

# Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward

James W. Kirchner<sup>1,2,3</sup>

Received 10 February 2008; revised 9 October 2008; accepted 22 October 2008; published 25 February 2009.

[1] Water fluxes in catchments are controlled by physical processes and material properties that are complex, heterogeneous, and poorly characterized by direct measurement. As a result, parsimonious theories of catchment hydrology remain elusive. Here I describe how one class of catchments (those in which discharge is determined by the volume of water in storage) can be characterized as simple first-order nonlinear dynamical systems, and I show that the form of their governing equations can be inferred directly from measurements of streamflow fluctuations. I illustrate this approach using data from the headwaters of the Severn and Wye rivers at Plynlimon in mid-Wales. This approach leads to quantitative estimates of catchment dynamic storage, recession time scales, and sensitivity to antecedent moisture, suggesting that it is useful for catchment characterization. It also yields a first-order nonlinear differential equation that can be used to directly simulate the streamflow hydrograph from precipitation and evapotranspiration time series. This single-equation rainfall-runoff model predicts streamflow at Plynlimon as accurately as other models that are much more highly parameterized. It can also be analytically inverted; thus, it can be used to “do hydrology backward,” that is, to infer time series of whole-catchment precipitation directly from fluctuations in streamflow. At Plynlimon, precipitation rates inferred from streamflow fluctuations agree with rain gauge measurements as closely as two rain gauges in each catchment agree with each other. These inferred precipitation rates are not calibrated to precipitation measurements in any way, making them a strong test of the underlying theory. The same approach can be used to estimate whole-catchment evapotranspiration rates during rainless periods. At Plynlimon, evapotranspiration rates inferred from streamflow fluctuations exhibit seasonal and diurnal cycles that agree semiquantitatively with Penman-Monteith estimates. Thus, streamflow hydrographs may be useful for reconstructing precipitation and evapotranspiration records where direct measurements are unavailable, unreliable, or unrepresentative at the scale of the landscape.

**Citation:** Kirchner, J. W. (2009), Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, *Water Resour. Res.*, 45, W02429, doi:10.1029/2008WR006912.

## 1. Introduction

[2] The spatial heterogeneity and process complexity of subsurface flow imply that any feasible hydrological model will necessarily involve substantial simplifications and generalizations. The essential question for hydrologists is which simplifications and generalizations are the right ones. Physically based rainfall-runoff models (see *Beven* [2001] for an overview) attempt to link catchment behavior with measurable properties of the landscape, but many properties controlling subsurface flow are only measurable at scales

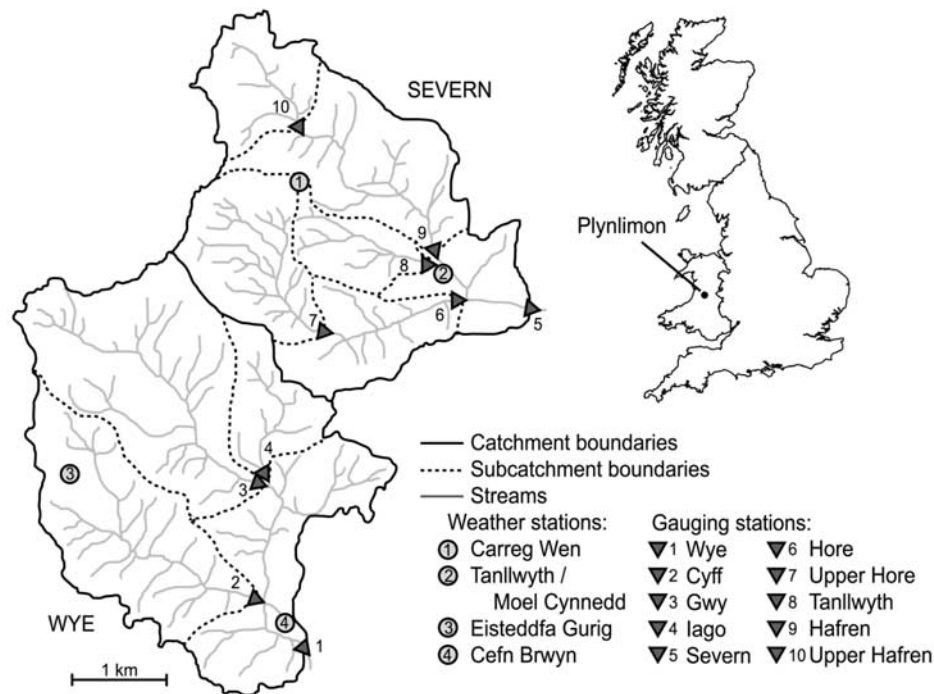
that are many orders of magnitude smaller than the catchment itself. Thus, although it seems obvious that catchment models should be “physically based,” it seems less obvious how those models should be based on physics. Many hydrologic models are based on an implicit premise that the microphysics in the subsurface will “scale up” such that the behavior at larger scales will be described by the same governing equations (e.g., Darcy’s law, Richards’ equation), with “effective” parameters that somehow subsume the heterogeneity of the subsurface [*Beven*, 1989]. It is currently unclear whether this upscaling premise is correct, or whether the effective large-scale governing equations for these heterogeneous systems are different in form, not just different in the parameters, from the equations that describe the small-scale physics [*Kirchner*, 2006].

[3] This observation raises the question of how we can identify the right constitutive equations to describe the macroscopic behavior of these complex heterogeneous systems. For decades, hydrologists have used characteristic

<sup>1</sup>Department of Earth and Planetary Science, University of California, Berkeley, California, USA.

<sup>2</sup>Swiss Federal Institute for Forest, Snow, and Landscape Research WSL, Birmensdorf, Switzerland.

<sup>3</sup>Department of Environmental Sciences, Swiss Federal Institute of Technology, ETH Zurich, Zurich, Switzerland.



**Figure 1.** Location map for the headwater catchments of the Severn and Wye rivers at Plynlimon, Wales ( $52^{\circ}27'N$ ,  $3^{\circ}43'W$ ), showing locations of automatic weather stations (circles) and gauging stations (triangles).

curves to describe the macroscopic behavior of blocks of soil, recognizing that these empirical functions integrate across the complex and heterogeneous processes that govern water movement at the pore scale. Likewise, one can pose the question of whether there are “characteristic curves” at the scale of small catchments, that can usefully integrate over the complexity and heterogeneity of the landscape at all scales below, say, a few square kilometers. And if such “characteristic curves” are meaningful and useful at the scale of small catchments, can they also be measured at that scale?

[4] Since at least the time of Horton [1936, 1937, 1941], a major theme in catchment hydrology has been the interpretation of streamflow variations in terms of the drainage behavior of hillslope or channel storage elements [e.g., Nash, 1957; Laurenson, 1964; Lambert, 1969, 1972; Mein *et al.*, 1974; Brutsaert and Nieber, 1977; Rodriguez-Iturbe and Valdes, 1979; van der Tak and Bras, 1990; Rinaldo *et al.*, 1991], whose parameter values are typically calibrated to the observed hydrograph (see Beven [2001] and Brutsaert [2005] for an overview). In some cases, these parameters can be interpreted as reflecting basin-scale hydraulic properties [e.g., Brutsaert and Nieber, 1977; Brutsaert and Lopez, 1998], and in others they can be correlated with catchment geomorphic characteristics [e.g., Nash, 1959], facilitating hydrologic prediction in ungauged catchments. However, the form of the constitutive relationship (the storage-discharge function) must normally be known in advance.

[5] Here I show that, if the catchment can be represented by a single storage element in which discharge is a function of storage alone, the form of this storage-discharge function can be estimated from analysis of streamflow fluctuations. In

contrast to conventional methods of recession analysis (see reviews by Hall [1968], Tallaksen [1995], and Smakhtin [2001], and references therein), this approach does not specify the functional form of the storage-discharge relationship a priori, instead determining it directly from data. (For further comparisons between previous work and the present approach, see section 15.1 below.) Using this approach, one can construct a first-order nonlinear differential equation linking precipitation, evapotranspiration, and discharge, with no need to account explicitly for changes in storage; these are instead inferred from the resulting changes in discharge. This single equation allows one to predict streamflow hydrographs from precipitation and evapotranspiration time series. It can also be inverted, allowing one to use streamflow fluctuations to infer precipitation and evapotranspiration rates at whole-catchment scale.

## 2. Field Site and Data

[6] The analysis presented here grew out of an exploration of rainfall-runoff behavior at the Plynlimon catchments in mid-Wales. Plynlimon has been a focal point of hydrological research for at least four decades, resulting in several hundred scientific publications [e.g., Calder, 1977; Kirby *et al.*, 1991; Beven and Binley, 1992; Sklash *et al.*, 1996; Neal *et al.*, 1997b; Kirchner *et al.*, 2000; Robinson and Dupeyrat, 2005; Marc and Robinson, 2007; Kirby *et al.*, 1997, and references therein]. The Plynlimon catchments comprise roughly  $20 \text{ km}^2$  of the headwaters of the Wye and Severn rivers (Figure 1); the Wye catchment is grassland, whereas the Severn catchment was dominated by conifer plantations during 1992–1996, the time period analyzed here. The Wye and Severn rivers flow from adjacent catchments on the

same upland massif, predominantly composed of Ordovician and Silurian mudstones, sandstones, shales, and slates, and generally considered to be watertight [Kirby *et al.*, 1991]. Although borehole observations have shown clear evidence for extensive groundwater circulation through fractures down to depths of tens of meters [Neal *et al.*, 1997a; Shand *et al.*, 2005], no evidence of substantial intercatchment groundwater flow has been reported. The soil mantles at both catchments are dominated by blanket peats >40 cm thick at higher altitudes, podzols at lower altitudes, and valley bottom alluvium, peat, and stagnohumic gleys along the stream channels [Kirby *et al.*, 1991].

[7] The climate of Plynlimon is cool and humid; monthly mean temperatures are typically 2–3°C in winter and 11°–13°C in summer, and annual precipitation is roughly 2500–2600 mm/a, of which approximately 500 mm/a is lost to evapotranspiration and 2000–2100 mm/a runs off as stream discharge (Table 1). Precipitation varies seasonally, averaging 280–300 mm/month during the winter (December/January/February) but only 135–155 mm/month during the summer (June/July/August). Rainfall is frequent; more than 1 mm of rainfall occurs on about 45% of summer days and over 60% of winter days. Frost can occur in any month of the year, but snow accounts for only about 5% of total annual precipitation, and persistent snow cover is rare [Kirby *et al.*, 1991].

[8] Precipitation and streamflow have been measured continuously at Plynlimon since the 1970s by the Centre for Ecology and Hydrology (formerly the Institute of Hydrology). In addition to a network of ground-level storage rain gauges that are read monthly, the Severn and Wye catchments are each outfitted with a pair of automatic weather stations, one near the bottom of each catchment and one near the top (circles, Figure 1). These weather stations provide hourly records of precipitation, as well as incoming solar and net radiation, wet and dry bulb temperature, and wind speed and direction, allowing estimation of potential evapotranspiration via the Penman-Monteith method. Streamflow is measured at 15-min intervals by a trapezoidal critical depth flume on the Severn and a Crump weir on the Wye, as well as by flumes on eight tributary streams (triangles, Figure 1).

[9] This paper uses data from the four automatic weather stations, the Severn triangular flume, and the Wye weir. Data from 1992 through 1996 were selected for analysis, because during this interval none of these instruments suffered extended outages, with the result that a continuous, consistent data set is available for the entire 5-year period. Nevertheless, as with any long-term environmental data set, anomalies occur in a small number of records (here, less than 1% of the total). Each discharge and weather station record was examined by eye for the entire 5-year period, and clearly anomalous measurements were replaced with interpolated values from adjacent reliable measurements, or when necessary by appropriately scaled averages from other stations. The 15-min discharge data were aggregated to hourly sums, synchronized with the hourly weather station data.

[10] Two brief extracts from the full 5-year record are shown in Figure 2. As one can see, the Severn and Wye rivers both respond promptly to rainfall inputs, but the Wye is visibly more “flashy” than the Severn. In both catch-

**Table 1.** Basic Physiographic and Hydrological Characteristics of the Plynlimon Catchments<sup>a</sup>

	Severn	Wye
Drainage area (km <sup>2</sup> )	8.70	10.55
Altitude range (m)	319–738	341–738
Forest cover (%)	67.5	1.2
Strahler stream order	4	4
Drainage density (km/km <sup>2</sup> )	2.40	2.04
Main channel length (km)	4.6	7.3
Main channel slope (m/km)	67	36
Mean water fluxes 1972–2004		
Precipitation (mm/a)	2553	2599
Streamflow (mm/a)	1987	2111
Evapotranspiration (mm/a)	566	488

<sup>a</sup>Sources: Kirby *et al.* [1991] and Marc and Robinson [2007]. Evapotranspiration is calculated from the difference between precipitation and streamflow averages.

ments, there is a clear correspondence between the intensity and duration of rainfall events, and the timing and intensity of storm runoff. Motivated by the rainfall-runoff behavior observed at Plynlimon, the analysis below presents a simple, analytically tractable, empirically testable framework for understanding the hydrologic behavior of small catchments. I now describe this analytical framework, and will return to its application to the Plynlimon catchments in section 5.

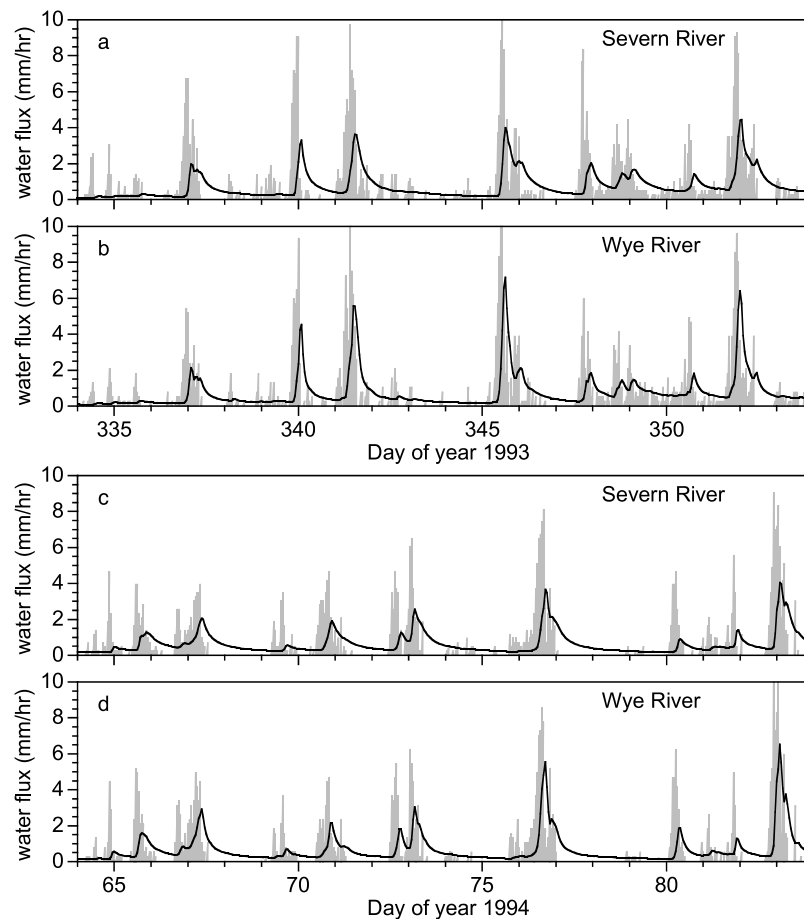
### 3. Catchment Hydrology as a First-Order Dynamical System

[11] This analysis begins, as most catchment-scale hydrological models do, with the conservation-of-mass equation,

$$\frac{dS}{dt} = P - E - Q, \quad (1)$$

where  $S$  is the volume of water stored in the catchment, measured in units of depth (e.g., mm of water), and  $P$ ,  $E$ , and  $Q$  are the rates of precipitation, evapotranspiration, and discharge, respectively, in units of depth per time (e.g., mm of water per hour).  $P$ ,  $Q$ ,  $E$ , and  $S$  are understood to be functions of time, and are understood to be averaged over the whole catchment.

[12] Applications of equation (1) should take account of how its individual terms are measured, and the spatial scales over which such measurements are applicable. Precipitation measurements are intrinsically local, because precipitation rates vary in space and time, and rain gauges are typically many orders of magnitude smaller than the catchments that they are used to represent. (New technologies such as precipitation radar can provide spatially distributed estimates of rainfall rates, but still must be benchmarked to rain gauge data.) Estimates of evapotranspiration, whether derived from Penman-Monteith methods, eddy correlation instruments, or evaporation pans, also have effective footprints that are orders of magnitude smaller than typical catchments. Estimates of changes in storage, as measured by piezometer wells and soil moisture probes, are likewise highly localized, and are also strongly dependent on spatially variable material properties of the subsurface. Of the four terms in equation (1), only discharge is an aggregated measurement for the entire catchment. Therefore the anal-



**Figure 2.** Time series of hourly rainfall (gray) and discharge (solid black curves) for headwaters of the Severn and Wye rivers during 20-day periods in (a, b) December 1993 and (c, d) March 1994. Rainfall time series recorded in the two catchments are similar but not identical. Wye flows are more responsive to storm events than Severn flows. Flows in both rivers generally increase when the catchment mass balance is positive (rainfall flux is higher than discharge) and decrease when the mass balance is negative (rainfall flux is lower than discharge). As a result, flow peaks in both streams occur at the end of rainfall events, as rainfall fluxes drop below runoff fluxes and the catchment mass balance turns negative. This behavior is consistent with the simple first-order dynamical system described in equations (1) and (2).

ysis presented here explores what one can learn about catchment processes from fluctuations in streamflow, without assuming that measurements of precipitation or evapotranspiration are spatially representative. The analysis also makes no use of direct measurements of changes in storage, because they are often unavailable.

[13] This analysis makes the fundamental assumption that the discharge in the stream,  $Q$ , depends solely on the amount of water stored in the catchment,  $S$ . That is, the analysis assumes that there is some storage-discharge function  $f(S)$  such that

$$Q = f(S). \quad (2)$$

This premise is not valid in every catchment, but in many cases it can be a useful approximation, and it is an essential assumption in the analysis that follows. Of course, in any catchment some fraction of stream discharge may be controlled by processes other than the release of water from storage. Two obvious examples are direct precipitation onto the stream surface itself, and precipitation onto areas

that are impermeable or saturated and are directly connected to the stream. These processes will route precipitation directly to discharge as bypassing flow, rather than adding it to subsurface storage. The analysis presented here does not require that bypassing flow is entirely absent, but assumes that it is not a dominant component of discharge. If, instead, discharge is dominated by bypassing flow, the approach presented here may fail, because processes such as channel routing (which are not treated in detail here) may dominate the runoff response. A method for assessing the quantitative significance of bypassing flow is presented in section 15.4.

[14] The premise that discharge depends on storage is broadly consistent with the smaller-scale governing equations that drive subsurface transport. For example, the flow of water downward through the unsaturated zone is controlled by its matric potential and hydraulic conductivity, which are both steep nonlinear functions of water content. Flow in the saturated zone depends on the slope of the water table, which varies with storage in the saturated zone, and on the saturated hydraulic conductivity, which varies as a function of depth; thus transmissivity also depends on the



total storage in the saturated zone. As a result, stream discharge is often a steep nonlinear function of groundwater levels in the surrounding catchment [e.g., *Laudon et al.*, 2004, Figure 6]. Many of the processes and rate coefficients that control water flow in the subsurface are strongly, and nonlinearly, dependent on storage.

[15] Nonetheless it is not clear how these nonlinear relationships, which may differ from point to point across the landscape, will combine to create a storage-discharge relationship for the catchment as a whole. For this reason, my approach assumes no particular functional form for the storage-discharge relationship  $f(S)$ , instead allowing both the form of  $f(S)$  and its coefficients to be estimated directly from runoff time series data. I assume only that  $Q$  is an increasing single-valued function of  $S$  ( $dQ/dS > 0$  for all  $Q$  and  $S$ ), and thus that the storage-discharge function is invertible. Thus the discharge in the stream provides an implicit measure of the volume of water stored in the catchment:

$$S = f^{-1}(Q). \quad (3)$$

Equations (1) and (2) form a first-order dynamical system, in which  $P$ ,  $Q$ ,  $E$ , and  $S$  are all understood to be functions of time. This dynamical system would be particularly simple if  $Q$  were a linear function of  $S$ . The properties of such linear systems have been extensively studied in hydrology, but in general  $Q$  will be a nonlinear function of  $S$ , resulting in a richer spectrum of possible behaviors. This more general nonlinear case is the focus of the analysis presented here.

[16] Regardless of the form that  $f(S)$  takes, the structure of the dynamical system directly yields an important inference concerning catchment storm response. Because  $Q$  is a function of  $S$  alone, storage (and thus discharge) will be rising whenever  $P - E > Q$ , and falling whenever  $Q > P - E$ . The peak discharge ( $dQ/dt = 0$ ) will coincide with the peak storage ( $dS/dt = 0$ ), which will occur when  $Q = P - E$ . During storm events, the time of peak rainfall will generally occur during the rising limb of the hydrograph (when  $P - E > Q$  and thus  $dS/dt > 0$  and  $dQ/dt > 0$ ). Because the peak rainfall corresponds to rising flow, which by definition will occur before the peak discharge, the peak flow will lag the peak rainfall, even in the absence of any travel time delays for pulses of stormflow to reach the weir. Furthermore, the peak flow will occur as the rainfall rate falls below discharge, and thus the mass balance (equation (1)) turns negative.

[17] The Severn and Wye rivers exhibit this pattern of behavior, as Figure 2 shows. The Wye is somewhat more responsive than the Severn to rainfall inputs, but both catchments behave as the dynamical system of equations (1) and (2) would predict: when rainfall fluxes exceed streamflow fluxes (and thus the catchment mass balance is positive), discharge increases, and when streamflow exceeds rainfall (and thus the mass balance is negative), discharge decreases. Peak flows occur as rainfall events are ending, when rainfall fluxes drop below streamflow fluxes (and thus the mass balance changes sign). Thus the lag to peak is determined primarily by the duration of storm events; it is not a fixed characteristic time scale of the catchment.

[18] This behavior is inherent in the structure of the dynamical system described by equations (1) and (2),

because the derivative in equation (1) creates a dynamical phase lag between fluctuations in precipitation and fluctuations in streamflow. If storm runoff were dominated by bypassing flow, and thus changes in catchment storage were unimportant in the storm response, this phase lag would be negligible. Figure 2 shows that this is not the case at Plynlimon. In addition to this dynamical lag, there may also be a travel time lag for stormflows to move downstream through the channel network. As shown in section 7 below, in the Severn and Wye catchments this travel time lag is roughly 1 h, which is less than the width of the black lines shown in Figure 2.

#### 4. Estimating Catchment Sensitivity to Changes in Storage: Theory

[19] Differentiating equation (2) with respect to time and substituting equation (1) directly yields the following differential equation for the rate of change of discharge through time:

$$\frac{dQ}{dt} = \frac{dQ}{dS} \frac{dS}{dt} = \frac{dQ}{dS} (P - E - Q). \quad (4)$$

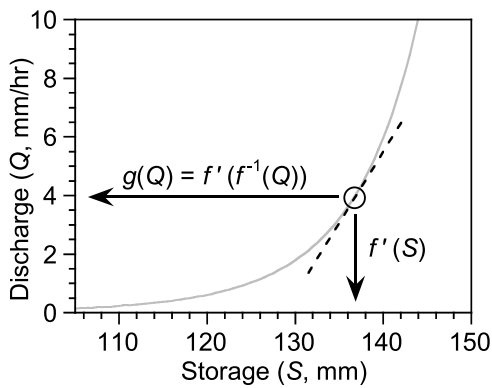
The term  $dQ/dS$  will be crucial in the analysis that follows; it is the derivative of the storage-discharge relationship  $f(S)$ , and represents the sensitivity of discharge to changes in storage. Normally, derivatives like  $dQ/dS$  would be expressed in terms of  $S$ , but  $S$  cannot be directly measured at the catchment scale for the reasons described in section 3. However, because  $S$  is assumed to be a single-valued function of  $Q$ ,  $dQ/dS$  can also be expressed as a function of  $Q$ , here defined as  $g(Q)$ :

$$\frac{dQ}{dS} = f'(S) = f'(f^{-1}(Q)) = g(Q). \quad (5)$$

The function  $g(Q)$  will be called the “sensitivity function” because it expresses the sensitivity of discharge to changes in storage. Mathematically, it is the implicit differential form of the storage-discharge relationship; it measures how changes in discharge are related to changes in storage, but it does so as a function of  $Q$  (which is directly measurable) rather than  $S$  (which is not). This makes it more useful than the conventional form  $f'(S)$  for the analysis that follows. Figure 3 illustrates the relationship between the sensitivity function  $g(Q)$  and the storage-discharge relationship  $f(S)$ . The function  $g(Q)$  can be estimated from observational data by combining equations (5) and (4) to yield

$$g(Q) = \frac{dQ}{dS} = \frac{dQ/dt}{dS/dt} = \frac{dQ/dt}{P - E - Q}, \quad (6)$$

which implies that the slope of the storage-discharge function  $f(S)$  can be determined from instantaneous measurements of precipitation ( $P$ ), evapotranspiration ( $E$ ), discharge ( $Q$ ), and the rate of change of discharge ( $dQ/dt$ ). Of the three fluxes ( $P$ ,  $E$ , and  $Q$ ), discharge can be measured more reliably than precipitation or evapotranspiration at the whole-catchment scale, for the reasons described in section 3 above. Therefore equation (6) can be most accurately estimated when precipitation and evapotranspiration fluxes



**Figure 3.** Explanatory diagram for the catchment sensitivity function  $g(Q)$ , the implicit differential form of the storage-discharge relationship  $f(S)$ . At any particular point along the storage-discharge relationship  $Q = f(S)$  (gray curve), the local sensitivity of discharge to changes in storage is expressed by the local derivative,  $dQ/dS$  (the slope of the dashed line). Normally, such a derivative is expressed as a function of the variable on the horizontal axis (i.e., as the derivative function  $f'(S)$ ). However, because the storage-discharge relationship is a monotonic function and therefore is invertible, the derivative  $dQ/dS$  can also be expressed as a function of discharge,  $g(Q) = f'(f^{-1}(Q))$ . This implicit form of the derivative is useful because discharge is directly measurable and storage is not.

are small compared to discharge ( $P \ll Q$  and  $E \ll Q$ ). Under these conditions, equation (6) is approximated by

$$g(Q) = \frac{dQ}{dS} \approx \left. \frac{-dQ/dt}{Q} \right|_{P \ll Q, E \ll Q}. \quad (7)$$

Equation (7) implies that one can estimate the sensitivity function  $g(Q)$  from the time series of  $Q$  alone. To do this, one must identify intervals of time when precipitation and evapotranspiration are small compared to discharge, but it is not necessary to measure either  $P$  or  $E$  accurately as long as their rough magnitude compared to  $Q$  is known. From the sensitivity function  $g(Q)$ , one can derive the storage-discharge relationship  $f(S)$  by first inverting equation (5),

$$\int dS = \int \frac{dQ}{g(Q)}, \quad (8)$$

thus obtaining  $S$  as a function of  $Q$ , and then by inverting this function to obtain  $Q$  as a function of  $S$ .

[20] Apart from the requirement that  $Q = f(S)$  must be an increasing function of  $S$  (and thus that  $g(Q)$  must always be positive), nothing in the approach outlined here requires  $f(S)$  or  $g(Q)$  to have any particular mathematical form. In practice,  $g(Q)$  will be an empirical function that is estimated from streamflow time series data, and it could potentially exhibit different functional forms in different catchments. A few simple functional forms of  $g(Q)$  can be integrated and inverted analytically to yield closed-form solutions for  $f(S)$ . For other functional forms, equation (8) can be solved by

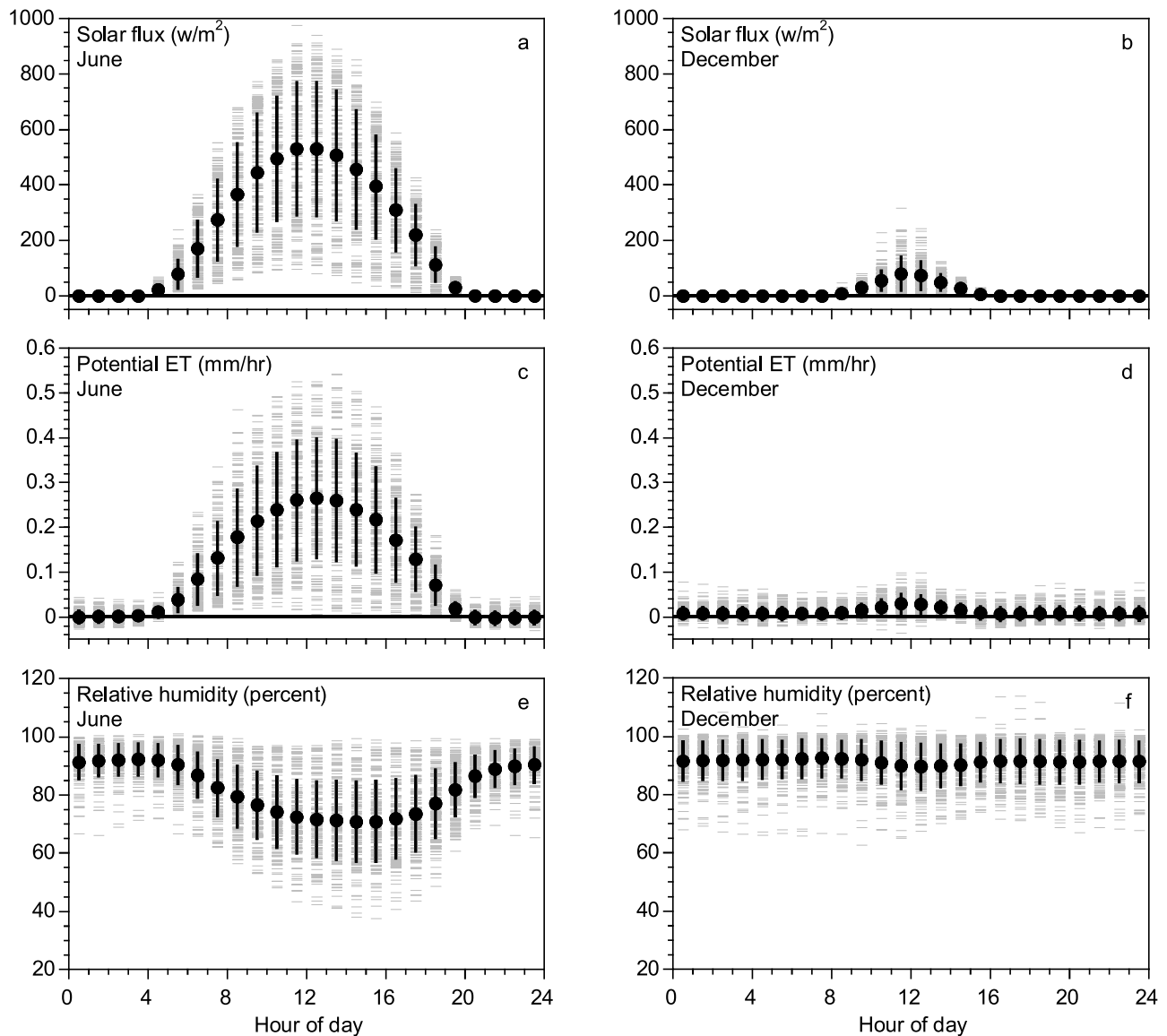
numerical integration in order to construct an empirical storage-discharge relationship.

## 5. Estimating Catchment Sensitivity to Changes in Storage: Practical Details

[21] Implementing this approach in practice requires identifying times when precipitation and evapotranspiration fluxes are small enough that equation (6) will be well approximated by equation (7). I used two different methods to identify these low-precipitation, low-ET periods at Plynlimon, and both yielded similar results. The first approach used the automatic weather station data to estimate potential evapotranspiration via the Penman-Monteith method. The estimated potential evapotranspiration does not need to accurately reflect actual evapotranspiration, but only its general magnitude, because equation (7) does not require estimating a mass balance for the catchment, but only identifying times when the mass balance is dominated by discharge. To implement this approach at Plynlimon, I selected the hourly records for which discharge was at least 10 times larger than both potential evapotranspiration and precipitation (as measured by the weather station rain gauges).

[22] The second approach assumes that potential evapotranspiration fluxes in humid catchments should be relatively small at night, because relative humidity is typically near 100% (and thus the vapor pressure deficit is small), and there is no solar radiation to drive transpiration fluxes (see Figure 4). To implement this approach at Plynlimon, I selected the hourly records for nighttime (defined as times for which solar flux was less than  $1 \text{ W/m}^2$  averaged over the hour in question, the previous hour, and the following hour), and during which there was also no recorded rainfall within the previous 6 h or the following 2 h. Selecting either these rainless night hours, or hours with negligible precipitation and potential evapotranspiration (as described above), yields roughly 1600 to 2000 h/a at Plynlimon. Although these two methods for identifying low-precipitation, low-evapotranspiration conditions do not result in exactly the same records being analyzed (only about half of the records overlap between the two approaches), they both yield similar results in the analysis that follows. The analysis shown below is based on the rainless night hours at the Severn and Wye catchments. Figure 5 shows an example of these rainless nighttime periods, for a short segment of the Severn River time series.

[23] From hourly streamflow records during periods when  $P \ll Q$  and  $E \ll Q$ , we can estimate  $g(Q)$  in equation (7) by plotting the flow recession rate ( $-dQ/dt$ ) as a function of discharge ( $Q$ ), as shown in Figure 6. Graphs like Figure 6, here termed “recession plots,” were proposed by Brutsaert and Nieber [1977] as an alternative to conventional recession curves, in which discharge is plotted as a function of time. Recession plots are particularly appropriate in the present case, because equation (7) requires low-precipitation, low-evaporation conditions, which usually form a highly discontinuous time series (as in Figure 5). Such a discontinuous time series would be ill suited to conventional recession analysis (although others have dealt with this problem by splicing short intervals together into pseudocontinuous recession curves; see Lamb and Beven [1997] for one such analysis). Recession plots such as



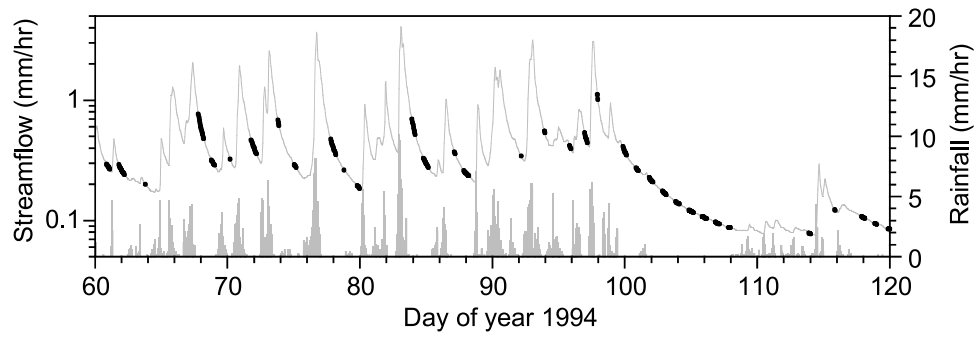
**Figure 4.** Solar flux, Penman-Monteith potential evapotranspiration, and relative humidity as a function of time of day for (left) June and (right) December, calculated from hourly measurements at the Cefn Brwyn automated weather station in the Wye catchment, 1992–1996. Black dots and lines indicate means and standard deviations. During hours of darkness, potential evapotranspiration is nearly zero, and relative humidity is close to 100%.

Figure 6 provide a general way to display and analyze recession behavior, without presupposing that the underlying data are continuous in time.

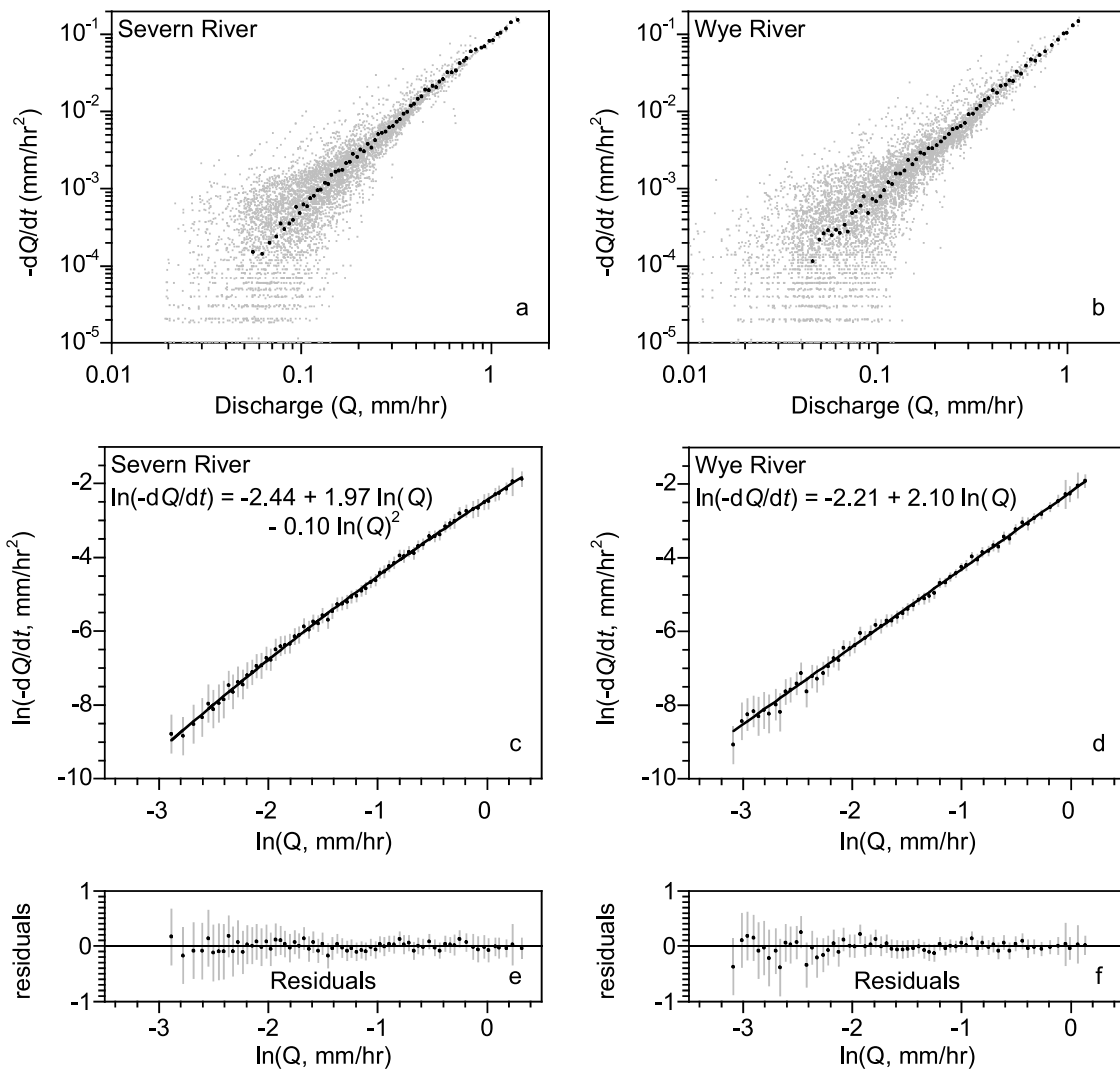
[24] Following *Brutsaert and Nieber* [1977], I estimate the rate of flow recession as the difference in discharge between two successive hours,  $-dQ/dt = (Q_{t-\Delta t} - Q_t)/\Delta t$ , and plot this as a function of the average discharge over the two hours,  $(Q_{t-\Delta t} + Q_t)/2$ . Estimating the terms in this way avoids any artifactual correlation between  $Q$  and  $-dQ/dt$ . Because  $Q$  and  $-dQ/dt$  will both typically span several orders of magnitude, their relationship to one another can be best viewed on log-log plots. Figures 6a and 6b show the relationship between discharge and flow recession for hourly measurements from the Severn and Wye rivers (gray dots, Figure 6). In both streams, the rate of flow recession is roughly a power law function of discharge. *Brutsaert and*

*Nieber* [1977] used plots like Figure 6 to define the lower envelope of  $-dQ/dt$  as a function of  $Q$ , under the assumption that these points would be least affected by evapotranspiration, but in practice, much of the spread in  $-dQ/dt$  at any particular value of  $Q$  may be due to stochastic variability and measurement noise [*Rupp and Selker*, 2006a], particularly over the short intervals between individual hourly measurements. The present approach instead seeks the best estimate of  $g(Q)$  as an average description of the behavior of the catchment. This requires estimating the central tendency of  $-dQ/dt$  rather than its lower bound.

[25] Accurately estimating  $g(Q)$  requires careful attention to several details. The function  $g(Q)$  must correctly describe the relationship between  $Q$  and  $-dQ/dt$  when they are both small, and log-log plots like Figure 6 expand this domain. The individual hourly data exhibit significant scatter on log



**Figure 5.** Severn catchment hourly rainfall (vertical gray bars) and Severn River streamflow (gray curve) for March and April 1994, with rainless nighttime intervals highlighted in black.



**Figure 6.** Recession plots for the (left) Severn and (right) Wye rivers. (top) Flow recession rates ( $-dQ/dt$ ) as a function of flow ( $Q$ ) for individual rainless nighttime hours (gray dots, approximately 8,000 points per plot) and for averages of  $-dQ/dt$ , binned as described in the text (black dots). (middle) The averages and their associated standard errors (gray bars show  $\pm 1$  standard error), with best fit lines calculated by least squares regression with inverse variance weighting. (bottom) Residuals from these best fit lines. The binned means (black dots) deviate from the fitted lines by less than their standard errors, suggesting that the fitted lines are a quantitatively adequate description of the mean recession behavior of these catchments.



axes, particularly at discharges below about 0.1 mm/h. This scatter could arise from at least four sources: (1) random measurement noise, (2) coarse graining due to the finite discretization of discharge measurements, and thus of calculated flow recession rates (as is visually evident from the horizontal stripes in Figures 6a and 6b), (3) effects of any precipitation or evapotranspiration that may occur but be too small to be directly measurable, and (4) differences between the structure of the real-world catchment and the idealized dynamical system hypothesized here. Noise arising from any of these sources should introduce more scatter in the log of  $-dQ/dt$  at times when  $Q$  and  $-dQ/dt$  are small, as Figures 6a and 6b show.

[26] On a log scale, this scatter can introduce a bias, since fluctuations toward zero are larger in log units than equivalent fluctuations away from zero. Indeed, at low  $Q$ , there are many points for which discharge is constant or increasing, and thus  $-dQ/dt$  for these points cannot be plotted on a log axis at all. It might seem logical to simply exclude such points from the analysis, under the assumption that any such points cannot correspond to flow recession. However, many such points may represent random fluctuations around an average recession trend. Therefore they should not be excluded, because preferentially excluding random deviations in one direction but not the other would lead to biased estimates of the average recession rate  $-dQ/dt$  at any given  $Q$ .

[27] Instead, the scatter at low  $Q$  must be properly taken into account in order to estimate the functional relationship between  $-dQ/dt$  and  $Q$ . In Figure 6, I do this by binning the individual hourly data points into ranges of  $Q$ , and then calculating the mean and standard error for  $-dQ/dt$  and  $Q$  within each bin (including values of  $-dQ/dt \leq 0$ , which cannot be displayed on log axes). These means are the black dots in Figure 6. Working from the highest values of  $Q$  to the lowest, I delimit bins that span at least 1% of the logarithmic range in  $Q$ , and that include enough points that the standard error of  $-dQ/dt$  within the bin is less than half of its mean. The criterion  $\text{std.err.}(-dQ/dt) \leq \text{mean}(dQ/dt)/2$  is a first-order Taylor approximation to the criterion  $\text{std.err.}(\ln(-dQ/dt)) \leq 0.5$ , which cannot be directly evaluated when  $dQ/dt$  has both positive and negative values. The binned averages reflect the average recession rate  $-dQ/dt$  at each flow rate  $Q$ , without being unduly influenced by the stochastic scatter in  $-dQ/dt$  when  $Q$  is small.

[28] I then fit smooth curves to the binned means (black dots) using least squares regression, weighted by inverse variance (that is, by the reciprocal of the square of the standard errors of each binned average). This approach keeps highly uncertain points from exerting too much influence on the regression. This approach also yields the maximum-likelihood estimator for the best fit curve, if the deviations of the black dots from the true relationship are approximately normal. This is likely to be the case, because according to the central limit theorem, the errors in the binned means (black dots) should be distributed almost normally even if the individual measurements (gray dots) are not, since each black dot is typically calculated by averaging many individual points. As the residual plots at the bottom of Figure 6 show, the best fit curves fall within one standard error of nearly all of the binned means, implying that they capture nearly all of the systematic relationship between  $\ln(-dQ/dt)$  and  $\ln(Q)$ . If, on the other hand, the best fit curves fell outside the error bars of many of

the binned means, this would indicate that the curves were incorrectly estimated or were not flexible enough to follow the structural relationship between  $\ln(-dQ/dt)$  and  $\ln(Q)$ .

[29] In the absence of a strong theoretical expectation for the storage-discharge relationship to have a particular functional form, one must choose an empirical function to fit to the binned means in Figure 6. To fit the black dots in Figure 6, I chose a quadratic curve because it is both flexible enough to follow the major features of the data and smooth enough to permit modest extrapolation beyond the range of the black dots. This quadratic function leads directly to an expression for  $g(Q)$  as a quadratic in logs,

$$\ln(g(Q)) = \ln\left(\frac{-dQ/dt}{Q}\right)_{P \ll Q, E \ll Q} \approx c_1 + c_2 \ln(Q) + c_3 (\ln(Q))^2, \quad (9)$$

with parameter values of  $c_1 = -2.439 \pm 0.017$ ,  $c_2 = 0.966 \pm 0.035$ , and  $c_3 = -0.100 \pm 0.016$  for the Severn River, and parameter values of  $c_1 = -2.207 \pm 0.028$ ,  $c_2 = 1.099 \pm 0.048$ , and  $c_3 = -0.002 \pm 0.018$  for the Wye River, obtained by polynomial least squares regression. The coefficient  $c_2$  is one less than the slope of the log-log plots in Figure 6, owing to the factor of  $Q$  in the denominator of equation (9).

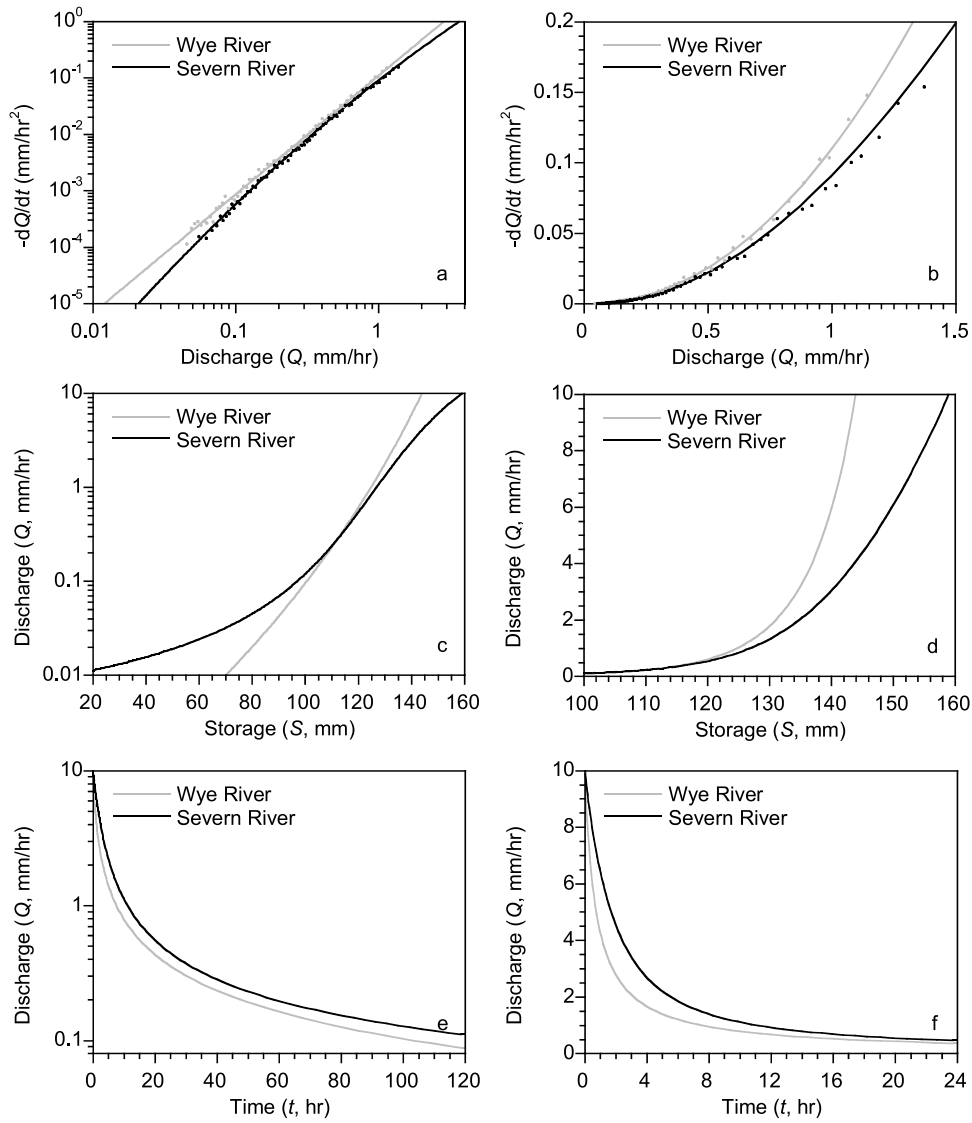
[30] The fitted curves for the Severn and Wye rivers look similar in Figure 6, although when they are overlain on one another, small differences are visually apparent (Figures 7a and 7b). Nonetheless, when these fitted curves are transformed to storage-discharge relationships, they are visually quite distinct (Figures 7c and 7d). Notably, the Wye River's storage-discharge relationship is more sharply curved than the Severn's, which is broadly consistent with the Wye's more abrupt response to precipitation, as shown in Figure 2. Integrating these storage-discharge relationships yields theoretical recession curves (discharge as a function of time); as Figures 7e and 7f show, the recession curves for the two catchments are visually similar, despite the obvious differences between their storage-discharge relationships. This observation suggests that conventional analyses of recession curves may not detect important differences in storage-discharge relationships between catchments. These differences are, however, apparent from the analysis outlined above.

## 6. Power Law Relationships Between $Q$ and $-dQ/dt$ : An Idealized Approximation

[31] Log-log recession plots such as Figure 6 are often approximately linear, suggesting a power law relationship between discharge  $Q$  and the recession rate  $-dQ/dt$ ,

$$-\frac{dQ}{dt} = aQ^b, \quad (10)$$

where  $b$  is the log-log slope of the best fit line. Following the fundamental contributions of *Horton* [1941] and *Brutsaert and Nieber* [1977], this power law recession behavior has been used to characterize catchments in a number of ways, usually based on a nonlinear reservoir model or a Boussinesq representation of flow in the subsurface [e.g., *Troch et al.*, 1993; *Brutsaert and Lopez*, 1998; *Tague and Grant*, 2004; *Rupp and Selker*, 2006b;



**Figure 7.** Comparison of recession behavior and storage-discharge relationships for the Severn and Wye catchments. Recession plots on (a) log-log and (b) linear axes illustrate differences between the two catchments' drainage characteristics. Data points are binned averages from Figure 6. The differences in the recession plots for the two catchments (Figures 7a and 7b) imply differences in their storage-discharge relationships as well, shown on (c) log linear and (d) linear axes. The different shapes of the inferred storage-discharge relationships are meaningful, but their relative placement is not, as equation (8) cannot determine absolute levels of storage. The two catchments' storage-discharge relationships are visibly different, but their recession curves, shown on (e) log linear and (f) linear axes, are almost indistinguishable.

*Lyon and Troch, 2007; Rupp and Woods, 2008*]. Power law recession relationships are also analytically tractable in the dynamical system outlined above, and imply an interesting family of storage-discharge relationships  $f(S)$ . It bears emphasis that these idealized power law functions are only a special case in the general analytical approach outlined in this paper, and I will return to the more general analysis in the following section.

[32] A power law relationship between  $Q$  and  $-dQ/dt$ , as in equation (10), would imply that  $g(Q)$  is

$$g(Q) = \frac{dQ}{dS} = \frac{dQ/dt}{-Q} = aQ^{b-1}. \quad (11)$$

Equation (8) thus becomes

$$\int dS = \int \frac{1}{a} Q^{1-b} dQ, \quad (12)$$

which can be solved as

$$S - S_o = \frac{1}{a} \frac{1}{2-b} Q^{2-b}, \quad (13)$$

where  $S_o$  is a constant of integration. Equation (13) can be inverted to obtain  $f(S)$ :

$$Q = f(S) = \sqrt[2-b]{(2-b)a(S - S_o)}. \quad (14)$$

In equation (10) and thus also in equation (14), the dimensions of the constant  $a$  will vary with  $b$ , as  $\text{length}^{(b-1)/(2-b)} \text{time}^{1/(2-b)}$ , for dimensional consistency. Equation (14) can also be rewritten in a more dimensionally straightforward form as

$$Q = f(S) = Q_{\text{ref}}((S - S_o)/k_1)^{1/(2-b)}, \quad (15)$$

where  $Q_{\text{ref}}$  is an arbitrary reference discharge, and the scaling constant  $k_1 = (Q_{\text{ref}}^{2-b})/[(2-b)a]$  has the same dimensions as storage.

[33] Equations (14) and (15) have three classes of solutions, and in each case the constant of integration  $S_o$  means something different. If  $b < 2$ , equation (14) yields  $Q$  as a power function of  $S$ , with  $S_o$  representing the residual storage remaining in the catchment when discharge drops to zero. In the special case where  $b = 1$ ,  $f(S)$  is linear and the conventional results for linear reservoirs (such as log linear recession curves) are obtained. As  $b$  increases from 1 toward 2,  $f(S)$  becomes an increasingly steep power function, with the exponent  $1/(2-b)$  in equation (15) approaching infinity as  $b$  approaches 2.

[34] When  $b = 2$ , the solution to equation (8) is an exponential function,

$$Q = f(S) = Q_{\text{ref}} e^{a(S-S_o)}, \quad (16)$$

where  $S_o$  now represents the value of storage when  $Q = Q_{\text{ref}}$ . Note that in equation (16), there will be some finite discharge at all values of  $S$ , allowing storage to decline indefinitely.

[35] When  $b$  is greater than 2, equations (14) and (15) become hyperbolic, and the meaning of  $S_o$  changes significantly. Values of  $b > 2$  imply that  $2 - b$  is negative, so equations (14) and (15) will yield imaginary values of  $Q$  unless  $S$  is less than  $S_o$ . Thus when  $b > 2$ ,  $S_o$  is no longer the lower bound to storage (at which discharge would decrease to zero); instead,  $S_o$  is the upper limit to storage, unreachable in practice, at which discharge would become infinite (for a different but mathematically equivalent interpretation, see *Rupp and Woods* [2008]). When  $b > 2$ , the behavior of equation (15) can be seen more clearly if it is rewritten as

$$Q = f(S) = \frac{Q_{\text{ref}}}{((S_o - S)/k_2)^{1/(b-2)}}, \quad (17)$$

where  $Q_{\text{ref}}$  is again an arbitrary reference discharge, and  $k_2 = -k_1 = (Q_{\text{ref}}^{2-b})/[a(b-2)]$  again has the same dimensions as storage. Equation (17) is equivalent to (15), but is easier to understand in this form because the scaling constant  $k_2$  and the exponent  $1/(b-2)$  are both positive when  $b > 2$ , whereas in equation (15) the scaling constant  $k_1$  and the exponent  $1/(2-b)$  would both be negative.

[36] The best fit values of  $b$ , obtained from Figure 6 by linear regression, are  $b = 2.168 \pm 0.017$  for the Severn River and  $b = 2.103 \pm 0.015$  for the Wye River. (These values differ somewhat from the linear terms in the polynomial regressions reported above, because of collinearity between the linear and quadratic terms in those polynomial expressions). These best fit values of  $b$  both exceed  $b = 2$  by more than six standard errors. Thus, to the extent that the Severn

and Wye catchments could both be approximated by power law recession plots, they would both appear to exhibit the hyperbolic behavior described by equation (17). Thus the hyperbolic solution represented by equation (17) may be more than just a mathematical oddity, and may be useful for understanding the behavior of flashy hydrologic systems.

[37] Figure 8 shows log-log recession plots (similar to Figure 6) for a range of exponents  $b$ , along with the corresponding storage-discharge relationships, and the resulting recession curves as functions of time. As Figure 8b illustrates, the storage-discharge relationship becomes dramatically more nonlinear as  $b$  increases. When  $b$  is greater than 2, discharge increases more than exponentially as a function of storage; that is, the log of  $Q$  curves upward as a function of  $S$  (Figure 8c). Figure 8d shows hypothetical recession curves of  $\log(Q)$  as a function of time, derived by integrating equations (1) and (2), or, alternatively, equations (4) and (11). As Figure 8d shows, these logarithmic recession curves become increasingly nonlinear as  $b$  increases, and are very sharply curved when  $b$  is greater than 2.

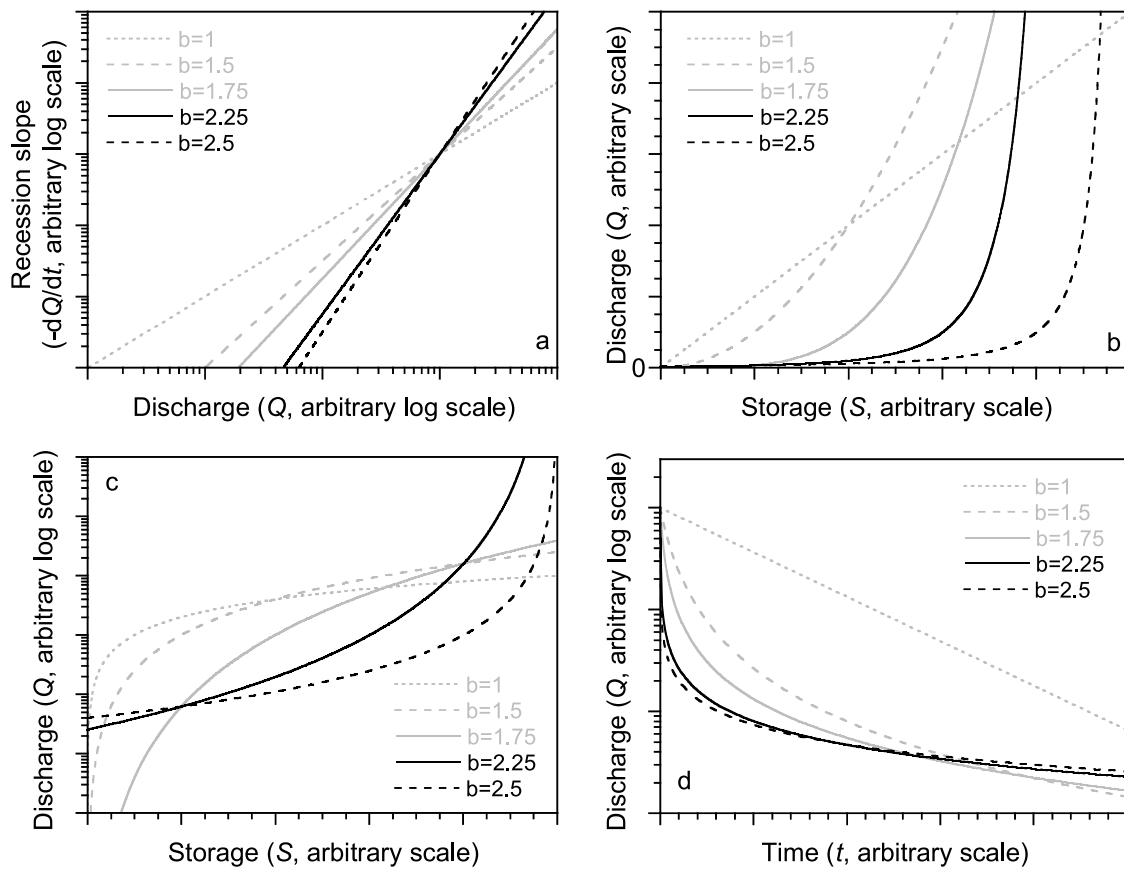
## 7. Simulating Hydrographs From Storage-Discharge Relationships

[38] From the preceding discussion, one can devise a straightforward strategy for rainfall-runoff modeling using the methods outlined above. The discharge sensitivity function  $g(Q)$  could be numerically integrated (or analytically integrated if its functional form is simple enough), yielding the storage-discharge relationship  $f(S)$ . One could then iteratively simulate the simple dynamical system formed by  $Q = f(S)$  and  $dS/dt = P - E - Q$ , initializing this system at some beginning time step using  $S = f^{-1}(Q)$ . From time series of  $P$  and  $E$ , one could then simulate the time series of  $Q$ .

[39] However, because  $Q$  is a differentiable and invertible function of  $S$ , the dynamical system of equations (1) and (2) can be solved in a more elegant way that does not require explicitly accounting for storage at all. Combining equations (4) and (5), one directly obtains

$$\frac{dQ}{dt} = \frac{dQ}{dS} \frac{dS}{dt} = g(Q)(P - E - Q), \quad (18)$$

which is a first-order nonlinear differential equation for  $Q$  that depends only on the values of  $P$  and  $E$  over time. Therefore one can simulate the streamflow hydrograph directly from time series of  $P$  and  $E$  by integrating equation (18) through time, given only a single value of  $Q$  to initialize the integration. This approach is more direct than explicitly solving equations (1) and (2), for two reasons. First, it avoids the need to know the antecedent moisture conditions at the beginning of a simulation. Second, and more significantly, it avoids the potentially difficult process of inferring the storage-discharge relationship  $f(S)$  from the sensitivity function  $g(Q)$ . Where, one might ask, has the storage variable gone? Note that in the conservation of mass equation, storage appears only as its time derivative; that is, one never needs to know the value of storage, but only its rate of change through time. Thus one can use a differential form of the storage-discharge function. If one uses the implicit differential form,  $g(Q)$ , one can eliminate  $S$  as an



**Figure 8.** (a) Idealized power law recession plots, with corresponding relationships between storage and discharge (on both (b) linear and (c) logarithmic scales) and (d) idealized recession curves on log linear axes. In Figures 8a and 8d curves correspond to equation (10), and in Figures 8b and 8c curves correspond to equations (15) and (17) for a range of exponents ( $b$ ); values of  $k$  are 1 in all cases. Curves for  $b < 2$  and  $b > 2$  are shown in gray and black, respectively. When  $b < 2$ , the storage-discharge relationship is a power function that declines to zero as  $S$  declines to the residual storage level  $S_o$ , which has been set at the left edge of Figures 8b and 8c. When  $b > 2$ , the storage-discharge relationship is hyperbolic, becoming infinitely steep as  $S$  rises toward the spillover level  $S_o$ , which has been set at the right edge of Figures 8b and 8c. Discharge grows more than exponentially as a function of storage when  $b > 2$  (black curves); that is, the storage-discharge relationship curves upward on log linear axes (Figure 8c) but curves downward for  $b < 2$  (gray curves). Logarithmic recession curves (Figure 8d) are nonlinear for  $b > 1$ , with the degree of curvature increasing as  $b$  increases.

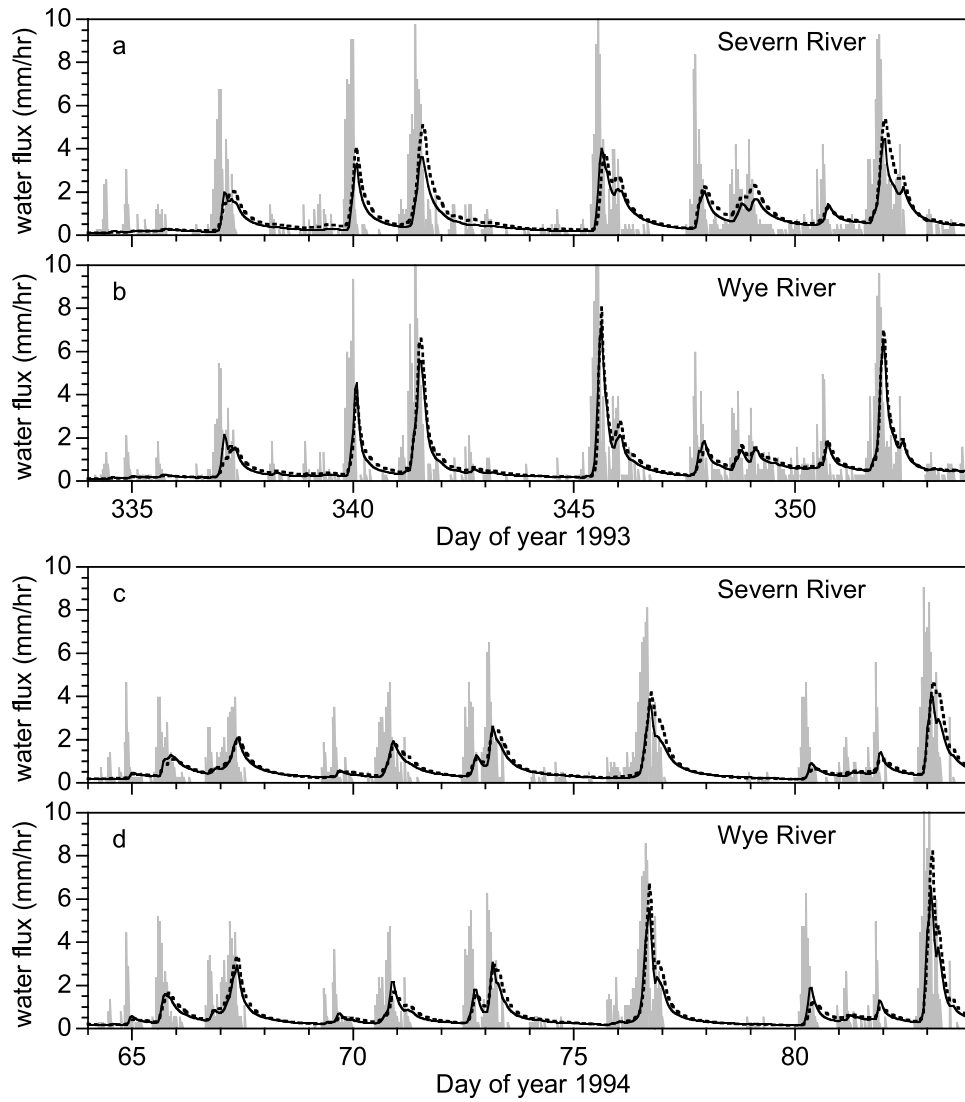
explicit variable completely. In other words, because discharge is a function of storage, changes in storage can be estimated from changes in discharge, as long as one accounts for the relationship between them, which is expressed by  $g(Q)$ .

[40] Implementing this approach requires attention to two practical details. The first detail concerns time lags in the catchment system. Owing to the time required for water to transit through the channel network, changes in discharge measured at the catchment outlet may lag behind changes in catchment storage. Field measurements show a typical flow velocity of roughly 1 m/s for the Severn [Beven, 1979], implying travel times of roughly 1 h between channel heads and the catchment outlet. Changes in subsurface storage may also lag behind precipitation inputs because of the time required for precipitation to infiltrate sufficiently to affect the hydraulic potentials that control stream discharge. Both of these time lags imply that changes in discharge, as observed at the outlet, may lag behind precipitation inputs

and thus behind the predictions of equation (18). These travel time lags are different from the phase lag that is inherent in this dynamical system (as described in section 3 above). The phase lag is captured in equation (18) but the travel time lags are not; thus they could potentially introduce timing errors in synthetic hydrographs. Any such travel time lags, however, will not affect the estimation of  $g(Q)$ , because that is based on  $Q$  and  $dQ/dt$ , which are measured simultaneously at the catchment outlet.

[41] A straightforward strategy for estimating the travel time lag can be inferred from the form of equation (18). Equation (18) implies that the rate of change of discharge,  $dQ/dt$ , should be correlated with the balance between precipitation, evapotranspiration, and discharge. Variations in  $P - E - Q$  will be dominated by variations in  $P$ , because precipitation is more variable than either evapotranspiration or discharge; for the Severn and Wye catchments, the variance of hourly  $P$  is over five times the variance of hourly  $Q$ , and over 50 times the variance of Penman-





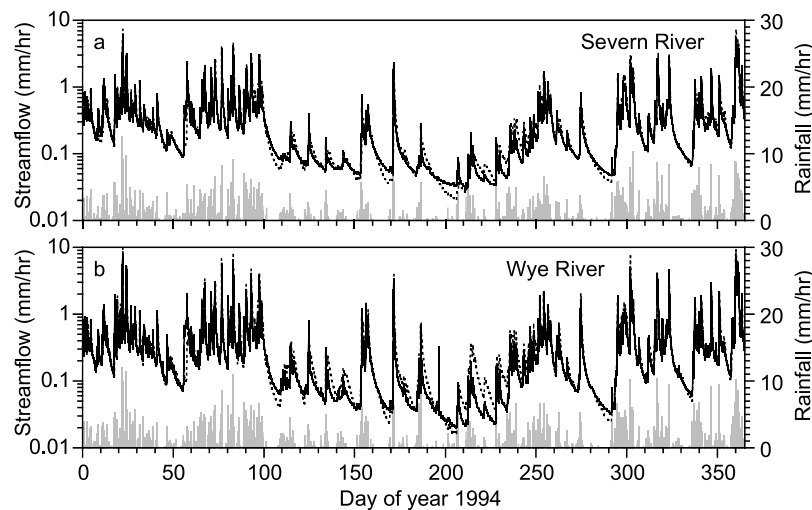
**Figure 9.** Synthetic hourly discharge time series (dotted black curves) predicted by equation (19), compared with measured discharge (solid black curves) and hourly rainfall (gray), for the Severn and Wye rivers during 20-day periods in (a, b) December 1993 and (c, d) March 1994. Predicted discharge is generally similar to observed discharge and mirrors the differences in storm response between the two catchments. Parameters of  $g(Q)$  were determined from Figure 6; no parameters were calibrated to the time series. Results are not sensitive to assumed evapotranspiration rates; predictions for  $E = E_o$  and  $E = 0$  differ by less than the width of the plotted lines.

Monteith estimates of hourly  $E$ . Therefore variations in  $dQ/dt$  should be correlated with variations in  $P$ , and any travel time lags should be apparent in the cross correlation between precipitation and  $dQ/dt$ . The cross correlation between  $P$  and  $dQ/dt$  peaks at lags of 1–2 h for both the Severn and Wye catchments, indicating a time lag of 1–2 h between changes in precipitation and changes in discharge as measured at the outlet. Lags this brief are of little consequence for simulating streamflow, since discharge is highly autocorrelated over such short time scales. Nonetheless, these lags can be taken into account straightforwardly by using appropriately lagged  $P$  and  $E$  time series in equation (18). The results shown in Figures 9 and 10 incorporate a 1-h lag; this is less than the widths of the lines in the graphs.

[42] The second detail that should be considered is the risk of numerical instabilities if equation (18) is integrated using Euler's method, because the term  $g(Q) \cdot Q$  is generally nonlinear, and  $Q$  typically varies by many orders of magnitude. Usually a better approach will be to integrate the log transform of equation (18),

$$\frac{d(\ln(Q))}{dt} = \frac{1}{Q} \frac{dQ}{dt} = \frac{g(Q)}{Q} (P - E - Q) = g(Q) \left( \frac{P - E}{Q} - 1 \right). \quad (19)$$

Because  $\ln(Q)$  will normally be locally much smoother than  $Q$  as a function of time, (19) will be easier than (18) to integrate.



**Figure 10.** Synthetic hourly hydrographs for the Severn and Wye rivers (dotted curves) generated by equation (19) compared with measured hourly hydrographs (solid curves) and hourly rainfall (gray). Streamflows are shown on logarithmic scales to emphasize low-flow behavior. Parameters of  $g(Q)$  were determined from Figure 6, not calibrated to the hydrographs. The only free parameter was the evapotranspiration scaling constant  $k_E$ , fitted to the entire 5-year period 1992–1996. Hydrographs for 1994 are shown; other years are similar.

[43] As Figures 9 and 10 show, this approach produces synthetic hydrographs that closely resemble the streamflow time series at Plynlimon. The hydrographs shown in Figures 9 and 10 were synthesized by iterating (19) on an hourly time step, using fourth-order Runge-Kutta integration. The  $g(Q)$  functions for the two catchments were obtained directly from Figure 6, and were not calibrated to the time series. The only calibration consisted of rescaling the Penman-Monteith potential evapotranspiration estimates  $E_o$  by an adjustable coefficient  $k_E$  to obtain the evapotranspiration time series  $E = k_E E_o$ ; a single value of  $k_E$  was fitted for the entire 5-year period 1992–1996. The analysis contains no other adjustable coefficients.

[44] As Figure 9 shows, the synthetic hydrographs correctly predict the general magnitude and timing of storm response at the two catchments, and generally reproduce the shape of the stormflow recessions. The synthetic hydrographs even reproduce the subtle differences in storm response between the two catchments; stormflow peaks in the Wye River are higher and narrower, with somewhat more rapid recessions. Note in particular that no parameters were adjusted to fit the stormflow periods shown in Figure 9. The two periods shown in Figure 9 correspond to relatively wet conditions, when the synthetic hydrographs (dashed lines) are insensitive to  $k_E$  and are therefore effectively free of any direct calibration. Nonetheless the results shown in Figure 9 compare well with much more complex models that have been applied to the Plynlimon catchments, with extensive parameter calibration [e.g., Rogers *et al.*, 1985; Bathurst, 1986].

[45] Catchment hydrologic models often perform relatively well in wet conditions, but break down during drier conditions. A model's low-flow characteristics are concealed when hydrographs are plotted on linear axes as in Figure 9, because flow variations spanning orders of magnitude (i.e., all except the highest flows) will appear as nearly horizontal lines close to the bottom of the plot. For this reason it is diagnostic to also compare synthetic and

measured hydrographs on logarithmic scales, as in Figure 10. As Figure 10 shows, the synthetic hydrographs reproduce the measured behavior in both catchments reasonably well during both wet periods and the drier intervals between them. The quantitative agreement between the synthetic and observed hydrographs on logarithmic scales, as measured by the Nash-Sutcliffe efficiency, is 0.91 and 0.86 for the Severn and Wye rivers, respectively, over the 5 years 1992–1996. These results compare favorably with the Nash-Sutcliffe efficiencies of other hydrologic models that are much more highly parameterized [Perrin *et al.*, 2001].

## 8. Cross Validation of Streamflow Predictions

[46] Although the recession plots in Figure 6 contain only fragmentary information about the original time series, it is reasonable to ask whether the approach outlined above is circular, given that it requires information from the hydrograph, which it then predicts. (The same question can be raised more pointedly for all hydrologic models that are calibrated directly to the hydrograph, which is to say virtually all hydrologic models.) A clear test, which is not circular, can be obtained from the following cross-validation exercise. I estimated  $g(Q)$  and  $k_E$  by the methods outlined above, but using streamflow data from just 1 year of the 5-year time series. I then used these estimates of  $g(Q)$  and  $k_E$  to generate synthetic hydrographs for each of the other 4 years of record. The 5 years encompass widely varying conditions, including the third wettest and third driest years in 33 years of precipitation records at the Severn [Marc and Robinson, 2007], with seasonal rainfall totals varying by more than a factor of two. Thus this can be considered a “differential split-sample test” in the terminology of Klemes [1986]. Such tests are still relatively uncommon in the modeling literature, and models often fail them [e.g., Seibert, 2003].

**Table 2.** Cross Validation: Parameter Values and Nash-Sutcliffe Efficiencies of Hourly Synthetic Hydrographs With Sensitivity Function  $g(Q)$  Estimated From Recession Plots for Individual Years<sup>a</sup>

Year(s) Tested Against	Year(s) Used to Estimate $g(Q)$ and $k_E$					
	1992	1993	1994	1995	1996	1992–1996
<i>Severn River N-S Efficiency</i>						
1992	<b>0.913</b>	0.934	0.906	0.843	0.910	0.914
1993	0.892	<b>0.906</b>	0.885	0.832	0.879	0.889
1994	0.928	0.940	<b>0.931</b>	0.885	0.927	0.930
1995	0.878	0.898	0.872	<b>0.785</b>	0.886	0.882
1996	0.924	0.938	0.922	0.850	<b>0.928</b>	0.927
1992–1996	0.911	0.929	0.908	0.846	0.911	<b>0.913</b>
<i>Severn River Parameter Values</i>						
$c_1$ (recession plot)	–2.381	–2.486	–2.408	–2.373	–2.502	–2.439
$c_2$ (recession plot)	1.076	0.780	1.023	1.132	0.750	0.966
$c_3$ (recession plot)	–0.068	–0.186	–0.082	0.0*	–0.165	–0.100
$k_E$ (calibrated)	0.487	0.560	0.574	0.331	0.530	0.525
<i>Wye River N-S Efficiency</i>						
1992	<b>0.864</b>	0.881	0.842	0.840	0.799	0.851
1993	0.905	<b>0.907</b>	0.903	0.891	0.868	0.897
1994	0.881	0.888	<b>0.913</b>	0.863	0.858	0.875
1995	0.830	0.865	0.763	<b>0.812</b>	0.752	0.815
1996	0.850	0.866	0.882	0.819	<b>0.795</b>	0.838
1992–1996	0.870	0.885	0.863	0.850	0.820	<b>0.859</b>
<i>Wye River Parameter Values</i>						
$c_1$ (recession plot)	–2.185	–2.278	–2.053	–2.200	–2.321	–2.206
$c_2$ (recession plot)	1.135	0.880	1.219	1.086	0.998	1.103
$c_3$ (recession plot)	0.0*	–0.079	0.0*	0.0*	0.0*	0.0*
$k_E$ (calibrated)	0.366	0.409	0.614	0.286	0.309	0.346

<sup>a</sup>Coefficients  $c_1$ ,  $c_2$ , and  $c_3$  in the empirical sensitivity function  $g(Q)$  (equation (9)) were estimated from quadratic linear regression on recession plots similar to Figure 6 for each individual year. The coefficient  $c_2$  is one less than the slope of the log-log plots in Figure 6, owing to the factor of  $Q$  in the denominator of equation (9). Where the quadratic parameter  $c_3$  was not statistically significant ( $p > 0.1$ ), it was set equal to zero (indicated in the table as 0.0\*), and ordinary linear regression was used to estimate  $c_1$  and  $c_2$ . The evapotranspiration scaling factor  $k_E$  was fitted by maximizing the goodness of fit (minimizing the sum of squared deviations) between the synthetic and observed hydrographs on logarithmic axes. Parameters  $c_1$ ,  $c_2$ ,  $c_3$ , and  $k_E$  were estimated for years corresponding to table columns, and model efficiencies were then calculated for years corresponding to table rows. Off-diagonal efficiencies (representing nontrivial cross validation) are quantitatively similar to on-diagonal efficiencies shown in bold (representing goodness of fit with the same time series that was used to estimate the parameters).

[47] The results of this exercise are shown in Table 2. The diagonal elements of the matrices indicate the Nash-Sutcliffe efficiencies for calibrations: that is, for cases where the function  $g(Q)$  and the coefficient  $k_E$  have been estimated for the same year that the predictions are subsequently tested against. Off-diagonal elements show model performance for nontrivial validation; that is, for cases where none of the test data have been used to estimate  $g(Q)$  and  $k_E$ . The off-diagonal and on-diagonal elements have similar values, indicating that this approach can successfully simulate hydrographs that it has not already been estimated from.

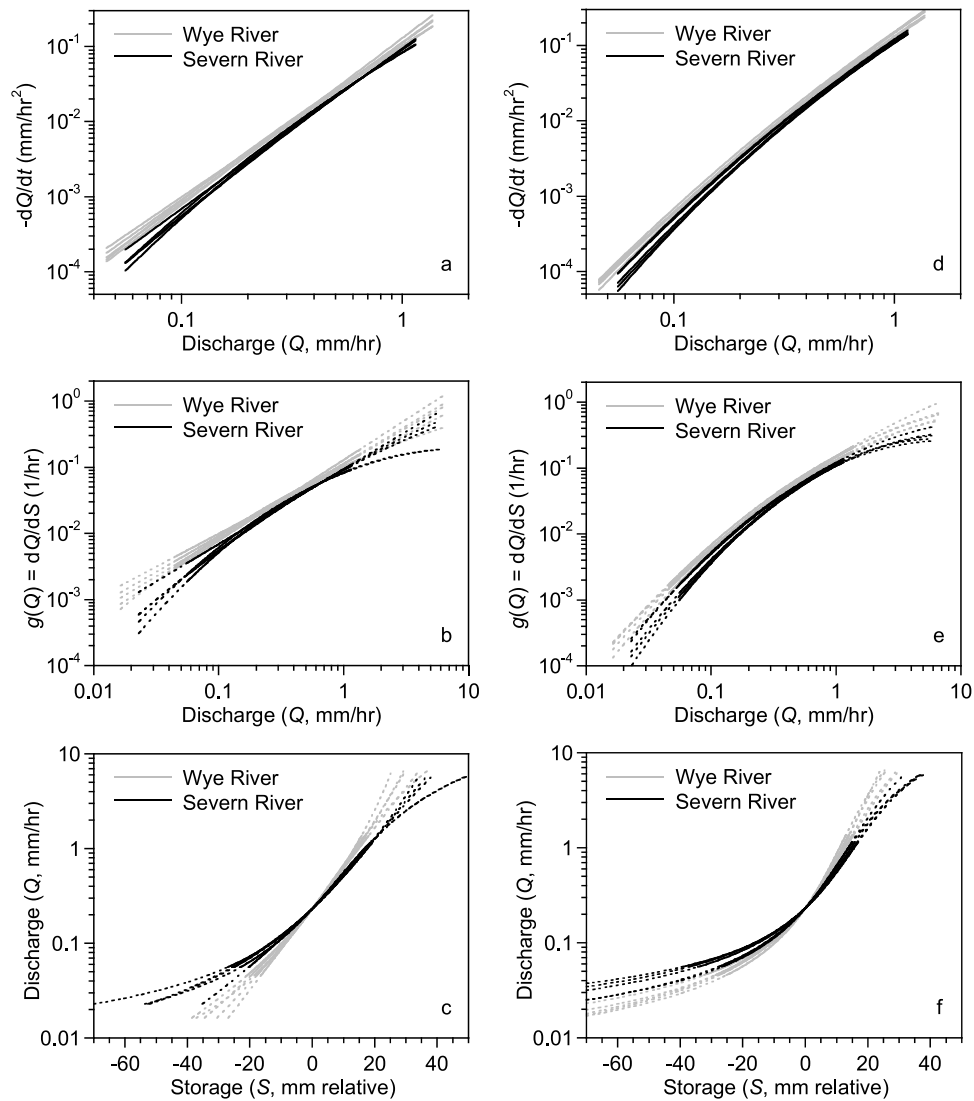
[48] Figure 11 shows the fitted curves derived from recession plots (as in Figures 6 and 7) for each of the 5 years, and the corresponding hydrological sensitivity functions  $g(Q)$  and the storage-discharge relationships that they imply. The recession plots, sensitivity functions, and storage-discharge relationships are roughly consistent from year to year (Figure 11). For example, the hydrological sensitivity functions for the Wye catchment are systematically greater than those of the Severn catchment across all of the years. Likewise, the storage-discharge relationships for the Wye catchment are distinctly steeper than those for the Severn catchment, regardless of which year's data are used to estimate them. The derived curves for the hydrologic sensitivity function and the storage-discharge relationship

diverge somewhat at the highest and lowest flows, as would naturally be expected because these flows involve extrapolations beyond the data in the recession plots.

[49] However, even with these extremes taken into account, Figure 11 implies that the sensitivity functions  $g(Q)$  and the storage-discharge relationships  $f(S)$  for the two catchments are reasonably stable characteristics of the catchments themselves, and are relatively insensitive to the idiosyncrasies of the particular data observed in any specific time interval. This is essential if we are to use  $g(Q)$  and  $f(S)$  for catchment characterization, or for operational forecasting of rainfall-runoff behavior. As Table 2 shows, parameter values estimated from one time period yield reasonable predictions of streamflow behavior for other periods with different climatic conditions, suggesting that this simple dynamical system may be useful for operational forecasting in some types of small catchments.

## 9. Direct Calibration to Rainfall-Runoff Time Series

[50] The recession plots shown in Figure 6 are an important tool for inferring the shape of the storage-discharge relationship  $f(S)$  or the catchment sensitivity function  $g(Q)$ . Nonetheless, if one is willing to take the functional form of  $g(Q)$  as given, equations (19) and (9) can be considered



**Figure 11.** (a) Fitted curves from recession plots and (b) their corresponding hydrological sensitivity functions and (c) storage-discharge relationships for 5 individual years (1992–1996) at Severn and Wye rivers (black and gray curves, respectively). Solid curves in Figures 11b and 11c indicate ranges of  $Q$  over which recession plots in Figure 11a were fitted (corresponding to black dots in Figure 6); dotted lines indicate extrapolations to annual average high and low flows. (d, e, f) Corresponding relationships for parameter values estimated by direct calibration to streamflow time series (see section 9).

as a simple four-parameter rainfall-runoff model that can be directly calibrated to time series of precipitation, evapotranspiration, and discharge. To test the utility of this approach, I jointly calibrated the three coefficients  $c_1$ ,  $c_2$ , and  $c_3$  in equation (9), along with the evapotranspiration scaling factor  $k_E$ , by minimizing the sum of squared deviations between the predicted and observed stream discharge on logarithmic axes. I calibrated a parameter set for each year individually, then used each parameter set to generate streamflow predictions for the other four years of record. This is the same cross-validation exercise conducted in section 8, except that here all of the parameters were determined by direct calibration against the time series of catchment discharge.

[51] The results of this cross-validation exercise show that this simple model performs almost equally well, both in verification tests against the same years that it was calibrated

with (the on-diagonal elements of Table 3) and in nontrivial validation tests against different years (the off-diagonal elements of Table 3). These results demonstrate that the model is reasonably robust.

[52] A comparison of Tables 2 and 3 shows that, unsurprisingly, the model fits the observed streamflow somewhat better when all four parameters are calibrated (Table 3) than when just one parameter ( $k_E$ ) is fitted to the streamflow time series (Table 2). What is perhaps surprising, however, is that the values of all four parameters in Table 3 are consistent across the different years of calibration. In many hydrological models, the parameters introduce more degrees of freedom than the data can adequately constrain, with the result that different combinations of parameters often give equally good fits to the calibration data (the equifinality problem [Beven and Binley, 1992]), and thus the parameter values are often sensitive to the idiosyncrasies of the



**Table 3.** Cross Validation: Parameter Values and Nash-Sutcliffe Efficiencies of Hourly Synthetic Hydrographs With Sensitivity Function  $g(Q)$  Estimated by Calibration to Time Series<sup>a</sup>

Year(s) Tested Against	Year(s) Used to Estimate $g(Q)$ and $k_E$					
	1992	1993	1994	1995	1996	1992–1996
<i>Severn River N-S Efficiency</i>						
1992	<b>0.951</b>	0.946	0.947	0.937	0.936	0.947
1993	0.929	<b>0.931</b>	0.928	0.913	0.910	0.924
1994	0.948	0.948	<b>0.950</b>	0.944	0.943	0.949
1995	0.887	0.880	0.894	<b>0.902</b>	0.902	0.900
1996	0.923	0.915	0.931	0.942	<b>0.942</b>	0.937
1992–1996	0.930	0.926	0.932	0.931	0.930	<b>0.934</b>
<i>Severn River Parameter Values</i>						
$c_1$ (calibrated)	–2.225	–2.168	–2.234	–2.120	–2.212	–2.197
$c_2$ (calibrated)	0.983	0.934	0.956	0.999	0.845	1.005
$c_3$ (calibrated)	–0.207	–0.247	–0.201	–0.166	–0.207	–0.174
$k_E$ (calibrated)	0.604	0.683	0.654	0.584	0.578	0.610
<i>Wye River N-S Efficiency</i>						
1992	<b>0.948</b>	0.946	0.938	0.938	0.936	0.945
1993	0.942	<b>0.944</b>	0.941	0.941	0.940	0.943
1994	0.935	0.936	<b>0.940</b>	0.933	0.939	0.938
1995	0.924	0.932	0.920	<b>0.935</b>	0.918	0.931
1996	0.928	0.931	0.939	0.928	<b>0.940</b>	0.936
1992–1996	0.937	0.940	0.938	0.938	0.937	<b>0.941</b>
<i>Wye River Parameter Values</i>						
$c_1$ (calibrated)	–2.019	–1.926	–1.901	–1.890	–2.074	–1.966
$c_2$ (calibrated)	1.024	1.071	1.177	1.027	1.021	1.068
$c_3$ (calibrated)	–0.159	–0.136	–0.098	–0.138	–0.125	–0.128
$k_E$ (calibrated)	0.710	0.676	0.766	0.643	0.767	0.708

<sup>a</sup>Coefficients  $c_1$ ,  $c_2$ , and  $c_3$  in the empirical sensitivity function  $g(Q)$  (equation (9)), along with the evapotranspiration scaling factor  $k_E$ , were jointly calibrated by maximizing the goodness of fit (minimizing the sum of squared deviations) between the synthetic and observed hydrographs on logarithmic axes. Parameters  $c_1$ ,  $c_2$ ,  $c_3$ , and  $k_E$  were estimated for years corresponding to table columns, and model efficiencies were then calculated for years corresponding to table rows. Off-diagonal efficiencies (representing nontrivial cross validation) are quantitatively similar to on-diagonal efficiencies shown in bold (representing calibration goodness of fit), and parameter values are broadly consistent for each site across all years, suggesting that the model is not overfitted to the calibration data sets.

calibration data. The values shown in Table 3 indicate that this is not the case, suggesting that the model is not overparameterized.

[53] The parameter values obtained by direct calibration (Table 3) are broadly consistent with those obtained from the recession plots (Table 2). Both sets of parameters indicate that the Wye catchment has higher  $g(Q)$  values and thus is more sensitive to changes in storage than the Severn (Figure 11). Both sets of parameters also indicate that  $g(Q)$  is more strongly curved in the Severn than the Wye, implying a somewhat shallower storage-discharge relationship  $f(S)$ . However, compared to the coefficients obtained from the recession plots, the parameter values obtained by time series calibration imply greater downward curvature (consistently more negative values of  $c_3$ ) in the sensitivity function  $g(Q)$  (compare Figures 11b and 11e). This in turn implies that the lower range of the storage-discharge relationship  $f(S)$  is flatter (compare Figures 11c and 11f), with the result that long-term streamflow recession will be slower under these parameter values.

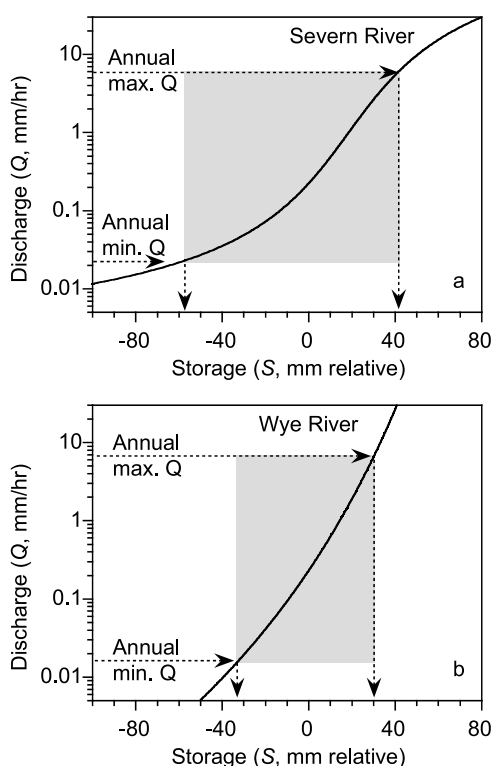
## 10. Catchment Characterization: Estimating Catchment Dynamic Storage

[54] Catchments can be usefully characterized by their dynamic storage, that is, their variation in storage between dry and wet periods [e.g., Kirby *et al.*, 1991; Uchida *et al.*, 2006; Spence, 2007]. The size of a catchment's dynamic

storage provides important insight into both vulnerability to flooding and sustainability of low flows. In principle, it should be straightforward to estimate dynamic storage by taking a running integral of the catchment mass balance (equation (1)). In practice, however, this integral will normally be subject to large errors, as small measurement biases and uncertainties in  $P$ ,  $Q$ , and  $E$  accumulate through time.

[55] If the catchment is characterized by a robust storage-discharge relationship, one can straightforwardly estimate the dynamic storage as the difference between two storage levels  $S_{\max}$  and  $S_{\min}$  corresponding to any two discharge rates  $Q_{\max}$  and  $Q_{\min}$ . The choice of  $Q_{\max}$  and  $Q_{\min}$  will depend on the time interval over which the dynamic storage is to be assessed. Here I estimate the dynamic storage on an annual time scale by using the averages of annual maximum and minimum flows; these are 0.023 and 5.81 mm/h at the Severn, and 0.016 and 6.54 mm/h at the Wye. One can then find the dynamic storage between these discharge values, either by inverting the storage-discharge relationship  $f(S)$ , or by integrating the reciprocal of the hydrologic sensitivity function  $g(Q)$ ,

$$\begin{aligned}
 S_{\max} - S_{\min} &= f^{-1}(Q_{\max}) - f^{-1}(Q_{\min}) = \int_{Q_{\min}}^{Q_{\max}} \frac{dS}{dQ} dQ \\
 &= \int_{Q_{\min}}^{Q_{\max}} \frac{1}{g(Q)} dQ.
 \end{aligned} \tag{20}$$



**Figure 12.** Dynamic catchment storage, estimated as the difference between storage levels corresponding to maximum and minimum discharges. Here the means of annual maximum and minimum flows are used and yield dynamic storage of approximately (a) 98 mm at the Severn and (b) 62 mm at the Wye. Solid black curves show storage-discharge relationships estimated from recession plots (Figure 6). Storage measures are relative rather than absolute; axes here show storage relative to storage at mean discharge.

As Figure 12 shows, this procedure yields an annual dynamic storage of approximately 98 mm at the Severn and 62 mm at the Wye, if the parameters of  $g(Q)$  are estimated from the recession plots (Figure 6). If the parameter values are estimated by direct calibration to the streamflow time series as in section 9 above, the annual dynamic storage estimates are somewhat greater (124 mm at the Severn and 107 mm at the Wye), because the inferred storage-discharge relationship is somewhat flatter.

[56] These estimates of dynamic storage roughly agree with estimates from field measurements made at Plynlimon during the 1970s and 1980s. Annual ranges of soil moisture, as measured by neutron probe methods, averaged  $58 \pm 30$  mm (mean  $\pm$  standard deviation) over the 8 years from 1974 through 1981; over the same time period, annual changes in geological storage, estimated by a running mass balance, averaged  $70 \pm 28$  mm (data extracted from Kirby *et al.* [1991, Figure 23, p. 55]). These field measurements argue for the general plausibility of the dynamic storage estimates derived above, but close quantitative agreement should not be expected, because the neutron probe measurements may not be representative of the whole catchment

[Kirby *et al.*, 1991], and running mass balances are vulnerable to accumulating errors, as described above.

[57] Over longer spans of time, wider ranges of climatic conditions may be encountered, leading to correspondingly wider ranges of storage levels and streamflows than would be encountered for any given year. For example, over the 27-year record from 1974 through 2000, flows at the Severn varied from 0.008 to 11.3 mm/h and flows at the Wye varied from 0.008 to 9.3 mm/h. Using these discharge ranges, equation (20) yields dynamic storage estimates of approximately 190 mm at the Severn and 95 mm at the Wye, roughly 1.5–2 times the range in storage that was calculated for an average year.

[58] Equation (20) can also be used to account for changes in catchment storage when estimating evapotranspiration by mass balance methods. In such applications  $Q_{\max}$  and  $Q_{\min}$  would be replaced by the discharges at the beginning and end of the interval over which cumulative precipitation and discharge have been measured, and for which cumulative evapotranspiration is to be estimated.

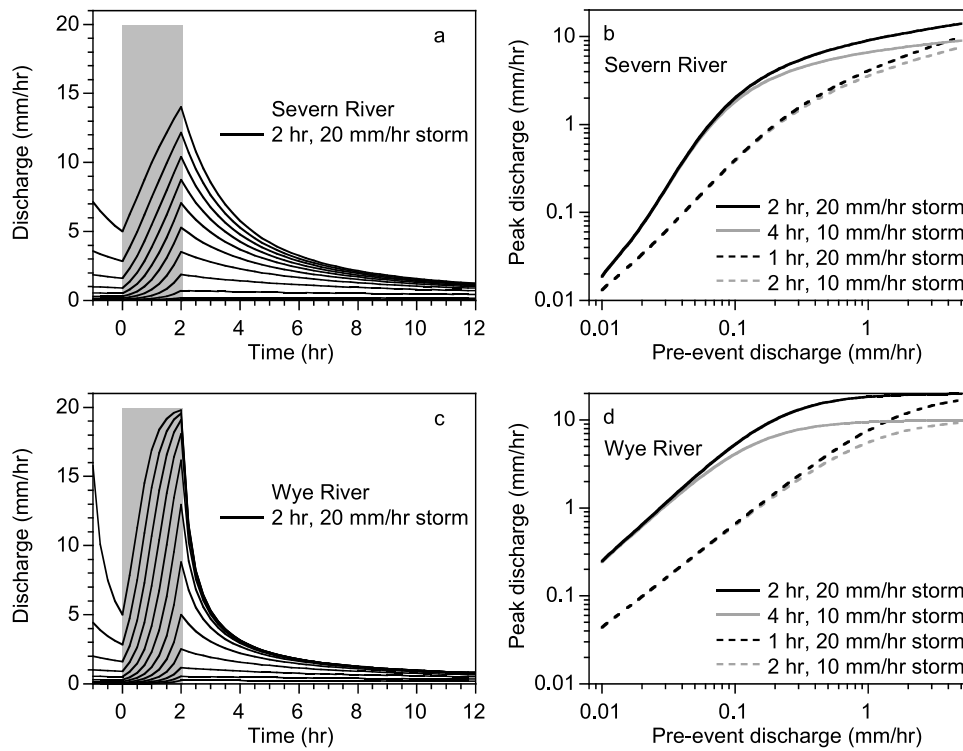
[59] The dynamic storage, as estimated here, will be less than the total storage because catchments can retain significant volumes of residual water, even under drought conditions. If this residual storage does not have a measurable effect on streamflow, it cannot be estimated from hydro-metric methods like those outlined here, but can only be estimated from conservative chemical or isotopic tracers. The strong damping of tracer fluctuations observed in streamflow relative to precipitation at Plynlimon implies either large volumes of residual storage or strong dispersive mixing in the subsurface [Neal and Rosier, 1990; Kirchner *et al.*, 2000, 2001].

## 11. Catchment Characterization: Estimating Sensitivity to Antecedent Moisture

[60] Hydrologists have long recognized that the antecedent moisture status of a catchment has a strong effect on its storm runoff response. Antecedent moisture has been a major challenge for hydrological prediction, for two reasons: (1) it has been difficult to accurately estimate the moisture status of a catchment through time, by either measurement or modeling, and (2) it has been difficult to quantify the functional relationship between this antecedent moisture and storm runoff.

[61] In catchments where discharge is a function of storage, the approach outlined above directly solves both of these problems. If discharge is a function of storage, then the catchment's antecedent moisture (i.e., storage) will be implicitly measured by stream discharge, and the catchment's response to a unit increase in storage will be directly quantified by  $g(Q)$ . The hydrologic sensitivity function  $g(Q) = dQ/dS$  directly expresses the effects of antecedent moisture, by quantifying the change in discharge ( $dQ$ ) that accompanies a given change in storage ( $dS$ ), at a given level of storage and its accompanying discharge ( $Q$ ).

[62] Because both discharge and  $g(Q)$  will change as a storm progresses, accurately estimating storm runoff will require integrating equation (18) or equation (19) through time. Figure 13 shows simulated hydrographs for a hypothetical 2-h, 20 mm/h storm, indicated by the gray shaded region in Figures 13a and 13c. Each trace in the "fan" of hydrographs corresponds to a different level of antecedent



**Figure 13.** (a, c) Simulated storm hydrographs for a 2-h, 20 mm/h storm under different levels of antecedent moisture and therefore different preevent discharge rates and (b, d) relationship between peak discharge and preevent discharge for storms of different durations and intensities. Hydrographs were simulated by fourth-order Runge-Kutta integration of equation (19). Gray shaded region in Figures 13a and 13b indicates assumed precipitation input.

moisture, and thus a different streamflow at the onset of the storm. The peak discharge varies systematically with the preevent discharge, and with the duration and intensity of the storm, as shown in Figures 13b and 13d. The dotted and solid curves in Figures 13b and 13d show results for storms with total volumes of 20 mm and 40 mm, respectively; the gray and black curves indicate rainfall intensities of 10 mm/h and 20 mm/h, respectively. Lookup tables or nomograph plots like those shown in Figure 13 could be used as an operational guide to estimating peak storm discharge. In this regard, one should recognize that the recession plots used to estimate  $g(Q)$  extend only to discharges of roughly 1–1.5 mm/h, and the higher peak flows in Figure 13 involve substantial extrapolation beyond this range. However, the synthetic hydrographs shown in Figure 9 indicate that equation (19) can simulate stormflows reasonably accurately, well beyond the range of the recession plots. Thus Figures 13b and 13d represent a proof-of-concept demonstration showing that peak storm discharge can be straightforwardly estimated from the sensitivity function  $g(Q)$ , using preevent discharge as an implicit measure of antecedent moisture.

## 12. Catchment Characterization: Estimating Characteristic Recession Time “Constants”

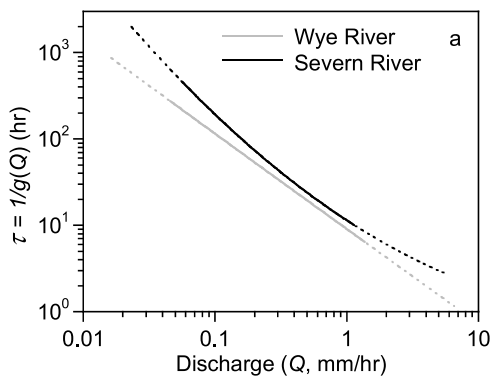
[63] Catchment recession behavior is often described by a characteristic time constant, the  $e$ -folding time of the exponential decay of discharge during recession. Conventional recession theory shows that if discharge is a linear

function of storage ( $Q = kS$ ), under recession conditions ( $P \approx 0$ ,  $E \approx 0$ ) discharge will decline exponentially as a function of time:  $Q = Q_0 e^{-kt} = Q_0 e^{-t/\tau}$ . The rate of the exponential decay can be measured by the “recession constant”  $k$ , which has dimensions of 1/time, or more intuitively by its reciprocal, the “recession time constant”  $\tau$ , which has dimensions of time. Graphically,  $k$  is the slope of the recession hydrograph on log linear axes, and  $\tau$  is its reciprocal. In the more general case where discharge is a nonlinear function of storage, a log linear plot of the recession hydrograph will no longer be a straight line, and the “constants”  $k$  and  $\tau$  will no longer be constant, but instead will vary as the catchment drains. Nonetheless, the concept of characteristic recession time remains useful for describing how rapidly streamflow declines during recession.

[64] From equation (19), we can see directly that the log linear slope of the recession hydrograph is simply

$$\frac{d(\ln(Q))}{dt} = g(Q) \left( \frac{P - E}{Q} - 1 \right) \Big|_{P \approx 0, E \approx 0} \approx -g(Q), \quad (21)$$

and thus the recession constant  $k$  will be equal to  $g(Q)$  and the characteristic recession time constant  $\tau$  will be equal to  $1/g(Q)$ . Obviously neither of these “constants” will be constant unless  $g(Q)$  is a constant (as it would be if the storage-discharge relationship were linear). Instead, as Figure 14 shows, the characteristic recession time “constant” varies by roughly 3 orders of magnitude within the



**Figure 14.** Characteristic time “constant” of recession, as a function of discharge, for Severn and Wye rivers. Solid curves indicate ranges of  $Q$  over which  $g(Q)$  was fitted to data in recession plots (Figure 6); dotted curves indicate extrapolations to annual average high and low flows.

annual range of flows at both the Severn and Wye, from hours at high flows, to thousands of hours at low flows. Even at low flows, the recession time “constant”  $\tau$  gives no indication of actually becoming constant.

[65] The sensitivity to antecedent moisture described in Figure 13 and the time-varying recession “constants” documented in Figure 14 are two interrelated consequences of the strong nonlinearity in the storage-discharge relationships at Plynlimon. This nonlinearity is inconsistent with many time series methods commonly used in hydrology, such as unit hydrographs and related transfer function methods. Such methods assume that streamflow is a lagged, linear additive function of precipitation inputs, or equivalently that streamflow responses to precipitation inputs are independent of antecedent moisture. These assumptions will not be met in catchments that are characterized by a nonlinear relationship between discharge and storage (as indicated, for example, by a slope that differs from 1 in recession plots like Figure 6). Hydrologists have long recognized that catchments’ hydrologic response is often nonlinear, [e.g., Linsley *et al.*, 1982; Tallaksen, 1995; Wittenberg, 1999; Brutsaert, 2005], calling into question the time series methods and conceptual hydrological models that are based on this premise.

### 13. Doing Hydrology Backward: Estimating Catchment-Averaged Precipitation Rates From Streamflow Fluctuations

[66] The analysis presented above has thus far focused on the conventional problem of hydrological prediction, namely constructing a synthetic streamflow hydrograph from specified precipitation and evapotranspiration time series. As shown above, reasonable results can be obtained by representing the catchment as a simple first-order nonlinear dynamical system (equation (18)), characterized by the sensitivity function  $g(Q)$ . The marked simplicity of this system also makes it potentially useful for an entirely different class of questions. The system is simple enough that it is invertible; therefore, it can be used to infer temporal patterns of precipitation and evapotranspiration at small-catchment scale, using measured streamflow fluctuations as input.

Rearranging the terms of equation (18), one directly obtains

$$P - E = \frac{dS}{dt} + Q = \frac{dQ/dt}{dQ/dS} + Q = \frac{dQ/dt}{g(Q)} + Q. \quad (22)$$

The entire right-hand side of equation (22) can be calculated directly from the streamflow time series if the function  $g(Q)$  is known. As outlined above,  $g(Q)$  can be estimated directly from the streamflow time series, without measurements of either  $P$  or  $E$ ; one needs only to identify periods when  $P$  and  $E$  are both small compared to discharge, but their exact values are unimportant. Therefore equation (22) can be used to calculate a time series of  $(P - E)$  directly from the streamflow time series, independent of measurements of  $P$  or  $E$ .

[67] How is this possible? Inferring rainfall patterns from streamflow fluctuations has typically been considered infeasible, because of the problem of accounting for changes in catchment storage. However, if discharge is a function of storage, then changes in discharge reflect changes in storage; they are related to one another through  $g(Q)$ , the local gradient of the storage-discharge relationship. Equation (22) is just the conservation of mass equation, in which the rate of change of storage has been reexpressed as the rate of change of discharge, divided by the sensitivity of discharge to changes in storage,  $g(Q)$ . As long as one knows the sensitivity of discharge to changes in storage, one knows how much storage must change to produce a particular measured change in discharge.

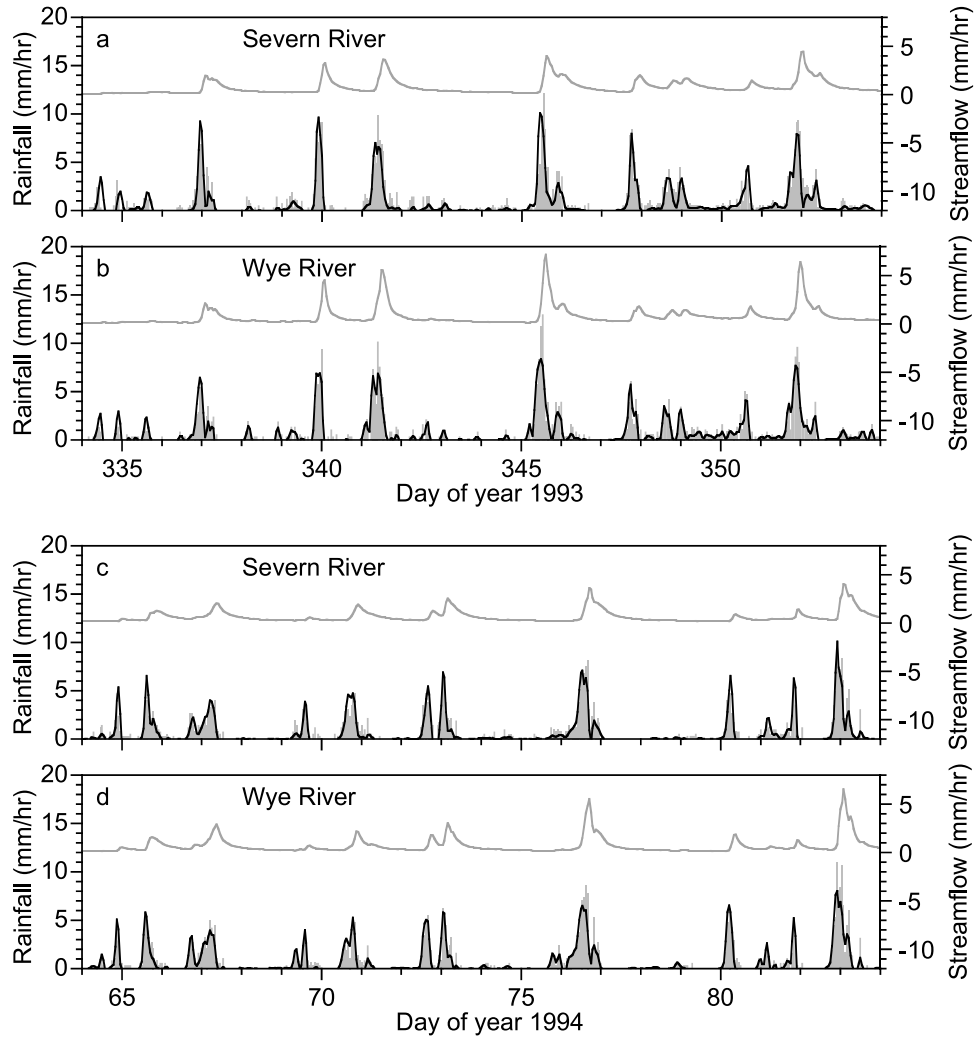
[68] Implementing equation (22) in practice will require attention to several details. As described in section 7, there is a time lag between changes in discharge from the hillslope and changes in streamflow at the weir; at Plynlimon, these lags are approximately 1 h on average. Because  $dQ/dt$  and  $g(Q)$  must be estimated from streamflow at discrete points in time, one must choose those points carefully in order to give the right time shift between precipitation and streamflow. Also, because discharge can change rapidly and  $g(Q)$  can be a steep function of  $Q$ , it is important to average  $g(Q)$  over the discharge measurements used to estimate  $dQ/dt$ . Otherwise, if (for example) discharge rises substantially from a very small initial value, but only the initial value of  $Q$  is used in the denominator of equation (22),  $g(Q)$  will be too small to adequately represent catchment sensitivity over the time interval, leading to unrealistically large inferred changes in storage and inferred precipitation rates. These considerations lead to the following formula for inferring  $(P - E)$ :

$$P_t - E_t \approx \frac{(Q_{t+\ell+1} - Q_{t+\ell-1})/2}{[g(Q_{t+\ell+1}) + g(Q_{t+\ell-1})]/2} + (Q_{t+\ell+1} + Q_{t+\ell-1})/2, \quad (23)$$

where  $\ell$  is the travel time lag for changes in discharge to reach the weir.

[69] Unfortunately, the left-hand side of equation (22) contains both  $P$  and  $E$ . Ideally one would like to be able to isolate  $P$  alone. This cannot be done exactly, but it can be approximated using the following argument. Whenever it is raining at humid catchments like Plynlimon, relative humidity should be high and evapotranspiration rates should





**Figure 15.** Hourly measured streamflow (gray curve and right axis) and rainfall rates inferred from streamflow fluctuations (solid black curves and left axis) using equation (24), compared with measured rainfall (gray shaded regions and left axis) for Severn and Wye rivers during 20-day periods in (a, b) December 1993 and (c, d) March 1994. Measured rainfall is the average of two automated weather stations in each catchment. Streamflow and rainfall axes are set to the same scale so that they can be compared but are offset for clarity. Even when rainfall events produce only a small streamflow response, streamflow fluctuations yield accurate estimates of rainfall timing and magnitude.

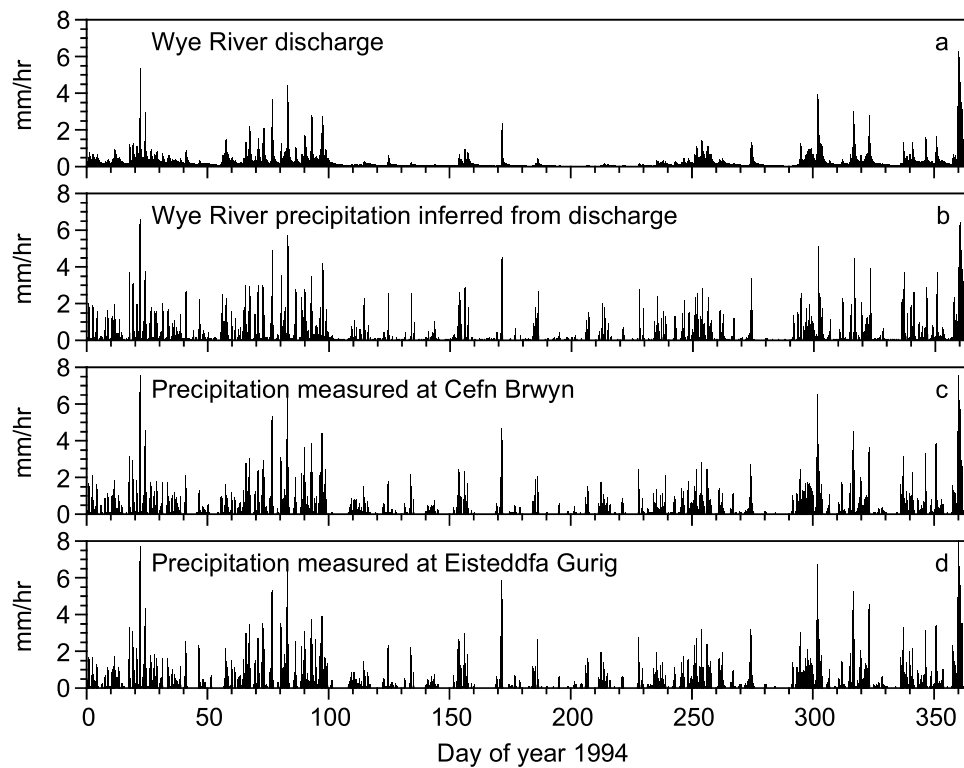
be correspondingly low. Thus to a first approximation, whenever  $P - E$  is greater than zero, one can assume that  $P - E \approx P$ . Thus we can estimate rainfall rates as

$$P_t \approx \max\left(0, \frac{(Q_{t+\ell+1} - Q_{t+\ell-1})/2}{[g(Q_{t+\ell+1}) + g(Q_{t+\ell-1})]/2} + (Q_{t+\ell+1} + Q_{t+\ell-1})/2\right). \quad (24)$$

Why would one want to be able to infer precipitation rates from streamflow fluctuations? Precipitation rates vary in space and time, and conventional rain gauges are much smaller than the catchments they are used to represent. For example, the Plynlmion catchments are of order  $10 \text{ km}^2$  in area, whereas each of the rain gauges used to measure precipitation rates at Plynlmion are only  $0.00000003 \text{ km}^2$  in area, more than 8 orders of magnitude smaller. Thus rainfall rates measured at any individual rain gauge, or over any

sparse network of rain gauges, may not accurately represent precipitation inputs to the catchment as a whole. By contrast, equations (22)–(24) allow one to estimate precipitation rates at the scale of the landscape rather than the scale of the rain gauge.

[70] Figure 15 illustrates the use of equation (24) to estimate precipitation patterns. As Figure 15 shows, equation (24) estimates the timing, magnitude, and duration of precipitation events at Plynlmion reasonably well. This is not just a case of precipitation inputs equaling stream outputs when the catchment is thoroughly wet; instead, as Figure 15 shows, the inferred rainfall rates are realistic even when streamflow is a small fraction of precipitation. Inferred rainfall rates can be many times higher than streamflow rates because when discharge is low, discharge is relatively insensitive to changes in storage (that is,  $g(Q)$  is small), and therefore a given increase in discharge implies relatively large changes in storage (and thus relatively high rainfall



**Figure 16.** (a) Six-hour average Wye River streamflow and (b) precipitation rates inferred from streamflow, compared to 6-h average precipitation rates measured by automated weather stations at (c) Cefn Brwyn and (d) Eisteddfa Gurig, located near the bottom and top of the Wye catchment, respectively. Data for 1994 are shown; other years are similar.

rates). Note also that in general, peaks in both inferred and observed rainfall rates correspond to the times of the most rapid increase in streamflow, reflecting the  $90^\circ$  phase lag between precipitation and streamflow that is intrinsic to the dynamical system formed by equations (1) and (2).

[71] Figure 16 shows precipitation rates inferred from streamflow in the Wye River for an entire year (in this case 1994; other years are qualitatively similar), along with precipitation rates measured at the two automated weather stations in the Wye catchment, Cefn Brwyn and Eisteddfa Gurig (circles numbered 3 and 4 in the location map in Figure 1). As Figure 16 shows, rainfall rates inferred from discharge are similar to rainfall rates measured in the two rain gauges. The inferred rainfall rates generally reflect the

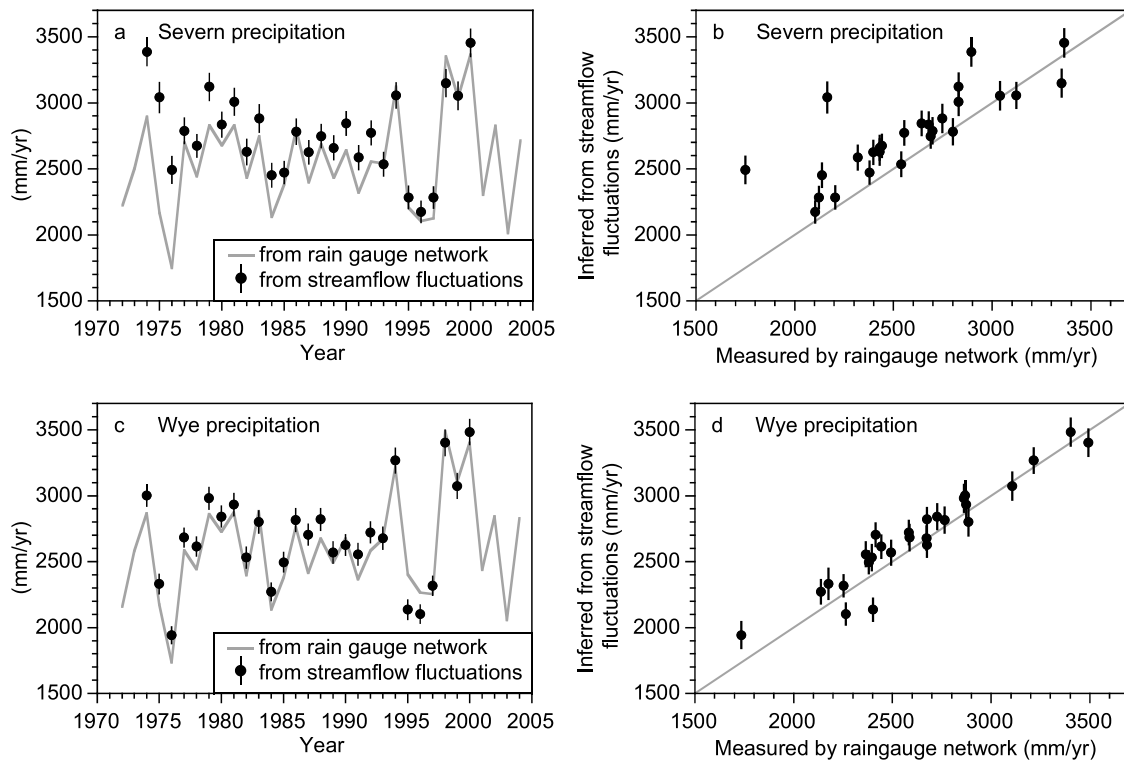
observed timing, magnitude, and duration of rainfall events, even during relatively dry periods in the summer, when streamflow is low for weeks at a time.

[72] The inferred rainfall rates do not exactly match either rain gauge, but neither do the two rain gauges exactly match each other. This observation suggests that the agreement between the two rain gauges could be used as a benchmark for assessing the agreement between the measured and inferred rainfall rates. Table 4 shows the correlation between the measured rainfall rate (i.e., the mean of the two rain gauge records) and the inferred rainfall rate in each catchment, averaged over time periods ranging from 1 h to 1 day. As the averaging period becomes longer, the correlation between inferred and measured rainfall rates becomes

**Table 4.** Correlations Between Observed Precipitation Rates and Those Inferred From Streamflow Fluctuations Compared With Correlations Between Precipitation Rates Observed at the Two Rain Gauges in Each Catchment<sup>a</sup>

Averaging Period	Severn River		Wye River	
	<i>P</i> Inferred From Streamflow Versus Average of Gauges	Two Rain Gauges: Carreg Wen Versus Tanllwyth	<i>P</i> Inferred From Streamflow Versus Average of Gauges	Two Rain Gauges: Cefn Brwyn Versus Eisteddfa Gurig
1 h	<b><i>r</i> = 0.811</b>	<i>r</i> = 0.787	<b><i>r</i> = 0.879</b>	<i>r</i> = 0.877
3 h	<b>0.891</b>	0.884	<b>0.942</b>	0.938
6 h	<b>0.920</b>	0.913	<b>0.953</b>	0.956
12 h	<b>0.943</b>	0.930	<b>0.964</b>	0.967
24 h	<b>0.955</b>	0.938	<b>0.970</b>	0.973

<sup>a</sup>Correlations with rates inferred from streamflow fluctuation are shown in bold.



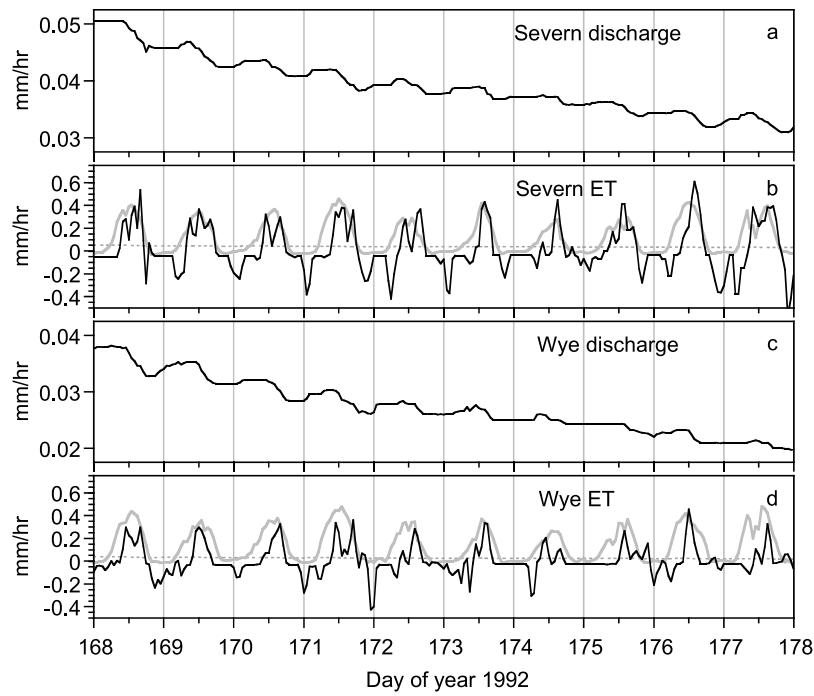
**Figure 17.** Annual precipitation totals inferred from streamflow fluctuations (black dots) for Severn and Wye catchments, compared to annual precipitation captured in a dense network of storage gauges [Marc and Robinson, 2007].

stronger, because small discrepancies in timing become less consequential. As a benchmark for comparison, Table 4 also shows the correlation between the precipitation rates measured at the two rain gauges in each catchment. In general, the correlations between the measured and inferred rainfall rates are similar to the correlations between the two rain gauges. In other words, the inferred rainfall rates agree with the rain gauges roughly as well as the two rain gauges agree with each other.

[73] The comparison between inferred and measured precipitation rates is a strong test of the underlying theory. Equations (22)–(24) are not calibrated in any way to the precipitation data that they are tested against, because the sensitivity function  $g(Q)$  is estimated from the streamflow time series alone. Thus Figures 15 and 16 and Table 4 are a completely independent test of the theory. Furthermore, in an information-theoretic sense, precipitation is a more information-rich time series than streamflow, which is smoother and thus more redundant with itself through time. Therefore the precipitation time series provides a more richly detailed set of observations for the theory to be tested against. The history of hydrology shows that many different rainfall-runoff models can successfully take an information-rich precipitation time series and smooth it out to make a realistic-looking, information-poor streamflow time series. It is less obvious that this process should be reversible, such that subtle fluctuations in the streamflow time series yield realistic estimates of precipitation rates through time. It is even less obvious that this should be possible without any calibration to the precipitation time series, yet this is what equations (22)–(24) do.

[74] Because precipitation at Plynlimon is relatively uniform in space (as indicated by the strong correlations between the pairs of rain gauges), the size of the effective “footprint” of the inferred precipitation estimate is unclear; does it encompass the entire catchment, or a more limited area adjacent to the stream network? In either case, the effective footprint of the precipitation estimates is orders of magnitude larger than conventional rain gauges, and more widely distributed across the landscape. Such spatially integrated precipitation estimates are potentially useful for “scaling up” individual rain gauge records to the scale of small catchments. Likewise, precipitation time series inferred from streamflow may provide useful “ground truth” for radar-based precipitation estimates, which have a typical pixel size on the same order as small catchments (but many orders of magnitude larger than conventional rain gauges). Precipitation estimates from streamflow may also be useful in estimating catchment inputs that are difficult to measure directly, such as snowmelt, canopy interception, or fog drip.

[75] Precipitation estimates derived from streamflow fluctuations may also be useful in reconstructing past precipitation, in cases where streamflow records are available but precipitation records are not. To explore this possibility, I used hourly streamflow records from 1974 through 2000 to reconstruct hourly precipitation estimates for the two catchments over those years, using equation (24). I then aggregated these precipitation estimates to annual totals, and compared them against annual precipitation totals from a dense network of monthly read storage rain gauges in the two catchments [Marc and Robinson, 2007]. The sensitivity functions  $g(Q)$  for the two catchments were the same ones



**Figure 18.** (a, c) Fluctuations in Severn and Wye River discharge during an extended rainless interval in June 1992. (b, d) Evapotranspiration rates inferred from these streamflow fluctuations (black curves) and Penman-Monteith evapotranspiration estimates calculated from automatic weather station data (gray curves). The discharge fluctuations are extremely small; the axes of Figures 18a and 18c are enlarged by a factor of 50 relative to Figures 18b and 18d. Dotted gray lines in Figures 18b and 18d show stream discharge drawn to scale. Vertical bars mark midnight.

estimated earlier from the recession plots (Figure 6). As Figure 17 shows, the inferred precipitation totals agree almost exactly with the rain gauge measurements in the Wye catchment. In the Severn catchment, the inferred precipitation totals are strongly correlated with the rain gauge measurements, but are 150–200 mm/a higher, on average. As before, the inferred precipitation estimates are not calibrated in any way to the rain gauge measurements. Thus the results shown in Figure 17 provide strong support for the underlying theory, and suggest that these methods may be useful for inferring precipitation rates where direct measurements are not available.

#### 14. Doing Hydrology Backward: Inferring Evapotranspiration Patterns From Streamflow Fluctuations

[76] Because streamflow fluctuations quantitatively reflect precipitation inputs to the catchment, as shown above, it is natural to ask whether streamflow fluctuations also reflect evapotranspiration losses. Hydrologists have developed several strategies over the past few decades for using discharge measurements during streamflow recession to infer catchment-scale evapotranspiration rates [e.g., *Tschinkel*, 1963; *Daniel*, 1976; *Brutsaert*, 1982; *Boronina et al.*, 2005; *Szilagyi et al.*, 2007]. In the dynamical system described by equations (1) and (2), precipitation and evapotranspiration have comparable but opposite effects on catchment storage and thus on streamflow. This raises the possibility that streamflow fluctuations can be used to infer temporal patterns of landscape-scale evapotranspiration as well as precipitation. It is therefore interesting to test how

well streamflow reflects evapotranspiration rates, even at Plynlimon, where evapotranspiration is a relatively small fraction of the water balance (Table 1).

[77] The catchment mass balance can be rewritten to express the rate of evapotranspiration as

$$E = P - Q - \frac{dS}{dt} = P - Q - \frac{dQ/dt}{g(Q)}. \quad (25)$$

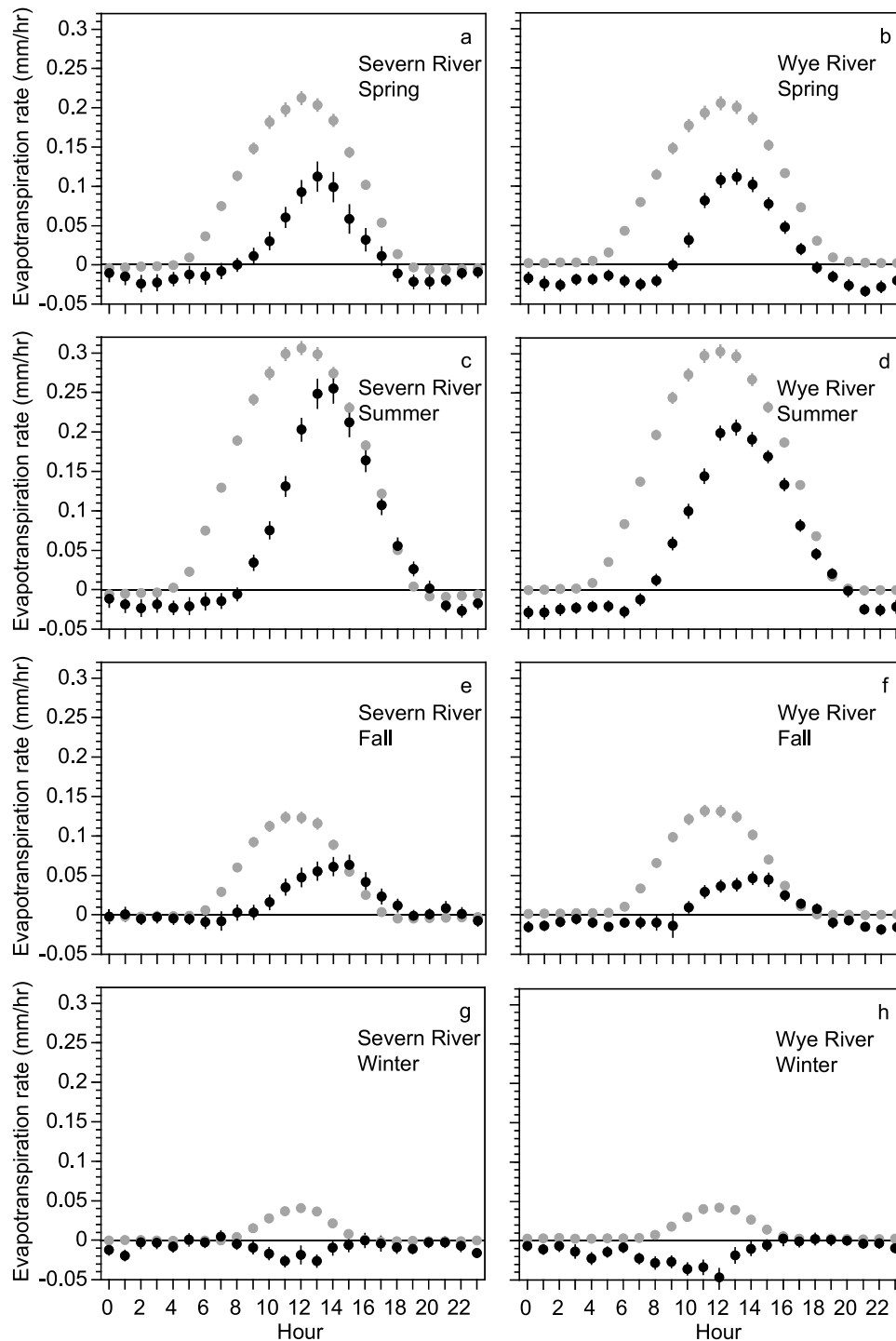
It would appear logical to use equation (25) to estimate evapotranspiration rates directly from measured rainfall rates and streamflow fluctuations. In practice, however, the uncertainty in  $P$  during rainfall events will be many times bigger than  $E$  (which is on the order of 0.1 mm/h), making the direct application of equation (25) impractical.

[78] Instead, the approach adopted here is to restrict the analysis to rainless periods (defined for Plynlimon as periods when no precipitation is recorded in any of the four rain gauges during the previous 6 h or the following 2 h, as in section 5 above). Under those conditions, one can assume that  $P$ , and thus the uncertainty in  $P$ , must be small. Then the same considerations used to derive equation (24) lead to the following expression for inferring evapotranspiration rates from streamflow fluctuations:

$$E_t|_{P=0} = -\frac{dQ/dt}{g(Q)} - Q \approx -\frac{(Q_{t+\ell+1} - Q_{t+\ell-1})/2}{[g(Q_{t+\ell+1}) + g(Q_{t+\ell-1})]/2} - (Q_{t+\ell+1} + Q_{t+\ell-1})/2, \quad (26)$$

where, as in equation (23),  $\ell$  is the travel time lag for changes in discharge to reach the weir.





**Figure 19.** (a–h) Hourly mean evapotranspiration rates estimated from streamflow fluctuations (black dots) and Penman-Monteith potential evaporation (gray dots) by season of year. Dots show means  $\pm 1$  standard error for all rainless hours from 1992 through 1996. In order to minimize the effects of outliers, averages for winter are calculated as 98% trimmed means (i.e., excluding the highest and lowest 1% of observations). (a, b) Spring is March–May, (c, d) summer is June–August, (e, f) fall is September–November, and (g, h) winter is December–February.

[79] Figure 18 shows evapotranspiration rates inferred from streamflow using equation (26) during an extended rainless interval at Plynlimon, illustrating both the potential and the pitfalls of this approach. Superimposed on a gradual discharge recession, one can see fluctuations in which flow typically declines from morning through afternoon, and

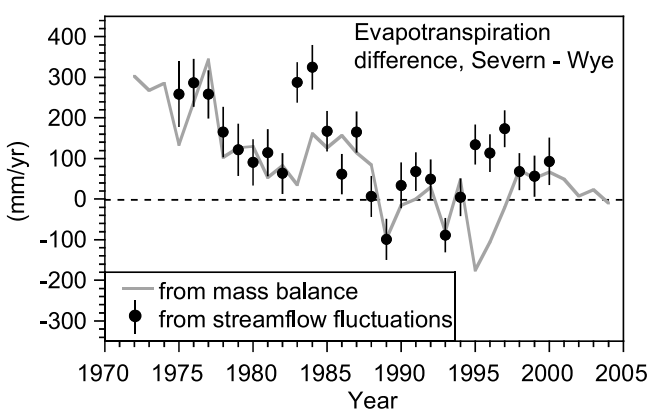
then remains constant or rises slightly during the evening and night (Figures 18a and 18c). These fluctuations are interpreted by equation (26) as reflecting diurnal variations in evapotranspiration rates (black lines, Figures 18b and 18d): evapotranspiration is greatest during the middle of the

day, leading to a decline in catchment storage and a corresponding reduction in streamflow.

[80] However, the discharge measurements must be very precise or else the discharge fluctuations will be obscured by measurement noise, because the fluctuations in discharge are tiny. In fact, to make the discharge fluctuations visible in Figure 18, the discharge axes have been expanded by fiftyfold relative to the evapotranspiration plots. The discharge time series are shown to scale as the gray dotted lines in Figure 18b and 18d; on that scale, the discharge fluctuations are invisible. Because discharge is low, the catchment sensitivity function  $g(Q)$  is small as well, reflecting the fact that when the catchment is dry, discharge is relatively insensitive to changes in storage. Thus small changes in discharge imply large changes in storage; in equation (26), this has the effect of amplifying the real fluctuations in discharge, but also amplifying the measurement noise. As a result, the inferred diurnal cycles of evapotranspiration are noisy. Nonetheless, they roughly agree with the magnitude and timing of the diurnal cycles in Penman-Monteith potential evaporation, shown as the gray lines in Figures 18b and 18d.

[81] Reducing the noise in the inferred evapotranspiration rates would require measuring subtle streamflow fluctuations very accurately, which the Plynlimon stream gauging stations were not designed to do. Given the noisy data that are available, however, one can still clarify the diurnal patterns in evapotranspiration by averaging over many daily cycles. Figure 19 shows the inferred evapotranspiration rates for each hour of the day in each season, averaged over all rainless hours from 1992 through 1996. The corresponding cycles in Penman-Monteith potential evaporation are shown in gray for comparison. As Figure 19 shows, the evapotranspiration cycles inferred from streamflow fluctuations are broadly consistent with the Penman-Monteith estimates, although both the timing and amplitude differ somewhat. It should be kept in mind that the Penman-Monteith estimates of potential evaporation are not themselves a “gold standard,” because it is not clear how they should be extrapolated to the scale of the landscape. Nonetheless, both the Penman-Monteith estimates and the inferred evapotranspiration rates reach their peak near the middle of the day and fall to near zero at night. Predicted and observed diurnal cycles vary similarly from one season to the next, with the exception of the winter, when the inferred evapotranspiration cycle is weak but appears to be inverted from the Penman-Monteith estimates. That is, on average there appears to be a small input of water to the catchment near noon; this could potentially reflect melting of frost or snow during the middle of the day.

[82] The strongest conclusion one can draw from Figures 18 and 19 is that equation (26) provides at least semiquantitative estimates of evapotranspiration rates. Nonetheless, the fact that this works at all, even at Plynlimon where evapotranspiration rates are relatively low and discharge fluctuations are correspondingly tiny, represents a strong test of the dynamical systems approach developed here. It is important to remember that the only parameters in equation (26) are those embedded in the sensitivity function  $g(Q)$ , and that  $g(Q)$  is estimated from recession plots that contain no information about evapotranspiration rates. Therefore evapotranspiration estimates like those shown in



**Figure 20.** Difference between annual evapotranspiration rates in the Severn and Wye catchments, inferred from streamflow fluctuations (black dots with standard errors), and from mass balances (gray curve). Mass balances were calculated from precipitation recorded in an extensive network of monthly read storage rain gauges and from streamflows at the Severn flume and Wye weir (data from *Marc and Robinson* [2007]). Black dots are trimmed means of hourly evapotranspiration rates (upper and lower 1% of data excluded to minimize the effect of outliers), calculated from streamflow fluctuations using equation (26) during rainless periods (defined as zero rainfall recorded at all automated weather stations from the previous 6 h through the following 2 h; periods with fewer than two operational weather stations were excluded). Inferred evapotranspiration rates were averaged for each season and then were aggregated to annual time scales; standard errors were calculated by first-order, second-moment error propagation.

Figure 18 and 19 are not calibrated to observed evapotranspiration rates in any way, and thus represent a completely independent test of the theory.

[83] Even in the most general qualitative sense, one sees the signature of the dynamical system in temporal patterns shown in Figure 18. One can see that in general, there is a  $90^\circ$  phase lag between the variations in evapotranspiration rates and the fluctuations in streamflow. Because evapotranspiration depletes catchment storage, which in turn regulates discharge to the stream, the midday peak in evapotranspiration rates corresponds to the fastest decline in streamflow (not the minimum streamflow, as it would if evapotranspiration were removing water directly from the stream). This is the same dynamical phase lag that was noted in sections 3 and 13 between precipitation and streamflow.

[84] Estimated evapotranspiration rates are determined by the balance between the two terms of equation (26), which have opposite signs during streamflow recession. Because the sensitivity function  $g(Q)$  determines the balance between the two terms in equation (26), inaccuracies in estimating  $g(Q)$  could have large effects on evapotranspiration rates inferred from streamflow. However, even in cases where equation (26) does not give quantitatively accurate estimates of absolute rates, it may still be useful in estimating relative changes in evapotranspiration rates through time. Estimates of changes in evapotranspiration

rates over time may be useful for assessing the effects of changes in climate or land use on the water cycle. For example, concerns over the hydrologic effects of changes in vegetation (particularly afforestation of moorland areas) were the original motivation for long-term research at Plynlimon. Painstaking water balance studies, based on networks of dozens of storage-type rain gauges that were read monthly, have shown that during the 1970s and 1980s, transpiration rates in the forested catchment (Severn) were somewhat higher than in the moorland catchment (Wye). However, over time the difference in evapotranspiration rates between the two catchments has decreased because of both the increasing age of the forest stand and the gradual reduction of the forested area through timber harvesting [Robinson and Dupeyrat, 2005; Marc and Robinson, 2007].

[85] To test whether this change in evapotranspiration rates could also be detected from streamflow fluctuations, I applied equation (26) to the hourly discharge records from the Severn and Wye rivers over the 26 years from 1975 through 2000. This analysis used the same  $g(Q)$  functions used elsewhere in this paper, estimated from the recession plots in Figure 6. As Figure 20 shows, the fluctuations in streamflow in the two catchments indicate a gradual decline in evapotranspiration rates in the Severn relative to the Wye through the 1970s and 1980s, reproducing the general trend observed in the mass balance studies at Plynlimon over the same period (gray line, Figure 20). It bears emphasis that the inferred evapotranspiration rates shown in Figure 20 are not calibrated in any way to either Penman-Monteith estimates or to catchment mass balances. Indeed, the average evapotranspiration rates inferred from streamflow fluctuations in both of the catchments are low, on average, compared to either Penman-Monteith or mass balance estimates. This could arise because equation (26) omits interception losses, which are included in evapotranspiration rates inferred from catchment mass balances. It could also arise if  $g(Q)$  is somewhat too high at low flows, leading to a persistent bias in the relative sizes of the two terms of equation (26). Nonetheless, equation (26) reproduces the difference in evapotranspiration rates between the two catchments, and the trend in that difference through time (Figure 20), suggesting that this approach may be useful for detecting how landscape-scale evapotranspiration rates respond to changes in vegetation or climate.

## 15. Discussion

### 15.1. Comparison With Previous Analyses

[86] In the long history of recession analysis in hydrology, there have been many attempts to relate the recession behavior of streams to a catchment-scale drainage function; see reviews by Hall [1968], Tallaksen [1995], and Smakhtin [2001], and references therein, as well as more recent work by Lamb and Beven [1997], Wittenberg and Sivapalan [1999], Wittenberg [1999, 2003], Szilagyi et al. [2007], and Rupp and Woods [2008]. The approach presented here differs from each of these previous efforts in one or more of the following four ways. First, in the present approach the storage-discharge relationship is expressed in its implicit differential form, the sensitivity function  $g(Q)$ . This is advantageous because  $g(Q)$  can be estimated directly from recession plots such as Figure 6, and because it allows the

dynamical system to be expressed as a single first-order differential equation. This equation is invertible, allowing one to estimate precipitation or evapotranspiration from the streamflow time series, as well as to predict the hydrograph from precipitation and evapotranspiration time series.

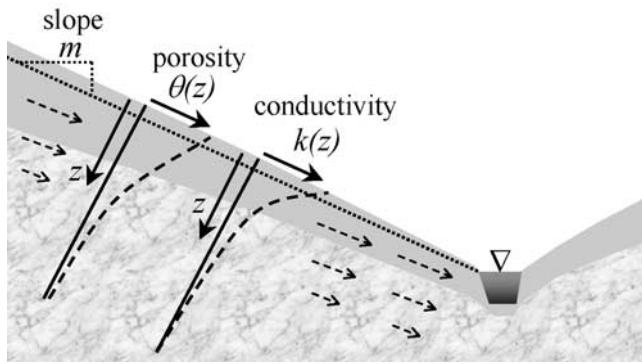
[87] Second, the present approach does not specify the form of the relationship between storage and discharge, or the corresponding sensitivity function  $g(Q)$ , instead allowing it to be determined directly from streamflow data. This makes the analysis more general and portable than one that assumes that the storage-discharge relationship must have a particular functional form. In this respect the approach is a not a “black box” model, but rather a “gray box” systems analysis tool (in the sense of Kirchner [2006]) because it uses the catchment’s behavior to reveal what its governing equations are. Because the sensitivity function is determined directly from streamflow data and encapsulates the drainage behavior of the catchment in concise form, it may also be a useful tool for catchment characterization.

[88] Third, the present approach attempts, as much as possible, to take account of the uncertainties and confounding factors that typically affect catchment data. As Tallaksen [1995] points out, many recession analyses (but see Lamb and Beven [1997] for an exception) pay too little attention to the confounding effects of evapotranspiration, even though its potential to significantly steepen recession curves has been known for decades [e.g., Federer, 1973]. The recession plots used to estimate  $g(Q)$ , by contrast, do not require continuous recession curves, so one can filter out data that are significantly affected by evapotranspiration. Similarly, the bounding lines that are typically fitted to recession plots of the Brutsaert-Nieber type [Brutsaert and Nieber, 1977] are susceptible to biases and artifacts [Rupp and Selker, 2006a], which the curve-fitting procedure in Figure 6 is designed to avoid.

[89] Fourth, the present approach makes no distinction between base flow and quick flow. Instead, it treats catchment drainage, from base flow to peak stormflow and back again, as a single continuum of hydrological behavior. This eliminates the need to separate the hydrograph into different components, and makes the analysis simple, general and portable.

### 15.2. Model Simplicity and Real-World Complexity

[90] The analysis presented here is based on a very simple model structure, but the flow paths and processes controlling runoff in real-world catchments are complex and spatially heterogeneous. Decades of field observations testify to the complexity of Plynlimon’s flow systems [e.g., Neal, 2004; Shand et al., 2005]. For example, abundant ephemeral springs and soil pipes at Plynlimon indicate a spatially heterogeneous and temporally dynamic flow system in the subsurface [e.g., Sklash et al., 1996]. Borehole measurements on a hillslope in the Severn catchment have also identified multiple subsurface flow paths and complex interactions between them, with stormflows creating downwelling and upwelling hydraulic gradients in adjacent borehole nests separated by only a few meters [Haria and Shand, 2004, 2006]. Thus the Plynlimon catchments are characterized by both spatial complexity and process heterogeneity. Yet at the scale of several square kilometers, these complexities and heterogeneities aggregate to a simple (albeit nonlinear) catchment-scale storage-discharge rela-



**Figure 21.** A hillslope cross section, illustrating a simple conceptual model for catchment-scale storage-discharge relationships. The profiles of conductivity  $k(z)$  and porosity  $\theta(z)$ , together with the hydraulic gradient  $m$ , control how storage and discharge vary with changes in water table depth (dotted line) and thus with changes in storage.

tionship. The fact that such a simple approach can capture the behavior of such a complex real-world system is a hopeful sign for its more general utility.

[91] The key to making such a simple approach workable is that it infers the form of the governing equations from the behavior of the system. In earlier exploratory work, for example, I tried modeling rainfall-runoff relationships in the Severn and Wye catchments with both linear and exponential reservoir models of various degrees of complexity, but they all exhibited obvious deviations from the data no matter what parameters were chosen. The failure of those models illustrates how the problem of model estimation in complex systems goes beyond the problem of parameter estimation: one needs to be fitting parameters to the right model equations in the first place, and knowing the right equations a priori is difficult. Indeed, a possible explanation for the problems of parameter identifiability and equifinality that often arise in hydrological models is that the governing equations could be wrong (in the sense that they could be inapplicable at the scales at which they are used), but their deficiencies could be masked by overparameterization [e.g., Beven, 1989; Beven, 2006; Kirchner, 2006]. The literature on parameter estimation in hydrology is large and growing, but the problem of identifying appropriate governing equations deserves more attention than it has received so far.

[92] One also needs to start with the right dynamical premises. For example, another way to “infer the form of the governing equations from the behavior of the system” would be through time series deconvolution of the rainfall and runoff signals, where the governing equation that is revealed is the convolution kernel. Such a linear convolution approach, however, would assume that streamflow is a superposition of time-shifted and rescaled precipitation records, and therefore could not explain the effects of antecedent moisture on runoff timing and magnitude. These considerations illustrate the importance of having an appropriate dynamical structure, and appropriate governing equations, before using data to estimate parameter values.

### 15.3. Physical Interpretation of Storage-Discharge Relationships

[93] One may object that the approach outlined here seems far removed from a physically based model of catchment processes. This approach is designed to capture a central physical process in catchment hydrology (the filling and draining of catchment storage), without being overly prescriptive concerning the physical details regulating that process at the small scale. This can be considered an advantage rather than a drawback, to the extent that those physical details would be “surplus content” that cannot usually be constrained by observational data. Indeed, many “physically based” models themselves have only a loose connection to the underlying physics, because their governing equations typically require parameters that cannot be measured at the relevant scales [e.g., Sherlock et al., 2000], and cannot be adequately constrained by calibration [Beven, 1989].

[94] Nonetheless it is not difficult to devise physical models, or perhaps more appropriately, physical rationalizations, for the storage-discharge relationships hypothesized here. For example, Figure 21 outlines a conceptual model that can be used to explain the storage-discharge relationship in terms of flow through the saturated zone at the hillslope scale. Porosity and saturated conductivity typically decrease nonlinearly with depth below the surface  $z$ , as shown schematically in Figure 21. In upland catchments, conductivity typically decreases by orders of magnitude over depths that are small compared to the hillslope relief, with the result that the water table is nearly parallel with the surface topography. In an idealized hillslope like the one shown in Figure 21, the increment of storage  $dS$  corresponding to any increment of depth  $dz$  is determined by the local drainable porosity  $\theta(z)$ ,

$$dS = \theta(z)dz, \quad (27)$$

and the increment of discharge  $dQ$  over an increment of depth  $dz$  is determined by the local conductivity  $k(z)$  and the slope of the water table  $m$ , which determines the hydraulic gradient,

$$dQ = m k(z)dz. \quad (28)$$

The total storage and discharge can be found by integrating over the saturated zone, starting from the water table depth  $z_0$ ,

$$S = \int_{z_0}^{\infty} \theta(z)dz, \quad Q = m \int_{z_0}^{\infty} k(z)dz, \quad (29)$$

and from these the storage-discharge relationship can be constructed. The sensitivity function  $g(Q)$  is, from (27) and (28), simply the ratio between conductivity and porosity at the water table depth  $z$ , scaled by the slope  $m$ :

$$g(Q) = \frac{dQ}{dS} = \frac{dQ/dz}{dS/dz} = \frac{m k(z)}{\theta(z)}. \quad (30)$$

Thus there is a simple correspondence between the variation in conductivity and porosity with depth, and



the storage-discharge relationship  $Q = f(S)$  and the sensitivity function  $g(Q)$ .

[95] This one-dimensional model ignores many issues that are important in real-world, three-dimensional catchments. In real-world catchments, porosity and conductivity are highly variable from point to point, and continuity relationships (including the effects of planform convergence and divergence) must be satisfied along each flow path, with the result that water table depth varies across the landscape. Thus one would not expect the sensitivity function  $g(Q)$ , as estimated from Figure 6, to agree with the conductivity and porosity profiles at any individual point in the catchment. Instead, the storage-discharge relationship characterizes how the catchment, as a whole, releases stored water to runoff. In this regard, rationalizing the sensitivity function  $g(Q)$  in terms of depth profiles in the catchment is analogous to subsuming the heterogeneity of the subsurface in an “effective” hydraulic conductivity in a typical physically based model.

[96] This is not the only physical model that is consistent with a functional relationship between catchment storage and discharge. For example, the Boussinesq models of groundwater flow [e.g., *Brutsaert and Nieber*, 1977] have recently been solved for power law conductivity profiles in both horizontal and sloping aquifers [*Van de Grind et al.*, 2002; *Rupp and Selker*, 2005, 2006b]. Similar power law transmissivity profiles have also been invoked in TOPMODEL [*Ambrose et al.*, 1996; *Duan and Miller*, 1997; *Iorgulescu and Musy*, 1997]. Both TOPMODEL and late-time solutions to the Boussinesq model yield recession plots characterized by power law behavior,  $-dQ/dt = aQ^b$  [*Brutsaert and Nieber*, 1977; *Duan and Miller*, 1997; *Iorgulescu and Musy*, 1997; *Rupp and Selker*, 2006b], similar to the behavior shown in Figure 6.

[97] However, for all such models (with the exception of *Rupp and Woods* [2008]), the exponent  $b$  can only take on values  $1 \leq b < 2$ , with the upper limit of  $b = 2$  corresponding to an exponential, rather than power law, profile. Equations (27)–(30) represent a more generalized framework for modeling storage-discharge behavior, because they are not constrained by particular functional forms describing how subsurface characteristics vary with depth. As discussed in section 6 above, cases where  $b > 2$  (such as the Severn and Wye rivers) require a hyperbolic relationship between discharge and storage. Such a hyperbolic relationship could arise on the idealized hillslope shown in Figure 21 if, for example, porosity  $\theta$  and conductivity  $k$  vary as hyperbolic functions of depth below the surface,  $\theta(z) = \theta_o/(z/z_o)^\alpha$  and  $k(z) = k_o/(z/z_o)^\beta$ . Any pair of profiles meeting the criteria  $0 \leq \alpha < 1$  and  $\beta = (b - 1 - \alpha)/(b - 2)$  will result in a storage-discharge relationship described by equation (17) and a recession plot described by equation (10) for  $b > 2$  (realism also requires that the water table never rises into the nonphysical region where  $\theta > 1$ ). This is just the simplest example of a family of density and porosity profiles that are consistent with power law recession plots with  $b > 2$ .

[98] The fundamental structural assumption in the analysis presented here is that discharge is controlled solely by catchment storage, and that both discharge and storage can be meaningfully averaged at the scale of the catchment. This does not mean that every point on the landscape needs to

obey the same drainage equation, but rather that the aggregate behavior of the catchment can be described by such a relationship. In reality, the drainage equations regulating different points on the landscape will differ, and the distribution of storage across the landscape will depend on the spatial arrangement of the local drainage equations (and thus may vary also as a function of total water storage). As a result, the storage-discharge relationship that characterizes the catchment’s behavior may not describe any individual point on the landscape.

#### 15.4. Estimating and Accounting for Bypassing Flow

[99] The simplifying assumption that discharge depends on catchment storage alone is of course an approximation, although in some cases it may be a quantitatively adequate one. In other cases a simple storage-discharge model can serve as a null hypothesis that can be compared with more structurally complex models. As an example, consider a catchment in which a fraction of precipitation bypasses catchment storage and is shunted directly to streamflow (via, for example, overland flow, macropore flow, or direct precipitation onto the wetted channel itself). In such a catchment, discharge would be the sum of drainage from catchment storage and bypassing flow, such that equation (2) would become

$$Q = f(S) + k_P P, \quad (31)$$

where  $k_P$  is the fraction of precipitation that bypasses storage. The sensitivity function  $g$  would be derived similarly to equation (5), with the further complication that it would now depend on  $Q - k_P P$ , the discharge that comes from draining of storage:

$$\frac{dQ}{dS} = f'(S) = f'(f^{-1}(Q - k_P P)) = g(Q - k_P P). \quad (32)$$

Because  $g$  is estimated during times when  $P = 0$ , its functional form and parameter values will be identical to those previously determined in section 5. Differentiating equation (31) with respect to time, one directly obtains a first-order nonlinear differential equation that, like equation (18), describes the evolution of discharge through time, driven by the time series of precipitation and evapotranspiration:

$$\frac{dQ}{dt} = \frac{dQ}{dS} \frac{dS}{dt} + k_P \frac{dP}{dt} = g(Q - k_P P) \cdot (P - E - Q) + k_P \frac{dP}{dt}. \quad (33)$$

One can estimate the importance of bypassing flow by determining the value of  $k_P$  that gives the best match to the streamflow time series. Using the methods of section 7 and determining both the bypassing constant  $k_P$  and the evapotranspiration scaling constant  $k_E$  by calibration gives best fit values of  $k_P = 0.008$  for the Severn and  $k_P = 0.007$  for the Wye, implying that less than 1% of precipitation bypasses catchment storage and is shunted directly to runoff. Adding this additional parameter improves the goodness of fit to the streamflow time series by only a trivial amount: Nash-Sutcliffe efficiencies increase from 0.913 to 0.915 at the Severn and from 0.859 to 0.861 at the Wye. Cross validation using the methods of section 8 yields similar results; the best

fit values of  $k_P$  are small across all individual years, ranging from 0.006 to 0.010 at the Severn and from 0.005 to 0.016 at the Wye. If the parameters that describe  $g(Q - k_P P)$  are also estimated by calibration, as in section 9, the best fit values of  $k_P$  are similarly small (ranging from 0.003 to 0.008 at the Severn and from 0.009 to 0.014 at the Wye), and adding the bypassing term alters the other parameter values by only a few percent or less.

[100] These results imply that bypassing flow makes only a small contribution to streamflow at Plynlimon, and that accounting for bypassing flow improves streamflow predictions only marginally. Adding this second runoff mechanism, however, substantially complicates the process of inferring precipitation patterns from the streamflow time series. Because equation (33) depends on both precipitation and its derivative, the simple methods of section 13 do not apply. However, adding this runoff mechanism does not complicate efforts to infer evapotranspiration rates from streamflow fluctuations, because (as outlined in section 14) those inferences are drawn when precipitation rates (and thus also their derivatives) are zero.

### 15.5. Data Requirements

[101] Any analysis is only as good as the data it is based on. It would therefore be useful to systematically assess how biases and uncertainties in discharge and weather measurements might affect the results of this approach, using synthetic data. Such a detailed assessment must await a future paper, however, as it is beyond the scope of this one.

[102] The analysis presented above has used hourly discharge measurements to estimate the sensitivity function  $g(Q)$  for the Severn and Wye rivers, excluding periods that were likely to be affected by rainfall or evapotranspiration. One can naturally ask whether the approach developed here can be used with the daily discharge data that are widely available online. At Plynlimon, hourly discharge and weather data have been used primarily because they permit relatively strict tests of the theory, but nothing in the mathematics of this approach fundamentally requires the use of any particular time step. As a test, I reestimated  $g(Q)$ , following the methods of section 5, with daily average discharges for the Severn and Wye rivers. I included only days for which precipitation and potential evapotranspiration were less than 10% of discharge, consistent with the criteria outlined in section 5. At Plynlimon, these criteria are met for only about 60 days in the 5 years of record. Nonetheless, the resulting  $g(Q)$  functions yield predictions of (hourly) discharge, and inferences of hourly precipitation from discharge fluctuations, that are similar to those presented in sections 8 and 13, with Nash-Sutcliffe efficiencies typically differing by less than 0.02 from those reported in the last column of Table 2, and correlations differing by less than 0.02 from those reported in the bold columns of Table 4.

[103] It is important, however, to exclude the effects of precipitation and evapotranspiration as much as possible from the discharge data used to estimate  $g(Q)$ . To test what could happen if this is not done, I reestimated  $g(Q)$  from the hourly discharge data without excluding periods with high potential evapotranspiration, excluding only periods with recent rainfall. High evapotranspiration can be expected to lead to steeper streamflow recessions, and thus higher values of  $-dQ/dt$  in recession plots such as Figure 6; this distortion will be largest at small values of  $Q$  and  $-dQ/dt$ .

As a result, the recession plots for the Severn and Wye are less steep, and more upward curving, if periods of high potential evapotranspiration are not excluded. The predicted streamflow time series and inferred precipitation rates are distorted as well, with Nash-Sutcliffe efficiencies that are typically about 0.1 lower than those shown in Table 2, and correlations with observed rainfall up to 0.06 lower than those shown in Table 4. These distortions would likely be larger at less humid catchments, where evapotranspiration is a bigger part of the water balance.

### 15.6. Limitations

[104] The structural simplicity of the approach outlined here obviously limits its generality. For example, it cannot be expected to give reasonable results in a catchment where Hortonian overland flow is a dominant runoff mechanism, without the addition of a bypassing mechanism like the one proposed above. Nonetheless, the methods outlined above can be used to test whether bypassing flow is important, and it is likely that there are many catchments where bypassing flow is only a small component of runoff.

[105] In snowmelt-dominated catchments, liquid water storage (and thus discharge) will respond to the rate of snowmelt rather than the rate of precipitation (i.e., snowfall) per se. Thus the “precipitation” rate inferred from streamflow fluctuations using the methods of section 13 will represent the rate of snowmelt and liquid precipitation reaching the ground surface. The fact that streamflow fluctuations reflect snowmelt rather than snowfall should perhaps be considered an advantage rather than a drawback, because this is one of the only methods by which spatially averaged snowmelt rates can be estimated.

[106] A bigger challenge is presented by catchments in which runoff is controlled by interconnected subsurface reservoirs with different storage-discharge relationships. In some catchments, for example, streamflow rises and falls in a daily oscillation driven by diurnal evapotranspiration cycles [Lundquist and Cayan, 2002; Czikowsky and Fitzjarrald, 2004]. Such behavior cannot be explained by a single reservoir with a single storage-discharge relationship, in which evapotranspiration losses should produce daily reductions in streamflow and discharge, but no rebound at night. Instead, a daily oscillation in streamflow would seem to imply a riparian zone that is continually recharged by drainage from upslope, with evapotranspiration losses leading to net declines in storage (and thus streamflow) during the day, and recharge from upslope leading to net increases in storage (and thus streamflow) at night. The challenge in any such multi-component system is that one cannot easily infer the properties of any individual component from the behavior of the system as a whole (as with the simple methods developed here).

[107] The methods outlined here also cannot be applied to ephemeral streams, because when discharge goes to zero, different levels of storage will correspond to the same (i.e., zero) discharge. Thus  $f(S)$  will be noninvertible and  $g(Q)$  will be ill defined. This sets a practical lower limit to the size of catchments where these methods can be applied, since they must be large enough to support permanent streams.

[108] It is equally clear that these methods must break down for catchments that are too large, but it is not yet clear

how big is too big. The catchments studied here are roughly  $10 \text{ km}^2$  in area. One can speculate that in significantly larger catchments (say,  $1000 \text{ km}^2$  in area), the lag times required for changes in discharge to propagate through the channel network would be so long, and so variable with distance from the outlet, that the methods presented here would not work. Also, if storage-discharge relationships are spatially heterogeneous at all scales, the methods presented here cannot be expected to work in catchments that are much larger than individual storm systems. In such a catchment, the runoff response to an individual storm will depend on the storage-discharge relationship and antecedent moisture in whatever part of the catchment the storm lands. Thus the apparent storage-discharge relationship, as viewed from the catchment outlet, could vary significantly from one storm to the next.

[109] By contrast, in catchments that are smaller than the scale of individual storms, each storm will typically cover the whole catchment area. For the reasons outlined earlier in this section, one would expect the aggregation of spatial heterogeneity in such a catchment to be repeatable from one storm to the next, yielding a catchment-scale storage-discharge relationship that is stable through time. Similar results would be obtained even in catchments that are larger than individual storms, if the local storage-discharge relationships are spatially heterogeneous on small scales, but relatively uniform, on average, at the scale of individual storms.

[110] In large river basins, rainfall-runoff behavior is determined more by the spatial distribution of precipitation and the routing of flood flows through the channel network, and less by the storage-discharge dynamics considered here. However, the approach presented here may still be helpful in large basins, by providing a small-catchment runoff “kernel” that can be aggregated through the channel network. Linking this dynamical systems approach with channel routing via a geomorphic instantaneous unit hydrograph (or similar approach) could help in understanding both the nonlinear response and travel time delays that characterize the hydrologic behavior of large basins.

[111] These considerations highlight the importance of understanding how storage-discharge relationships vary across the landscape. Much could be learned by using the methods outlined here to measure  $g(Q)$  across nested networks of small gauged catchments. Work on this is currently underway.

[112] Because  $g(Q)$  captures the integrated drainage behavior of catchments (at least those that are well approximated by the simple structural assumptions envisioned here), it may provide a useful framework for catchment characterization. Little is known yet about how consistent  $g(Q)$  is from one catchment to the next, or how much (and how systematically) it varies with a catchment’s near-surface geology, its soil characteristics, its geomorphic properties, its climatic setting, and so forth. *Tague and Grant* [2004], however, have shown that the log-log slope and intercept of recession plots for streams in the Cascade Mountains of Oregon are correlated with the fraction of highly conductive Plio-Pleistocene volcanic bedrock in their catchments. If  $g(Q)$  can be estimated from some combination of catchment characteristics, then it may help in solving the problem of hydrologic prediction in ungauged basins. If, on the other hand,  $g(Q)$  cannot be estimated reliably from catchment characteristics, it may

imply that catchments’ storage-discharge behavior depends on catchment properties that cannot be readily measured (such as, for example, the variation in hydraulic conductivity with depth) and thus that the ungauged basins problem cannot be solved. In that case, the most efficient way forward may be to simply gauge catchments long enough to estimate their sensitivity functions  $g(Q)$ , as an implicit measure of the hidden catchment properties that control their fundamental hydrologic behavior.

## 16. Summary and Conclusions

[113] In catchments where discharge  $Q$  is a function of storage  $S$ , the storage-discharge relationship  $Q = f(S)$  can be combined with the conservation-of-mass equation to form a nonlinear first-order dynamical system (equations (1) and (2)). This dynamical system becomes particularly simple if the storage-discharge relationship is expressed in its implicit differential form, the hydrologic sensitivity function  $g(Q) = dQ/dS$  (equation (5)). Both the mathematical form and the parameters of this sensitivity function can be estimated directly from recession plots (Figure 6) showing how the rate of streamflow recession  $-dQ/dt$  varies with discharge  $Q$ , under conditions where precipitation and evapotranspiration rates are negligible compared to discharge.

[114] Using the sensitivity function  $g(Q)$ , the relationship between precipitation, evapotranspiration, and runoff through time can be encapsulated in a single first-order nonlinear ordinary differential equation (equation (18)). This equation can be numerically integrated to straightforwardly model rainfall-runoff behavior through time (Figures 9 and 10). The performance of this single-equation rainfall-runoff model, as measured by its Nash-Sutcliffe efficiency, is similar to models that are much more complex and much more highly parameterized. Estimating the model parameters using different years of data, whether via recession plots or via direct parameter calibration, yields consistent parameter values and consistently high model performance (Tables 2 and 3). The consistency of parameter values and model performance across different years with different weather conditions indicates that this approach meets a crucial criterion of mathematical modeling: the constants stay constant while the variables vary. Cross-validation tests show that model performance is similar when tested against years that were used for parameter estimation, and years that were not (Tables 2 and 3), indicating that the model is more than just a mathematical marionette that can only “dance” to the tune of its own calibration data.

[115] The sensitivity function  $g(Q)$  is a useful tool for catchment characterization. From  $g(Q)$ , one can directly estimate a catchment’s dynamic storage (section 10), its sensitivity to antecedent moisture (section 11), and its recession time constants (section 12). More generally, because  $g(Q)$  compactly summarizes how catchments store and release water, it may provide a tool for linking hydrologic behavior to measurable catchment characteristics.

[116] Because the dynamical system that links precipitation, evapotranspiration, runoff, and storage can be expressed in a single ordinary differential equation, it can also be inverted, to express the balance between precipitation and evapotranspiration as a function only of discharge and its time derivative (equation (22)). Thus one can use streamflow fluctuations to estimate time series of spatially



averaged precipitation and evapotranspiration. Precipitation time series inferred from streamflow accurately reproduce the timing, duration, and intensity of rainfall events observed at Plynlimon (Figures 15 and 16), as well as long-term variations in annual rainfall totals (Figure 17). Precipitation rates inferred from streamflow at the Plynlimon catchments agree with direct rain gauge measurements roughly as well as the two rain gauges in each catchment agree with each other (Table 4).

[117] Streamflow fluctuations yield noisy, semiquantitative estimates of evapotranspiration rates (Figure 18) because of the strong sensitivity of the evapotranspiration estimates to small measurement errors in discharge. Nonetheless, evapotranspiration rates inferred from streamflow show diurnal cycles that resemble those in Penman-Monteith estimates of potential evaporation (Figure 19), and long-term estimates of the difference in evapotranspiration rates between the Severn and Wye catchments generally agree with independent estimates from mass balance studies (Figure 20). Considered together, Figures 15–20 suggest that streamflow fluctuations can yield useful estimates of precipitation and evapotranspiration rates at the scale of small catchments, particularly where direct measurements are unavailable, unreliable, or unrepresentative at the catchment scale. “Doing hydrology backward” also provides an important test of the underlying theoretical framework, because its only parameters are those embedded in  $g(Q)$ , which are derived from streamflow data alone. Thus precipitation and evapotranspiration rates inferred from streamflow fluctuations are not calibrated to the observed precipitation or evapotranspiration data, making them a particularly strong test of the underlying theory.

[118] Characterizing a catchment by a single nonlinear storage-discharge relationship, or its implicit differential counterpart, the sensitivity function  $g(Q)$ , involves dramatically simplifying (and possibly oversimplifying) the complex and spatially heterogeneous processes and properties that control water fluxes at the catchment scale. Thus the range of applicability of the approach presented here is not yet clear. Nonetheless, the analyses presented above demonstrate that, at least for some catchments, this approach provides a useful quantitative tool for predicting runoff from rainfall, and also for inferring rainfall and evapotranspiration from runoff.

[119] More broadly, however, this approach can be viewed as a way to characterize and diagnose the functioning of hydrologic systems at small-catchment scale. As shown above, it is demonstrably useful for both hydrologic prediction and hydrologic inversion. It also provides direct information about three important catchment characteristics: the size of the dynamic moisture store, the sensitivity of storm runoff to antecedent moisture, and the characteristic time scales of catchment drainage. Perhaps more importantly from a scientific standpoint, the approach is also falsifiable in multiple ways, because the theory is not “tuned” to match the data in any of the comparisons presented above, except Table 3.

[120] Finally, the approach outlined here is also mathematically straightforward, involving nothing more complex than a single first-order nonlinear differential equation, and is not difficult to apply in practice. The calculations can all be done on spreadsheets, and most importantly, they are

based on widely available hydrologic data (i.e., predominantly streamflow data, with rain gauge measurements and weather data as ancillary inputs). The approach does not require data that are hard to get (such as, for example, measurements of subsurface material properties or moisture status). As an inferential tool for understanding catchment behavior, and as a predictive tool for rainfall-runoff modeling, the approach outlined here performs at least as well as other methods that are considerably more complex, and correspondingly more difficult to apply. With its simple mathematics and minimal data requirements, this approach was developed specifically so that it could be readily applied in practice, and readily extended to new field situations. The next key task is to assess its applicability to diverse hydrologic settings.

[121] **Acknowledgments.** I thank Christina Tague, Sarah Godsey, Bob Moore, Keith Beven, and Colin Neal for helpful discussions concerning this work, Andrea Rinaldo for his suggestions, and four anonymous reviewers for their comments. I particularly thank the Plynlimon field staff and Ken Blyth, Jim Hudson, and Mark Robinson of the Centre for Ecology and Hydrology, Wallingford, for providing the field data analyzed here. This work was supported by the National Science Foundation (EAR-0125550), the Berkeley Water Center, and the Miller Institute for Basic Research.

## References

- Ambrose, B., K. Beven, and J. Freer (1996), Toward a generalization of the TOPMODEL concepts: Topographic indices of hydrological similarity, *Water Resour. Res.*, **32**, 2135–2145, doi:10.1029/95WR03716.
- Bathurst, J. C. (1986), Sensitivity analysis of the Systeme Hydrologique European for an upland catchment, *J. Hydrol.*, **87**, 103–123, doi:10.1016/0022-1694(86)90117-4.
- Beven, K. (1979), Generalized kinematic routing method, *Water Resour. Res.*, **15**, 1238–1242, doi:10.1029/WR015i005p01238.
- Beven, K. (1989), Changing ideas in hydrology—The case of physically based models, *J. Hydrol.*, **105**, 157–172, doi:10.1016/0022-1694(89)90101-7.
- Beven, K. J. (2001), *Rainfall-Runoff Modelling: The Primer*, 360 pp., John Wiley, Chichester, U.K.
- Beven, K. (2006), Searching for the Holy Grail of scientific hydrology:  $Q_t = H(S, R, \Delta t)A$  as closure, *Hydrol. Earth Syst. Sci.*, **10**, 609–618.
- Beven, K. J., and A. M. Binley (1992), The future of distributed models: Model calibration and predictive uncertainty, *Hydrol. Processes*, **6**, 279–298, doi:10.1002/hyp.3360060305.
- Boronina, A., S. Golubev, and W. Balderer (2005), Estimation of actual evapotranspiration from an alluvial aquifer of the Kouris catchment (Cyprus) using continuous streamflow records, *Hydrol. Processes*, **19**, 4055–4068, doi:10.1002/hyp.5871.
- Brutsaert, W. H. (1982), *Evaporation Into the Atmosphere: Theory, History, and Applications*, 299 pp., D. Reidel, Dordrecht, Netherlands.
- Brutsaert, W. (2005), *Hydrology: An Introduction*, 605 pp., Cambridge Univ. Press, Cambridge, U.K.
- Brutsaert, W., and J. P. Lopez (1998), Basin-scale geohydrologic drought flow features of riparian aquifers in the southern Great Plains, *Water Resour. Res.*, **34**, 233–240, doi:10.1029/97WR03068.
- Brutsaert, W., and J. L. Nieber (1977), Regionalized drought flow hydrographs from a mature glaciated plateau, *Water Resour. Res.*, **13**, 637–643, doi:10.1029/WR013i003p00637.
- Calder, I. R. (1977), Model of transpiration and interception loss from a spruce forest in Plynlimon, central Wales, *J. Hydrol.*, **33**, 247–265, doi:10.1016/0022-1694(77)90038-5.
- Czikowsky, M. J., and D. R. Fitzjarrald (2004), Evidence of seasonal changes in evapotranspiration in eastern U.S. hydrological records, *J. Hydrometeorol.*, **5**, 974–988, doi:10.1175/1525-7541(2004)005<0974:EOSCI>2.0.CO;2.
- Daniel, J. F. (1976), Estimating groundwater evapotranspiration from streamflow records, *Water Resour. Res.*, **12**, 360–364, doi:10.1029/WR012i003p00360.
- Duan, J., and N. L. Miller (1997), A generalized power function for the subsurface transmissivity profile in TOPMODEL, *Water Resour. Res.*, **33**, 2559–2562, doi:10.1029/97WR02186.



- Federer, C. A. (1973), Forest transpiration greatly speeds streamflow recession, *Water Resour. Res.*, 9, 1599–1604, doi:10.1029/WR009i006p01599.
- Hall, F. R. (1968), Base flow recessions—A review, *Water Resour. Res.*, 4, 973–983, doi:10.1029/WR004i005p00973.
- Haria, A. H., and P. Shand (2004), Evidence for deep sub-surface flow routing in forested upland Wales: Implications for contaminant transport and stream flow generation, *Hydrol. Earth Syst. Sci.*, 8, 334–344.
- Haria, A. H., and P. Shand (2006), Near-stream soil water-groundwater coupling in the headwaters of the Afon Hafren, Wales: Implications for surface water quality, *J. Hydrol.*, 331, 567–579, doi:10.1016/j.jhydrol.2006.06.004.
- Horton, R. E. (1936), Natural stream channel-storage, *Eos Trans. AGU*, 17, 406.
- Horton, R. E. (1937), Natural stream channel-storage (second paper), *Eos Trans. AGU*, 18, 440.
- Horton, R. E. (1941), Virtual channel-inflow graphs, *Eos Trans. AGU*, 22, 811.
- Iorgulescu, I., and A. Musy (1997), Generalization of TOPMODEL for a power law transmissivity profile, *Hydrol. Processes*, 11, 1353–1355, doi:10.1002/(SICI)1099-1085(199707)11:9<1353::AID-HYP585>3.0.CO;2-U.
- Kirby, C., M. D. Newson, and K. Gilman (Eds.) (1991), *Plynlimon Research: The First Two Decades*, Inst. of Hydrol., Wallingford, U.K.
- Kirby, C., C. Neal, H. Turner, and P. Moorhouse (1997), A bibliography of hydrological, geomorphological, sedimentological, biological, and hydrochemical references to the Institute of Hydrology experimental catchment studies in Plynlimon, *Hydrol. Earth Syst. Sci.*, 1, 755–763.
- Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, 42, W03S04, doi:10.1029/2005WR004362.
- Kirchner, J. W., X. Feng, and C. Neal (2000), Fractal stream chemistry and its implications for contaminant transport in catchments, *Nature*, 403, 524–527, doi:10.1038/35000537.
- Kirchner, J. W., X. Feng, and C. Neal (2001), Catchment-scale advection and dispersion as a mechanism for fractal scaling in stream tracer concentrations, *J. Hydrol.*, 254, 82–100, doi:10.1016/S0022-1694(01)00487-5.
- Klemes, V. (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J.*, 31, 13–24.
- Lamb, R., and K. Beven (1997), Using interactive recession curve analysis to specify a general catchment storage model, *Hydrol. Earth Syst. Sci.*, 1, 101–113.
- Lambert, A. O. (1969), A comprehensive rainfall-runoff model for an upland catchment, *J. Inst. Water Eng.*, 23, 231–238.
- Lambert, A. O. (1972), Catchment models based on ISO functions, *J. Inst. Water Eng.*, 26, 413–422.
- Laudon, H., J. Seibert, S. Köhler, and K. Bishop (2004), Hydrological flow paths during snowmelt: Congruence between hydrometric measurements and oxygen 18 in meltwater, soil water, and runoff, *Water Resour. Res.*, 40, W03102, doi:10.1029/2003WR002455.
- Laurenson, E. M. (1964), A catchment storage model for runoff routing, *J. Hydrol.*, 2, 141–163, doi:10.1016/0022-1694(64)90025-3.
- Linsley, R. K., M. A. Kohler, and J. L. H. Paulhus (1982), *Hydrology for Engineers*, 508 pp., McGraw-Hill, Toronto, Ont., Canada.
- Lundquist, J. D., and D. R. Cayan (2002), Seasonal and spatial patterns in diurnal cycles in streamflow in the western United States, *J. Hydrometeorol.*, 3, 591–603, doi:10.1175/1525-7541(2002)003<0591:SASPID>2.0.CO;2.
- Lyon, S. W., and P. A. Troch (2007), Hillslope subsurface flow similarity: Real-world tests of the hillslope Péclet number, *Water Resour. Res.*, 43, W07450, doi:10.1029/2006WR005323.
- Marc, V., and M. Robinson (2007), The long-term water balance (1972–2004) of upland forestry and grassland at Plynlimon, mid-Wales, *Hydrol. Earth Syst. Sci.*, 11, 44–60.
- Mein, R. G., E. M. Laurenson, and T. A. McMahon (1974), Simple non-linear model for flood estimation, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, 100, 1507–1518.
- Nash, J. E. (1957), The form of the instantaneous unit hydrograph, *IASH Publ.*, 45, 114–121.
- Nash, J. E. (1959), Systematic determination of unit hydrograph parameters, *J. Geophys. Res.*, 64, 111–115, doi:10.1029/JZ064i001p00111.
- Neal, C. (2004), The water quality functioning of the upper River Severn, Plynlimon, mid-Wales: Issues of monitoring, process understanding and forestry, *Hydrol. Earth Syst. Sci.*, 8, 521–532.
- Neal, C., and P. T. W. Rosier (1990), Chemical studies of chloride and stable oxygen isotopes in two conifer afforested and moorland sites in the British uplands, *J. Hydrol.*, 115, 269–283, doi:10.1016/0022-1694(90)90209-G.
- Neal, C., et al. (1997a), The occurrence of groundwater in the lower Palaeozoic rocks of upland Wales, *Hydrol. Earth Syst. Sci.*, 1, 3–18.
- Neal, C., J. Wilkinson, M. Neal, M. Harrow, H. Wickham, L. Hill, and C. Morfitt (1997b), The hydrochemistry of the River Severn, Plynlimon, *Hydrol. Earth Syst. Sci.*, 1, 583–617.
- Perrin, C., C. Michel, and V. Andreassian (2001), Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, *J. Hydrol.*, 242, 275–301, doi:10.1016/S0022-1694(00)00393-0.
- Rinaldo, A., A. Marani, and R. Rigon (1991), Geomorphological dispersion, *Water Resour. Res.*, 27, 513–525, doi:10.1029/90WR02501.
- Robinson, M., and A. Dupeyrat (2005), Effects of commercial timber harvesting on streamflow regimes in the Plynlimon catchments, mid-Wales, *Hydrol. Processes*, 19, 1213–1226, doi:10.1002/hyp.5561.
- Rodriguez-Iturbe, I., and J. B. Valdes (1979), The geomorphologic structure of hydrologic processes, *Water Resour. Res.*, 15, 1409–1420, doi:10.1029/WR015i006p01409.
- Rogers, C. C. M., K. J. Beven, E. M. Morris, and M. G. Anderson (1985), Sensitivity analysis, calibration and predictive uncertainty of the Institute of Hydrology distributed model, *J. Hydrol.*, 81, 179–191, doi:10.1016/0022-1694(85)90175-1.
- Rupp, D. E., and J. S. Selker (2005), Drainage of a horizontal Boussinesq aquifer with a power law hydraulic conductivity profile, *Water Resour. Res.*, 41, W11422, doi:10.1029/2005WR004241.
- Rupp, D. E., and J. S. Selker (2006a), Information, artifacts, and noise in  $dQ/dt - Q$  recession analysis, *Adv. Water Resour.*, 29, 154–160, doi:10.1016/j.advwatres.2005.03.019.
- Rupp, D. E., and J. S. Selker (2006b), On the use of the Boussinesq equation for interpreting recession hydrographs from sloping aquifers, *Water Resour. Res.*, 42, W12421, doi:10.1029/2006WR005080.
- Rupp, D. E., and R. A. Woods (2008), Increased flexibility in base flow modelling using a power law transmissivity profile, *Hydrol. Processes*, 22, 2667–2671, doi:10.1002/hyp.6863.
- Seibert, J. (2003), Reliability of model predictions outside calibration conditions, *Nord. Hydrol.*, 34, 477–492.
- Shand, P., A. H. Haria, C. Neal, K. J. Griffiths, D. C. Gooddy, A. J. Dixon, T. Hill, D. K. Buckley, and J. E. Cunningham (2005), Hydrochemical heterogeneity in an upland catchment: Further characterisation of the spatial, temporal, and depth variations in soils, streams, and groundwaters of the Plynlimon forested catchment, Wales, *Hydrol. Earth Syst. Sci.*, 9, 621–644.
- Sherlock, M. D., N. A. Chappell, and J. J. McDonnell (2000), Effects of experimental uncertainty on the calculation of hillslope flow paths, *Hydrol. Processes*, 14, 2457–2471, doi:10.1002/1099-1085(20001015)14:14<2457::AID-HYP106>3.0.CO;2-I.
- Sklash, M. G., K. J. Beven, K. Gilman, and W. G. Darling (1996), Isotope studies of pipeflow at Plynlimon, Wales, UK, *Hydrol. Processes*, 10, 921–944, doi:10.1002/(SICI)1099-1085(199607)10:7<921::AID-HYP347>3.0.CO;2-B.
- Smakhtin, V. U. (2001), Low flow hydrology: A review, *J. Hydrol.*, 240, 147–186, doi:10.1016/S0022-1694(00)00340-1.
- Spence, C. (2007), On the relation between dynamic storage and runoff: A discussion on thresholds, efficiency, and function, *Water Resour. Res.*, 43, W12416, doi:10.1029/2006WR005645.
- Szilagyi, J., Z. Gribovszki, and P. Kalicz (2007), Estimation of catchment-scale evapotranspiration from baseflow recession data: Numerical model and practical application results, *J. Hydrol.*, 336, 206–217, doi:10.1016/j.jhydrol.2007.01.004.
- Tague, C., and G. E. Grant (2004), A geological framework for interpreting the low-flow regimes of Cascade streams, Willamette River Basin, Oregon, *Water Resour. Res.*, 40, W04303, doi:10.1029/2003WR002629.
- Tallaksen, L. M. (1995), A review of baseflow recession analysis, *J. Hydrol.*, 165, 349–370, doi:10.1016/0022-1694(94)02540-R.
- Troch, P. A., F. P. De Troch, and W. Brutsaert (1993), Effective water table depth to describe initial conditions prior to storm rainfall in humid regions, *Water Resour. Res.*, 29, 427–434, doi:10.1029/92WR02087.
- Tschinkel, H. M. (1963), Short-term fluctuation in streamflow as related to evaporation and transpiration, *J. Geophys. Res.*, 68, 6459–6469.
- Uchida, T., J. J. McDonnell, and Y. Asano (2006), Functional intercomparison of hillslopes and small catchments by examining water source, flowpath and mean residence time, *J. Hydrol.*, 327, 627–642, doi:10.1016/j.jhydrol.2006.02.037.
- Van de Grind, A. A., J. J. De Vries, and E. Seyhan (2002), Groundwater discharge from areas with variable specific drainage resistance, *J. Hydrol.*, 259, 203–220, doi:10.1016/S0022-1694(01)00583-2.
- van der Tak, L. D., and R. L. Bras (1990), Incorporating hillslope effects into the geomorphologic instantaneous unit hydrograph, *Water Resour. Res.*, 21, 743–754.

- Wittenberg, H. (1999), Baseflow recession and recharge as nonlinear storage processes, *Hydrol. Processes*, 13, 715–726, doi:10.1002/(SICI)1099-1085(19990415)13:5<715::AID-HYP775>3.0.CO;2-N.
- Wittenberg, H. (2003), Effects of season and man-made changes on baseflow and flow recession: Case studies, *Hydrol. Processes*, 17, 2113–2123, doi:10.1002/hyp.1324.
- Wittenberg, H., and M. Sivapalan (1999), Watershed groundwater balance estimation using streamflow recession analysis and baseflow separation, *J. Hydrol.*, 219, 20–33, doi:10.1016/S0022-1694(99)00040-2.

---

J. W. Kirchner, Department of Earth and Planetary Science, University of California, Berkeley, CA 94720-4767, USA. (kirchner@berkeley.edu)