Linearity and nonlinearity of basin response as a function of scale: Discussion of alternative definitions

M. Sivapalan, C. Jothityangkoon, and M. Menabde

Centre for Water Research, Department of Environmental Engineering, University of Western Australia, Crawley, Australia

Received 12 May 2001; revised 16 October 2001; accepted 16 October 2001; published 16 February 2002

[1] Two uses of the terms "linearity" and "nonlinearity" appear in recent literature. The first definition of nonlinearity is with respect to the dynamical property such as the rainfall-runoff response of a catchment, and nonlinearity in this sense refers to a nonlinear dependence of the storm response on the magnitude of the rainfall inputs [*Minshall*, 1960; *Wang et al.*, 1981]. The second definition of nonlinearity [*Huang and Willgoose*, 1993; *Goodrich et al.*, 1997] is with respect to the dependence of a catchment statistical property, such as the mean annual flood, on the area of the catchment. They are both linked to important and interconnected hydrologic concepts, and furthermore, the change of nonlinearity with area (scale) has been an important motivation for hydrologic research. While both definitions are correct mathematically, they refer to hydrologically different concepts. In this paper we show that nonlinearity in the dynamical sense and that in the statistical sense can exist independently of each other (i.e., can be unrelated). If not carefully distinguished, the existence of these two definitions can lead to a catchment's response being described as being both linear and nonlinear at the same time. We therefore recommend separating these definitions by reserving the term "nonlinearity" for the classical, dynamical definition with respect to rainfall inputs, while adopting the term "scaling relationship" for the dependence of a catchment hydrological property on catchment area. *INDEX TERMS:* 1821 Hydrology: Floods; 1824 Hydrology: Geomorphology (1625); 1860 Hydrology: Runoff and streamflow; 1869 Hydrology: Stochastic processes; *KEYWORDS:* Scaling, Nonlinearity, Floods, Unit Hydrographs, Geomorphology

1. Introduction

[2] Modeling of rainfall-runoff responses at the catchment scale involves the analytical and/or numerical derivation of both the dynamical and statistical properties of the discharge at the catchment's outlet in terms of the rainfall and other climate, soil, vegetation, and geomorphological properties of the catchment. In this respect, two uses of the terms "linearity" and "nonlinearity" have appeared in recent literature.

[3] The first, classical definition of nonlinearity is with respect to the dynamical rainfall-runoff response of a catchment. A dynamical rainfall-runoff relationship attempts to describe, based on physical considerations, the discharge hydrograph, $Q(t)$ $[L^3/T]$ for a specified climatic input, namely the rainfall hyetograph, $R(t)[L/T]$, and for given catchment properties. This dependence can be symbolically written in the functional form

$$Q(t) = F\{R(t), A, \alpha, \beta \ldots\}, \tag{1}$$

where, α, β,... are various parameters, characterizing the climate, soil, vegetation, and geomorphological properties of the catchment, and (1) also includes an explicit representation of catchment area A. The dependence on $R(t)$ of the right-hand side of (1) is, in general, nonlinear, and also time and location dependent. The dynamical definition of nonlinearity refers to a nonlinear dependence of the runoff response on the magnitude of the rainfall inputs $R(t)$.

[4] A linear approximation to (1), widely used in hydrological practice, is given by the following convolution integral:

$$Q(t) = \int_0^t G(t-\tau)R(\tau)d\tau, \tag{2}$$

where $G(t)$ is the so-called instantaneous unit hydrograph (IUH) of the catchment. The function Gencompasses within it all of the climatic, soil, vegetation, and geomorphological properties of the catchment (including catchment area A), remains invariant in time and space, and is independent of $R(t)$. Equation (2) is often taken as a definition of the linearity of catchment response. There has been a number of studies in the past that have suggested that catchment responses in the sense of (1), i.e., the functional forms of F, are nonlinear for small catchments (small A) and that the nonlinearity decreases and catchments become more linear with increasing catchment area A [*Minshall*, 1960; *Wang et al.*, 1981]. On the other hand, *Robinson et al.* [1995], using numerical simulations, showed that nonlinearity at small scales is dominated by the hillslope response, that nonlinearity at large scales is dominated by channel network hydrodynamics, and that nonlinearity does not really disappear at any scale.

[5] The second definition of nonlinearity that has been offered recently [*Huang and Willgoose*, 1993; *Goodrich et al.*, 1997] is with respect to statistical properties of $Q(t)$. For example, when the mean annual peak discharge Q_p or mean annual runoff volume V are considered random variables, this refers to the relationships of their probability distributions, or moments, as functions of an underlying catchment geomorphological property, such as area A. Clearly, these relationships are not dynamical ones relating to input-output behavior for individual events but are integrated statistical responses over some time period, such as the year.

0043-1397/02/2001WR000482$09.00

[6] By way of illustration, some empirically observed properties of catchment response often exhibit a dependence on the most important geomorphological characteristic of a catchment, namely its drainage area A. For example,

$$Q_T = c_1 A^{\theta}, \tag{3}$$

$$V_m = c_2 A^{\beta}, \tag{4}$$

where Q_T $[L^3/T]$ is the annual peak discharge with a return period T and V_m $[L]$ is the mean annual runoff volume. Empirical values of the exponents θ and β, often called scaling exponents, have been found to lie in the range $0.5 < \theta < 1$ and $0.6 < \beta < 1$ [*Leopold et al.*, 1964; *Benson*, 1962, 1964; *Alexander*, 1972]. Indeed, *Benson* [1962, 1964] found that the exponent θ tended to decrease with increasing aridity (e.g., $\theta = 0.85$ for humid New England, and $\theta = 0.59$ for semiarid Texas and New Mexico). Benson also found, later supported by *Alexander* [1972] with worldwide data, that θ actually decreased with increasing catchment area. *Goodrich et al.* [1997] recently examined peak runoff Q_T and annual water yield V_m over a range of catchment sizes on the Walnut Gulch watershed in Arizona and found that θ changed from 0.90 for small subcatchments to 0.55 for larger subcatchments, while β varied from 0.97 to 0.82 over the same area ranges.

[7] Both of the above definitions conform to the standard mathematical definition of linearity: function $y = f(x)$ is linear with respect to the input variable x if and only if $f(c_1x_1 + c_2x_2) = c_1 f(x_1) + c_2 f(x_2)$, where c_1 and c_2 are arbitrary constants. However, both the input and output variables in the above definitions are different. In the first case it is a pair of dynamical variables $R(t)$ and $Q(t)$, while in the second case it is a pair of variables A and Q_T (or V_m), where A is a geomorphological characteristic of the catchment, whereas Q_T and V_m are statistical characteristics. As it follows from (3) and (4), these relations are nonlinear except for the particular cases of $\theta = 1$ and $\beta = 1$.

[8] The aim of this technical note is to distinguish between the two definitions of hydrological linearity/nonlinearity presented above. We aim to illustrate, based on a simple linear model of flood response, that the two definitions are hydrologically unrelated (one is not a consequence of the other) and that nonlinearities in the dynamical sense and in the statistical sense (as defined above) can exist independently of each other. Both types of nonlinear behavior are important, interconnected issues in hydrology, and moreover, the change of nonlinearity (by whatever definition) with size or scale has been an important motivation for much hydrologic research and speculation [*Wang et al.*, 1981; *Robinson et al.*, 1995; *Goodrich et al.*, 1997]. We will therefore argue that different terminology should be used to describe these so as to avoid the confusion that can result from usage that does not distinguish between these two definitions.

2. Flood Peak Estimation Using a Simple Storage Model

[9] To throw light on the differences between the two definitions of nonlinearity, we introduce a simple, lumped, and linear model of catchment response. Catchment response consists of two independent component processes: runoff generation on hillslopes and flow routing in the channel network. For simplicity, these are represented through two linear stores, arranged in series. The hillslope store receives rainfall input at rate i $[L/T]$ over a storm duration t_r and delivers runoff to the network store, located downstream of it, at a rate $q_h[L/T]$: the latter is assumed to be a linear function of the volume of water stored in the hillslope, denoted by $S_h[L]$. The outflow from the network store, $q[L/T]$, which is the same as the outflow from the catchment at its outlet, is assumed to be a linear function of the water stored in the network, denoted by $S_n[L]$. Figure 1 provides a schematic of the simple linear model. The water balance equations for these two parts of the model are expressed as

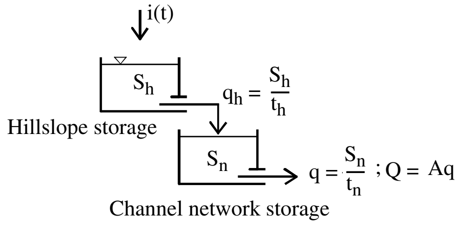

Figure 1. Schematic of a simple, linear, and lumped model of catchment response.

$$dS_h/dt = i(t) - q_h(t), \tag{5a}$$

$$q_h(t) = S_h(t)/t_h, \tag{5b}$$

$$dS_n/dt = q_h(t) - q(t), \tag{6a}$$

$$q(t) = S_n(t)/t_n, \tag{6b}$$

where t_h is the mean hillslope response time and t_n is the mean channel response time, both of which are assumed to be constants for the catchment in question. All of the quantities in (5a), (5b)–(6a), (6b) are normalized by catchment area and are expressed in units of $[L]$ or $[L/T]$. We also define an unscaled discharge by converting $q(t)$ to $Q(t)$ $[L^3/T]$ by multiplying by the catchment area A $[L^2]$, i.e., $Q(t) = Aq(t)$.

[10] For simplicity, we look at the case where the rainfall input to the catchment falls at a constant (in time and space) rate i and where the catchment is dry initially, meaning the stores S_h and S_n are empty at the beginning of the storm. The effects of antecedent conditions have been studied previously in detail within the same context by *Robinson and Sivapalan* [1997b] and will not be repeated here. The outflow hydrograph for the simple case can be derived in a straightforward manner. Here we present only the final expressions for the peak of the resulting hydrograph, denoted by $q_p[L/T]$ and $Q_p[L^3/T]$:

$$\frac{q_p}{i} = \frac{t_n}{t_n - t_h}\left\{1 - \exp\left(-\frac{t_r}{t_n}\right)\right\} - \frac{t_h}{t_n - t_h} \cdot \left\{1 - \exp\left(-\frac{t_r}{t_h}\right)\right\}, \qquad t_h > 0,\ t_n \neq t_h, \tag{7a}$$

$$\frac{q_p}{i} = \frac{1}{2}\left(\frac{t_r^2}{t_n^2}\right) - \frac{1}{3}\left(\frac{t_r^3}{t_n^3}\right) + \ldots, \qquad t_h > 0,\ t_n = t_h, \tag{7b}$$

$$\frac{q_p}{i} = \left\{1 - \exp\left(-\frac{t_r}{t_n}\right)\right\}, \qquad t_h = 0. \tag{7c}$$

The unscaled peak discharge Q_p is simply given by

$$Q_p = A\, q_p. \tag{8}$$

[11] Equations (7a), (7b), and (7c) are used to compute the magnitudes of the dimensionless flood peaks, q_p/i, for different values of the timescales t_r, t_h, and t_n. The results are presented in Figure 2a as families of curves relating q_p/i to t_n for different values

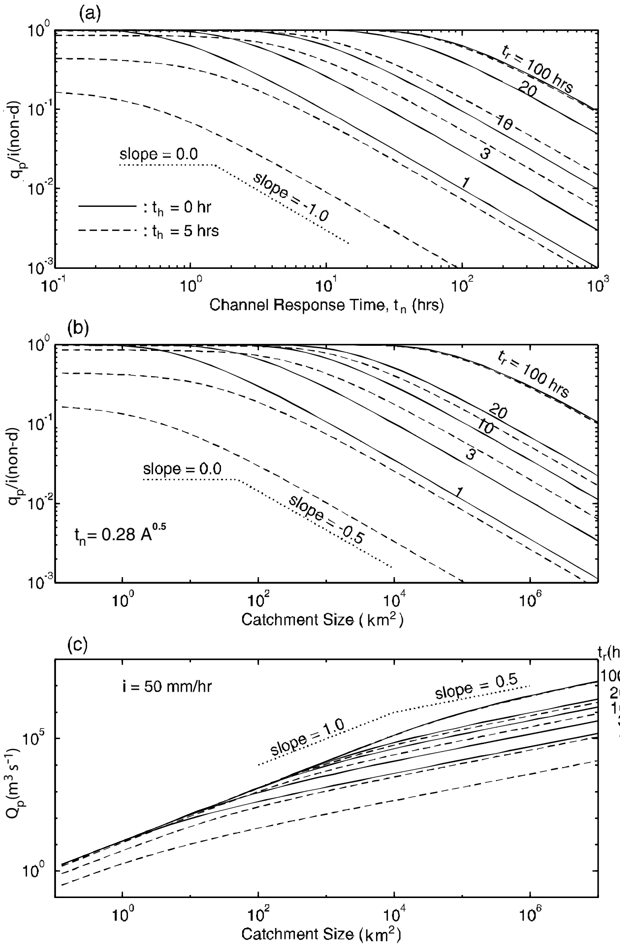

Figure 2. (a) Plot of nondimensional flood peak q_p/i as a function of channel network response time t_n for different values of t_r and t_h; (b) plot of nondimensional flood peak q_p/i as a function of catchment area A for different values of t_r and t_h; and (c) plot of flood peak Q_p(m^3/s) as a function of catchment area A for different values of t_r and t_h.

of t_r and t_h. They show that the flood peaks, q_p/i $[L/T]$, remain constant for small values of t_n (scaling exponent with respect to t_n is zero), while for larger values of t_n they decrease linearly with increasing values of t_n (scaling exponent is -1). The effect of t_h is to smooth and thus reduce the magnitude of the flood peaks, without changing the scaling exponents with respect to t_n. This can easily be seen by asymptotic analysis of (7a), (7b), (7c) for $t_n \to 0$ and $t_n \to \infty$.

[12] Since we are interested in the relationship of the flood peaks q_p or Q_p catchment area A, we assume, for simplicity, that (1) t_h is independent of catchment area (a reasonable assumption used often; see *Robinson et al.* [1995]) and (2) t_n is a scaling function of catchment size A [*Robinson and Sivapalan*, 1997b] of the form

$$t_n = \tau A^{\eta}, \qquad (9)$$

where t_n is in hours and A is in km^2. In this case, for a start, we assume that $\tau = 0.28$, and $\eta = 0.5$. The results shown in Figure 2a can now be converted to a relationship of q_p/i and of Q_p (using equations (7a), (7b), (7c) and (8) with an assumed value of $i = 50$ mm/h), against catchment area A. The results are presented in Figures 2b and 2c. In the case of Figure 2b the results are similar to those shown in Figure 2a, except that the scaling exponents with respect to catchment area A are 0.0 and -0.5, i.e., $-\eta$.

[13] Focusing on Figure 2c, we can clearly see that the scaling exponent of Q_p with respect to catchment area A changes from 1.0 (unity) for small catchments to 0.5 for large catchments, with a transition zone in the middle with variable exponents. This is similar in character to the "nonlinear" relationship, in the sense of (3), exhibited in previous studies, including that at Walnut Gulch [*Goodrich et al.*, 1997]. It should be noted here that for general values of η the exponent for our simple case changes from 1 for small catchments to $1 - \eta$ for large catchments.

[14] Note that while the catchment response, in the sense of (1) or (2), remains linear for all catchment sizes, the observed change in scaling exponent is caused by a change of runoff process, as shown by *Robinson and Sivapalan* [1997a] and *Gupta and Waymire* [1998]. In small catchments, storm duration is long compared to the catchment's residence time, and consequently the catchment reaches steady state, with the whole catchment area contributing to the flood peak. As long as this remains true, the flood peak increases linearly with catchment area, and thus the scaling exponent remains 1 (unity). On the other hand, in "large catchments," storm duration is smaller than the catchment's residence time. The fraction of the catchment area contributing to the flood peak is proportional to the ratio of storm duration to catchment residence time. This ratio, following (9), decreases at the rate of $A^{-\eta}$ with an increase of catchment area A. Thus the partial area contributing to flood peak increases only at the rate of $A^{1-\eta}$, leading to the exponent $1 - \eta$, which is less than unity.

[15] The above demonstration has been based on the response to a single storm. We have not chosen to present results on the scaling behavior of the mean annual flood or of the flood peak with a specified return period, in the true sense of (3). This would require much more detailed statistical analyses or random simulations, involving the derived flood frequency method. Such extensive analyses with more complex models have indeed been carried out and do support the conclusions reached here [*Gupta and Waymire*, 1998; *Blöschl and Sivapalan*, 1997; *Robinson and Sivapalan*, 1997a, 1997b]. For example, for the case corresponding to $t_h = 0$, *Robinson and Sivapalan* [1997a] have shown that the scaling exponents of mean annual flood changed from 1.0 to $1 - \eta$, as it was for the simple case presented here earlier.

[16] In summary, what we have demonstrated is that a simple, linear model in the sense of (2) can still lead to the "nonlinear" scaling type of behavior in the sense of (3) and (4). While the scaling exponent changed from 1 to $1 - \eta$ with increasing area A, the dynamical catchment rainfall-runoff response has remained linear. Indeed, our model confirms the (partial) explanation given by *Gupta and Waymire* [1998], *Robinson and Sivapalan* [1997a], and *Blöschl and Sivapalan* [1997] for the observed decrease of scaling exponents with increasing catchment area, namely the interaction between mean storm duration and catchment residence time.

3. Hack's Law and the Connection to Network Residence Time

[17] Now we take a look at the relationship (equation (9)) between mean network residence time t_n and catchment area A. Mean residence time in the network depends on two factors. One is the mean stream length, which is determined by the planar geometry of the catchment and its associated network. The other is the velocity of travel, which is determined by the hydraulics of flow, governed by stream slope and the elevation geometry of the

channel network, and both at-a-site and downstream hydraulic geometry variations. In general, stream velocity could vary in space and in time, governed by the spatial and temporal variability of rainfall inputs and catchment characteristics. In other words, (9) itself could be a nonlinear function, in the sense of (1).

[18] For simplicity, however, we will consider here a linear response function for t_n, by assuming that the flow velocity remains constant in time and space, irrespective of rainfall intensity or antecedent conditions. In this case, (9) could be derived by combining the constant velocity assumption with a geometric relationship, which connects the catchment's mainstream length L and drainage area A. The latter is known as Hack's law [*Hack*, 1957]. Equation (9), with $\eta = 0.5$, is therefore compatible with a mean stream length that is proportional to A^{η}, i.e., $t_n = \lambda A^{\eta}/v$, where v is the constant mean velocity and λ is a constant catchment property. In many catchments, however, observed values of the exponent are slightly greater than 0.5, about 0.58 on average, which is often attributed to the "fractality" of the channel network structure [see, e.g., *Rodriguez-Iturbe and Rinaldo*, 1997].

[19] By the definition given in (3) and (4) the relationship (9) between mean network residence time and catchment area A must also be "nonlinear" because the exponents differ from unity. Yet, for the case of constant velocity, (9) arises from a linear model of network response, combined with a purely geometrical property of the network. The fractional exponent ($\eta < 1$) is therefore not necessarily linked to some nonlinearity of the network's dynamical response, in the sense of (1); in the simple linear case it reflects a geometrical property of the catchment.

4. Connection Between Nonlinearity and Scaling

[20] So far we have shown that a simple linear, dynamical model in the sense of (2) can still give rise to "nonlinearity" with respect to catchment area, in the sense of (3) and (4). It is, of course, quite possible that when nonlinearity in the sense of (1) is also present, it can further impact on the nonlinearity with respect to catchment area in the sense of (3) and (4). For example, (9) could be nonlinear in the sense of (2), owing to the nonlinearity of channel flow hydrodynamics. Similarly, on the basis of empirical evidence and numerical simulations with a rainfall-runoff model, *Goodrich et al.* [1997] have shown that in semiarid catchments, such as Walnut Gulch, the rainfall-runoff volume response is strongly influenced by (1) increasing channel losses with increasing catchment area in semiarid catchments and (2) the likelihood of partial storm coverage, thunderstorms being the main mechanism of storm rainfall in this region. *Goodrich et al.* [1997] showed that these factors can also clearly contribute to the "nonlinearity," in the sense of (3) and (4), exhibited in Walnut Gulch.

[21] *Goodrich et al.* [1997], following on from their definition of nonlinearity in the sense of (3) and (4), concluded that the fact that the exponents changed from near unity for small catchments to less than unity for large catchments ($\theta = 0.55$ and $\beta = 0.82$) was an indication that nonlinearity of catchment response increased with increasing catchment area. They contrasted these observations with the earlier results of *Minshall* [1960] and *Wang et al.* [1981], who suggested that the nonlinearity decreased with increasing catchment area. *Goodrich et al.* [1997], on the basis of their investigations in Walnut Gulch, tried to explain this apparent contradiction by suggesting that the previous conclusions were based on work on humid catchments and may not hold in semiarid catchments, where channel losses and partial storm coverage can cause increasing "nonlinearity" with increasing catchment area.

[22] However, *Goodrich et al.* [1997] failed to explicitly highlight in their comparisons of the two results that two different definitions were involved. Despite the obvious differences in behavior between humid and semiarid catchments, the cited differences in nonlinear behavior were, in fact, due to the two different definitions of nonlinearity used. The nonlinearity highlighted in the Walnut Gulch results is of the statistical type, as given by (3) and (4). The nonlinearity described by *Minshall* [1960] and *Wang et al.* [1981] is the dynamical type in the sense of (1), with respect to rainfall inputs. As we have demonstrated in this paper, the traditionally accepted definition of nonlinearity, in the sense of (1), is distinctly different from the definition of the nonlinearity in the sense of (3) or (4) and can exist independently of the other. Thus the apparent contradictions cited above will not arise if the two definitions are clearly distinguished and not confused.

[23] We therefore submit that this ambiguity henceforth be resolved by reserving the term "nonlinearity," as per the common and widely accepted practice, for the nonlinear dependence of the dynamical hydrological response on rainfall inputs, including losses, partial storm coverage, and river hydrodynamics. In contrast, the statistical dependence of a hydrological property on catchment area is best described by the widely used term "scaling relationship."

5. Conclusions

[24] Two uses of the terms "linearity" and "nonlinearity" have appeared in recent hydrology literature. The first definition of nonlinearity is with respect to the dynamical property such as the rainfall-runoff response of a catchment, and nonlinearity in this sense refers to a nonlinear dependence of the runoff response on the magnitude of the rainfall inputs. The second definition of nonlinearity is with respect to statistical dependence of a catchment property, such as the mean annual flood, on catchment area. Both definitions are mathematically correct but hydrologically different. Indeed, we have shown that nonlinearity in the dynamical sense and that in the statistical sense can exist independently of each other. If not carefully distinguished, they will lead to a catchment's response being described as being both linear and nonlinear at the same time, depending on the definition one uses. To avoid ambiguity, we recommend the continued use of the term nonlinearity for the dynamical response to rainfall inputs, while the dependence on catchment area A should be described as a scaling relationship. The effects of dynamical nonlinearity on statistical scaling behavior remains an issue, and our clarification of the definitions is not to suppress genuine interest in this important area of research.

[25] **Acknowledgments.** This work was partially funded by an Australian Research Council Large Grant awarded to the first author. It was also supported by a Royal Thai Government Scholarship, awarded to the second author. CWR Reference ED 1555 CJ.

References

Alexander, G. N., Effect of catchment area on flood magnitude, *J. Hydrol.*, *16*, 225–240, 1972.

Benson, M. A., Factors influencing the occurrence of floods in a humid region of diverse terrain, *U. S. Geol. Surv. Water Supply Pap.*, *1580-B*, 64 pp., 1962.

Benson, M. A., Factors affecting the occurrence of floods in the south-west, *U. S. Geol. Surv. Water Supply Pap.*, *1580-D*, 72 pp., 1964.

Blöschl, G., and M. Sivapalan, Process controls on regional flood frequency: Coefficient of variation and basin scale, *Water Resour. Res.*, *33*(12), 2967–2980, 1997.

Goodrich, D. C., L. J. Lane, R. M. Shillito, S. N. Miller, K. H. Syed, and

D. A. Woolhiser, Linearity of basin response as a function of scale in a semi-arid watershed, *Water Resour. Res.*, *33*(12), 2951–2965, 1997.

Gupta, V. K., and E. Waymire, Scale invariance and regionalization of floods, in *Scale Invariance and Scale Dependence in Hydrology*, edited by G. Sposito, pp. 88–135, Cambridge Univ. Press, New York, 1998.

Hack, J. T., Studies of longitudinal stream profiles in Virginia and Maryland, *U. S. Geol. Surv. Prof. Pap.*, *294-B*, 1957.

Huang, H. Q., and G. Willgoose, Flood frequency relationships dependent on catchment area: An investigation of causal relationships, paper presented at Towards the 21st Century: Engineering for Water Resources Conference, Inst. Eng. Aust., Newcastle, New South Wales, 1993.

Leopold, L. B., M. G. Wolman, and J. P. Miller, *Fluvial Processes in Geomorphology,* W. H. Freeman, New York, 1964.

Minshall, N. E., Predicting storm runoff on small experimental watersheds, *J. Hydraul. Div. Am. Soc. Eng.*, *86*(HYB), 17–38, 1960.

Robinson, J. S., and M. Sivapalan, An investigation into the physical causes of scaling and heterogeneity of regional flood frequency, *Water Resour. Res.*, *33*(5), 1045–1059, 1997a.

Robinson, J. S., and M. Sivapalan, Temporal scales and hydrological regimes: Implication for flood frequency scaling, *Water Resour. Res.*, *33*(12), 2981–2999, 1997b.

Robinson, J. S., M. Sivapalan, and J. D. Snell, On the relative roles of hillslope processes, channel routing and network geomorphology in the hydrologic response of natural catchments, *Water Resour. Res.*, *31*(12), 3089–3101, 1995.

Rodriguez-Iturbe, I., and A. Rinaldo, *Fractal River Basins: Chance and Self-Organization*, Cambridge Univ. Press, 547 pp., New York, 1997.

Wang, C. T., V. K. Gupta, and E. Waymire, A geomorphologic synthesis of nonlinearity in surface runoff, *Water Resour. Res.*, *17*(3), 545–554, 1981.

C. Jothityangkoon, M. Menabde, and M. Sivapalan, Centre for Water Research, Department of Environmental Engineering, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia. (sivapalan@cwr.uwa.edu.au)