Annual hydroclimatology of the United States

A. Sankarasubramanian

International Research Institute for Climate Prediction, Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, USA

Richard M. Vogel

WaterSHED Center, Department of Civil and Environmental Engineering, Tufts University, Medford, Massachusetts, USA

Received 13 April 2001; revised 21 January 2002; accepted 21 January 2002; published 25 June 2002.

[1] An overview of the annual hydroclimatology of the United States is provided. Time series of monthly streamflow, temperature, and precipitation are developed for 1337 watersheds in the United States. This unique data set is then used to evaluate several approaches for estimating the long-term water balance and the interannual variability of streamflow. Traditional relationships which predict either actual evapotranspiration or the interannual variability of streamflow from an aridity index $\phi = \overline{\mathrm{PE}}/\overline{P}$ are shown to perform poorly for basins with low soil moisture storage capacity. A water balance model is used to formulate new relationships for predicting actual evapotranspiration and the interannual variability of streamflow. These relationships depend on both the aridity index $\phi = \overline{\mathrm{PE}}/\overline{P}$ and a new soil moisture storage index. A physically based approach for estimating the soil moisture storage index is introduced which requires monthly time series of precipitation, potential evapotranspiration, and an estimate of maximum soil moisture holding capacity. The net results are improved expressions for the long-term water balance and the interannual variability of streamflow which do not require either calibration or streamflow data. *INDEX TERMS:* 1833 Hydrology: Hydroclimatology; 1836 Hydrology: Hydrologic budget (1655); 1818 Hydrology: Evapotranspiration; 1860 Hydrology: Runoff and streamflow; 1878 Hydrology: Water/energy interactions; *KEYWORDS:* hydroclimatology, water balance, aridity, soil moisture, humidity, evapotranspiration

1. Introduction

[2] Hydroclimatologic models describe the interactions between land surface and atmospheric processes at different spatial and temporal scales. The simplest hydroclimatologic model which is valid across all spatial and temporal scales is the lumped form of the continuity equation applied to a watershed:

$$\frac{dS}{dt} = P - E - Q - G\ , \tag{1}$$

where P, E, and Q are the average depth of precipitation, actual evapotranspiration, and runoff respectively, G is the net amount of groundwater that leaves aquifer storage, and dS represents the change in storage in the basin over the time interval dt. Equation (1) is commonly used as the basis for describing the annual hydroclimatology of a region [*Schreiber*, 1904; *Ol'dekop*, 1911; *Budyko*, 1974; *Milly*, 1994a, 1994b; *Wolock and McCabe*, 1999; *Zhang et al.*, 2001].

[3] A common approach for summarizing the long-term hydroclimatology of a region is to plot the ratio of mean annual evapotranspiration to precipitation $\overline{E}/\overline{P}$ versus the ratio of mean annual potential evapotranspiration to precipitation $\overline{\mathrm{PE}}/\overline{P}$. The indices $\overline{E}/\overline{P}$ and $\overline{\mathrm{PE}}/\overline{P}$ are termed the evapotranspiration and aridity ratios, respectively. *Budyko* [1974] hypothesized a generalized functional relationship between the evapotranspiration ratio and the aridity index:

$$\frac{\overline{E}}{\overline{P}} = \psi\left(\frac{\overline{\mathrm{PE}}}{\overline{\Pi}}\right). \tag{2}$$

Figure 1 summarizes relationships of the form given in (2) introduced by *Schreiber* [1904], *Ol'dekop* [1911], *Budyko* [1974], and *Pike* [1964]. Asymptotes A and B in Figure 1 correspond to the upper limits on the evapotranspiration ratio or the lower limits on runoff. For regions with unlimited moisture supply the actual evapotranspiration approaches potential evapotranspiration, which is illustrated by line B in Figure 1. Similarly, when potential evapotranspiration exceeds precipitation, then actual evapotranspiration approaches precipitation (line A in Figure 1). Taking expectations in equation (1) and assuming negligible net changes in basin storage over the long-term leads to

$$\overline{Q} = \overline{P} - \overline{E}, \tag{3}$$

where $\overline{Q}$, $\overline{P}$, and $\overline{E}$ represent the long-term mean annual runoff, precipitation, and actual evapotranspiration, respectively. The primary challenge associated with estimating relationships of the type given in equations (1)–(3) is in obtaining precise estimates of actual evapotranspiration $\overline{E}$. Accurate estimates of actual evapotranspiration $\overline{E}$ may be obtained from equation (3) only when reliable estimates of $\overline{P}$ and $\overline{Q}$ are available from observed records or from maps or digital grids of precipitation and runoff. Relationships of the type shown in Figure 1 are quite useful because they provide

0043-1397/02/2001WR000619$09.00

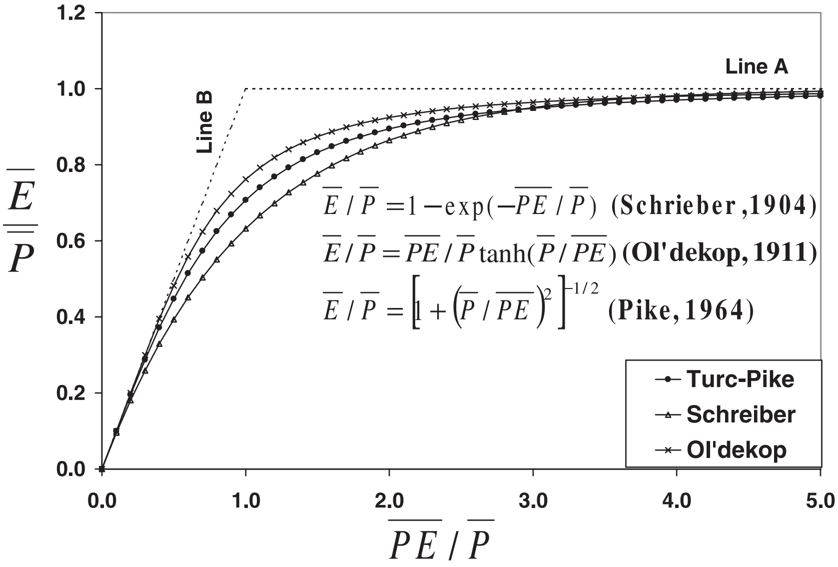

Figure 1. Empirical relationships between evapotranspiration ratio $\overline{E}/\overline{P}$ and aridity index $\phi = \overline{PE}/\overline{P}$.

approximations to the actual evapotranspiration from measurements of rainfall and potential evapotranspiration, thereby avoiding the need for streamflow measurements. One goal of this study is to evaluate relationships of the form illustrated in Figure 1 using current hydroclimatologic data sets available for the continental United States. Another goal is to develop a physical model for estimating the mean annual water balance and the interannual variability of streamflow, which does not require either calibration and/or streamflow data. Our approach involves deriving the mean and variance of annual streamflow from an existing physical watershed model. We exploit the "abcd" water balance model [*Thomas*, 1981; *Alley*, 1984; *Fernandez et al.*, 2000] to derive new relationships which enable prediction of the evapotranspiration ratio $\overline{E}/\overline{P}$ and the runoff variability ratio σ_Q/σ_P as a function of the aridity index and a new soil moisture storage index. The performances of the derived relationships are evaluated using a unique database of monthly streamflow, temperature, and precipitation records for 1337 watersheds within the continental United States.

[4] Section 2 describes the national databases of temperature, precipitation, and streamflow. Section 3 evaluates empirical relationships of the form given in equation (2) for predicting actual evapotranspiration from estimates of precipitation and potential evapotranspiration. Section 4 documents the ability of the abcd water balance model to model the annual hydrology of the continental United States. Sections 5 and 6 evaluate the performance of the new relationships derived here for estimating the mean annual water balance and the interannual variability of streamflow, respectively. Section 7 introduces and evaluates a new approach for describing the long-term water balance and interannual variability of streamflow, which uses the aridity index and a soil moisture index but does not require calibration or streamflow data.

2. National Databases of Temperature, Precipitation, and Streamflow

[5] During the 1990s, advances in computer technology combined with the introduction of many new national digital data sets have led to improvements in our ability to describe temporal and spatial variations in hydrology and climate. The following sections describe a unique new national climate and streamflow data set which was employed in this study.

2.1. Streamflow Database

[6] The national streamflow database, termed the hydroclimatologic data network (HCDN), was developed by *Slack et al.* [1993]. This data set, available on CD-ROM from the U.S. Geological Survey, consists of records of average daily streamflow at 1553 sites located throughout the United States. The data meet certain measurement accuracy criterion outlined by *Slack et al.* [1993]. The HCDN contains river flows from 1874 to 1988, with an average record length of 44 years. The streamflow data included in the HCDN is purported to be relatively free from anthropogenic influences and the accuracy ratings of these records are at least "good" as per USGS standards. We used a subset of the HCDN streamflow data from the 1337 watersheds shown in Figure 2 with at least 10 years of record length.

2.2. Temperature and Precipitation Database

[7] Thirty-seven year time series of monthly precipitation and average minimum and average maximum daily temperature for the continental United States were obtained for the 1337 HCDN watersheds using 0.5° time series grids based on the precipitation-elevation regressions on independent slopes model (PRISM) climate analysis system [*Daly et al.*, 1994]. PRISM uses a precipitation-elevation regression relationship to distribute point measurements to evenly spaced grid cells. PRISM is considered an improvement over other spatial interpolation methods such as inverse distance weighting or kriging because it attempts to account for orographic effects by using precipitation-elevation regression functions. PRISM also employs adiabatic lapse rate corrections in its temperature interpolations.

[8] The monthly climate time series grids were spatially averaged over each HCDN basin using a geographic information system (GIS) and a digital elevation map (DEM) of the United States. A DEM of the United States was used to

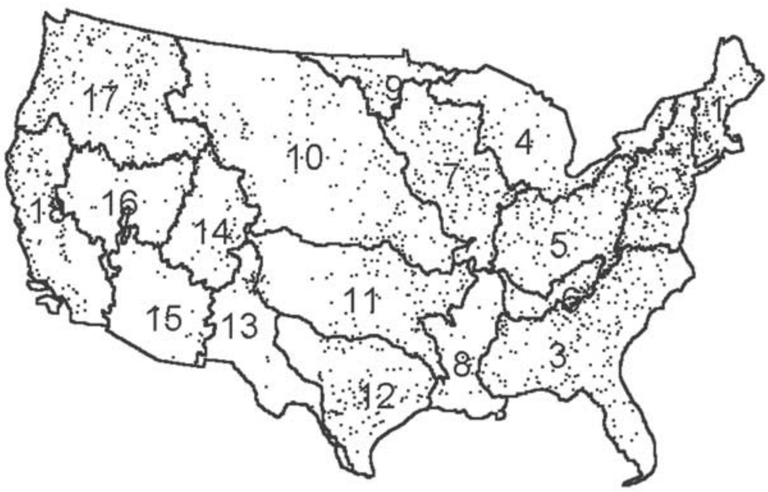

Figure 2. Location of stream gauges [from *Slack et al.*, 1993] within each of the 18 major U.S. water resource regions.

delineate the watershed boundaries for each of the HCDN river sites. A relatively coarse DEM (1 km resolution raster grid) was employed in this analysis owing to the computational challenge of delineating 1337 watersheds. The end result is a unique national time series data set of monthly precipitation and temperature measurements over the period 1951–1988 corresponding to each of the 1337 watersheds. Using the monthly time series of average minimum and average maximum temperature data along with extraterrestrial solar radiation, estimates of monthly potential evapotranspiration were obtained using a method introduced by *Hargreaves and Samani* [1982]. Extraterrestrial solar radiation was estimated for each HCDN basin by computing the solar radiation over 0.1° grids using the method introduced by *Duffie and Beckman* [1980] and then summing those estimates over the entire basin. The Hargreaves method was the highest ranked temperature-based method for computing potential evapotranspiration reported in American Society of Civil Engineers (ASCE) Manual 70 analysis [*Jensen et al.*, 1990].

3. Estimation of Actual Evapotranspiration

[9] Figure 3 illustrates observations of the evapotranspiration ratio $\overline{E}/\overline{P}$ versus the aridity index $\overline{\mathrm{PE}}/\overline{P}$ for the 1337 watersheds illustrated in Figure 2. Actual evapotranspiration is estimated using $\overline{E} = \overline{P} - \overline{Q}$, which is assumed to be the best available estimate. Almost all observations in Figure 3 fall below the asymptotes A and B with the exception of few basins falling outside asymptote B. This is probably due to the overall long-term sensible heat flux being negative in these cases. Figure 3 also compares observations of the

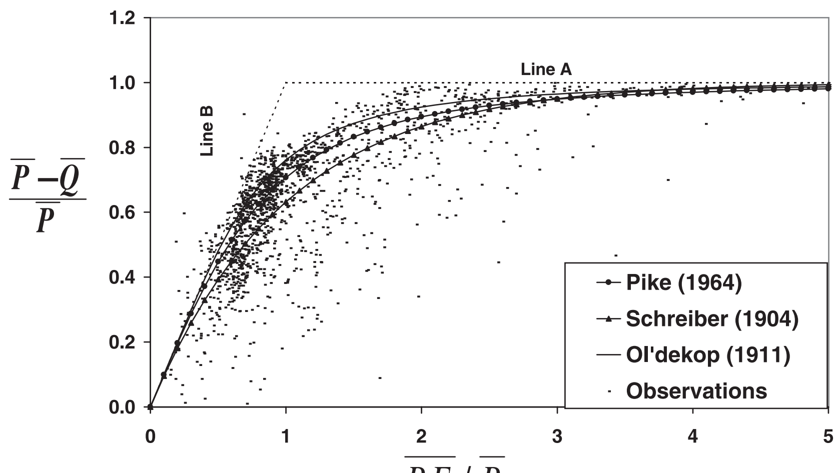

Figure 3. Comparison of observations of evapotranspiration ratio with estimates based on empirical Budyko-type relationships at 1337 hydroclimatologic data network (HCDN) watersheds.

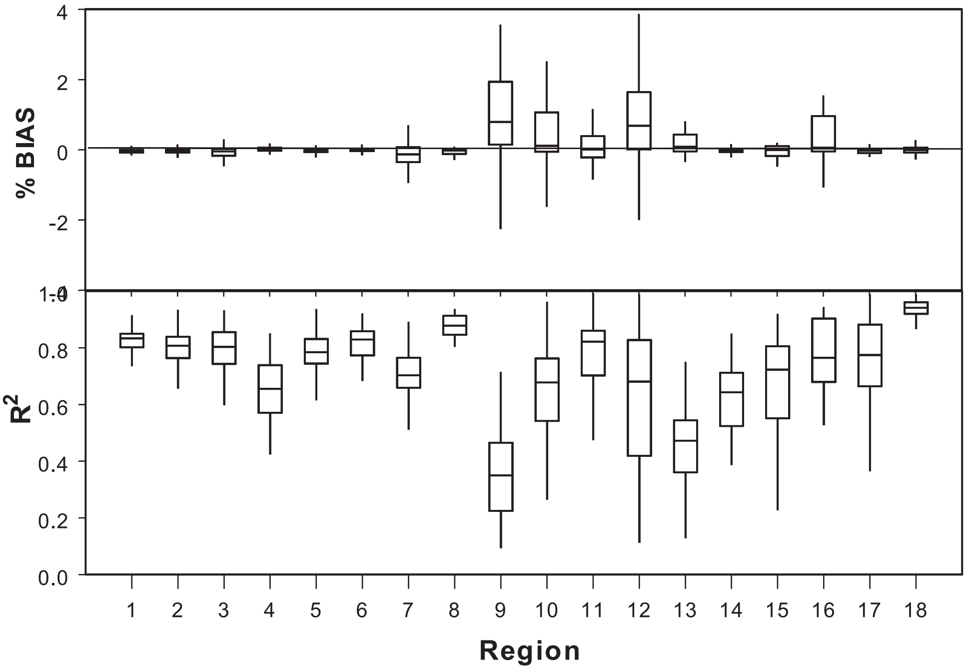

Figure 4. Regional box plots of goodness of fit criteria % BIAS and R^2 corresponding to the calibration of abcd model to 1337 HCDN watersheds.

evapotranspiration ratio with the empirical relationships summarized earlier in Figure 1. Many of the observations shown in Figure 3 depart significantly from any of the empirical relationships shown; hence the aridity index alone is not a good predictor of the evapotranspiration ratio. The results illustrated in Figure 3 imply that additional factors are necessary to explain the long-term water balance. This motivated us to derive alternative relations with the goal of improving our ability to predict the evapotranspiration ratio. Our results in Figure 3 are similar to the results reported recently by *Zhang et al.* [2001], who document that differences in watershed vegetation regimes can lead to departures from the empirical relationships between evapotranspiration ratio and the aridity index shown in Figures 1 and 3.

[10] Efforts to improve upon the empirical relationships reported in Figures 1 have exploited the structure of watershed models. For example, *Dooge* [1992] employed a monthly water balance model with seasonal parameters to derive the evapotranspiration ratio $\overline{E}/\overline{P}$ as a function of the aridity index to explain the sensitivity of runoff to climate change. *Milly* [1994a, 1994b] derived a relationship for the evapotranspiration ratio based on an uncalibrated lumped watershed model and compared his results to observed annual runoff at 0.5° spatial resolution for the eastern portion of the United Sates. *Milly* [1994b] shows that the evapotranspiration ratio is a function of a number of variables in addition to the aridity index, including the soil moisture holding capacity, the number of precipitation events per year, and seasonality parameters. *Wolock and McCabe* [1999] describe the spatial variability of average annual runoff for 344 climate divisions in the conterminous United States using a simple water balance model. Accounting for soil moisture storage capacity and seasonality parameters led to improvements in their ability to estimate mean annual runoff particularly for watersheds which exhibit seasonal moisture and energy changes which are out of phase with each other. Other investigators have modeled average annual watershed runoff using runoff maps or statistical regression techniques (see *Vogel et al.* [1999] for a review). For example, *Vogel et al.* [1999] developed remarkably precise regional regression models which relate the mean and variance of annual streamflow to geomorphic and climatic characteristics for all regions of the continental United States.

4. Calibration of a Water Balance Model

[11] The idea here is to employ a watershed model structure to derive steady state values of various hydrologic processes. For this purpose, we employ the abcd water balance model, introduced by *Thomas* [1981] and later tested and recommended by *Alley* [1984] and *Fernandez et al.* [2000]. *Fernandez et al.* [2000] review previous applications of the abcd model. We begin by testing the ability of the abcd model to characterize the annual hydrology of the continental United States. Appendix A describes the structure of the abcd model and the physical significance of the parameters *a*, *b*, *c*, and *d*. We calibrate the abcd model using annual time series of precipitation, potential evapotranspiration, and streamflow at each of the 1337 basins using the shuffled complex evolution (SCE) global optimization algorithm. The SCE algorithm is a general-purpose global optimization algorithm designed to handle the multilevel or nested local optimal solutions typically encountered in the calibration of nonlinear simulation models. The algorithm combines a randomized search strategy with "simplex"-like searches. It has proven to be effective for calibration of hydrologic watershed models by many investigators [*Duan et al.*, 1992].

[12] We illustrate the goodness of fit of the annual abcd model calibrations in Figure 4 using the two criteria: percentage bias in estimation of mean annual streamflow, % BIAS, and the coefficient of determination, R^2, between the observed and calibrated annual flows. The % BIAS is

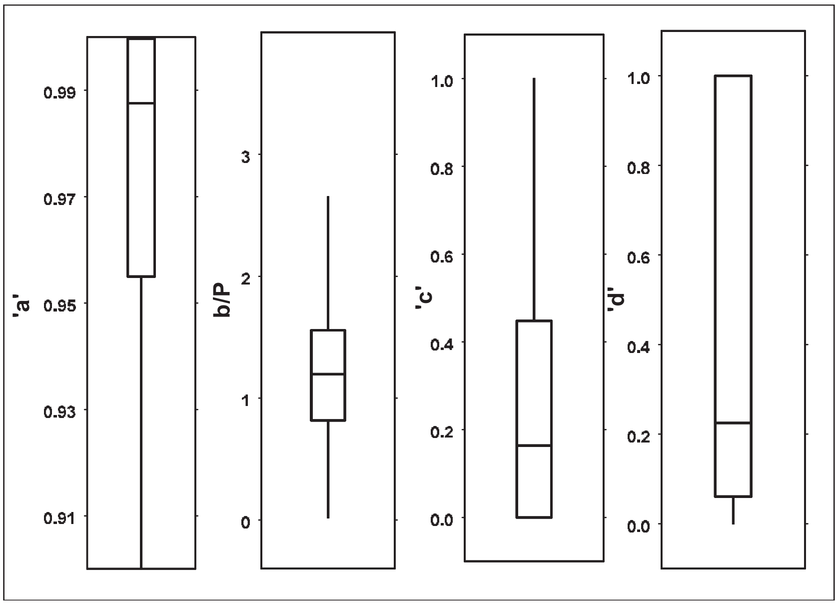

Figure 5. Box plots of calibrated values of annual abcd watershed model parameters.

almost always <2%, and the median value of % BIAS is approximately zero everywhere except regions 9 and 12. Figures 2 and 4 together illustrate that the goodness of fit of the abcd model tends to be best (highest) in the more humid eastern and northwestern regions of the United States. Even though the overall goodness of fit is marginal, our primary concern is with the ability of the abcd model to reproduce the mean annual runoff. Since % BIAS is generally low, we consider the abcd model to be acceptable for computing the mean annual streamflow at over 85% of the basins where $R^2 > 0.6$. Note that *Alley* [1984] and *Fernandez et al.* [2000] have obtained much more accurate calibrations using the abcd model with a monthly time step.

[13] Figure 5 illustrates box plots of the distribution of the calibrated values of the abcd model parameters. Interestingly, calibrated values of model parameter a were nearly always in excess of 0.90 with 50% of the values of $a > 0.99$.

5. Long-Term Water Balance

[14] The goal of this section is to use the abcd model to derive expressions for the evapotranspiration ratio which lead to better agreement with observations than the empirical Budyko-type curves shown earlier in Figure 3. Since the abcd model contains a soil moisture accounting component, it enables us to incorporate the impact of soil moisture variations on the long-term water balance. Appendix B uses the abcd model to derive the following expression for the evapotranspiration ratio $\overline{E}/\overline{P}$ as a function of the aridity index $\phi = \overline{\mathrm{PE}}/\overline{P}$ and a new soil moisture storage index $\gamma = b/\overline{P}$:

$$\frac{\overline{E}}{\overline{P}} = \frac{1}{2}\left\{1 + \gamma(1-R) - \left[1 - 2\gamma(1-R) + \gamma^2\left(1 - 2R + R^2\right)\right]^{0.5}\right\}, \tag{4}$$

where $R = \exp(-\phi/\gamma)$. Figure 6 compares estimates of the evapotranspiration ratio $\overline{E}/\overline{P}$ based on observations $\overline{E} = \overline{P} - \overline{Q}$, with both equation (4) and the *Pike* [1964] model illustrated earlier in Figures 1 and 3. Pearson's correlation coefficient between the observed evapotranspiration ratios and those computed using equation (4) is 0.95, whereas the correlation coefficient between the observed evapotranspiration ratios and those computed using the Pike relationship is only 0.73.

[15] Figure 7 illustrates the ability of our new formulation in equation (4) to capture the impact of the soil moisture storage (using $\gamma = b/\overline{P}$) on the evapotranspiration ratio $\overline{E}/\overline{P}$. As expected, for basins in the same hydroclimatologic regime (same aridity index $\overline{E}/\overline{P}$), those basins with larger values of soil moisture storage correspond to the higher evapotranspiration ratios. As the soil moisture storage index $\gamma = b/\overline{P}$ approaches zero, the evapotranspiration ratio drops owing to rapid runoff. As the soil moisture storage term $\gamma = b/\overline{P}$ increases, equation (4) approaches the Budyko-type curves shown in Figure 1. Figure 7 demonstrates that the inability of the *Pike* [1964] and other Budyko-type relationships to predict variations in $\overline{E}/\overline{P}$ is primarily due to the fact that they ignore the impact of land surface characteristics on the evapotranspiration process. The challenge associated with use of equation (4) relates to estimation of γ. In section 7 we document an estimator of γ based only on observations of precipitation and potential evapotranspiration, thereby avoiding the need for model calibration.

6. Interannual Variability of Streamflow

[16] Unlike the mean annual water balance, little research has addressed the interannual variability of streamflow. Here, as in the work of *Koster and Suarez* [1999], we define interannual variability of streamflow using the runoff variability ratio σ_Q/σ_P, or the ratio of the standard deviation of annual runoff to the standard deviation of annual precipitation. *Koster and Suarez* [1999] showed that the runoff variability ratio can be related to the evapotranspiration variability ratio

$$\frac{\sigma_Q}{\sigma_P} = 1 - \frac{\sigma_E}{\sigma_P}, \tag{5}$$

where σ_E/σ_P denotes the standard deviation of actual evapotranspiration σ_E to the standard deviation of precipitation σ_P. *Koster and Suarez* [1999] derived the following

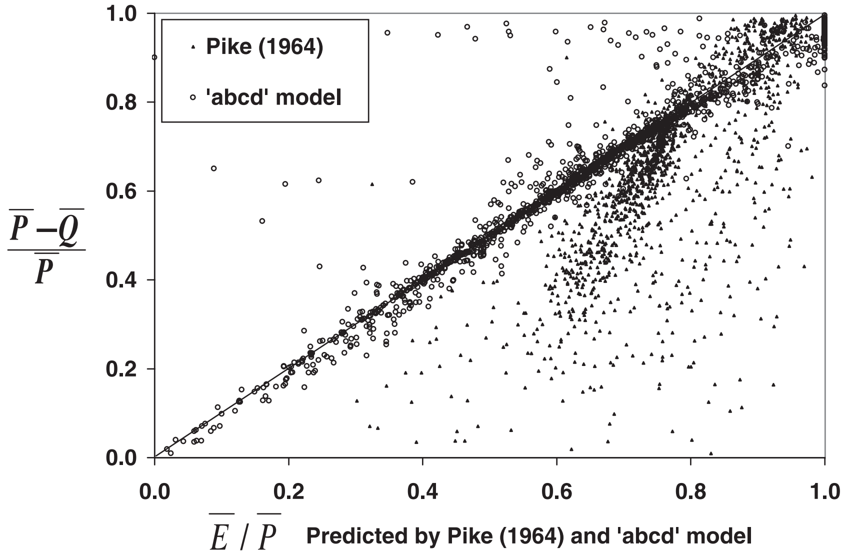

Figure 6. Comparison of evapotranspiration ratios based on equation (4) with estimates based on the *Pike* [1964] relationship at 1337 HCDN watersheds.

relationship for σ_E/σ_P as a function of the aridity index ϕ, assuming negligible interannual variability in potential evapotranspiration (radiative fluxes) and negligible covariance between precipitation and potential evaporation:

$$\frac{\sigma_E}{\sigma_P} = F(\phi) - F'(\phi), \tag{6}$$

where $F(\phi) = [\phi \tanh \phi^{-1} (1 - \cosh \phi + \sinh \phi)]^{0.5}$. *Koster and Suarez* [1999] used general circulation model output to validate equation (6). Figure 8 evaluates the validity of equations (5) and (6) using the 1337 HCDN sites. The correlation coefficient between the observed and predicted values of σ_Q using equations (5) and (6) is 0.91, and the slope of a regression line fitted between observed and predicted values of σ_Q is 0.71, instead of unity. Equations (5) and (6) perform poorly for basins with high streamflow variability because they neglect variations in soil moisture storage. One expects basins with higher soil moisture storage capacity to reduce or buffer the streamflow variability.

[17] Appendix C derives the following improved expression for predicting the interannual variability of streamflow:

$$\frac{\sigma_Q}{\sigma_P} = \left.\frac{\partial Q}{\partial P}\right|_{\phi,\gamma,a}, \tag{7}$$

which is a function of the aridity index ϕ, our new soil index $\gamma = b/\overline{P}$, and the model parameter a. Recalling Figure 5

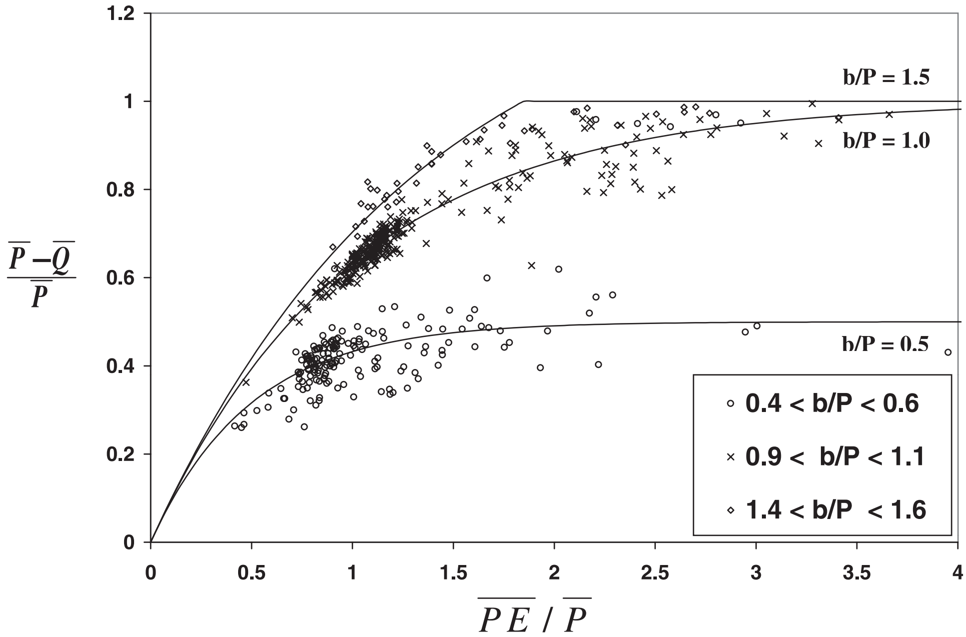

Figure 7. Estimated evapotranspiration ratio based on relationship derived from the abcd model as a function of aridity index $\phi = \overline{PE}/\overline{P}$ and soil moisture index $\gamma = b/\overline{P}$.

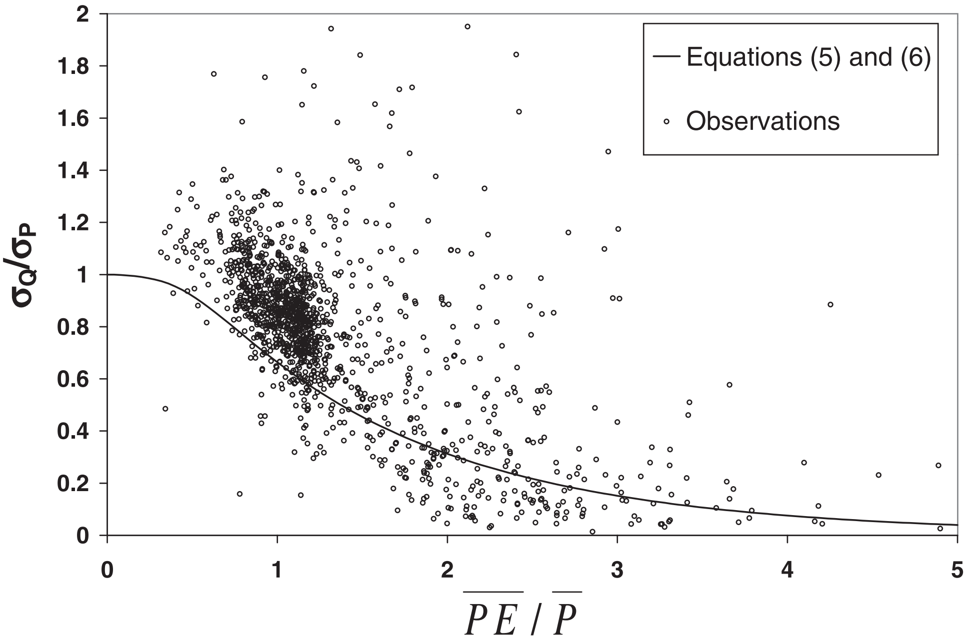

Figure 8. Comparison of interannual variability of streamflow using the *Koster and Suarez* [1999] relationship given in equation (6) with observations at HCDN watersheds.

where we observed most values of model parameter a near unity, hence we assume $a = 1$ here. Figure 9 compares the observed and predicted values of σ_Q based on equation (7). The correlation coefficient between the observed and predicted σ_Q is 0.93, and the slope of the fitted line is now 0.98, a significant improvement over the earlier results based on the work of *Koster and Suarez* [1999] displayed in Figure 8. We found that including the variance of PE and the covariance between P and PE did not improve our ability to predict σ_Q. Figure 10 illustrates how our new approach in equation (7) leads to improvements over *Koster and Suarez* [1999] because of inclusion of the soil moisture index γ in addition to the aridity index ϕ. For comparison, Figure 10 also illustrates the *Koster and Suarez* [1999] relationship.

7. Physical Models of Annual Hydroclimatology

[18] Sections 5 and 6 document how important knowledge of the soil moisture index ϕ is if one wishes to estimate either the mean annual water balance or the interannual variability in streamflow. In this section we develop a

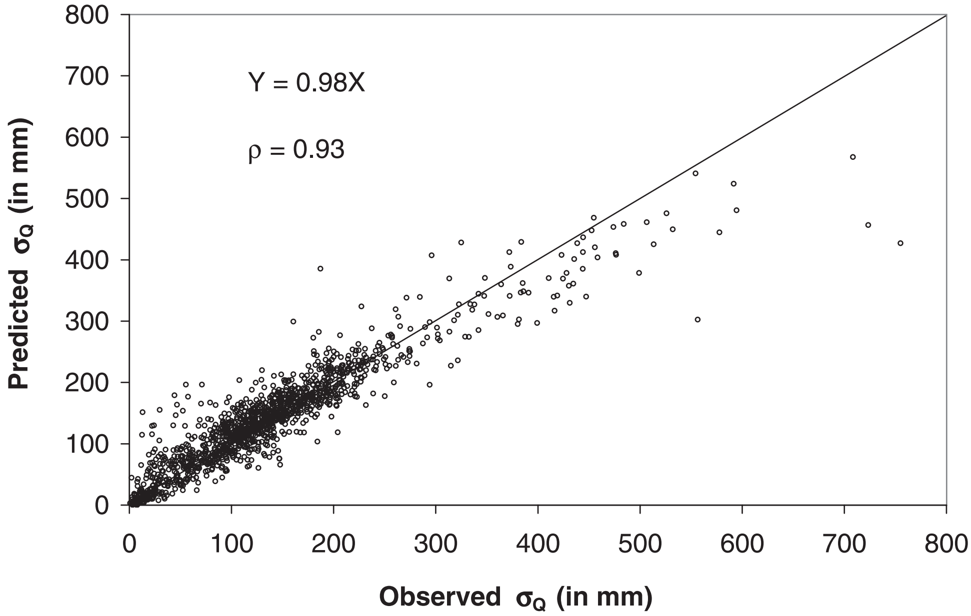

Figure 9. Comparison of interannual variability of streamflow using the relationship in equation (7) derived from the abcd model with observations at HCDN watersheds.

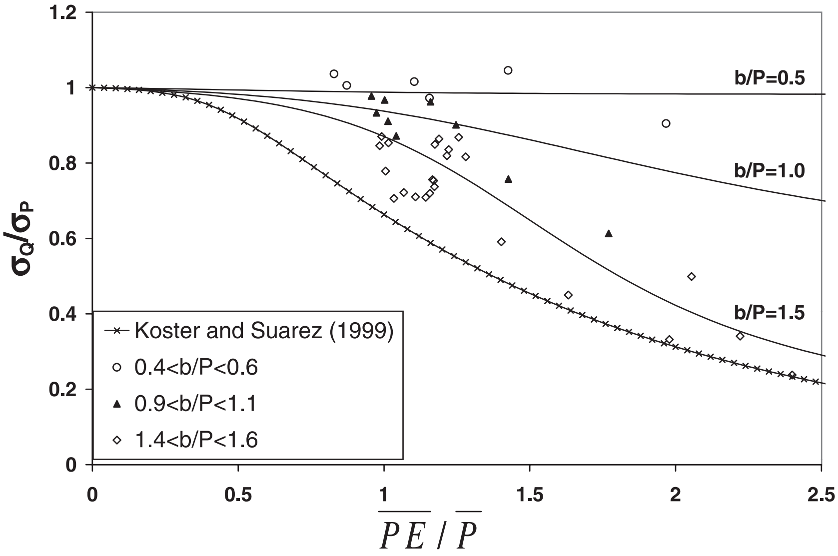

Figure 10. Runoff variability ratio σ_Q/σ_P in equation (7) as a function of aridity index $\phi = \overline{PE}/\overline{P}$ and soil moisture storage index $\gamma = b/\overline{P}$.

physically based approach for estimating the model parameter b from the observations of precipitation, potential evapotranspiration, and maximum soil moisture holding capacity θ of the basin.

[19] Appendix A documents that $b = \max(E_t + S_t)$. We consider the maximum value of b as one possible estimator of b. The maximum value of b results when

$$b_{\max} = \max(E_t) + \max(S_t). \quad (8)$$

The maximum of the state variable S_t is the maximum soil moisture holding capacity θ of the basin. The maximum of the state variable E_t can be obtained from the asymptotes of Budyko's framework:

$$\max(E_t) = \begin{cases} P_t & \text{if } P_t \leq \mathrm{PE}_t \\ \mathrm{PE}_t & \text{otherwise} \end{cases}. \quad (9)$$

The maximum monthly actual evapotranspiration is obtained from equation (9) for each month from the monthly time series of P and PE and aggregated to obtain the maximum annual evapotranspiration. This procedure was repeated for each year. The maximum soil moisture holding capacity for each basin was obtained from 0.5° grid estimates of maximum soil moisture holding capacity given by *Dunne and Wilmott* [1996]. The 0.5° grid of maximum soil moisture holding capacity was spatially averaged over each HCDN basin to obtain the average maximum soil moisture holding capacity for each basin. Using the average maximum soil moisture holding capacity and maximum annual evapotranspiration of the basin obtained from equation (9), we estimated the model parameter b using equation (8).

[20] To evaluate equation (8), we only consider those basins for which the calibration of the abcd model led to $R^2 > 0.85$. This resulted in a total of 458 basins out of the 1337 basins. Figure 11 compares the relationship between the calibrated model parameter b and b estimated using equation (8). Although the correlation $\rho = 0.68$ is relatively weak, equation (8) does enable one to estimate the mean annual water balance and interannual variability of streamflow without calibration of a watershed model. In other words, equations (4) and (7) may be used with equation (8) to describe the annual water balance and interannual variability of streamflow at "ungaged" basin.

7.1. Evaluation of Model for Long-Term Water Balance at Ungaged Sites

[21] Here our concern is how well our approach for estimation of the evapotranspiration ratio using equations (4) and (8) performs at an ungaged site. Figure 12 compares the evapotranspiration ratio estimated from equations (4) and (8) with the observed ratio based on $(\overline{P} - \overline{Q})/\overline{P}$ and the *Pike* [1964] relationship. Figure 12 illustrates that our new approach based on equations (4) and (8) reproduces the long-term water balance better than the *Pike* [1964] relationship. The *Pike* [1964] relationship is biased upward as was demonstrated earlier in Figure 3, particularly for basins with low evapotranspiration ratios. Interestingly, our new approach based on equations (4) and (8) performs best for basins with low evapotranspiration ratios.

7.2. Evaluation of Model for Interannual Variability of Streamflow at Ungaged Sites

[22] Here our concern is how well our approach for estimation of interannual variability of streamflow using equations (7) and (8) performs at an ungaged site. Estimates of b obtained from equation (8) are substituted into equation (7) to obtain an estimate of the variance of streamflow σ_Q for the 458 basins. Figures 13a and 13b compare estimates of σ_Q based on the *Koster and Suarez* [1999] relationship in equation (6) with our estimates based on equations (7) and (8). The slope of the fitted line in Figure 13b is 1.03 based on equations (7) and (8), whereas the slope of the fitted line in Figure 13a based on the *Koster and Suarez* [1999] relationship is only 0.74. It is evident from Figure 13 that the relationship introduced here for estimation of σ_Q is an improvement over the approach

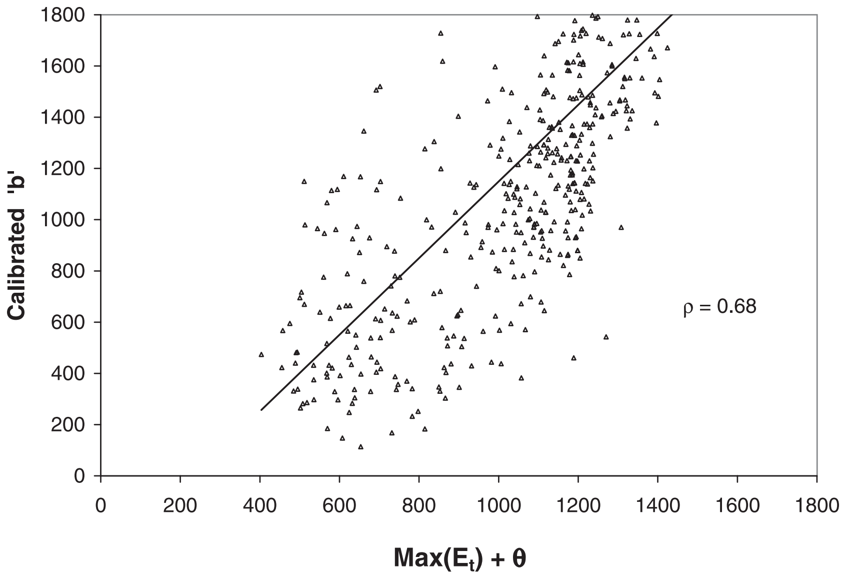

Figure 11. Comparison of calibrated model parameter *b* with estimates of *b* based on equation (8) at an ungaged site.

suggested by *Koster and Suarez* [1999], particularly for basins with high streamflow variability.

8. Conclusions

[23] This study has sought to develop physical models which are representative of the annual hydroclimatology of the continental United States and which do not require either calibration or streamflow data for their application. Our comparisons in Figure 3 revealed that simple Budyko-type relationships of the type introduced by *Schreiber* [1904], *Ol'dekop* [1911], *Budyko* [1974], and *Pike* [1964] are unable to reproduce actual evapotranspiration observations for most regions of the continental United States. We document that those relations perform poorly, particularly for basins with low evapotranspiration ratios $\overline{E}/\overline{P}$, because they do not account for influence of variations in land surface characteristics on actual evapotranspiration. This result is analogous to the recent finding by *Zhang et al.* [2001], who found that in addition to the aridity index, variations in watershed land cover have an important influence on the long-term water balance. In this study, a new estimator of actual evapotranspiration is derived from a water balance model, resulting in equation (4). This expression relates the evapotranspiration ratio $\overline{E}/\overline{P}$ to both an aridity index $\phi = \overline{\mathrm{PE}}/\overline{P}$ and soil moisture storage index $\gamma = b/\overline{P}$, where *b* characterizes the soil moisture storage capacity of the watershed. We documented that incorporation of this new soil moisture soil index led to improvements in our ability to characterize the long-term

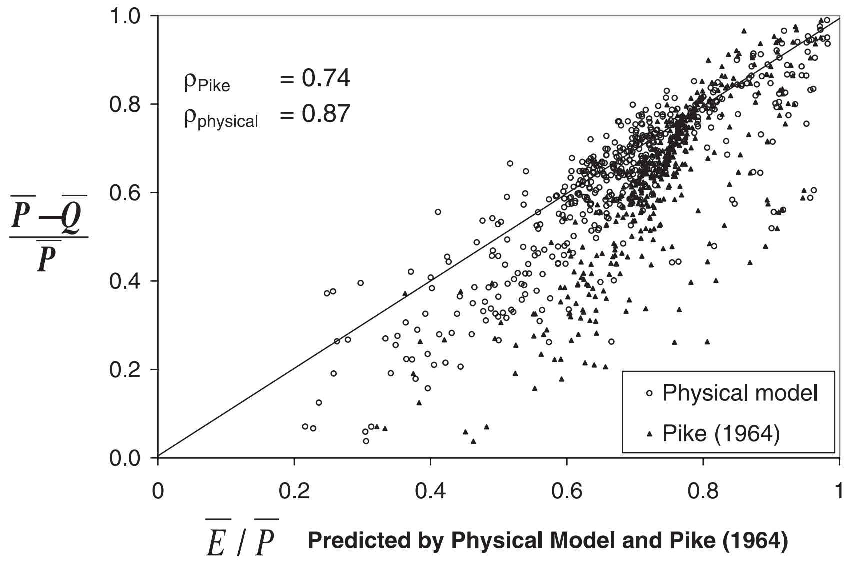

Figure 12. Comparison of actual evapotranspiration computed from equations (4) and (8) and from the *Pike* [1964] model with observations of actual evapotranspiration at HCDN sites.

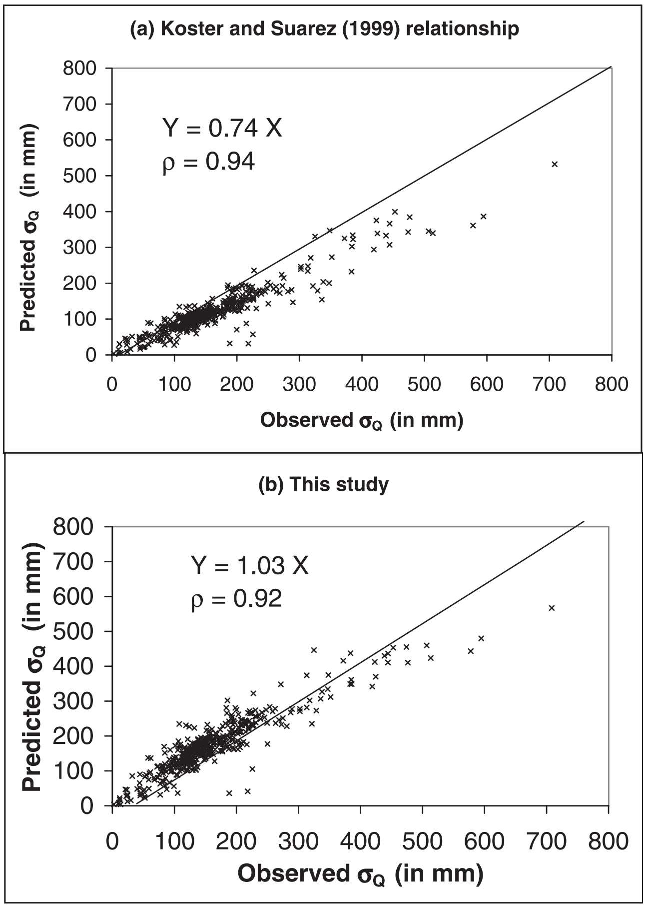

Figure 13. Comparison of observed interannual variability of streamflow values computed from (a) the *Koster and Suarez* [1999] relationship in equation (6) and with (b) physical model developed here based on equations (7) and (8).

water balance even when streamflow observations are unavailable.

[24] Analogous to expressions which describe the long-term water balance as a function of an aridity index, *Koster and Suarez* [1999] introduced a new approach for characterizing the interannual variability of streamflow as a function of the aridity index. Analogous to our results for the long-term water balance, we found that the existing method for describing the interannual variability of streamflow introduced by *Koster and Suarez* [1999] could be improved by integrating our new soil moisture storage capacity index. Using the structure of a water balance model, we derived an expression for the interannual variability of streamflow as a function of aridity index ϕ and soil moisture index γ. The new relationship offers an improvement over the approach introduced by *Koster and Suarez* [1999] even for basins without any streamflow observations.

Appendix A: The "abcd" Model

[25] The abcd model is a nonlinear water balance model which accepts precipitation and potential evaporation as input, producing streamflow as output. The abcd model was originally introduced by *Thomas* [1981] at an annual time step and later recommended by others at a monthly time step [*Alley*, 1984; *Fernandez et al.*, 2000]. The abcd model defines two state variables, W_t, termed "available water," and Y_t, which we term "evapotranspiration opportunity." Available water is defined as

$$W_t = P_t + S_{t-1}, \tag{A1}$$

where P_t is precipitation during period t and S_{t-1} is soil moisture storage at the beginning of period t. Evapotranspiration opportunity is maximum water that can leave the basin as evapotranspiration at any given time t and is defined as

$$Y_t = E_t + S_t, \tag{A2}$$

where E_t represents actual evaporation during period t and S_t represents soil moisture storage at the end of period t. Evapotranspiration opportunity Y_t is postulated as a nonlinear function of available water W_t using

$$Y_t(W_t) = \frac{W_t + b}{2a} - \sqrt{\left(\frac{W_t + b}{2a}\right)^2 - \frac{W_t b}{a}}. \tag{A3}$$

This function simply assures that $Y_t \leq W_t$, $dW(0)/dY = 1$, and $dW(\infty)/dY = 0$. Allocation of available water W_t between E_t and S_t is accomplished by assuming that the rate of loss of soil moisture to evaporation is proportional to the potential evapotranspiration, so that $dS/dt = [-\mathrm{PE}(S/b)]$. Solving this differential equation and assuming $S_{t-1} = Y_t$ leads to

$$S_t = Y_t \exp(-PE_t/b). \tag{A4}$$

The difference between available water and evapotranspiration opportunity $W_t - Y_t$ is the sum of groundwater recharge and direct runoff. The parameter c allocates the quantity $W_t - Y_t$ between groundwater recharge $c(W_t - Y_t)$ and direct runoff $(1 - c)(W_t - Y_t)$. Finally, groundwater discharge to the stream channel is modeled as dG_t, where d is the fourth model parameter and G_t is groundwater storage at the end of period t. Groundwater storage at the end of period t is equal to previous groundwater storage plus groundwater recharge less groundwater outflow so that

$$G_t = G_{t-1} + c(W_t - Y_t) - dG_t. \tag{A5}$$

The streamflow computed by the model for the given parameters a, b, c, and d is

$$Q_t = (1 - c)(W_t - Y_t) + dG_t. \tag{A6}$$

The abcd model has four parameters a, b, c, and d, each having a physical interpretation. The parameter a $(0 \leq a \leq 1)$ reflects the "propensity of runoff to occur before the soil is fully saturated" [*Thomas et al.*, 1981]. The parameter b is an upper limit on the sum of evapotranspiration and soil moisture storage. The parameter c is equal to the fraction of streamflow, which arises from groundwater, so that it is equivalent to the base flow index discussed in hydrology textbooks. The reciprocal of the parameter d is equal to the groundwater residence time, so that d is proportional to the base flow recession constant also discussed in hydrology textbooks.

Appendix B: Derivation of the Ratio of Actual Evapotranspiration to Precipitation for the abcd Model

[26] Given estimates of the model parameters a, b, mean annual precipitation $\overline{P}$, and mean annual potential evapotranspiration $\overline{\mathrm{PE}}$, the ratio of actual evapotranspiration $\overline{E}$ to precipitation can be derived. Under steady state conditions the subscripts of the state variables can be dropped, leading to

$$Y = \frac{W+b}{2a} - \sqrt{\left(\frac{W+b}{2a}\right)^2 - \frac{Wb}{a}}, \tag{B1}$$

where $W = P + S$ with $S = YR$, where $R = \exp(-\mathrm{PE}/b)$. Substitution of W and S into (B1) results in a quadratic equation in Y. Since $Y < b$, the only possible root of the quadratic equation is

$$Y = \frac{1}{2(a-R)}\left\{P + b(1-R) - \left[P^2 + 2Pb(1+R-2a) + b^2\left(1-2R+R^2\right)\right]^{0.5}\right\}. \tag{B2}$$

A zero-order approximation of $\overline{E}$ is $\overline{E} = \overline{Y}(1-\overline{R})$, leading to

$$\overline{E} = \frac{1-\overline{R}}{2\left(a-\overline{R}\right)}\left\{\overline{P} + b\left(1-\overline{R}\right) - \left[\overline{P}^2 + 2\overline{P}b\left(1+\overline{R}-2a\right) + b^2\left(1-2\overline{R}+\overline{R}^2\right)\right]^{0.5}\right\}, \tag{B3}$$

where $\overline{R} = \exp(-\overline{\mathrm{PE}}/b)$. Defining $\phi = \overline{\mathrm{PE}}/\overline{P}$ and $\gamma = b/\overline{P}$, we obtain

$$\frac{\overline{E}}{\overline{P}} = \frac{1-R}{2(a-R)}\left\{1 + \gamma(1-R) - \left[1 + 2\gamma(1+R-2a) + \gamma^2\left(1-2R+R^2\right)\right]^{0.5}\right\}, \tag{B4}$$

where R may also be expressed using $R = \exp(-\phi/\gamma)$. Recall from Figure 5 that the approximation $a \cong 1$ is quite reasonable, leading to the simplified expression given in equation (4).

Appendix C: Derivation of Runoff Variability Ratio σ_Q/σ_P

[27] Since the abcd model is a nonlinear model, it is not possible to derive the variance of streamflow exactly; hence we derive the first-order approximation

$$\sigma_Q^2 = \sigma_P^2\left(\left.\frac{\partial Q}{\partial P}\right|_{\overline{P},\overline{\mathrm{PE}}}\right)^2 + \sigma_{\mathrm{PE}}^2\left(\left.\frac{\partial Q}{\partial PE}\right|_{\overline{P},\overline{\mathrm{PE}}}\right)^2 + 2C_{P\mathrm{PE}}\left.\frac{\partial Q}{\partial P}\right|_{\overline{P},\overline{\mathrm{PE}}}\left.\frac{\partial Q}{\partial \mathrm{PE}}\right|_{\overline{P},\overline{\mathrm{PE}}}, \tag{C1}$$

where the derivatives are evaluated at the mean values $\overline{P}$ and $\overline{\mathrm{PE}}$ and $C_{P\mathrm{PE}}$ is the covariance between P and PE. The above two derivatives can be expressed using the chain rule

$$\frac{\partial Q}{\partial P} = \frac{\left(1 - \frac{\partial Y}{\partial W}\right)}{\left(1 - \frac{\partial S}{\partial Y}\frac{\partial Y}{\partial W}\right)} \tag{C2}$$

$$\frac{\partial Q}{\partial \mathrm{PE}} = \frac{\frac{\partial S}{\partial \mathrm{PE}}\left(1 - \frac{\partial Y}{\partial W}\right)}{\left(1 - \frac{\partial S}{\partial Y}\frac{\partial Y}{\partial W}\right)}. \tag{C3}$$

Expressing the partial derivatives $\partial S/\partial Y$ and $\partial S/\partial \mathrm{PE}$ as a function of $\phi = \overline{\mathrm{PE}}/\overline{P}$ and $\gamma = b/\overline{P}$, we obtain

$$\partial S/\partial Y = \exp\left(-\overline{\mathrm{PE}}/b\right) = R, \tag{C4}$$

$$\partial S/\partial \mathrm{PE} = \frac{\overline{Y}}{b}\exp\left(-\overline{\mathrm{PE}}/b\right) = \frac{\overline{Y}}{\overline{P}}\frac{R}{\gamma}, \tag{C5}$$

$$\frac{\partial Y}{\partial W} = \frac{1}{2a} - \frac{1}{2}\left[\left(\frac{w+b}{2a}\right)^2 - \frac{Wb}{a}\right]^{-0.5}\left[\left(\frac{w+b}{2a^2}\right) - \frac{b}{a}\right]. \tag{C6}$$

Evaluating the above derivative at $\overline{W}$ and after some algebra, equation (C6) can be simplified to

$$\frac{\partial Y}{\partial W} = \frac{1}{2a} - \frac{1}{2}\left[\frac{1}{(2a)^2}\left(1 + \frac{\overline{Y}}{\overline{P}}R + \gamma\right)^2 - \frac{\gamma}{a}\left(1 + \frac{\overline{Y}}{\overline{P}}R\right)\right]^{-0.5} \cdot \left[\frac{1}{(2a)^2}\left(1 + \frac{\overline{Y}}{\overline{P}}R + \gamma\right) - \frac{\gamma}{a}\right]. \tag{C7}$$

From equation (B4), the term $\overline{Y}/\overline{P}$ can be simplified as

$$\frac{\overline{Y}}{\overline{P}} = \frac{1}{2(a-R)}\left\{1 + \gamma(1-R) - \left[1 + 2\gamma(1+R-2a) + \gamma^2\left(1-2R+R^2\right)\right]^{0.5}\right\}, \tag{C8}$$

where $\phi = \overline{\mathrm{PE}}/\overline{P}$, $\gamma = b/\overline{P}$, and $R = \exp(-\phi/\gamma)$.

[28] Substituting equations (C4), (C5), (C7), and (C8) in (C2) and (C3), one can evaluate the derivatives $\partial Q/\partial P$ and $\partial Q/\partial \mathrm{PE}$ as a function of ϕ, γ, and a. Substituting these derivatives in equation (C1), we get the first-order approximation of the variance of streamflow as a function of ϕ, γ, a, σ_P, σ_{PE}, and $C_{P\mathrm{PE}}$. *Koster and Suarez* [1999] showed that it is reasonable to assume σ_{PE} and $C_{P\mathrm{PE}}$ as zero. In this case, equation (C1) reduces to a simplified form given in equation (7).

[29] **Acknowledgments.** Although the research described in this article has been funded in part by the United States Environmental Protection Agency through STAR grant R 824992-01-0 to Tufts University, it has not been subjected to the agency's required peer and policy review and therefore does not necessarily reflect the views of the agency, and no endorsement should be inferred. The authors are indebted to Chris Daly for providing us with the monthly time series grids of temperature and precipitation and to Ian Wilson for his assistance in processing those grids. We are also indebted to the two anonymous reviewers whose comments led to significant improvements in our manuscript.

References

Alley, W. M., On the treatment of evapotranspiration, soil moisture accounting, and aquifer recharge in monthly water balance models, *Water Resour. Res.*, *20*(8), 1137–1149, 1984.

Budyko, M. I., *Climate and Life*, translated from Russian by D. H. Miller, Academic, San Diego, Calif., 1974.

Daly, C., R. P. Neilson, and D. L. Philips, A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteorol.*, *33*(2), 140–158, 1994.

Dooge, J. C. I., Sensitivity of runoff to climate change: A Hortonian approach, *Bull. Am. Meteorol. Soc.*, *73*(12), 2013–2024, 1992.

Duan, Q., S. Sorooshian, and V. Gupta, Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*(4), 1015–1031, 1992.

Duffie, J. A., and W. A. Beckman, *Solar Engineering of Thermal Processes*, 109 pp., John Wiley, New York, 1980.

Dunne, K. A., and C. J. Wilmott, Global distribution of plant extractable water capacity of soil, *Int. J. Climatol.*, *16*, 841–859, 1996.

Fernandez, W., R. M. Vogel, and A. Sankarasubramanian, Regional calibration of a watershed model, *Hydrol. Sci. J.*, *45*(5), 689–707, 2000.

Hargreaves, G. H., and Z. A. Samani, Estimating potential evapotranspiration, *J. Irrig. Drainage Eng.*, *108*(3), 225–230, 1982.

Jensen, M. E., R. D. Burman, and R. G. Allen, Evapotranspiration and irrigation water requirements, *Manuals and Rep. on Eng. Practice*, *70*, 350 pp., Am. Soc. of Civ. Eng., New York, 1990.

Koster, R. D., and M. J. Suarez, A simple framework for examining the interannual variability of land surface moisture fluxes, *J. Clim.*, *12*, 1911–1917, 1999.

Milly, P. C. D., Climate, interseasonal storage of soil water and the annual water balance, *Adv. Water Resour.*, *17*, 19–24, 1994a.

Milly, P. C. D., Climate, soil water storage, and the average annual water balance, *Water Resour. Res.*, *30*(7), 2143–2156, 1994b.

Ol'dekop, E. M., On evaporation from the surface of river basins, *Transactions on Meteorological Observations*, Lur-evskogo, Univ. of Tartu, Tartu, Estonia, 1911.

Pike, J. G., The estimation of annual runoff from meteorological data in a tropical climate, *J. Hydrol.*, *2*, 116–123, 1964.

Schreiber, P., Ueber die Beziehungen zwischen dem Niederschlag und der Wasseruhrung der Flusse in Mitteleuropa, *Meteorol. Z.*, *21*, 441–452, 1904.

Slack, J. R., A. M. Lumb, and J. M. Landwehr, Hydroclimatic data network (HCDN): A U.S. Geological Survey streamflow data set for the United States for the study of climate variation, 1874–1988, *Water Resour. Invest. Rep.*, *93-4076*, 1993.

Thomas, H. A., Improved methods for national water assessment, report, *Contract WR 15249270*, U.S. Water Resour. Council, Washington, D. C., 1981.

Vogel, R. M., I. Wilson, and C. Daly, Regional regression models of annual streamflow for the United States, *J. Irrig. Drainage Eng.*, *125*(3), 148–157, 1999.

Wolock, D. M., and G. M. McCabe, Explaining spatial variability in mean annual runoff in the conterminous United States, *Clim. Res.*, *11*, 149–159, 1999.

Zhang, L., W. R. Dawes, and G. R. Walker, Response of mean annual evapotranspiration to vegetation changes at catchment scale, *Water Resour. Res.*, *37*(3), 701–708, 2001.

A. Sankarasubramanian, International Research Institute for Climate Prediction, Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964-8000, USA. (sankar@iri.columbia.edu)

R. M. Vogel, WaterSHED Center, Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155, USA. (rvogel@tufts.edu)