Prediction of regional water balance components based on climate, soil, and vegetation parameters, with application to the Illinois River Basin

Jeffrey D. Niemann

Department of Civil Engineering, Colorado State University, Fort Collins, Colorado, USA

Elfatih A. B. Eltahir

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Received 24 October 2003; revised 17 December 2003; accepted 31 December 2003; published 9 March 2004.

[1] This paper presents a framework for studying regional water balance in which the physical processes are first described at the local instantaneous scale and then integrated to the annual, basin-wide scale. The integration treats the relative soil saturation (i.e., the soil moisture divided by the porosity) and precipitation intensities as stochastic variables in space and time. A statistical equilibrium characterizes the annual water balance, resulting in a specific relation that predicts the space-time average of soil saturation in terms of soil, climate, and vegetation parameters. Specific relationships are proposed to relate the space-time average soil saturation to runoff, groundwater recharge, and evapotranspiration. This framework is applied to the Illinois River Basin. The shape of the spatial and temporal distributions of soil saturation are determined from observations. The other parameters are determined from the physical characteristics of the basin and calibration procedures. The resulting model is able to reproduce an observed relation between the space-time average soil saturation and precipitation. It is also able to reproduce observed relations between space-time average soil saturation and space-time average evapotranspiration, surface runoff, and groundwater runoff. *INDEX TERMS:* 1836 Hydrology: Hydrologic budget (1655); 1866 Hydrology: Soil moisture; 1854 Hydrology: Precipitation (3354); 1860 Hydrology: Runoff and streamflow; *KEYWORDS:* Illinois River, probabilistic methods, soil moisture, water balance

Citation: Niemann, J. D., and E. A. B. Eltahir (2004), Prediction of regional water balance components based on climate, soil, and vegetation parameters, with application to the Illinois River Basin, *Water Resour. Res.*, *40*, W03103, doi:10.1029/2003WR002806.

1. Introduction

[2] A longstanding objective in hydrology is understanding the hydrologic budgets of large regions [e.g., *Thornthwaite and Mather*, 1955; *Gleick*, 1987]. The partitioning of precipitation among surface runoff, groundwater runoff, evaporation, and transpiration involves numerous physical processes and affects many applications in hydrologic practice and research. A key element in the description of the water balance is scale. At the local spatial scale and instantaneous timescale, many physical characteristics and fluxes can be considered as constants and treated deterministically. At large scales, such as the Illinois River Basin (69,264 km^2), variability is often substantial and nontrivial. In fact, such variability is fundamental to the water balance.

[3] Proper consideration of spatial and temporal variability is critical to realistic hydrologic modeling [*Sharma and Luxmoore*, 1979; *Sharma et al.*, 1980; *Owe et al.*, 1982, 1992; *Gelhar*, 1993]. For example, *Troendle* [1983] suggested that the majority of the streamflow of the western United States is produced by 10–15% of the land area. To account for spatial variability, *Moore* [1985] proposed a "probability-distributed" model of runoff production. This model uses stochastic variability of soil saturation (i.e., the soil moisture divided by the porosity) at independent soil elements to determine the runoff from a series of precipitation events. Moore also discussed the interaction of the storage elements to represent the redistribution of water due to horizontal flows. Spatial variability has been considered in the development of land surface schemes in general circulation models (GCMs) as well. *Entekhabi and Eagleson* [1989] used an approach that resembles Moore's to describe the spatial variability within grids cells of GCMs. *Entekhabi and Eagleson* [1991] followed this work and investigated the sensitivity of soil saturation to climate change under equilibrium conditions. *Eagleson* [1978a, 1978b, 1978c, 1978d, 1978e, 1978f, 1978g] considered the temporal dynamics of a water balance in detail. He used stochastic inputs over time (i.e., Poisson arrival of precipitation events) to force deterministic physical processes. His model is one dimensional, considering fluxes only in the vertical direction and neglecting the effects of spatial variability.

[4] In this paper, the primary objective is to develop a framework that integrates simple but physically-based descriptions of the water balance processes from the local instantaneous scale to the annual, basin-wide scale. We seek

0043-1397/04/2003WR002806$09.00

a physically-based model that predicts all the water balance components (runoff, recharge, evapotranspiration) based on information about climate, soil, and vegetation. We then apply this general framework to the study of the water balance of the Illinois River Basin. This river basin occupies a large portion of the State of Illinois and is a good case study for large-scale water balance in part because its hydrology is relatively simple. The topographic relief is low and snowfall and snowpack are less important than at most locations in North America [*Yeh et al.*, 1998]. More importantly, Illinois has good availability of data including measurements of soil moisture throughout the state [*Hollinger and Isard*, 1994]. *Findell and Eltahir* [1997] used this data to examine the feedback between soil moisture conditions and rainfall in Illinois. *Yeh et al.* [1998] estimated the evapotranspiration in Illinois using both atmospheric and soil water balances. *Eltahir and Yeh* [1999] examined the connections between precipitation, soil moisture, and groundwater, and their roles in determining the persistence of floods and droughts in Illinois.

[5] In applying our general model to the Illinois River Basin, we take advantage of the availability of soil moisture data. These data are used to identify appropriate probability density functions for soil saturation in space and time (section 2), which are used to relate the small and large spatial and temporal scales. Spatial and temporal variability of precipitation are also included in the model. The model development (section 3) builds on the approach used by *Entekhabi and Eagleson* [1989] in handling spatial variability in hydrologic processes, and goes further by addressing the issue of temporal variability in the spatial averages of hydrologic fields. The inclusion of temporal variability relies on the soil saturation observations and a key result from *Eltahir and Bras* [1993], which allows inclusion of the temporal variability of precipitation. The main elements of the proposed model are calibrated independently using soil saturation data. The model is then applied as a whole (section 4) to determine relationships between annual precipitation, average soil saturation, and total runoff.

2. Variability of Soil Saturation

[6] Beginning in 1981, high quality soil moisture data have been collected at 19 stations in Illinois using neutron probes [*Hollinger and Isard*, 1994]. The probes indirectly measure the average soil moisture over depth increments of 0–10 cm, 10–30 cm, 30–50 cm, and so on to a maximum depth of 2 meters. The data are reported as daily values but typically fewer than 5 stations are measured on any given day. In this paper, all the soil moisture data have been transformed into relative soil saturation by dividing the soil moisture by the porosity [*Hollinger and Isard*, 1994]. Only the average soil saturation between 0 and 10 cm is analyzed here. This range was selected because a simple correlation analysis indicates that it is most closely related to the level of runoff production in the basin. Some implications of this selection are discussed in more detail later in the paper.

2.1. Spatial Variability of Soil Saturation

[7] The spatial variability of soil saturation is considered first. For simplicity, the soil saturations at different stations are treated as independent and identically distributed. The assumption of independence is reasonable if the spatial scale of the region is large relative to any spatial correlation scale in the soil saturation. As a partial test of this assumption, we have calculated the correlations between the soil saturation observations at different stations within the same week. The collection of data into weeks is necessary because few data are collected simultaneously. Although certain pairs of stations exhibit significant dependence due to their close proximity, the average correlation between two stations was found to be about -0.06, which suggests that the spatial correlation of soil saturation is not significant at the regional scale.

[8] If the soil saturation data are approximately independent and identically distributed, then the spatial pattern of soil saturation can be completely described by a single probability density function. We further hypothesize that the soil saturation is drawn from an Erlang distribution:

$$f_s = \frac{k^k}{\bar{s}^k (k-1)!} s^{k-1} e^{-ks/\bar{s}} \tag{1}$$

where s is soil saturation, $\bar{s}$ is the spatial mean soil saturation, and k is an integer shape parameter. The shape parameter is the squared inverse of the coefficient of variation, so $k \equiv \bar{s}^2/\sigma_s^2$ where σ_s is the standard deviation. A larger k indicates that the mean is larger relative to the standard deviation. An exponential distribution is equivalent to an Erlang distribution with $k = 1$, and a gamma distribution is an extension in which k can be noninteger. *Entekhabi and Eagleson* [1989] assumed that a gamma distribution describes the spatial variation of soil saturation, but they did not test this assumption. The Erlang distribution is convenient because it can be manipulated analytically, but it is inconvenient because it allows $0 \leq s < \infty$ when in reality $0 \leq s \leq 1$. We interpret locations with $s > 1$ as points at which the soil is saturated and water is standing on the surface. Numerical values of $s > 1$ are not used in the equations that follow; instead, the portion of the distribution with $s > 1$ is replaced with a probability mass at $s = 1$ for consistency with our physical interpretation.

[9] We can evaluate whether the spatial distribution of soil saturation conforms to an Erlang distribution using the Kolmogorov-Smirnov (KS) test [*Tadikamalla*, 1990; *Bain and Engelhardt*, 1992]. This test compares the cumulative distribution function (CDF) estimated from the data to the CDF of the hypothesized distribution. The maximum deviation between the CDFs is compared to the one expected at a given probability. Because $\bar{s}$ and k can change in time, this test must be performed separately with each spatial pattern of soil saturation. Only samples with 10 or more data points are tested. Samples with fewer data can be considered, but the estimates of k are unreliable because k depends on the sample variance. At a 10% level of significance, the KS tests accept the Erlang distribution for 89% of the samples (Table 1). The mean observed value of k is 11. As a

Table 1. Percent of the Observed Spatial Distributions of Soil Saturation (s) That Pass the KS Tests

Distribution	5% Level	10% Level	20% Level
Exponential	12	3	1
Erlang	92	89	78
Erlang with $k = 11$	98	94	84

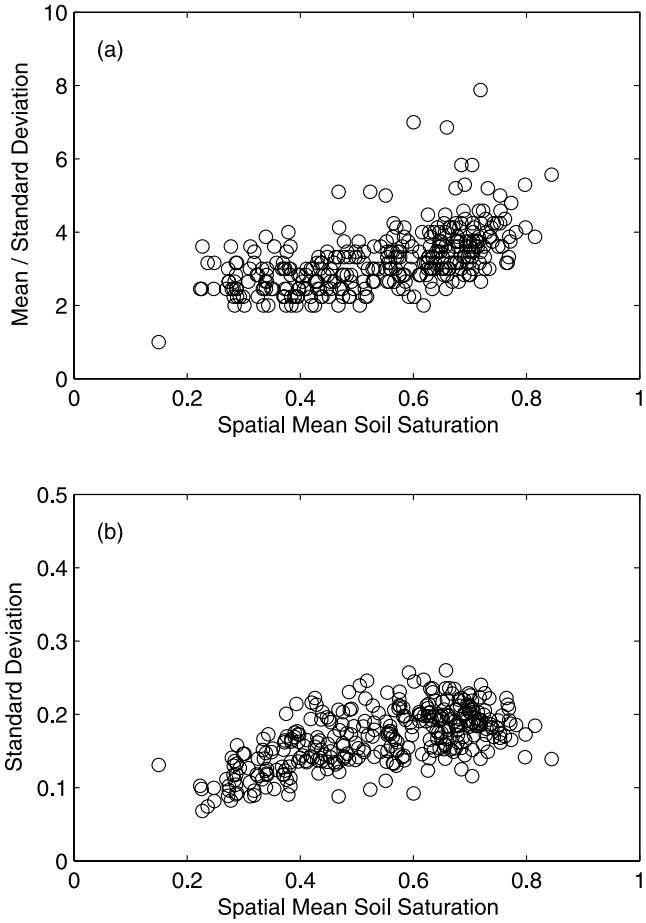

Figure 1. Plots showing how the characteristics of the spatial distribution of soil saturation s vary in time. Each circle represents the spatial distribution of soil saturation at a particular time. The x coordinate is the spatial mean soil saturation $\overline{s}$, and the y coordinate is another distribution parameter as indicated in the plot. Together the two coordinates completely define the shape of the Erlang distribution at the given time.

comparison, an exponential distribution is accepted for only 3% of the samples.

[10] In the analysis above, we have allowed both $\overline{s}$ and k to vary independently for each snapshot, but such flexibility may not be necessary. We next examine whether the standard deviation (σ_s) or the ratio of the mean and standard deviation ($\sqrt{k}$) is approximately constant among the snapshots. A constant value of $\sqrt{k}$ (or k) implies that the spatial distribution has a consistently higher variance during wetter periods. Figure 1 plots $\sqrt{k}$ and σ_s as functions of the mean soil saturation $\overline{s}$. $\sqrt{k}$ usually falls between 2 and 5, and it exhibits relatively little dependence on the mean soil saturation when the mean is between 0.2 and 0.6. In contrast, σ_s typically falls between 0.1 and 0.3, and tends to increase with spatial mean soil saturation when the mean is between 0.2 and 0.6. It remains relatively constant when the mean soil saturation is between 0.6 and 0.8. The coefficient of variation (COV) can be used to measure the variability of each quantity relative to their mean value. The COV for $\sqrt{k}$ is 0.26, whereas the COV for σ_s is slightly smaller at 0.23. Any linear dependence on $\overline{s}$ can be measured with the correlation coefficient. The correlation between $\sqrt{k}$ and $\overline{s}$ is about 0.55, whereas the correlation between σ_s and $\overline{s}$ is about 0.60. These results suggest that k is approximately constant except when the soil saturation is high, in which case the variance is more nearly fixed.

[11] On the basis of these observations we test the hypothesis that every sample of soil saturation is drawn from an Erlang distribution with $k = 11$. At the 10% level, 94% of the samples conform to this hypothesis (Table 1). This result is somewhat surprising because a higher percentage of snapshots pass the KS test when the parameter k is held constant than when it is estimated from the snapshot data. Such a result is possible because the KS test measures only the maximum deviation between the modeled and empirical CDFs, whereas the estimation procedure for k minimizes a measure of error that includes all points in the CDF. Because such a high percentage of snapshots conform to this simplified model, we conclude that the spatial distribution of soil saturation can be well approximated by an Erlang distribution with variable mean, $\overline{s}$, but a constant shape parameter of $k = 11$.

2.2. Temporal Variability of Mean Soil Saturation

[12] It is still necessary to characterize the temporal variability of $\overline{s}$ within each year. To do so, we follow the approach from the previous section. It is assumed that $\overline{s}$ can be characterized by a single distribution $f_{\overline{s}}$ within each year, and it is hypothesized that the population conforms to a beta distribution. Specifically:

$$f_{\overline{s}}(\overline{s}) = \frac{1}{B(b,c)}\overline{s}^{b-1}(1-\overline{s})^{c-1} \qquad 0 \le \overline{s} \le 1 \tag{2}$$

where B is the beta function, and b and c are parameters. The mean $E[\overline{s}]$ and standard deviation $\sigma_{\overline{s}}$ can be found from the parameters b and c:

$$E[\overline{s}] = \frac{b}{b+c} \tag{3}$$

$$\sigma_{\overline{s}} = \frac{1}{(b+c)}\sqrt{\frac{bc}{b+c+1}}. \tag{4}$$

This hypothesis can be evaluated using the KS test for each year. At the 10% level of confidence, 100% of the distributions pass the test (Table 2). For comparison, an Erlang distribution would be accepted for 95% of the years and an exponential distribution is never accepted.

[13] In the analysis of the beta distribution, b and c are estimated from the mean $E[\overline{s}]$ and the variance $\sigma_{\overline{s}}$ from each year. We next consider four hypotheses that can each determine $\sigma_{\overline{s}}$ for different values of $E[\overline{s}]$: (1) b is constant, (2) c is constant, (3) $E[\overline{s}]/\sigma_{\overline{s}}$ is constant (similar to the spatial distribution), and (4) $\sigma_{\overline{s}}$ is constant. Any of these hypoth-

Table 2. Percent of the Observed Temporal Distributions of Spatial Mean Soil Saturation ($\overline{s}$) That Pass the KS Tests

Distribution	5% Level	10% Level	20% Level
Exponential	0	0	0
Erlang	95	95	81
Beta	100	100	90
Beta with $\sigma_{\overline{s}} = 0.16$	100	90	81

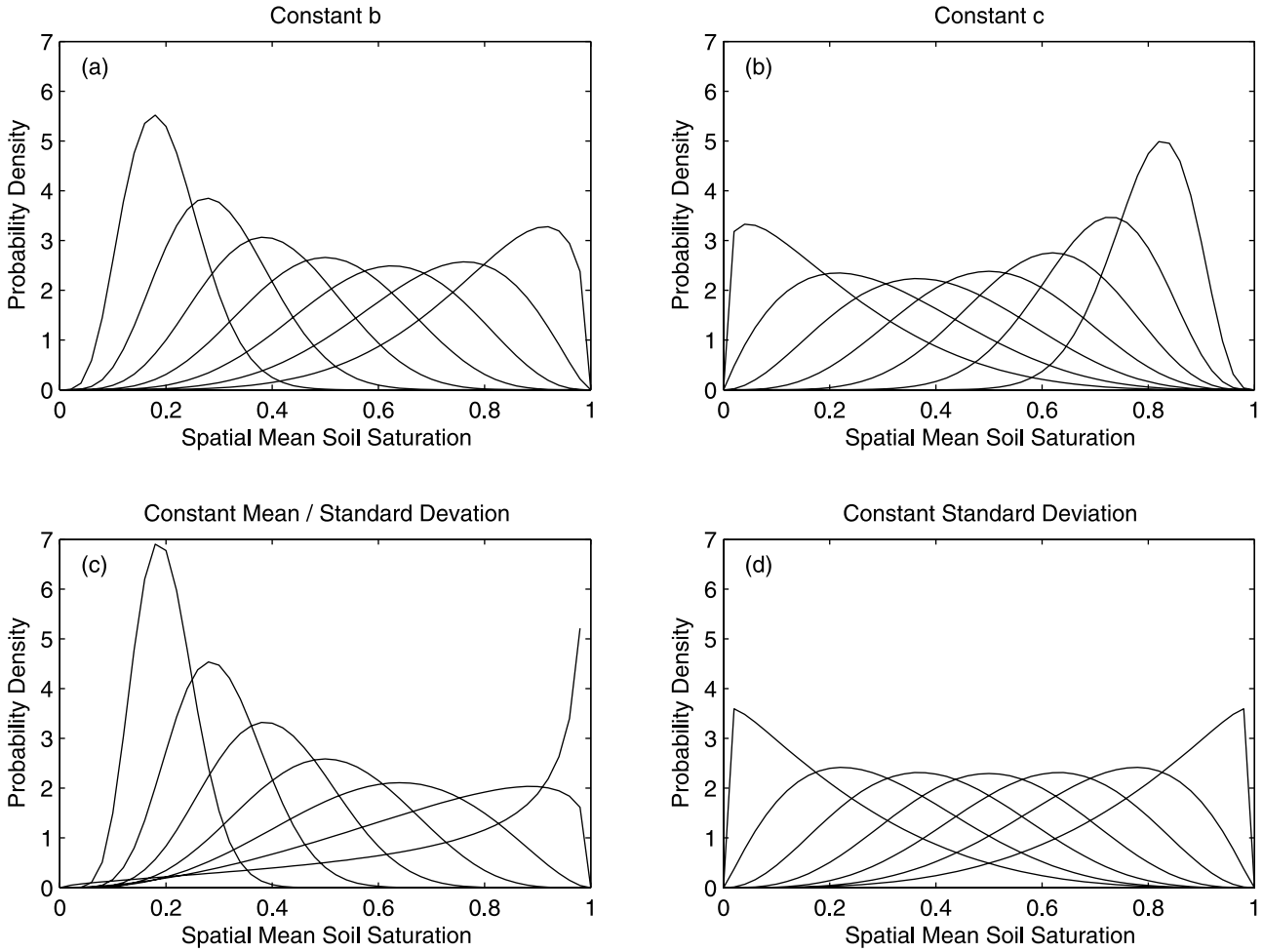

Figure 2. Hypothesized beta probability density functions describing the temporal variability of the spatial mean soil saturation $\overline{s}$. In each plot the mean $E[\overline{s}]$ is allowed to vary, which produces a different line, but a specified distribution characteristic is held constant: b is constant in Figure 2a, c is constant in Figure 2b, $E[\overline{s}]/\sigma_{\overline{s}}$ is constant in Figure 2c, and $\sigma_{\overline{s}}$ is constant in Figure 2d.

eses, if valid, would simplify the statistical description of the temporal variability of $\overline{s}$. Figure 2 shows the implications of these four assumptions on the probability density function of $\overline{s}$ as $E[\overline{s}]$ is varied. If b is held constant, the distribution has a prominent tail when the mean is large but no such tail when the mean is small. This form of distribution would imply that some periods are quite dry even when the space-time average soil saturation is large. A similar but more pronounced behavior is observed if $E[\overline{s}]/\sigma_{\overline{s}}$ is constant, and the opposite behavior is observed if c is constant. If $\sigma_{\overline{s}}$ is constant, a tail is observed for both extremes.

[14] Figure 3 evaluates the four hypotheses by plotting the estimated value of the four parameters (b, c, $E[\overline{s}]/\sigma_{\overline{s}}$, and $\sigma_{\overline{s}}$) for each year in the Illinois soil saturation dataset. If one of these parameters is constant, the plot of that parameter should exhibit no scatter and no dependence on the space-time average soil saturation. The coefficients of variation are 0.54, 0.56, 0.51, and 0.19 for constant b, c, $E[\overline{s}]/\sigma_{\overline{s}}$, and $\sigma_{\overline{s}}$, respectively. The correlations with $E[\overline{s}]$ are 0.04, −0.32, 0.33, and 0.07, for constant b, c, $E[\overline{s}]/\sigma_{\overline{s}}$, and $\sigma_{\overline{s}}$, respectively. $\sigma_{\overline{s}}$ has the smallest variability and a low correlation with $E[\overline{s}]$. b has the lowest correlation, but a relatively high variability. Overall, the results best support the assumption that $\sigma_{\overline{s}}$ is constant.

[15] The KS test can also be used to evaluate whether the observed variation of $\overline{s}$ is compatible with a distribution in which $\sigma_{\overline{s}}$ is constant. In this test, $E[\overline{s}]$ is estimated from each year of data, but the standard deviation is fixed at the average value for all the years ($\sigma_{\overline{s}}$ = 0.16). At the 10% confidence level, 90% of the years pass the KS test. Thus $\sigma_{\overline{s}}$ can be treated as a constant for the temporal distribution of $\overline{s}$ irrespective of the mean value $E[\overline{s}]$.

3. Model Development

[16] The model we propose is one dimensional and focuses on the vadose zone. Local instantaneous descriptions of the hydrologic processes are described first. We classify locations in the basin as either recharge or discharge locations (Figure 4). Most points in the basin are recharge locations. At these points, precipitation is partitioned between surface runoff and infiltration. Infiltrating water contributes to the soil saturation, which is considered constant with depth in the vadose zone. Water leaves the vadose zone through evapotranspiration to the atmosphere and recharge to the groundwater. Water leaves the groundwater reservoir only through groundwater runoff, which occurs at the discharge locations (mainly rivers and other

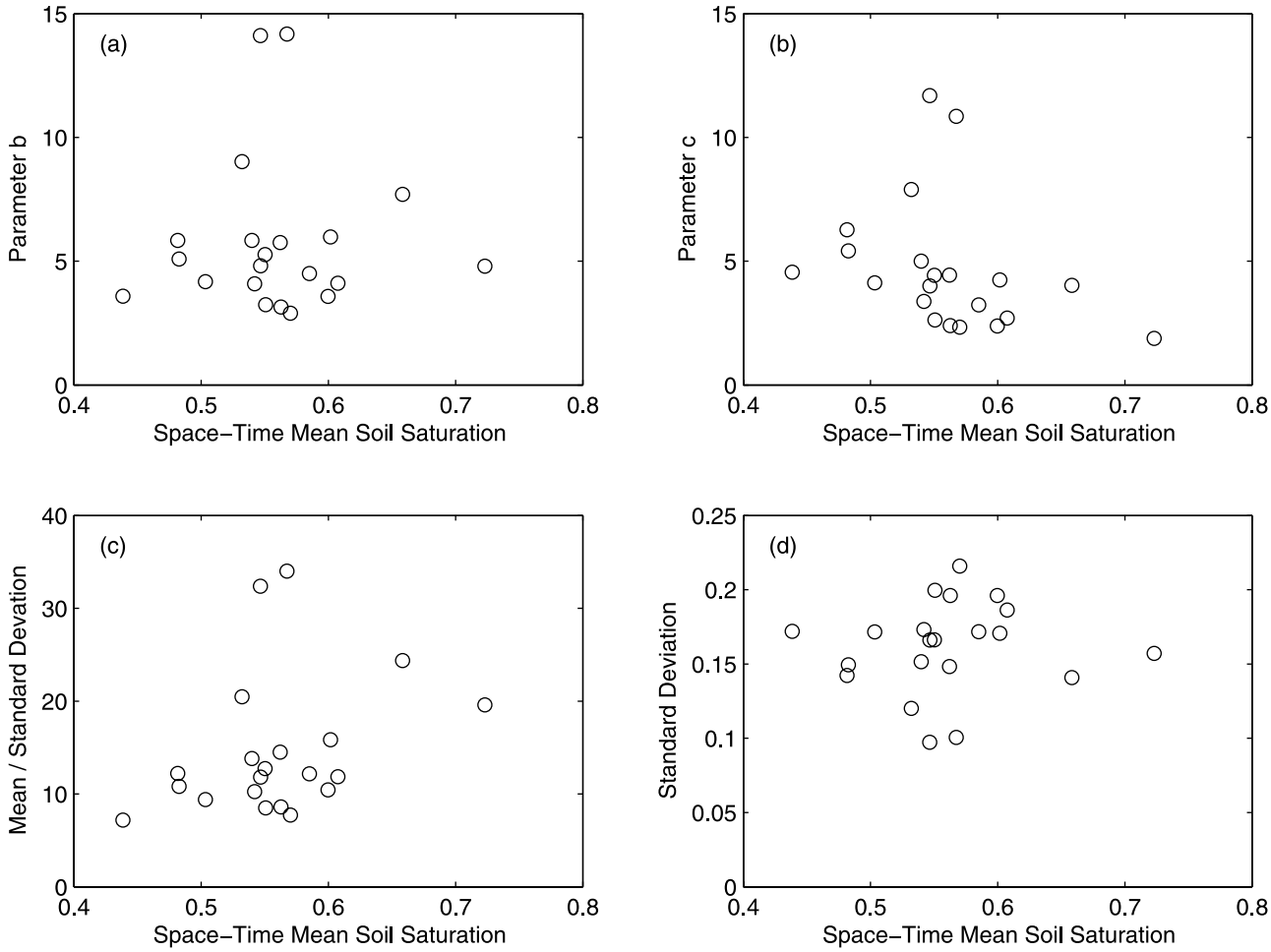

Figure 3. Plots showing how the characteristics of the temporal distribution of the spatial mean soil saturation $\overline{s}$ vary between years. Each circle corresponds to a single year. The x coordinate is always $E[\overline{s}]$ for the particular year, and the y coordinate is another distribution parameter as indicated in the plot. The two coordinate values completely define the shape of the beta distribution for that year.

water bodies). At discharge locations, recharge and infiltration do not occur. Once the local instantaneous hydrologic fluxes are characterized, they are integrated over space and time to determine the space-time average fluxes. Simple models for the spatial and temporal variability of precipitation and soil saturation are included in this integration process. A statistical equilibrium condition, which requires the space-time average inflows and outflows from the vadose zone to balance, is then satisfied. The sections below describe the models of the hydrologic processes and the statistical water balance condition in detail.

3.1. Infiltration and Surface Runoff

[17] At the local instantaneous scale, the surface runoff rate R can be described as the difference between the precipitation intensity P and the infiltration capacity F. Specifically,

$$R = \begin{cases} P - F & \text{If } P > F \\ 0 & \text{If } P \le F \end{cases}. \tag{5}$$

Note that F designates the infiltration capacity rate rather than a cumulative infiltration capacity. This model describes only the most basic type of runoff production. It neglects any additional runoff or infiltration as water moves toward the basin outlet, and it neglects the roles of snowfall, snowpack, and interception.

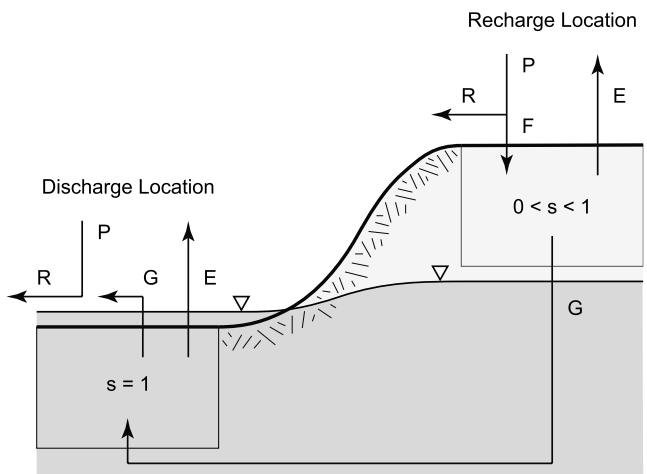

Figure 4. Schematic diagram of the model.

[18] The infiltration capacity depends on the soil saturation, time after ponding, hydraulic conductivity, lateral redistribution of water, and other factors. Numerous models have been proposed for infiltration capacity using varying degrees of empiricism [*Green and Ampt*, 1911; *Horton*, 1935; *Philip*, 1957]. One class of models describes the infiltration capacity as a function of the storage capacity available in the soil, which depends on the soil saturation [*Holtan*, 1961, 1965; *Entekhabi and Eagleson*, 1989]. We use a model that is similar to *Entekhabi and Eagleson* [1989] but explicitly describes both recharge and discharge locations:

$$F = \begin{cases} \alpha(1-s) + K_h & \text{Recharge Locations} \\ 0 & \text{Discharge Locations} \end{cases} \quad (6)$$

where α is an infiltrability parameter, s is the soil saturation, and K_h is the saturated hydraulic conductivity. The soil saturation is assumed to be constant with depth in the vadose zone. At recharge locations, the infiltration capacity is equivalent to a percolation rate when the soil is saturated (see below). When it is unsaturated, the infiltration capacity is higher because water fills the voids in the vadose zone (a pressure gradient may also increase the infiltration capacity). α controls the increase in infiltration capacity with reduced soil saturation. At discharge locations, $F = 0$ because the hydraulic gradient must be oriented to allow groundwater to return to the surface. Together, equations (5) and (6) describe the local instantaneous surface runoff rate if the precipitation intensity, soil saturation, and parameters (α and K_h) are known.

[19] To model the water balance of the Illinois River Basin, we are interested in the space-time average surface runoff. The average runoff differs from the local instantaneous runoff because the precipitation rate, soil saturation, and parameters vary in space and time. We treat the precipitation rate and soil saturation as random variables in space and time. For simplicity, spatial and temporal correlations of precipitation and soil saturation are neglected. Although stochastic variability of the parameters is not fully considered, it is indirectly expressed through the variability of soil saturation.

[20] Under these simplifying assumptions, the spatial average surface runoff ($\overline{R}$) can be calculated from the integral:

$$\overline{R} = \int \int_{P=F}^{\infty} (P - F) f_P dP f_s ds \quad (7)$$

where f_P and f_s are the spatial probability density functions of precipitation and soil saturation, respectively. Following *Entekhabi and Eagleson* [1989], the spatial distribution of precipitation can be written:

$$f_P = (1-\mu)\delta(P) + \frac{\mu^2}{\overline{P}} e^{-\mu P/\overline{P}} \quad (8)$$

where μ is the fraction of the basin receiving precipitation, $\delta(\cdot)$ is a dirac delta function, and $\overline{P}$ is the spatial average precipitation rate including the areas with zero precipitation. The distribution in equation (8) is exponential with a mass at zero that accounts for the probability of zero precipitation ($1 - \mu$). Using the expression for f_P, the spatially averaged surface runoff can be rewritten as:

$$\overline{R} = \overline{P} \int_{s=0}^{1} e^{-\mu F/\overline{P}} f_s ds. \quad (9)$$

[21] *Eltahir and Bras* [1993] have shown that μ increases linearly with $\overline{P}$, so the ratio $i \equiv \overline{P}/\mu$ is approximately constant. The variable i is the average precipitation rate in the portion of the basin where precipitation occurs. This result is important to the model development because it replaces two variables ($\overline{P}$ and μ) with one constant (i). Thus equation (9) can be rewritten:

$$\overline{R} = \overline{P} \int_{s=0}^{1} e^{-F/i} f_s ds. \quad (10)$$

f_s must be specified to determine $\overline{R}$. Section 2 demonstrated that f_s conforms to an Erlang distribution. However, an Erlang distribution allows $0 \leq s < \infty$ when $s > 1$ is not physically meaningful. *Entekhabi and Eagleson* [1989] accounted for this inconsistency by treating the portion of the distribution above one as a probability mass at $s = 1$. We follow this approach and further assume that the portion of the distribution with $s > 1$ is approximately the proportion of discharge locations within the basin. As an example, when $k = 11$ and $\overline{s} = 0.5$, about 0.35% of the basin is classified as discharge locations. If $\overline{s}$ increases to 0.7, the proportion increases to 8.7%. While these proportions seem reasonable, there are no direct observations of the size of discharge regions with which we can compare. Our approach is the simplest one possible, aside from neglecting discharge locations altogether, which is less realistic. If this portion of the distribution is used to represent discharge locations, the infiltration capacity at these sites is zero, and the integral in equation (10) becomes:

$$\overline{R} = \overline{P} \int_{s=0}^{1} e^{-[\alpha(1-s)+K_h]/i} f_s ds + \overline{P} \int_{s=1}^{\infty} f_s ds \quad (11)$$

After inserting equation (1) into equation (11) and integrating, one obtains an analytical expression for the spatially-averaged surface runoff:

$$\overline{R} = \overline{P}(k/\overline{s})^k \cdot \left\{ \frac{e^{-(K_h+\alpha)/i}}{(k/\overline{s} - \alpha/i)^k} + \sum_{\rho=0}^{k-1} \frac{e^{-k/\overline{s}}}{(k-1-\rho)!} \left[(\overline{s}/k)^{\rho+1} - \frac{e^{-K_h/i}}{(k/\overline{s} - \alpha/i)^{\rho+1}} \right] \right\}. \quad (12)$$

[22] Finally, the space-time average runoff $E[\overline{R}]$ can be calculated by integrating over the temporal distribution of $\overline{s}$. If a spatially averaged runoff coefficient is defined as $\overline{r} \equiv \overline{R}/\overline{P}$, then:

$$E[\overline{R}] = \int_{\overline{s}=0}^{1} \int_{\overline{P}=0}^{\infty} \overline{P} f_{\overline{P}} \overline{r} f_{\overline{s}} d\overline{s} = E[\overline{P}] \int_{\overline{s}=0}^{1} \overline{r} f_{\overline{s}} d\overline{s}. \quad (13)$$

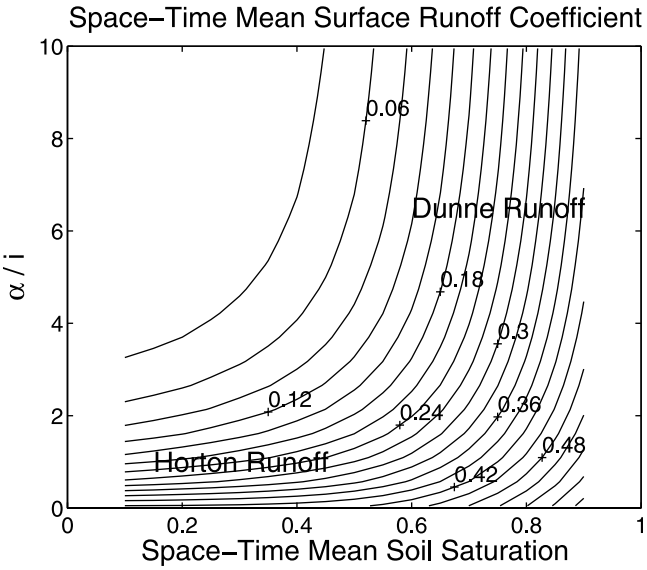

Figure 5. Space-time average surface runoff coefficient as a function of $E[\bar{s}]$ and α/i when $k = 11$, $\sigma_{\bar{s}} = 0.16$, and $K_h = 2.9 \times 10^{-5}$ cm/s.

Note that the temporal distribution of precipitation is not required to calculate $E[\bar{R}]$. Equation (13) can also be written in terms of the space-time average runoff coefficient:

$$E[\bar{r}] \equiv \frac{E[\bar{R}]}{E[\bar{P}]} = \int_{\bar{s}=0}^{1} \bar{r} f_{\bar{s}} d\bar{s}. \tag{14}$$

In section 2, $f_{\bar{s}}$ was shown to conform to a beta distribution with constant $\sigma_{\bar{s}}$ (see equation (2)). Unfortunately, equation (14) cannot be integrated analytically with this distribution, but it can be evaluated numerically.

[23] Figure 5 shows the dependence of $E[\bar{r}]$ on α/i and $E[\bar{s}]$. When $\alpha/i < 2$, the runoff coefficient increases substantially with decreasing α/i but exhibits little sensitivity to $E[\bar{s}]$ when $E[\bar{s}]$ is small. Low values of α/i indicate that the infiltration capacity is low relative to the precipitation intensity. Thus surface runoff is dominated by infiltration excess or "Horton" runoff [*Freeze*, 1974]. When $\alpha/i > 2$, the runoff coefficient increases with increasing $E[\bar{s}]$ but exhibits less sensitivity to α/i. When $E[\bar{s}]$ is large, many locations are saturated. Thus saturation excess or "Dunne" runoff dominates [*Hewlett*, 1961; *Hewlett and Hibbert*, 1965; *Dunne and Black*, 1970]. The magnitude of the runoff coefficient is controlled by K_h. Higher values of K_h allow more infiltration at recharge locations and reduce the surface runoff.

3.2. Evapotranspiration

[24] Evapotranspiration rates depend on the availability of moisture, energy to vaporize the moisture, and wind for transportation. One can define the potential evapotranspiration E_p as the evapotranspiration that would occur given unlimited moisture. E_p therefore accounts for the availability of energy and wind conditions but not moisture availability. In this analysis, the potential evapotranspiration is assumed to be constant in space and time because of data limitations. This assumption is important because potential evapotranspiration is known to vary significantly through the year (the role of temporal variability will be discussed at the end of this section).

[25] Efficiency of evapotranspiration ε can be defined as the ratio of the actual evapotranspiration E to the potential evapotranspiration ($\varepsilon \equiv E/E_p$). Clearly, ε depends on the soil saturation, but the precise nature of this dependence is not known. ε also depends on the relative magnitudes of transpiration and evaporation [*Lowry*, 1959; *Rodriguez-Iturbe et al.*, 1990]. When transpiration dominates, ε becomes an efficiency of water extraction by plant roots. *Molz* [1981] reviewed several relationships for ε in this case. When the soil is bare, evaporation dominates. *Eagleson* [1978d] and *Entekhabi and Eagleson* [1989] developed relationships for ε for bare soils. *Sud and Fennessey* [1982] derived an empirical relationship for ε that includes contributions from both evaporation and transpiration.

[26] In order to keep our model parsimonious, we lump evaporation and transpiration together. We propose:

$$\varepsilon = \begin{cases} s/\beta & \text{If } 0 \le s \le \beta \\ 1 & \text{If } \beta < s \le 1 \end{cases} \tag{15}$$

where β is the point that separates moisture-limited and energy/transport limited evapotranspiration. β is expected to be lower for more thickly vegetated soils [*Lowry*, 1959; *Rodriguez-Iturbe et al.*, 1990]. Similar models have been used by other authors for efficiency of water extraction by roots [*Feddes et al.*, 1976; *Molz*, 1981; *Entekhabi and Eagleson*, 1989].

[27] The spatially averaged efficiency of evapotranspiration $\bar{\varepsilon}$ can be determined by following the approach that was used to determine $\bar{r}$. In particular, $\bar{\varepsilon}$ is found by integrating ε over the spatial distribution of soil saturation:

$$\bar{\varepsilon} = \int_{s=0}^{1} \varepsilon f_s ds = \int_{s=0}^{\beta} \frac{s}{\beta} f_s ds + \int_{s=\beta}^{\infty} f_s ds. \tag{16}$$

As in the previous section, the portion of the Erlang distribution with $s > 1$ is treated as a probability mass at $s = 1$. After integrating equation (16), one obtains:

$$\bar{\varepsilon} = \frac{\bar{s}}{\beta} - \sum_{\rho=1}^{k} \frac{\rho (k/\bar{s})^{k-1-\rho} \beta^{k-1-\rho} e^{-k\beta/\bar{s}}}{(k-\rho)!}. \tag{17}$$

The space-time average efficiency $E[\bar{\varepsilon}]$ can be found from another integration:

$$E[\bar{\varepsilon}] = \int_{\bar{s}=0}^{1} \bar{\varepsilon} f_{\bar{s}} d\bar{s}. \tag{18}$$

This integration cannot be performed analytically, but numerical evaluation is not difficult.

[28] Figure 6 shows the results of the numerical integration. $E[\bar{\varepsilon}]$ increases with decreasing β or increasing $E[\bar{s}]$. When β is small, the average efficiency of evapotranspiration can reach one when $E[\bar{s}]$ is large. When β is large, the efficiency always remains below one. A large β implies that the soil and vegetation require high soil saturations in order

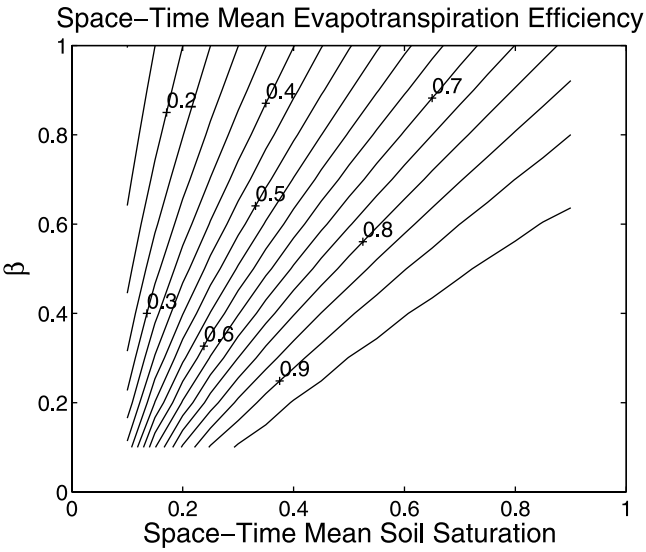

Figure 6. Space-time average efficiency of evapotranspiration as a function of $E[\overline{s}]$ and β when $k = 11$ and $\sigma_{\overline{s}} = 0.16$.

to evaporate and transpire efficiently. So, a significant proportion of locations have an evapotranspiration rate below the potential rate even when $E[\overline{s}]$ is large.

[29] In the derivation above, E_p was treated as a constant, which implicitly neglects any correlation between the variability of soil saturation and that of potential evapotranspiration. For the Illinois basin, seasonal cycles of these variables result in a correlation coefficient of about -0.55. The role of the correlation can be included in the model because the temporal integration is performed numerically. The impact of the correlation is most significant when β is large or the space-time average soil saturation is small. Under these circumstances, the evapotranspiration rate is reduced by less than 10% when the correlation is included. Given this relatively small impact, we neglect the correlation to maintain the simplicity of the model.

3.3. Recharge and Groundwater Runoff

[30] Recharge is controlled by a number of factors including soil saturation, water table level, and topography. A simple model for the local instantaneous recharge G can be developed by considering Darcy's Law, which states that:

$$G = -K_h(s)\frac{dh}{dz'} \tag{19}$$

where $K_h(s)$ is the unsaturated hydraulic conductivity, h is the hydraulic head, and z' is the depth below the ground surface. If the gradient in hydraulic head is dominated by the change in elevation, then equation (19) can be rewritten as:

$$G = -K_h(s)\frac{dz}{dz'} = K_h(s). \tag{20}$$

This is the well known percolation model, which is most applicable to equilibrium recharge when the water table is far below the surface. *Campbell* [1974] suggested that the unsaturated hydraulic conductivity can be described as:

$$K_h(s) = K_h s^{\gamma} \tag{21}$$

where K_h is the saturated hydraulic conductivity and γ is a parameter that depends on soil type. Finer textured soils tend to have larger values of γ [*Clapp and Hornberger*, 1978]. If a recharge efficiency is defined as $g \equiv G/K_h$, then g is:

$$g = s^{\gamma}. \tag{22}$$

This simple model for local instantaneous recharge neglects lateral convergence or divergence of water, variations of soil saturation with depth, the level of the water table, and other complicating factors. However, it requires only one parameter, which can be determined from the soil type.

[31] The spatially averaged recharge efficiency $\overline{g}$ can be found by integrating g over the spatial distribution of soil saturation:

$$\overline{g} = \int_{s=0}^{1} g f_s ds = \int_{s=0}^{1} s^{\gamma} f_s ds. \tag{23}$$

Notice that the integral excludes $s > 1$ because this portion of the distribution represents the discharge locations. Evaluating the integral, one finds:

$$\overline{g} = \frac{(k-1+\gamma)!}{(k-1)!}(k/\overline{s})^{-\gamma}\left[1 - e^{-k/\overline{s}}\sum_{\rho=0}^{k-1+\gamma}\frac{(k/\overline{s})^{k-1+\gamma-\rho}}{(k-1+\gamma-\rho)!}\right]. \tag{24}$$

The space-time recharge efficiency $E[\overline{g}]$ can be found from:

$$E[\overline{g}] = \int_{\overline{s}=0}^{1} \overline{g} f_{\overline{s}} d\overline{s} \tag{25}$$

which must be calculated numerically.

[32] Ultimately, we are more interested in the space-time average groundwater runoff than the recharge. However, the two quantities are related. Consider the water balance of the groundwater reservoir at the annual timescale. Inflow occurs through recharge and outflow occurs by evaporation and groundwater runoff. One can reasonably neglect the change in storage at the annual timescale. Although evaporation from groundwater is important during the summer months [*Yeh*, 2002], it is less important over the annual timescale. If the evaporation is negligible, then the space-time average recharge must balance the space-time average groundwater runoff. Thus $E[\overline{g}]$ is also a space-time average efficiency of groundwater runoff production.

[33] Figure 7 shows contours of space-time average groundwater runoff efficiency as a function of γ and $E[\overline{s}]$. Because γ is so large, the efficiency is strongly dependent on $E[\overline{s}]$. Little to no recharge or groundwater runoff is expected when $E[\overline{s}]$ is small. Increasing γ slightly reduces the efficiency because it reduces the recharge when the soil is unsaturated (see equation (22)). This dependence implies that finer textured soils such as silt and clay (which have

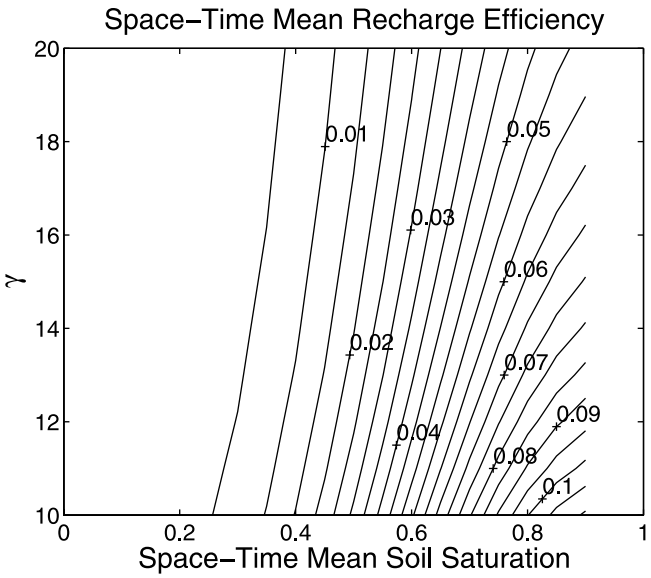

Figure 7. Space-time average efficiency of recharge or groundwater runoff as a function of $E[\overline{s}]$ and γ when $k = 11$ and $\sigma_{\overline{s}} = 0.16$.

larger values of γ) have a lower groundwater runoff efficiency than coarser soils such as sands.

3.4. Water Balance Condition

[34] The previous sections have related space-time average soil saturation to the space-time average surface runoff, evapotranspiration, and groundwater runoff. If $E[\overline{s}]$ is known, these relations can be used independently to determine their respective fluxes. However, $E[\overline{s}]$ is typically unknown, so an additional condition is required to use these models.

[35] An additional condition is available through a statistical form of the soil moisture water balance, which can be written:

$$\eta \frac{dE[\overline{s}]}{dt} = E[\overline{P}] - E[\overline{R}] - E[\overline{E}] - E[\overline{G}] \tag{26}$$

where η is the storage capacity of the vadose zone (the porosity times an acceptable soil depth) and t is time. Recall that all the temporal averages include only intra-annual variability. Thus it is possible for $E[\overline{s}]$ to change in time. This relation states that changes of $E[\overline{s}]$ in time correspond to imbalances in the space-time average fluxes. At the annual timescale, however, the space-time average soil saturation is approximately constant. Thus the left side of equation (26) becomes zero. If one also uses the definitions for space-time average runoff coefficient, efficiency of evapotranspiration, and efficiency of groundwater runoff, one can write:

$$0 = E[\overline{P}] - E[\overline{r}]E[\overline{P}] - E[\overline{\varepsilon}]E_p - E[\overline{g}]K_h \tag{27}$$

or

$$1 = E[\overline{r}] + \frac{E[\overline{\varepsilon}]}{D_E} + \frac{E[\overline{g}]}{D_K} \tag{28}$$

where $D_E \equiv E[\overline{P}]/E_p$ is a potential evapotranspiration-related humidity index and $D_K \equiv E[\overline{P}]/K_h$ is a conductivity-related humidity index.

[36] $E[\overline{s}]$ can be found from equation (28) because $E[\overline{r}]$, $E[\overline{\varepsilon}]$, and $E[\overline{g}]$ all depend on $E[\overline{s}]$. Figure 8 plots the dependence of $E[\overline{s}]$ on the two humidity indices. When either index is small, that index controls the soil saturation. For example, when the potential evapotranspiration is large, D_E is small, and D_K has virtually no influence on E[$\overline{s}$] (the contours are vertical in Figure 8). A similar argument holds for D_K, but it is never able to completely control the soil moisture. The plot suggests that the soil moisture can be classified into three states: exclusively D_E controlled, D_E dominated, and D_K dominated.

[37] Figure 8 can also be used in conjunction with Figures 5–7 to illustrate the use of the proposed model when $E[\overline{s}]$ is unknown. Given values for D_E and D_K, $E[\overline{s}]$ can be found from Figure 8. Then the space-time average fluxes can be found from Figures 5–7. Notice Figures 5–7 depend on specific parameter values (see the figure captions), so the plots are not applicable to all regions.

4. Model Application in Illinois

[38] In this section, the model is applied to the Illinois River Basin for the years between 1981 and 1999. During this period, soil saturation, precipitation, and runoff data are all available. The first subsection describes the data sources and estimates several required variables from the data and literature ($E[\overline{P}]$, E_p, K_h, i). Annual values are used for $E[\overline{P}]$; all the others are constants. In the second subsection, the models for $E[\overline{\varepsilon}]$, $E[\overline{r}]$, and $E[\overline{g}]$ are applied using data for $E[\overline{s}]$ to calibrate α, β, γ. Finally, in the third subsection, the models are applied together using the water balance condition to investigate the relationships between annual precipitation, soil saturation, and runoff.

4.1. Parameter Estimates

[39] $E[\overline{P}]$ can be estimated using daily precipitation data from the National Climatic Data Center. 129 stations are

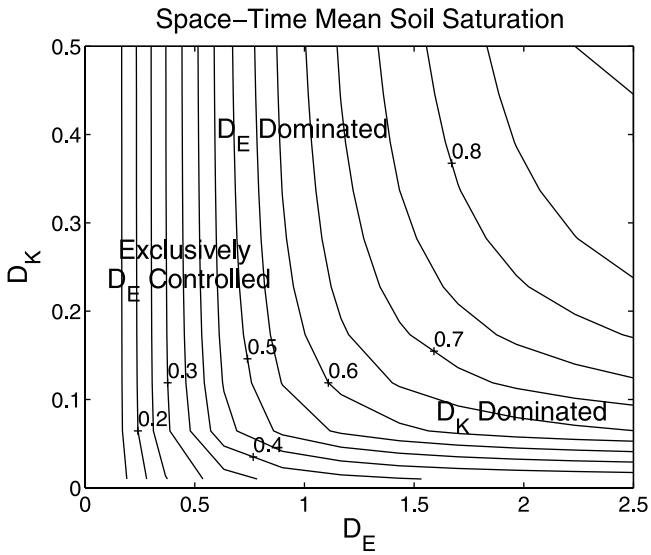

Figure 8. Space-time average soil saturation as a function of the two humidity indices D_E and D_K. Figure 8 was generated using $k = 11$, $\sigma_{\overline{s}} = 0.16$, $K_h = 2.9 \times 10^{-5}$ cm/s, $i = 3.2 \times 10^{-5}$ cm/s, $\alpha = 1.0 \times 10^{-4}$ cm/s, $\beta = 0.87$, and $\gamma = 19$.

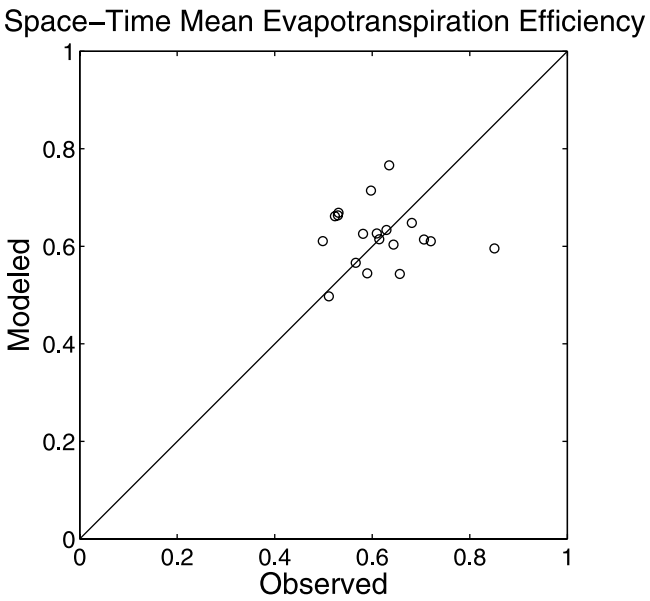

Figure 9. Comparison of the observed values of $E[\bar{\varepsilon}]$ with those from the model (using $k = 11$, $\sigma_{\bar{s}} = 0.16$, and $\beta = 0.87$).

available in Illinois, and based on Theissen polygons, 71 stations were found to represent the river basin. These stations were used to estimate $E[\overline{P}]$ for 1981 to 1999. $E[\overline{P}]$ is found to change substantially from year to year. For example, $E[\overline{P}] = 675$ mm/yr in 1988, and $E[\overline{P}] = 1264$ mm/yr in 1993.

[40] E_p was calculated from a standard method based on the Penman-Monteith equation [*Shuttleworth*, 1993]. The necessary climatic data are available at a daily resolution from the National Climatic Data Center at first-order climate stations. We used the station at Peoria, which is near the center of the Illinois River Basin. Unfortunately, this station has sufficient climatic data to calculate E_p only from 1984 to 1994. To apply the method, certain parameters must also be known (e.g., albedo), and these parameters were estimated by calibration. Average monthly potential evapotranspiration rates for Illinois are available in the literature from April to October [*Yeh et al.*, 1998]. These estimates are based on pan evaporation measurements reported by *Farnsworth and Thompson* [1982]. We selected the parameters by minimizing the disagreement between our average monthly estimates from April to October and those from *Yeh et al.* [1998]. In the end, we calculated $E_p = 958$ mm/yr. The maximum and minimum annual values were 850 mm/yr and 1150 mm/yr, so some mild variation of potential evapotranspiration is observed from year to year.

[41] An estimate of K_h is also required. *Hollinger and Isard* [1994] characterized the soil as silt loam at 17 of 19 stations where the soil moisture was measured, and the USDA classified the soil in Illinois as fine silt [*Food and Agricultural Organization*, 1975]. For silt, *Freeze and Cherry* [1979] report typical hydraulic conductivities between 1×10^{-7} cm/s and 1×10^{-3} cm/s. Unfortunately, such a broad range does not adequately constrain the value of K_h, so more information is needed. In the proposed model, K_h controls the relative contributions of the surface and groundwater to the total runoff. Although a significant amount of uncertainty exists, it has been estimated that about 50% of the runoff comes from groundwater [*Yeh*, 2002]. In order for the model to reproduce the appropriate relative contribution from groundwater, K_h is set to 2.9×10^{-5} cm/s (9,000 mm/yr), which is reasonable for silts.

[42] The average rainfall intensity i must also be estimated. *Eltahir and Bras* [1993] calculated i for five locations in different climates. For the humid-temperate locations, they reported i values between 3.1×10^{-5} cm/s and 9.4×10^{-5} cm/s with the most common values occurring in the lower part of this range. Because i enters the surface runoff equation only in ratios with K_h and α, which are both uncertain, a rough estimate of i is adequate. We selected $i = 3.2 \times 10^{-5}$ cm/s (10,000 mm/yr) for the Illinois River Basin.

4.2. Calibration of α, β, and γ

[43] β can be calibrated by comparing $E[\overline{E}]$ from the model and observations. The model's estimate of $E[\overline{E}]$ can be calculated from Figure 6 when $E[\bar{s}]$ and E_p are known and β is guessed. Notice that this estimate of $E[\overline{E}]$ is independent of the surface and groundwater runoff models. The estimate of $E[\overline{E}]$ can then be compared to the evapotranspiration determined from observations. For a large basin at an annual timescale, the evapotranspiration is well estimated by the difference between the precipitation and runoff. The annual runoff can be calculated from USGS daily discharge data at the basin outlet, which is near Valley City. The appropriate value of β is the one that produces the best agreement between the $E[\overline{E}]$ from the model and the observations for all the years considered. Using this procedure, we found $\beta = 0.87$. Figure 9 compares the efficiencies of evapotranspiration from the model and the water balance using this value of β. No systematic bias is observed in the model results, but some scatter is visible. The scatter indicates that a significant portion of the variability in evapotranspiration between years is not represented by the model. This limitation may be due to temporal variations in the potential evapotranspiration and β, which were assumed constant.

[44] The parameters α and γ can be calibrated using a similar procedure. Annual runoff is known for each year from the USGS data. The total runoff can also be calculated from the surface and groundwater models (Figures 5 and 7). Given $E[\bar{s}]$, $E[\overline{P}]$, and K_h, and guessing α and γ, the total runoff can be calculated as the sum of the surface and groundwater components. The best values of α and γ are the ones that minimize the disagreement between the observed and modeled runoff. This optimization procedure gives $\alpha = 1.0 \times 10^{-4}$ cm/s (30,000 mm/yr) and $\gamma = 19$. γ was not allowed to be any larger than 19 because, for silt loam, *Clapp and Hornberger* [1978] estimated a mean value of $\gamma = 13.6$ and a standard deviation 3.9. Thus $\gamma = 19$ is approximately 1.5 standard deviations from the mean, which we considered the limit of acceptability.

[45] Figure 10 compares the observed and modeled runoff coefficients using the calibrated values of α and γ. The modeled runoff coefficients reproduce the observations well for the more extreme conditions, but significant scatter is observed when the observed runoff coefficient is between 0.30 and 0.45. There is also a slight tendency to overestimate high runoff coefficients and underestimate low runoff coefficients. The behavior becomes more apparent in the model as the contribution of groundwater to the total runoff

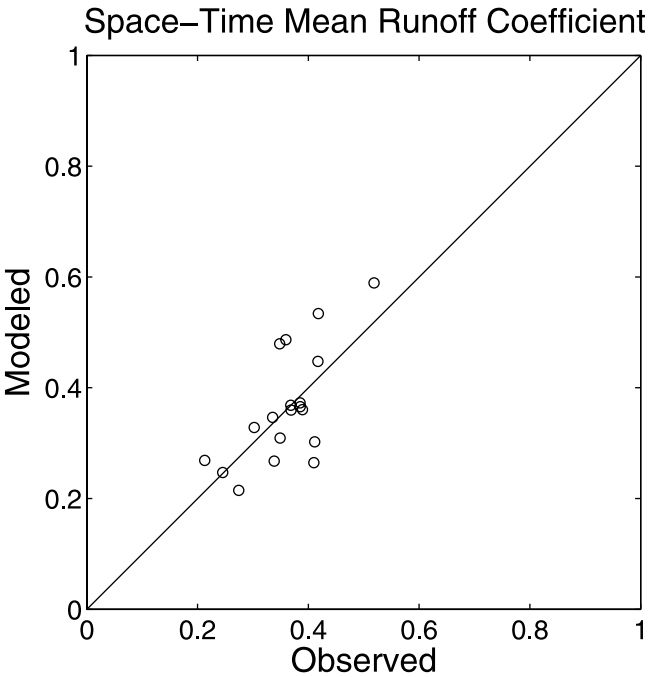

Figure 10. Comparison of observed and modeled values of space-time average runoff coefficient. The model results were generated using $k = 11$, $\sigma_{\bar{s}} = 0.16$, $K_h = 2.9 \times 10^{-5}$ cm/s, $i = 3.2 \times 10^{-5}$ cm/s, $\alpha = 1.0 \times 10^{-4}$ cm/s, and $\gamma = 19$.

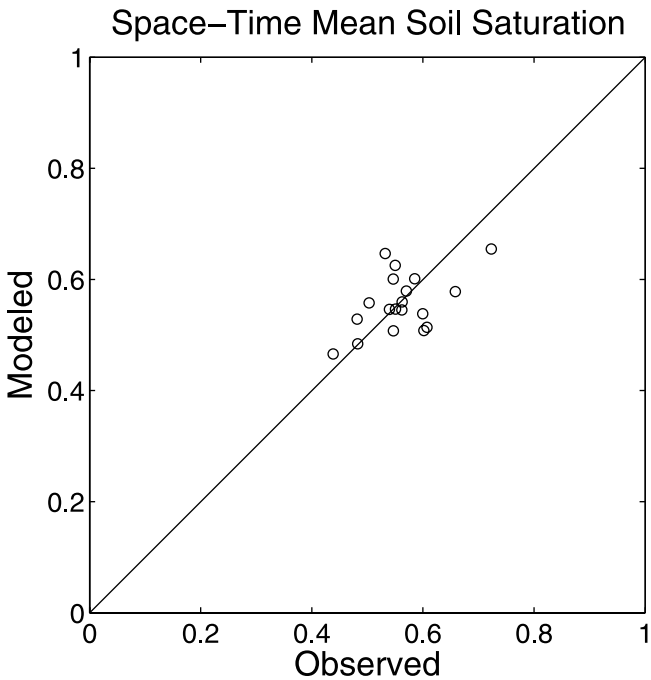

Figure 11. Comparison of annual values of observed and modeled $E[\bar{s}]$. The model results were generated using $k = 11$, $\sigma_{\bar{s}} = 0.16$, $E_p = 958$ mm, $K_h = 2.9 \times 10^{-5}$ cm/s, $i = 3.2 \times 10^{-5}$ cm/s, $\alpha = 1.0 \times 10^{-4}$ cm/s, $\beta = 0.87$, and $\gamma = 19$.

increases. One possible cause for this over-sensitivity is the use of soil saturation between 0 and 10 cm depth in the model. This range of depths is well-suited to describe surface runoff production, but recharge is probably better represented by the average saturation over the entire vadose zone. Soil saturation at larger depths is less variable than the soil saturation at smaller depths. For example, k was found to be 11 for 0–10 cm. If 0–2 m was used, the average estimate of k would be 29 indicating about one-third the variability with respect to the spatial mean soil saturation. This suggests that the groundwater runoff production would be less variable if the saturation was averaged over a larger soil depth. Thus the sensitivity of total runoff to changes in soil saturation would be reduced.

4.3. Model Results

[46] In the previous subsection, the evapotranspiration model and the surface and groundwater runoff models were calibrated separately. In this section, all the components of the model are applied together. Instead of supplying the data for $E[\bar{s}]$ to determine each of the fluxes independently, $E[\overline{P}]$ is supplied for each year, and $E[\bar{s}]$ is calculated from the statistical water balance condition in Figure 8.

[47] Figure 11 compares the calculated and observed values of $E[\bar{s}]$ for each year. Although some scatter is apparent, relatively good agreement is observed between the model and observations. These results suggest that the model can produce realistic estimates of $E[\bar{s}]$ from $E[\overline{P}]$ data, which supports the water balance condition used to generate Figure 8.

[48] Figure 12 compares the modeled and observed values of the space-time average runoff coefficient. Relatively good agreement is also observed here. In fact, the agreement is better in Figure 12 than in Figure 10, which compares the same quantities when $E[\bar{s}]$ data is used to drive the model instead of $E[\overline{P}]$. This improvement occurs because precipitation is a better statistical predictor of runoff than soil saturation. Alternatively stated, a fitted relation

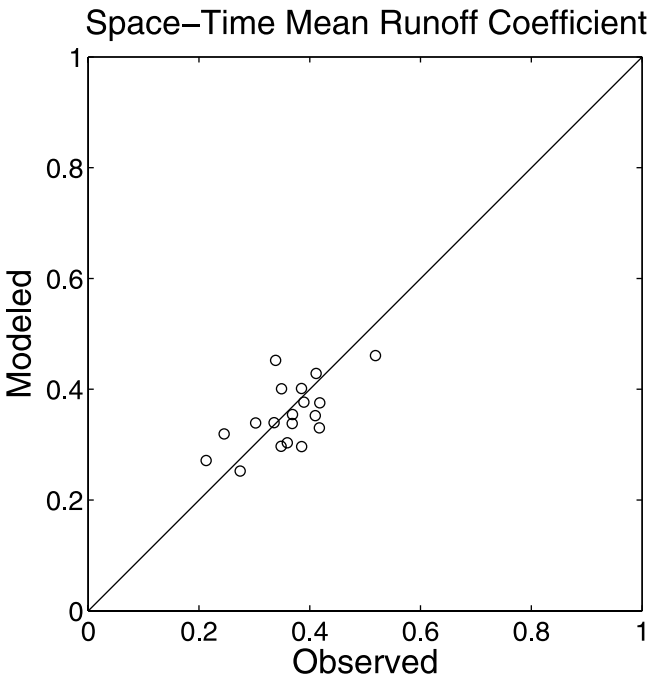

Figure 12. Comparison of annual values of observed and modeled runoff coefficient. The model results were generated using $k = 11$, $\sigma_{\bar{s}} = 0.16$, $E_p = 958$ mm, $K_h = 2.9 \times 10^{-5}$ cm/s, $i = 3.2 \times 10^{-5}$ cm/s, $\alpha = 1.0 \times 10^{-4}$ cm/s, $\beta = 0.87$, and $\gamma = 19$.

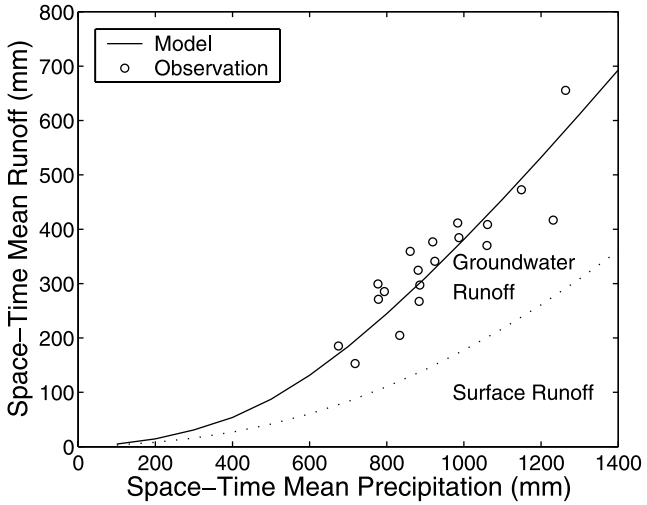

Figure 13. The observed and modeled relationships between space-time average runoff and space-time average precipitation. The distance below the dashed line indicates the model's surface runoff coefficient. The distance between the two lines indicates the model's groundwater runoff coefficient. The model results were generated using $k = 11$, $\sigma_{\bar{s}} = 0.16$, $E_p = 958$ mm, $K_h = 2.9 \times 10^{-5}$ cm/s, $i = 3.2 \times 10^{-5}$ cm/s, $\alpha = 1.0 \times 10^{-4}$ cm/s, $\beta = 0.87$, and $\gamma = 19$.

between precipitation and runoff explains more of the variability in runoff than a fitted relation between soil saturation and runoff. This interesting characteristic suggests that (1) runoff has some dependence on precipitation which is independent of soil saturation, (2) soil saturation has some variability that does not impact runoff production, or (3) both.

[49] Figure 13 shows the relationship between space-time average runoff and precipitation for the Illinois River Basin. The modeled relationship is nearly linear in the range where observations are available. The linearity is interesting because this relationship is not imposed externally. Rather, it is derived from a relationship between precipitation and soil saturation and relationships between soil saturation and surface and groundwater runoff. The model reproduces the overall trend in the runoff data, but scatter is also observed around the model predictions. The model cannot reproduce this scatter because the only variable that is changed between years is the precipitation. Thus, for a given space-time average precipitation, the model always gives the same space-time average runoff. The scatter may be due to variations in potential evapotranspiration between years or temporal variability in other parameters (e.g., β).

[50] Figure 14 shows the relationship between the space-time average runoff coefficient and space-time average soil saturation. A rather complex form of this relationship is observed. When $E[\bar{s}]$ is low, evapotranspiration dominates, and both the surface and groundwater runoff coefficients are low. As $E[\bar{s}]$ passes 0.4, both types of runoff production increase in strength and the total runoff coefficient line curves upward. When $E[\bar{s}]$ is greater than 0.6, the surface runoff coefficient continues to increase, but the groundwater runoff coefficient remains approximately constant. The groundwater's behavior is controlled by two opposing factors. As $E[\bar{s}]$ increases, groundwater recharge becomes more efficient (see equation (22)). However, smaller portions of the region contribute recharge because the proportion of discharge locations increases to convey the groundwater runoff. Thus the groundwater runoff coefficient remains approximately constant when $E[\bar{s}]$ increases past 0.6. The model and observations are in good agreement, but the results of the model at very low and high soil saturations have not been tested by data.

5. Conclusions

[51] 1. A general analytical framework has been developed to describe regional water balance including the most basic aspects of the spatial and temporal variabilities of soil saturation and precipitation. In this framework, simple local instantaneous descriptions of infiltration, surface runoff, evapotranspiration, and recharge are integrated over the probability density functions of precipitation and soil saturation to determine the space-time average fluxes. The variability of soil saturation and precipitation play key roles in determining the dependence of the space-time average water balance components on space-time average soil saturation.

[52] 2. The spatial distribution of soil saturation in Illinois was found to be consistent with an Erlang distribution. The distribution's shape parameter, k, which is the squared inverse of the coefficient of variation, was found to be relatively constant near $k = 11$. This result indicates that the soil saturation has more variability in space during wetter periods than in dry periods.

[53] 3. The temporal distribution of spatially averaged soil saturation in Illinois is consistent with a beta distribution, and the standard deviation remains approximately

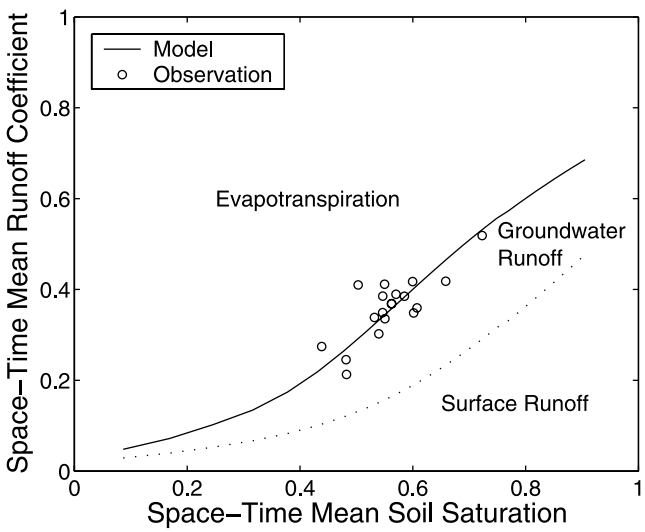

Figure 14. The observed and modeled relationships between space-time average runoff and space-time average soil saturation. The distance below the dashed line indicates the model's surface runoff. The distance between the two lines indicates the model's groundwater runoff. The distance above the solid line indicates the evapotranspiration. The model results were generated using $k = 11$, $\sigma_{\bar{s}} = 0.16$, $E_p = 958$ mm, $K_h = 2.9 \times 10^{-5}$ cm/s, $i = 3.2 \times 10^{-5}$ cm/s, $\alpha = 1.0 \times 10^{-4}$ cm/s, $\beta = 0.87$, and $\gamma = 19$.

constant ($\sigma_{\bar{s}} = 0.16$). This result implies that the temporal variability of spatial mean soil saturation within a year does not depend on the annual mean. This behavior is in contrast with that of the spatial distribution.

[54] 4. The modeling framework was applied to the Illinois River Basin and was calibrated to reproduce relationships between the space-time average soil saturation and efficiency of evapotranspiration as well as the space-time average soil saturation and runoff coefficient. The model is capable of describing the general behavior of the data, although some of the variability remains unexplained. The observed scatter may be explained by variations of the potential evapotranspiration in time and variations of soil saturation with depth, which were both neglected for simplicity.

[55] 5. A relationship between space-time average precipitation and runoff was inferred from the calibration and compared with the empirical relationship for the Illinois River Basin. The model is able to reproduce the general trend of the observations, but some scatter remains. Such scatter cannot be reproduced by the model, which always produces the same runoff for a given precipitation. This disagreement might suggest that the parameters vary from year to year (for example, due to changes in the density of vegetation) or that the runoff has some memory of previous precipitation values.

[56] 6. The model is also able to reproduce an observed relation between the space-time mean soil saturation and runoff coefficient for the Illinois River Basin. According to the model, the observed increase in the runoff coefficient with increasing soil saturation occurs primarily due to a higher efficiency of surface runoff production.

[57] Two major avenues are available for future research. The first avenue seeks to improve the model's representation of the physical processes and stochastic variability. Including appropriate spatial and temporal variability in the potential evapotranspiration and hydraulic conductivity would be beneficial, although data limitations are expected to make such improvements difficult. In addition, characterization of the spatial and temporal correlations of precipitation and soil saturation may provide important modifications to the relationships derived in this paper. A second avenue for future research is the application of this model to understand more about the water balance of the Illinois River Basin. For example, the model can be applied to the long term or climatological water balance of the Illinois River Basin. It can also be used to examine the sensitivity of the water balance to precipitation, soil saturation variability, and other characteristics.

[58] **Acknowledgments.** We gratefully thank two anonymous reviewers for their thoughtful comments, which substantially improved this manuscript. We also appreciate the financial support provided by the Army Research Office.

References

Bain, L. J., and M. Engelhardt (1992), *Introduction to Probability and Mathematical Statistics*, PWS-KENT, Boston, Mass.

Campbell, G. S. (1974), A simple method for determining unsaturated conductivity from moisture retention data, *Soil Sci.*, *117*(6), 311–314.

Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resour. Res.*, *14*(4), 601–604.

Dunne, T., and R. D. Black (1970), Partial area contributions in a small New England watershed, *Water Resour. Res.*, *6*(5), 1296–1311.

Eagleson, P. S. (1978a), Climate, soil, and vegetation: 1. Introduction to water balance dynamics, *Water Resour. Res.*, *14*(5), 705–712.

Eagleson, P. S. (1978b), Climate, soil, and vegetation: 2. The distribution of annual precipitation derived from observed storm sequences, *Water Resour. Res.*, *14*(5), 713–721.

Eagleson, P. S. (1978c), Climate, soil, and vegetation: 3. A simplified model of soil moisture movement in the liquid phase, *Water Resour. Res.*, *14*(5), 722–730.

Eagleson, P. S. (1978d), Climate, soil, and vegetation: 4. The expected value of annual evapotranspiration, *Water Resour. Res.*, *14*(5), 731–740.

Eagleson, P. S. (1978e), Climate, soil, and vegetation: 5. A derived distribution of storm surface runoff, *Water Resour. Res.*, *14*(5), 741–748.

Eagleson, P. S. (1978f), Climate, soil, and vegetation: 6. Dynamics of the annual water balance, *Water Resour. Res.*, *14*(5), 749–764.

Eagleson, P. S. (1978g), Climate, soil, and vegetation: 7. A derived distribution of annual water yield, *Water Resour. Res.*, *14*(5), 765–776.

Eltahir, E. A. B., and R. L. Bras (1993), Estimation of fractional coverage of rainfall in climate models, *J. Clim.*, *6*(4), 639–644.

Eltahir, E. A. B., and P. J. Yeh (1999), On the asymmetric response of aquifer water level to floods and droughts in Illinois, *Water Resour. Res.*, *35*(4), 1199–1217.

Entekhabi, D., and P. S. Eagleson (1989), Land surface hydrology parameterization for atmospheric general circulation models, *J. Clim.*, *2*, 816–831.

Entekhabi, D., and P. S. Eagleson (1991), Climate and the equilibrium state of land surface hydrology parameterizations, *Surv. Geophys.*, *12*(1–3), 205–220.

Farnsworth, R. K., and E. S. Thompson (1982), Mean monthly, seasonal, and annual pan evaporation for the United States, *Tech. Rep. 34*, Natl. Weather Serv., Natl. Oceanic and Atmos. Admin., Silver Spring, Md.

Feddes, R. A., P. J. Kowalik, K. K. Malinka, and H. Zaradny (1976), Simulation of field water uptake by plants using a soil water dependent root extraction function, *J. Hydrol.*, *31*(1–2), 13–26.

Findell, K. L., and E. A. B. Eltahir (1997), An analysis of the soil moisture-rainfall feedback, based on direct observations from Illinois, *Water Resour. Res.*, *33*(4), 725–735.

Food and Agricultural Organization (1975), Soil map of the world, scale 1:5,000,000, UNESCO Publ., Paris.

Freeze, R. A. (1974), Streamflow generation, *Rev. Geophys*, *12*(4), 627–647.

Freeze, R. A., and J. A. Cherry (1979), *Groundwater*, Prentice-Hall, Old Tappan, N. J.

Gelhar, L. W. (1993), *Stochastic Subsurface Hydrology*, Prentice-Hall, Old Tappan, N. J.

Gleick, P. H. (1987), The development and testing of a water balance model for climate change impact assessment: Modeling the Sacramento Basin, *Water Resour. Res.*, *23*(6), 1049–1061.

Green, W. H., and G. A. Ampt (1911), Studies on soil physics, part I, The flow of air and water through soils, *J. Agric. Sci.*, *4*(1), 1–24.

Hewlett, J. D. (1961), Watershed management, report, pp. 61–66, Southeastern For. Exp. Stn., Asheville, N. C.

Hewlett, J. D., and R. A. Hibbert (1965), Factors affecting the response of small watersheds to precipitation in humid areas, paper presented at the International Symposium on Forest Hydrology, Pa. State Univ., University Park.

Hollinger, S., and S. Isard (1994), A soil moisture climatology of Illinois, *J. Clim.*, *7*, 822–833.

Holtan, H. N. (1961), A concept for infiltration estimates in watershed engineering, *Rep. ARS 41-51*, Agric. Res. Serv., U.S. Dep. of Agric., Washington, D. C.

Holtan, H. N. (1965), A model for computing watershed retention from soil parameters, *J. Soil Water Conserv.*, *20*(3), 91–94.

Horton, R. E. (1935), Surface runoff phenomena: Part I, Analysis of the hydrograph, *Horton Hydrol. Lab. Publ. 101*, Edwards Bros., Ann Arbor, Mich.

Lowry, W. P. (1959), The falling rate phase of evaporative soil moisture loss: A critical evaluation, *Bull. Am. Meteorol. Soc.*, *40*, 605.

Molz, F. J. (1981), Models of water transport in the soil-plant system: A review, *Water Resour. Res.*, *17*(5), 1245–1260.

Moore, R. J. (1985), The probability-distributed principle and runoff production at point and basin scales, *Hydrol. Sci. J.*, *30*(2), 273–297.

Owe, M., E. B. Jones, and T. J. Schmugge (1982), Soil moisture variation patterns observed in Hand County, South Dakota, *Water Resour. Bull.*, *18*(6), 949–954.

Owe, M., A. A. Van de Griend, and A. T. C. Chang (1992), Surface moisture and satellite microwave observations in semiarid southern Africa, *Water Resour. Res.*, *28*(3), 829–839.

Philip, J. R. (1957), The theory of infiltration, *Soil Sci.*, *83*, 345–357.

Rodriguez-Iturbe, I., D. Entekhabi, and R. L. Bras (1990), Non-linear dynamics of soil moisture at climate scales: Stochastic analysis, *Cent. Global Change Sci. Rep. 5*, Mass. Inst. of Technol., Cambridge.

Sharma, M. L., and R. J. Luxmoore (1979), Soil spatial variability and its consequences on simulated water balance, *Water Resour. Res.*, *15*(6), 1567–1573.

Sharma, M. L., G. A. Gander, and C. G. Hunt (1980), Spatial variability of infiltration in a watershed, *J. Hydrol.*, *45*, 101–122.

Shuttleworth, W. J. (1993), Evaporation, in *Handbook of Hydrology*, edited by D. R. Maidment, McGraw-Hill, New York.

Sud, Y. C., and M. J. Fennessey (1982), An observational-data based evapotranspiration function for general circulation models, *Atmos. Ocean*, *20*(4), 301–316.

Tadikamalla, P. R. (1990), Kolmogorov-Smirnov type test-statistics for the gamma, Erlang-2 and the inverse Gaussian distributions when the parameters are unknown, *Commun. Stat. Simul. Comput.*, *19*(1), 305–314.

Thornthwaite, C. W., and J. R. Mather (1955), The water balance, *Publ. Climatol.*, *8*(1), 104 pp.

Troendle, C. A. (1983), The potential for water yield augmentation from forest management in the Rocky Mountain region, *Water Resour. Bull.*, *19*(3), 359–373.

Yeh, P. J.-F. (2002), The representation of ground water table dynamics in land-surface hydrology parameterization schemes, Ph.D. thesis, Dep. of Civ. and Environ. Eng., Mass. Inst. of Technol.

Yeh, P. J., M. Irizarry, and E. A. B. Eltahir (1998), Hydroclimatology of Illinois: A comparison of monthly evaporation estimates based on atmospheric water balance and soil water balance, *J. Geophys. Res.*, *103*(D16), 19,823–19,837.

E. A. B. Eltahir, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

J. D. Niemann, Department of Civil Engineering, Campus Delivery 1372, Colorado State University, Fort Collins, CO 80523, USA. (jniemann@engr.colostate.edu)