, as is shown in Figures 9a–9c.

4.5. Dew Point Temperature

[48] An adjustment factor $\tilde{A}(d)$ was introduced in section 2.5, given by equation (13). A comparison of the factors estimated for a complete simulation using (13) and data from a simulation when both the average potential evaporation, $\overline{E}_p(d)$, and the daily temperature fluctuation amplitude, ΔT_d, are known is illustrated in Figure 10. As can be seen, the use of $\tilde{A}(d)$ is satisfactory for estimating $\overline{E}_p(d)$. However, the definition of $\tilde{A}(d)$ is not suitable for adjusting $\Delta T_d(d-1)$, and therefore the parameter $\tilde{A}(d)$ was not used for such an adjustment in the case of data for Albuquerque (New Mexico).

[49] The values of P^*_{ann} depend on wetness of a particular month and are determined iteratively by comparing the mean observed and simulated monthly dew point temperatures. The estimated cycle for the location of Albuquerque: $P^*_{\text{ann}} = 65$ for January, 60 for February, 60 for March, 55 for April, 70 for May, 70 for June, 120 for July, 160 for August, 105 for September, 60 for October, 50 for November, 55 for December.

[50] Since dew point temperatures exhibit little variability on any given day, only monthly statistics are presented. Figures 9d and 9f show the mean daily dew point temperatures as well as the mean daily maximum and minimum dew point temperatures simulated for each month. Since the dew point temperature model is rather simple, the simulated and observed series are in less agreement as compared to the results of the air temperature simulation.

4.6. Wind Speed

[51] The mean, variance, lag-1 value of autocorrelation, and skewness of wind speed ($\overline{W}_s$, σ^2_s, $\rho_s(1)$, and γ_s) are estimated from wind speed data using conventional methods. For Albuquerque they are 3.491, 2.452, 0.755, and 1.688, respectively.

[52] As was discussed in section 2.6, the frequency distribution of wind speed data is positively skewed. Both the skewness properties and the first two moments of the distribution are preserved with the model (14). Figure 11 illustrates the wind speed histograms computed from the observed and simulated data for Albuquerque.

5. Covariation of Hydrometeorological Variables

[53] The weather generator explicitly couples a number of simulated variables. Although the cross-correlation properties are not directly accounted for as, for example, in the

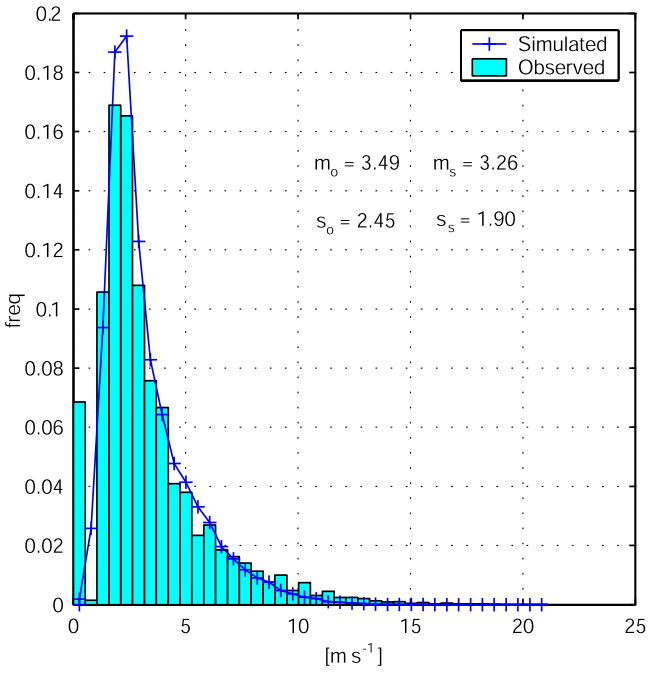

Figure 11. Histogram of hourly wind speed from the observed and simulated data (Albuquerque, New Mexico). Symbols m and s are the mean and standard deviation values, correspondingly, for the observed (subindex "o") and simulated (subindex "s") data.

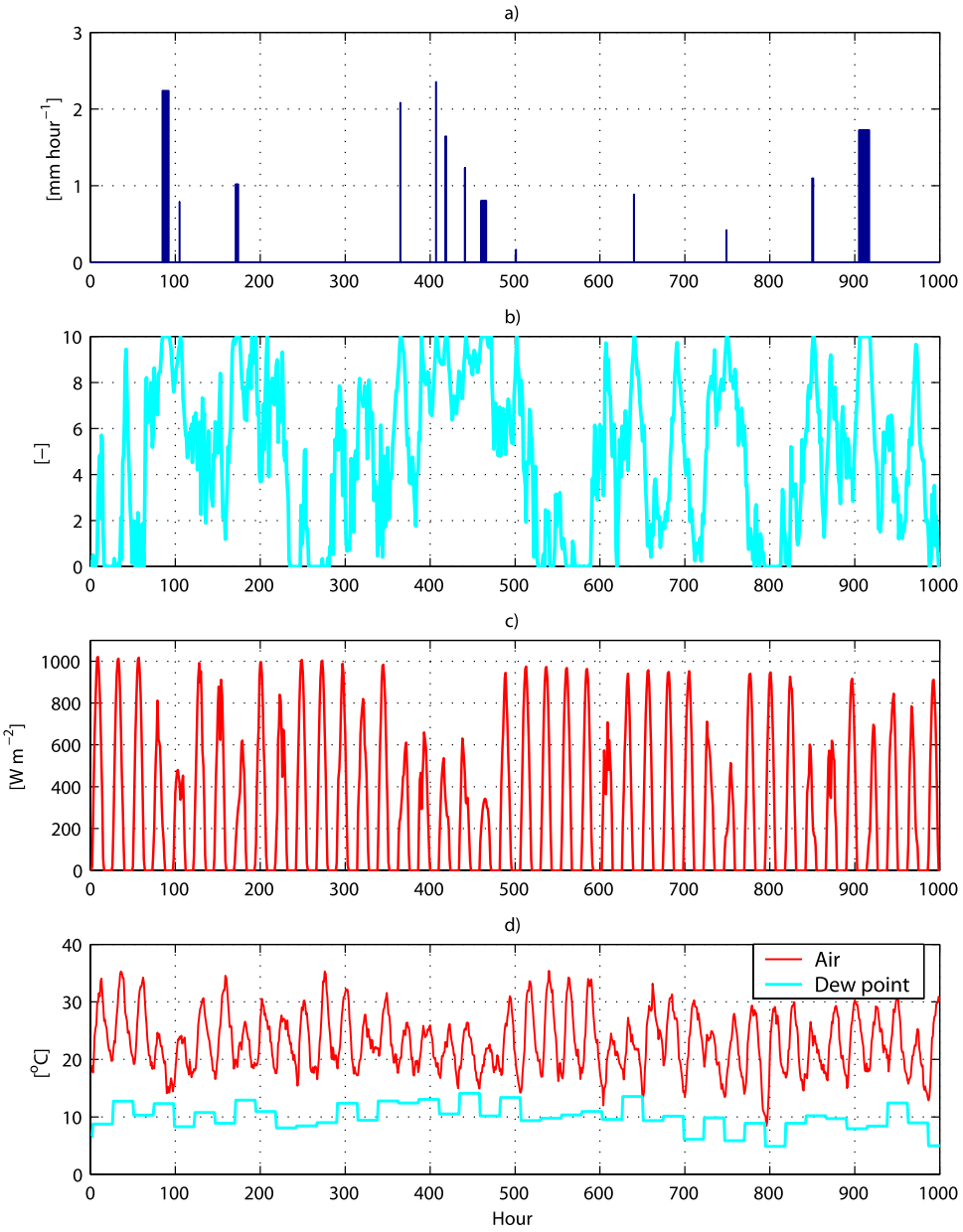

Figure 12. Simulated hourly hydrometeorological variables based on parameters derived for the location of Albuquerque (New Mexico) (start on 1 August): (a) rainfall; (b) cloud cover; (c) incoming shortwave radiation; and (d) air and dew point temperature.

WGEN model of *Richardson* [1981], one may expect that the weather variables should exhibit consistent covariation. Figure 12 illustrates such interdependencies in a qualitative manner. Simulations start in August and extend through half of September. As can be seen, the cloudiness dynamics respond to precipitation events and the incoming shortwave is correspondingly affected by the presence of clouds. The air temperature series exhibit both lower magnitude and diurnal variability during the days with precipitation. The dew point temperatures become less differentiated from the air temperatures during wet periods and show a substantial deviation from the minimum daily temperatures during drier hot days.

[54] Figure 13 illustrates the dependence of the mean monthly cloud cover on rainfall occurrence. As can be seen, there is a good correspondence between the simulated and observed data. The cloud cover model slightly overestimates the mean observed values, which can be attributed to both (1) overestimation of the rainfall occurrence for some months due to the introduced seasonality in rainfall model parameters and (2) some inadequacy of the exponential form of the transition function $J(t)$ to describe the cloud cover dynamics during transition to/from fair weather periods.

[55] Figure 14 shows the mean maximum and minimum air temperatures on rainy and rainless days for different months, derived from the simulated and observed data. While the air temperature model accounts for precipitation occurrence only implicitly (via $K(t)$ and $q(t)$ in equation (9)), a generally good agreement can be observed.

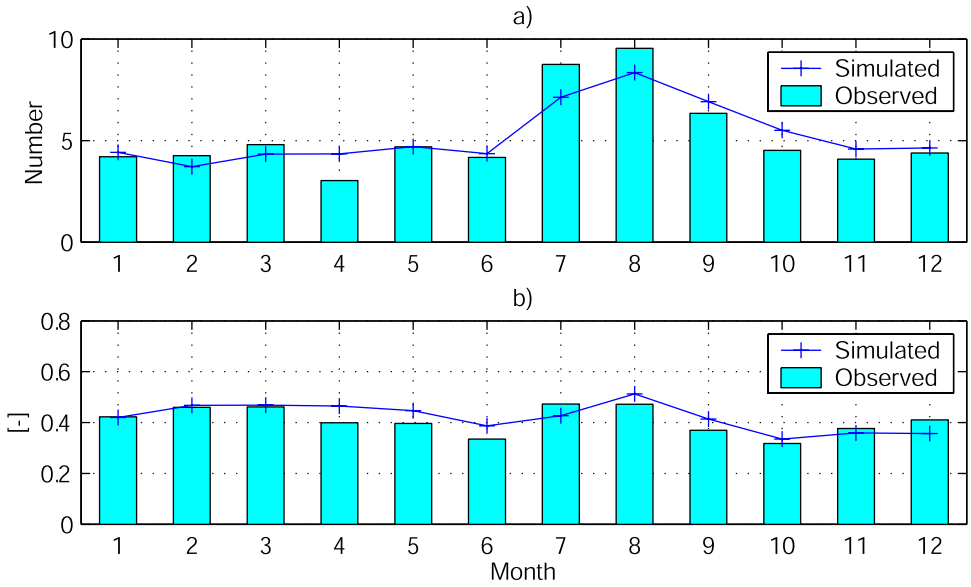

Figure 13. Observed and simulated mean monthly precipitation occurrence and cloud cover (Albuquerque, New Mexico): (a) mean number of storms and (b) mean cloudiness.

[56] Figure 15 shows sample cross-correlation functions between the mean daily cloud cover and air temperature amplitude for observed and simulated data. As can be seen, the highest cross correlation between the two variables is at zero lag, which is well reproduced by the model. The observed data also show that cloud cover somewhat leads temperature amplitude, i.e., a nonzero cross correlation exists at the lead time of 1 day (the asymmetry observed in Figure 15). This effect is not reproduced by the discussed model. Note, however, that the relationships between variables are causal, not statistical. Changes in temperature at the hourly timescale are related to the hourly cloud cover values. One would therefore expect that zero lag cross correlations should be reasonable; however, the asymmetry observed at longer lead times may not be necessarily well reproduced by the construct.

6. Summary

[57] This paper discusses a weather generator that allows one to synthetically simulate several hydrometeorological variables at the hourly scale: rainfall, total cloud cover, the incoming shortwave radiation, air temperature, humidity, and wind speed. The weather simulator of *Curtis and*

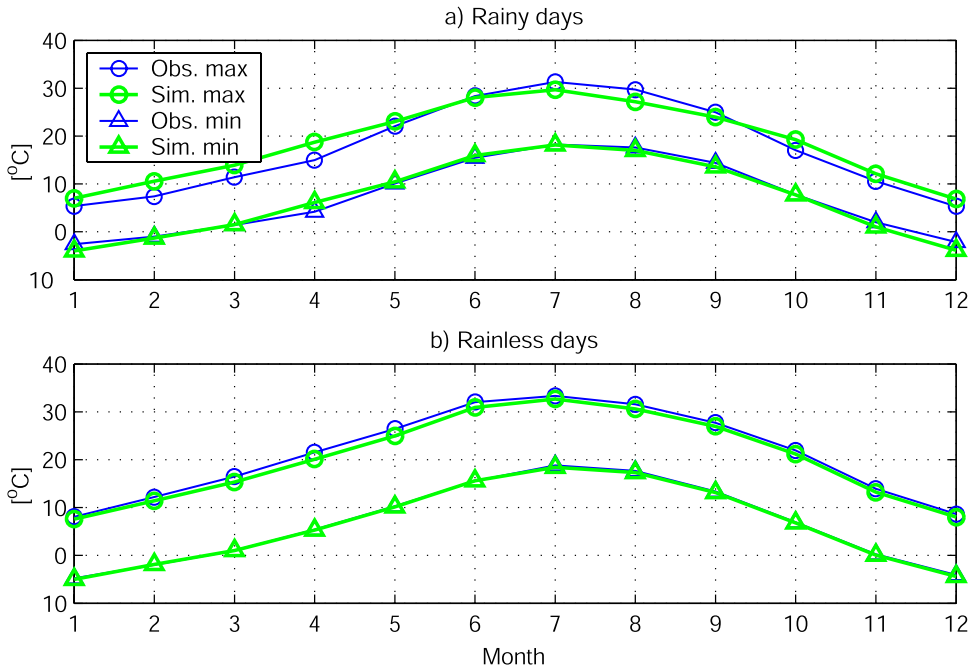

Figure 14. Mean maximum and minimum air temperatures derived from the observed and simulated data (Albuquerque, New Mexico) on (a) rainy days and (b) rainless days.

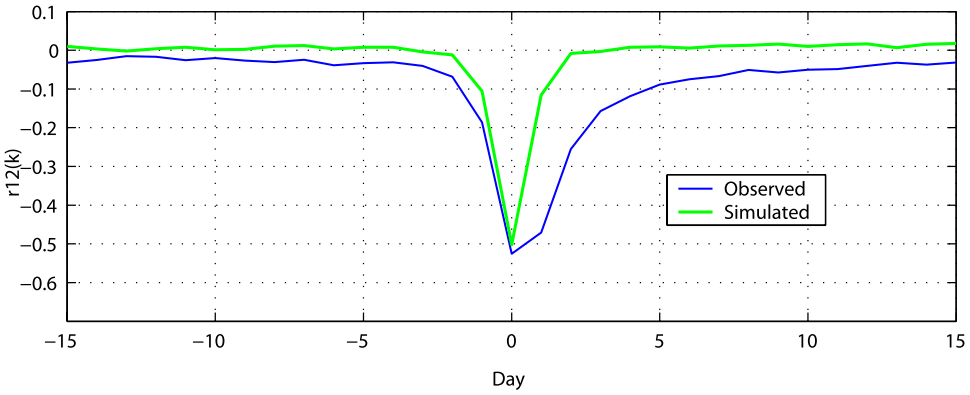

Figure 15. Sample cross-correlation functions derived between the mean daily cloud cover (variable 1) and air temperature amplitude (variable 2) derived from the observed and simulated data (Albuquerque, New Mexico).

Eagleson [1982] was used as the core framework for the model. A new shortwave radiation model was introduced based on work of *Gueymard* [1989], allowing one to represent separately the atmospheric radiative transfer for the two essential bands, VIS and NIR. Other necessary modifications were also implemented, which lead to a better or more efficient representation of the simulated statistics.

[58] Precipitation is considered to be the key driver of simulated hydrometeorological conditions, capturing the essential relationships among the processes of cloud cover, shortwave radiation, and air temperature and humidity. Consistent time series of hydrometeorological quantities are thus obtained: Although the cross-correlation properties are not directly accounted for, all major weather quantities exhibit proper covariation. While the Poison storm arrival model with rectangular pulses was used here, in principle, any rainfall generation routine, e.g., accounting for intra-storm/diurnal variability of the precipitation process, can be used as the driver of the simulated hydrometeorological series. The weather generator can also be forced using real rainfall observations, which are more commonly available, to create a consistent realization of climate for a location of interest.

[59] The generator was calibrated and validated with the long-term (30–35 years) data for three meteorological stations located in New Mexico (results are given in this paper), Arizona, and Oklahoma. The results show a consistent behavior. Further studies are needed to apply the model for wetter and colder climates.

[60] Standard micrometeorological observations of the simulated variables at the hourly scale are used for model parameterization. At many locations, however, certain climate variables are not observed. Only partial measures can thus be used, when parameterizing the described simulator. Remote sensing observations can be utilized to derive the arrival characteristic and magnitude of precipitation (e.g., Tropical Rainfall Measuring Mission data) as well as the properties and dynamics of cloud cover (e.g., MODIS sensor data). Methods combining remote sensing observations and topography can be used to derive estimates of near-surface atmospheric humidity [e.g., *Han et al.*, 2005]. Shortwave model and wind model can be fairly well parameterized using reasonable parameter values for the geographic location of interest. Air temperature data are fairly abundant and can be extrapolated from the nearest location. Obviously, a careful decision is needed in every particular situation since the lack of data may preclude using the described model on a consistent basis.

[61] The weather generator operates at the point scale, which can be assumed to be sufficiently representative for creating forcing for systems such as a hillslope or a small basin. The set of simulated variables can serve as input to a number of models of environmental systems, including hydrological, geomorphological, ecological, water resources, and agricultural applications. The model is also suitable for creating scenarios of climate regimes (e.g., dryer and wetter, warmer and colder climates) applicable in a variety of sensitivity studies. Other potential applications may involve gap filling in observed meteorological series. A variety of associated problems can be dealt with provided the simulator is extended to include the spatial aspect, for example, to create meteorological forcing in conditions of complex topography or larger areas or to address the subgrid variability of large-scale weather predictions. The extension to include spatial variability has not been done. (The source code of the weather generator, the manual, and the test applications are publicly available at http://hydrology.mit.edu.)

[62] **Acknowledgments.** The model presented here was based on the doctoral dissertation of David Curtis, who was supervised by Peter S. Eagleson. We thank Peter Eagleson for reading the manuscript and supporting this evolution of the model and its publication. We also thank four anonymous reviewers for their helpful comments. The original work of Curtis and Eagleson was supported by the National Oceanic and Atmospheric Administration, the National Weather Service, and the National Science Foundation. More recently, the work was performed with the support of the National Aeronautics and Space Administration (contract NAG57475), the National Oceanic and Atmospheric Administration (contract NA97WH0033), the NWS (Office of Hydrology)-MIT Cooperative Agreements, the Army Research Office, and the CNR (Italy)-MIT Cooperative Agreement.

References

Aguiar, R., and M. Collares-Pereira (1992), TAG–A time-dependent, auto-regressive, Gaussian model for generating synthetic hourly radiation, *Sol. Energy*, *49*(3), 167–174.

Ahmed, J. (1974), Optimization of water-use efficiency in crop production systems by dynamic simulation of crop behavior under stochastic regimes, Ph.D. thesis, Tex. A & M Univ., College Station.

Becker, S. (2001), Calculation of direct solar and diffuse radiation in Israel, *Int. J. Clim.*, *21*(12), 1561–1576.

Benjamin, J. R., and C. A. Cornell (1970), *Probability, Statistics and Decision for Civil Engineers*, McGraw-Hill, New York.

Bonta, J. V. (2004), Stochastic simulation of storm occurrence, depth, duration and within-storm intensities, *Trans. Am. Soc. of Agric. Eng.*, *47*(5), 1573–1584.

Bryan, J. G. (1964), Short range hour-by-hour forecast of temperature by projecting the characteristic curve with constants fitted to immediately preceding data, report, Travelers Res. Cent., Inc., Hartford, Conn.

Buishand, T. A. (1978), Some remarks on use of daily rainfall models, *J. Hydrol.*, *36*(3–4), 295–308.

Carlin, J., and J. Haslett (1982), The probability-distribution of wind power from a dispersed array of wind turbine generators, *J. Appl. Meteorol.*, *21*(3), 303–313.

Chia, E., and M. F. Hutchinson (1991), The beta-distribution as a probability model for daily cloud duration, *Agric. For. Meteorol.*, *56*(3–4), 195–208.

Chin, E. H. (1977), Modeling daily precipitation occurrence process with Markov-Chain, *Water Resour. Res.*, *13*(6), 949–956.

Curtis, D. C., and P. S. Eagleson (1982), Constrained stochastic climate simulation, *Tech. Rep. 274*, Mass. Inst. of Technol., Dep. of Civ. and Environ. Eng., Ralph M. Parsons Lab., Cambridge, Mass.

Darula, S., R. Kittler, and C. A. Gueymard (2005), Reference luminous solar constant and solar luminance for illuminance calculations, *Sol. Energy*, *79*(5), 559–565.

Falls, L. W. (1974), Beta-distribution—Statistical model for world cloud cover, *J. Geophys. Res.*, *79*(9), 1261–1264.

Foufoula-Georgiou, E., and W. Krajewski (1995), Recent advances in rainfall modeling, stimation, and forecasting, U.S. Natl. Rep, Int. Union Geod. Geophys. 1991–1994, *Rev. Geophys.*, *33*, 1125–1137.

Gabriel, R., and J. Neuman (1962), A Markov chain model for daily rainfall occurrence at Tel Aviv, Israel, *Q.J.R. Meteorol. Soc.*, *88*, 90–95.

Georgakakos, K., and M. L. Kavvas (1987), Precipitation analysis, modeling, and prediction in hydrology, *Rev. Geophys.*, *25*, 163–178.

Glassy, J., and S. W. Running (1994), Validating diurnal climatology logic of the MT-CLIM logic across a climatic gradient of Oregon, *Ecol. Appl.*, *4*, 248–257.

Grace, R. A., and P. S. Eagleson (1967), A model for generating synthetic sequences of short-time interval depths, paper presented at International Hydrology Symposium, Colo. State Univ., Fort Collins, Colo.

Grayman, W. M., and P. S. Eagleson (1969), Streamflow record length for modeling catchment dynamics, *Tech. Rep. 114*, Mass. Inst. of Technol., Dep. of Civ. and Environ. Eng., Ralph M. Parsons Lab., Cambridge.

Gueymard, C. (1989), A 2-band model for the calculation of clear sky solar irradiance, illuminance, and photosynthetically active radiation at the Earth's surface, *Sol. Energy*, *43*(5), 253–265.

Han, K. S., A. A. Viau, Y. S. Kim, and J. L. Roujean (2005), Statistical estimate of the hourly near-surface air humidity in eastern Canada in merging NOAA/AVHRR and GOES/IMAGER observations, *Int. J. Remote Sens.*, *26*(21), 4763–4784.

Han, Q., W. B. Rossow, and A. A. Lacis (1994), Near-global survey of effective droplet radii in liquid water clouds using ISCCP data, *J. Clim.*, *7*(4), 465–497.

Hanson, C. L., K. A. Cumming, D. A. Woolhiser, and C. W. Richardson (1994), Microcomputer program for daily weather simulations in the contiguous USA, Rep. ARS-114, 38 pp., U.S. Dep. of Agric., Agric. Res. Serv., Washington, D. C.

Hanson, C. L., G. L. Johnson, and W. L. Frymire (2002), The GEM (Generation of weather Elements for Multiple applications) weather simulation model, paper presented at 13th AMS Conference on Applied Climatology, Portland, Oreg., 13–16 May.

Hennessey, J. P. (1977), Some aspects of wind power statistics, *J. Appl. Meteorol.*, *16*(2), 119–128.

Ivanov, V. Y. (2006), Effects of dynamic vegetation and topography on hydrological processes in semi-arid areas, Ph.D. Thesis, Mass. Inst. of Technol., Cambridge.

Kimball, J. S., S. W. Running, and R. Nemani (1997), An improved method for estimating surface humidity from daily minimum temperature, *Agric. For. Meteorol.*, *85*(1–2), 87–98.

Maass, A., M. M. Hufschmidt, R. Dorfman, H. A. Thomas Jr., S. A. Marglin, and G. M. Fair (1962), *Design of water-resources systems*, 620 pp., Harvard Univ. Press, Cambridge, Mass.

Marien, J. L., and G. L. Vandewiele (1986), A point rainfall generator with internal storm structure, *Water Resour. Res.*, *22*(4), 475–482.

Morissey, M. L., and W. F. Krajewski (1993), A point-process model for tropical rainfall, *J. Geophys. Res.*, *98*, 16,639–16,652.

Nicks, A. D., and G. A. Gander (1993), Using CLIGEN to stochastically generate climate data inputs to WEPP and other water resources models, *U.S. Geol. Surv. Water Resour. Invest.*, *93–4018*, 443.

Nicks, A. D., and G. A. Gander (1994), CLIGEN: A weather generator for climate inputs to water resource and other models, paper presented at Fifth International Conference of Computers in Agriculture, Am. Soc. of Agric. Eng., Orlando, Fa., 903–909.

Nicks, A. D., C. W. Richardson, J. R. Williams (1990), Evaluation of the EPIC model weather generator, in *Erosion/Productivity Impact Calculator, 1. Model Documentation*, A. N. Sharpley and J. R. Williams, Eds., *USDA ARS Tech. Bull. 1768*, 235 pp. (Available from U.S. Govt. Print. Off., Washington, D. C.)

Olseth, J. A., A. Skartveit, and H. Zou (1995), Spatially continuous mapping of solar resources in a complex high latitude topography, *Sol. Energy*, *55*(6), 475–485.

Parlange, M. B., and R. W. Katz (2000), An extended version of the Richardson model for simulating daily weather variables, *J. Appl. Meteorol.*, *39*(5), 610–622.

Priestley, C. H. B., and R. J. Taylor (1972), On the assessment of surface heat flux and evaporation using large-scale parameters, *Mon. Weather Rev.*, *100*, 81–92.

Restrepo-Posada, P. J., and P. S. Eagleson (1982), Identification of independent rainstorms, *J. Hydrol.*, *55*(1–4), 303–319.

Richardson, C. W. (1981), Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resour. Res.*, *17*, 182–190.

Richardson, C. W., and D. A. Wright (1984), WGEN: A model for generating daily weather variables, *Publ. ARS-8*, 83 pp., U.S. Dep. of Agric., Agric. Res. Serv., Washington, D. C.

Rossow, W. B., and R. A. Schiffer (1991), ISCCP cloud data products, *Bull. Am. Meteorol. Soc.*, *72*, 2–20.

Running, S. W., R. R. Nemani, and R. D. Hungerford (1987), Extrapolation of synoptic meteorological data in mountainous terrain, and its use for simulating forest evapotranspiration, *Can. J. For. Res.*, *17*, 472–483.

Sariahmed, A., and C. C. Kisiel (1968), Synthesis of sequences of summer thunderstorm volumes for the Atterbury watershed in the Tucson area, in *The Use of Analog and Digital Computers in Hydrology*, *Proc. IAHS Symp.*, *80*, 439–447.

Schnorbus, M., and Y. Alila (2004), Generation of an hourly meteorological time series for an alpine basin in British Columbia for use in numerical hydrologic modeling, *J. Hydrometeorol.*, *5*(5), 862–882.

Schoof, J. T., A. Arguez, J. Brolley, and J. J. O'Brien (2005), A new weather generator based on spectral properties of surface air temperatures, *Agric. For. Meteorol.*, *135*, 241–251.

Slingo, A. (1989), A GCM parameterization for the shortwave radiative properties of water clouds, *J. Atmos. Sci.*, *46*(10), 1419–1427.

Slingo, A., and H. M. Schrecker (1982), On the shortwave radiative properties of stratiform water clouds, *Q. J.R. Meteorol. Soc.*, *108*, 407–426.

Smith, R. E., and H. A. Schreiber (1973), Point processes of seasonal thunderstorm rainfall: 1. Distribution of rainfall events, *Water Resour. Res.*, *9*(4), 871–884.

Stephens, G. L. (1978), Radiation profiles in extended water clouds: 2. Parameterization schemes, *J. Atmos. Sci.*, *35*(11), 2123–2132.

Thomas, G. E., and K. Stamnes (1999), *Radiative Transfer in the Atmosphere and Ocean*, 517 pp., Cambridge Univ. Press, Cambridge, U. K.

Todorovic, P. (1968), A mathematical study of precipitation phenomena, *Rep. CER 67-68PT65*, Eng. Res. Cent., Colo. State Univ., Fort Collins.

Todorovic, P., and V. Yevjevich (1969), Stochastic processes of precipitation, *Hydrol. Pap. 35*, 161 pp., Colo. State Univ., Fort Collins.

Wilks, D. S., and R. L. Wilby (1999), The weather generation game: a review of stochastic weather models, *Progr. Phys. Geogr.*, *23*(3), 329–357.

Wilson, E. G., and M. M. Hillerty (1931), Distribution of chi square, *Proc. Natl. Acad. Sci. U.S.A.*, *17*, 684–688.

Woolhiser, D. A. (1992), Modeling daily precipitation–Progress and problems, in *Statistics in the Environmental and Earth Sciences*, edited by A. T. Walden and P. Guttorp, pp. 71–89, Edward Arnold, London.

Woolhiser, D. A., and H. B. Osborn (1985), A stochastic model of dimensionless thunderstorm rainfall, *Water Resour. Res.*, *21*(4), 511–522.

R. L. Bras, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 15 Vassar Street, Cambridge, MA 02139, USA. (rlbras@mit.edu)

D. C. Curtis, Carlton Engineering, Inc., 3883 Ponderosa Road, Shingle Springs, CA 95682, USA. (dcurtis@carlton-engineering.com)

V. Y. Ivanov, Center for the Environment, Harvard University, 24 Oxford Street, Room 404, Cambridge, MA 02138, USA. (viva@mit.edu)