SIMILARITY AND SCALE IN CATCHMENT STORM RESPONSE

Eric F. Wood
Water Resources Program
Princeton University
Princeton, New Jersey

Murugesu Sivapalan
Centre for Water Research
University of Western Australia
Nedlands

Keith Beven
Centre for Research on Environmental Systems
University of Lancaster, Lancaster, England

Abstract. Until recently, very little progress had been made in understanding the relationship between small-scale variability of topography, soil, and rainfalls and the storm response seen at the catchment scale. The work reviewed here represents the first attempt at a systematic theoretical framework for such understanding in the context of surface runoff generation by different processes. The parameterization of hydrological processes over a range of scales is examined, and the concept of the "representative elementary area" (REA) is introduced. The REA is a fundamental scale for catchment modeling at which continuum assumptions can be applied for the spatially variable controls and parameters, and spatial patterns no longer have to be considered explicitly. The investigation of scale leads into the concept of hydrologic similarity in which the effects of the environmental controls on runoff generation and flood frequency response be investigated independently of catchment scale. The paper reviews the authors' initial results and hopefully will motivate others to also investigate the issues of hydrologic scale and similarity.

1. INTRODUCTION

The way in which storm rainfall reaching the ground surface causes a response in the flow of a river is a fascinating but most difficult area of study. Scientists, as well as artists and writers, have long been inspired to examine and speculate upon the flow of water both at the surface in streams and rivers and underground. Early measurements in the 17th century by the French physicists Edme Mariotte and Pierre Perrault and by the English astronomer Edmund Halley confirmed the basic concepts of the hydrological cycle, and in particular that rainfalls over a catchment area are sufficient to maintain continuous streamflows even during dry periods. Over the next two centuries, significant advances were made in the understanding of hydrological processes, notably Dalton's work on evaporation [Dalton, 1802], Darcy's law of porous media flow [Darcy, 1856], Manning's work in open channel flow [Manning, 1891], Richards' development of the unsaturated flow equation [Richards, 1931], and Horton's work in infiltration and runoff production [Horton, 1933] and river basin geomorphology [Horton, 1945].

These, and other major figures in hydrology, have contributed much toward our understanding of catchment storm responses, yet recent work suggests that there are still many problems left to explore. The problem is, in some ways, deceptively simple. We know that in any particular storm only part of the incident rainfall reaches the stream to become discharge during the (often rapid) rise and fall of the river known as the storm hydrograph. The remainder is absorbed into the soil (referred to as infiltration) and may be returned back to the atmosphere as evapotranspiration or contribute to the stream as a subsurface flow over an extended period of time. It is therefore important to be able to estimate the proportion of the rainfall equivalent to the volume of the stream hydrograph. This proportion is commonly called the "effective rainfall." We also know that the flow processes within the catchment have both retardation and diffusive or filtering effects, so that it is also important to know the time transformation between effective rainfall and the discharge hydrograph (see Figure 1).

Quite apart from the pure scientific interest of understanding the flow processes, being able to predict catchment storm responses is of considerable practical interest in the proper management of water resources and estimation of risk to life and property during major flood events. Hydrologists have not shirked from the requirement for quantitative prediction of catchment responses. Today there is a plethora of operational computer models available, few of which reflect more than a cursory understanding of the flow processes involved. Perhaps the most widely used model in engineering hydrology remains...
the unit hydrograph concept, developed by Sherman [1932] and often combined with Horton's concept that storm response is due primarily to surface runoff generated when rainfall rates exceed the infiltration capacity of the soil. The unit hydrograph represents a simple linear transfer function, assumed constant for a particular catchment, that provides the time transformation between the estimated effective rainfall and the resulting storm hydrograph in the stream channel. Thus these concepts have the two elements, a loss function for estimating effective rainfall and a time transformation, that are required to predict catchment storm response.

The difficulty in using the unit hydrograph approach lies primarily in the prediction of the effective rainfall. The proportion of storm rainfall that is equivalent to the volume of the storm hydrograph depends on many factors, notably the antecedent moisture status of the catchment, the nature of the soils in the catchment, and the storm rainfall intensities. As Freeze [1974, p. 627] has pointed out, the complexity of this response tended to promote the development of generalized regional analyses that used entire drainage basins as black box response units. This type of hydrologic prediction model treats the rainfall-runoff relationship in an empirical statistical manner without considering the causal mechanisms. This approach has great power in satisfying the needs of engineering design but does not provide any insight into the internal mechanisms of the hydrologic cycle.

Ideally, a prediction of the effective rainfall should be based on a proper understanding of the processes involved, but study of these processes has revealed patterns of real complexity.

During the International Hydrological Decade (1964–1974) there was a burst of activity aimed at a better understanding of hydrological processes through field experimentation. The focus of much of this research was concerned with the generation of streamflow in headwater tributaries during both storm and interstorm periods. This research is reviewed by Freeze [1974] and is more fully described by Kirkby [1978].
development resulting from this work was a reevaluation of the Hortonian concept of storm runoff as rainfall in excess of infiltration capacity of the soil (this type of runoff is often referred to as infiltration excess runoff generation).

Runoff generation is now known to result from a complexity of mechanisms, as shown in Figure 2. During any particular storm, different mechanisms may generate runoff from different parts of a catchment. Surface runoff from these (partial) contributing areas may be generated by either the infiltration excess mechanism on low-permeability soils, or from rainfall on areas of soil saturated by a rising water table even in high-permeability soils (referred to as saturation excess runoff generation). These saturated contributing areas expand and contract during and between storm events. It is also now known, as a result of using natural tracers to determine the source of storm discharges in streams, that in many catchments a significant proportion of the storm hydrograph is derived from subsurface water that is displaced from soil and groundwater by the incoming rainfall as subsurface stormflow.

Runoff generation has thus been revealed to be a highly nonlinear and spatially variable process, involving both surface and subsurface flow pathways. It is clear that predictions of runoff generation need to deal with the spatial patterns of antecedent moisture conditions, soil hydraulic characteristics, and rainfall intensities, with the expectation that under some conditions these patterns will be highly variable in space and time. And yet, many hydrological models continue to be based on a point scale model of catchment response, with the only recognition of areal variability being a multiplication by the catchment area. In fact, if the rainfall and discharge variables are expressed in units of volume or depth per unit area then even this last step is not necessary.

Overall, very little progress has been made in relating the small-scale complexity that is apparent from observations and experimental studies of flow processes to the relative simplicity (loss function and time transformation)

---

**Figure 2.** Mechanisms of runoff production: $P$ is precipitation, $P_c$ is channel precipitation, $f$ is infiltration, $of$ is overland flow, $rf$ is return flow, $i$ is interflow, $uf$ is unsaturated zone flow, and $wi$ is water table [from Beven, 1986b].
required for practical predictions at the catchment scale (see, for example, Pilgrim [1982]). It is this relationship that is the subject of this review. Dooge [1982] posed what we feel is a central issue facing hydrology today: What is the effect of spatial variability on the parameterization of hydrologic processes at a range of scales? As pointed out by Gupta et al. [1986, p. vii],

The scale problems in hydrology stem from the recognition that the mathematical relationships describing the physical relationships are manifest at different space-time scales. The broad scientific problem then is to identify and formulate suitable relationships at the scales of practical interest, test them experimentally and seek consistent analytical connections between these relationships and those known at other scales.

The research reviewed in this paper represents the very first attempts to investigate the effects of spatial variability and scale on the quantification and parameterization of catchment storm response. As will be seen, the research to date has concentrated on the dynamic nature of surface runoff generation during storm events, perhaps the most intriguing and central problem in hydrology for both theoretist and field hydrologist alike. Section 2 of this paper presents a semidistributed catchment model that incorporates the major mechanisms of runoff generation. The model takes into account the variability of topography, soils, and rainfall inputs in a physically realistic but computationally efficient way that has allowed a wide range of numerical experiments to be performed. As a result of these experiments, in section 3 the influence of spatial variability and scale are addressed in terms of a representative elementary area (REA) concept along with a macroscale model that is valid at scales larger than the REA. The first attempt at a theory of catchment hydrological similarity over a range of scales is presented for these models in section 4. Section 5 looks at similarity and scale within the context of flood frequency analysis with some results for idealized catchments. The results are drawn primarily from the papers of Wood and Hebson [1986], Sivapalan et al. [1987], Beven et al. [1988], Wood et al. [1988a], and Sivapalan et al. [1989]. Related papers included those of Beven [1986a, b], Hebson and Wood [1986], and Sivapalan and Wood [1986].

We feel that the first steps have now been taken toward a theoretical understanding of the role of small-scale variability and complexity in contributing to the responses at different catchment scales. The results have implications for a wide range of hydrological problems, including flood frequency analysis and understanding differences in the response of catchments within the same geographical region. In addition, the research has particular importance for the formulation of an appropriate hydrological land surface parameterization in climate modeling where it is necessary to have models at the global circulation model (GCM) grid scale that are consistent with the variability of hydrological processes at the subgrid scale. The sensitivity of GCM results to this lower boundary condition provided by the land surface hydrology may be especially important in studies of the hydrological impacts of climate change [Environmental Protection Agency, 1989]. We expect that the results presented will be refined and improved but hope that the work is a pointer to a productive direction for research that will be central to the progress of hydrology.

2. A MODEL OF RUNOFF PRODUCTION BASED ON CATCHMENT TOPOGRAPHY

A number of seminal field studies during the International Hydrological Decade (for example, those by Betson [1964], Hewlett and Hibbert [1963, 1967], Hewlett and Nutter [1970], Dunne and Black [1970], and Weyman [1970]) had a significant impact in the rethinking of runoff generation as perceived in the Horton infiltration excess overland flow–unit hydrograph approach. This work led to an increasing recognition that subsurface flows play a very important part in the redistribution of soil moisture between storm events in setting up the initial wetness conditions that govern streamflow generation for the subsequent rain event. It was also realized that for areas of relatively shallow soil, a dominant control on these subsurface flows is the local topography. In particular, areas of higher antecedent wetness, and therefore greater likelihood of generating runoff, should be expected in areas of convergent flow in plan and concave slopes in profile such as are commonly found in hillslope hollows and above the heads of the smallest ("first order") stream channels.

However, there are few hydrological models that take explicit account of the variability of catchment topography and soil characteristics. One approach that does was initiated by Freeze and Harlan [1969] and is based on partial differential equations describing surface and subsurface flow processes. These models are now being implemented, with some simplifications, at the catchment scale [e.g., Bathurst, 1986; Beven et al., 1987]. The flow equations must be solved numerically, using finite element or finite difference approximations, which at the catchment scale involves a very large number of calculation nodes and large matrices to be solved for many time steps. A recent study of a small headwater catchment involving a 12,000-node simulation over a period of 150 days required a run time of 50 hours on a Cyber 205 supercomputer [see Binley and Beven, 1989]. Such a computational burden has generally limited the application of these detailed models to single storm simulations and together with the number of parameter values incorporated into distributed models, makes calibration to real catchments (which may involve multiple runs) very difficult.

There are other problems with this approach. While these models have a firm theoretical foundation, the
physical theory on which they are based is theory for laboratory systems with well-defined boundary conditions. For real catchment systems with heterogeneous small-scale features like variable inputs at ground level due to the effects of a vegetation cover, preferential flow pathways through the soil due to the effects of soil structure and rills and microtopography affecting surface flow, there is no assurance that the same equations should be good descriptors of flow processes in the field, or that the parameters measured in the field will have values that are appropriate at the model grid or element scale within a particular model structure (see discussion by Beven [1989]).

However, since soil and topography are thought to be primary controls on water flows, and the topographic form of a catchment is usually readily available for a particular catchment from maps or remote sensing techniques, it should be possible to formulate a model, computationally simpler than those described above, but complex enough to retain the essence of the controls of soil, vegetation, and topography on runoff production. With this aim, a decade ago Beven and Kirkby [1979] developed a model (TOPMODEL) for transforming rainfall into runoff that utilizes a spatially variable topographic index to predict the occurrence and extent of saturated areas which generate surface runoff and subsequently streamflow. They provided a theory for local hydrological similarity in which all locations within the catchment having the same value of the topographic index were assumed to respond in a similar manner to similar inputs. The calculation of the topography index using maps was very time consuming and only recently, with the advances in interactive graphics and increasing availability of digital elevation models (DEMs), can it be calculated routinely [Band, 1986; Band and Wood, 1988]. The availability of DEMs and remotely sensed data (from Système Probatoire d’Observation de la Terre (SPOT), Thematic Mapper (TM), and Synthetic Aperture Radar (SAR)) at the 20–50-m scale, allows for a distributed version of TOPMODEL where the spatial variability in topography, soils and vegetation can be properly represented.

In what follows we describe a version of the model that predicts patterns of soil saturation and their relationship to both saturation excess and infiltration excess surface runoff generation making use of information on the variability of topography and soil characteristics. For simplicity in this initial study of scale and similarity effects, we concentrate on the prediction of surface runoff generation rates and volumes only during individual storm periods. The routing of the generated runoff over hillslopes and along stream channels is not included here but has been treated elsewhere in related work (see Beven, 1986a, b; Wood et al., 1988a; Sivapalan et al., 1989). Reinfiltiration of surface runoff on downslope unsaturated areas is also not treated explicitly. A final simplification is that subsurface storm runoff production and subsurface flows during interstorm periods are not explicitly treated, although the importance of subsurface stormflow is recognized and its prediction is possible within the TOPMODEL framework [Beven, 1986a, b].

2.1. Model of Saturation Excess Runoff Production

The model starts from the premise that prior to a storm the antecedent water table within the soil can be represented as the result of a steady average vertical recharge rate \( r \) so that at any location \( i \), downslope flow \( q_i \) will be given by

\[
q_i = ar
\]

We also assume there is relatively shallow soil so that the water table will be nearly parallel to the soil surface resulting in the local hydraulic gradient being close to the local slope angle, \( \tan \beta \), and that the transmissivity of the soil is an exponential function of depth, so that the downslope flow rate is also given by

\[
q_i = T_i \exp (\frac{-z_i}{\tan \beta}) \tan \beta
\]

Here \( f \) is a parameter, assumed constant in the catchment, that describes the rate of decline of soil transmissivity with depth. Combining (1) and (2) and integrating over the catchment area to obtain a catchment average depth to the water table, it can be shown that the relationship between this average depth \( \bar{z} \) and a local depth \( z_i \) is given by

\[
z_i - \bar{z} = \frac{1}{f} [\lambda - \ln \left(\frac{T_i}{\bar{T}} \tan \beta\right)]
\]

where \( \ln (T_i) \) is the areal average value of \( \ln (T_i) \) across the catchment and \( \lambda \) is the expected value of the topographic variable \( a/T_i \tan \beta \) which is a constant for a particular catchment. Beven [1982] has shown that the assumption that soil permeabilities show an exponential decline with depth is reasonable for a wide range of soil types. A very similar approach to predicting areas of saturated soil has been derived independently by O’Loughlin [1981, 1986] for more general profiles of saturated hydraulic conductivity, but without some of the analytical advantages of the exponential assumption. Equation (3) implies that all locations in the catchment with the same value of the combined topographic-soil index \( aT_i/(T_i \tan \beta) \) will have the same relationship between local depth to water table and the mean depth, i.e., it is an index of local hydrological similarity. Inspection of this equation shows that locations with high values of the topographic-soil index relative to the catchment constant \( \lambda \) will have smaller water table depths and will consequently have a greater propensity to saturate to the surface. In fact, for any particular value of \( \bar{T} \), knowledge of the pattern of soil and topography and therefore of the topographic-soil index, allows prediction of those areas where \( z_i \leq 0 \) and therefore the saturated contributing area. In fact, we should strictly define the area of saturation where \( z_i \leq \psi_c \), where \( \psi_c \) is the depth of the capillary fringe which is the depth of soil above the
water table that is saturated but within which water is held at a negative capillary potential. Prediction of the changes through time of $\bar{T}$ then allows the dynamic expansion and contraction of this contributing area to be modeled.

We may separate the topographic and soil contributions to the combined index by rewriting (3) as

$$f(z) = -[\ln (a/\tan \beta) - \lambda] + [\ln (T_c) - \ln (T_s)]$$

(4)

The variations in the catchment of the topographic variable $(a/\tan \beta)$ and the soil transmissivity $T_s$ may both be represented as distribution functions. Figure 3 shows the cumulative distribution functions for the Kings Creek catchment in Kansas. The right-hand side of (4) is then clearly the sum of two deviations of the local values of $\ln (a/\tan \beta)$ and $T_s$ from their catchment mean values, whereas the left-hand side represents the local deviation in water table depth from the catchment mean value scaled by the parameter $f$. Reference to Figure 3 shows that the expected deviations in the topographic variable are far greater than for the local values of transmissivity. Thus variability in the transmissivity will have a relatively small effect on the distribution of the combined index and consequently on the predicted patterns of water table depths and saturated contributing area for a given value of $z$.

The mean transmissivity will, however, have a major control on drainage during interstorm periods and consequently on the value of $\bar{T}$ prior to any storm. Within this theoretical framework it can be shown that subsurface outflows $Q_s$ during interstorm periods can be described by

$$Q_s = Q_o \exp (-\bar{T})$$

(5)

where $Q_o = A T_s \exp (-\lambda)$ and $A$ is the catchment area [Beven, 1986a]. Given an initial discharge prior to a storm, (5) can be inverted to estimate a value of $\bar{T}$ and hence using (3) to give the initial saturated area and pattern of local water table depths. Rainfall onto the saturated area will become surface runoff, while during the storm, infiltration takes place at locations where $z > \psi^e$. Where the rainfall is great enough to fill any storage deficit above the initial water table position, additional surface runoff will be generated and the contributing area expands out from its initial position.

2.2. Model of Infiltration Excess Runoff Production

Away from the initial contributing area, rainfall rates may be such as to exceed the infiltration capacity of the soil and produce infiltration excess runoff, even though the soil profile remains in part unsaturated. This will occur most readily on areas of low-permeability soil but will also depend on the initial moisture content at the soil surface which controls the local moisture deficit and vertical hydraulic gradient. The infiltration excess component of the model presented here is based on the Philip equation [Philip, 1969], which is an approximate analytical solution to the nonlinear Richards' equation, modified to take into account the time taken for the rainfall to bring the soil surface just to saturation (called the time to ponding). The unsaturated soil moisture characteristics are described by the Brooks and Corey [1964] equations which provide the functional relationship between suction pressure $\psi$ and soil moisture $\theta$ and between $\psi$ and hydraulic conductivity $K$. The Philip equation is not strictly compatible with the TOPMODEL framework in that it does not allow for the type of exponential decline in soil permeability with depth assumed within TOPMODEL. Beven [1984] has derived an infiltration model which is consistent with TOPMODEL, but for the current purpose it is not analytically tractable. We do not feel that this will have a significant effect on conclusions drawn from the study.

With these assumptions the infiltration rate $g(t)$ for any particular value of the surface saturated hydraulic conductivity of the soil, $K_s$, is given by

$$g(t) = p$$

(6a)

$$g(t) = C K_s + 1 \frac{1}{2} \frac{1}{2} \psi^e (t_f - t_p)^{-1/2}$$

(6b)

where $t_p$ is the time to ponding (the time when the infiltration rate falls below the rainfall rate $p$). Relationships for the variables $t_p$, $t_c$, $C$, and $S$ are functions of soil
characteristics, the rainfall rate, and the depth to the water table (which influences the moisture content of the column). The relationships are given by Sivapalan et al. [1987]. At each location characterized by the topographic-soil index \(aT/\tan \beta\) and the saturated hydraulic conductivity at the surface, \(K_o\), infiltration will take place at the rate \(g(t)\), with runoff at the rate \((p - g)\) until the time at which the cumulative infiltrated volume \(G(t)\) exceeds the initial storage deficit. Beyond this time, infiltration is assumed to be zero, and runoff occurs by the saturation excess mechanism at the rate \(p\).

Results from the application of this model show that the predicted surface runoff production by the saturation excess mechanism follows the pattern of catchment topography, with the greatest likelihood of runoff in valley bottoms and convergent headwater areas. Infiltration excess runoff production may also be affected by topographically controlled initial moisture profiles. Predictions show that both infiltration excess and saturation excess runoff can be produced within a catchment depending on position on the hillslope, soil, and rainfall characteristics (see Figure 4).

3. RUNOFF VARIABILITY AND SCALE

For hydrologists interested in land-surface processes, the variability in runoff (and how this variability changes with catchment scale) is the raison d'être for much current research. Field observations have shown that the major sources of heterogeneity leading to spatial differences in runoff are topography, soils, and rainfall (reviews of such field work can be found in the work by Kirkby [1978]). Figure 3 suggests that the variability in our topographic index, and thereby variability in topography, hillslope forms, and subcatchment shapes, will play a significant role in defining the patterns of runoff at small scales as predicted by our model. This model behavior is consistent with reported field observations; for example, see the
contributions by Dunne (chapter 7) and Kirkby (chapter 9) in the work by Kirkby [1978]. Until recently it was very cumbersome to model the influence of topographic variability on small-scale runoff processes because maps had to be used for the calculation of any topographic index. Now the wide availability of digital elevation models (DEMs) of terrain and the graphics workstations needed to process this data removes any computational barriers that may have existed.

Several studies have looked at the influence of soil heterogeneity on storm runoff, including the modeling work of Smith and Hebbert [1979] and Freeze [1980]. In many cases we feel that topographic variability will have a greater influence on the spatial variability of runoff at the subcatchment scale than will variability in soil properties. One reason is shown in Figure 3 where variability in our topographic index dominates over soil transmissivity variability at the small subcatchment scale. If our model of a catchment's response to rainfall is reasonably correct, then this implies that soil variability will have a smaller influence on the spatial pattern of water depths and subsequent storm responses than will the local topography. Wolock [1988] calculated the distribution of our topographic and topographic-soil indices for 145 catchments in the northeastern part of the United States; his results support Figure 3 in that variability in topography dominates the variability in the topographic-soil index and subsequently the predicted spatial patterns of storm runoff. One can certainly imagine catchments where soil variability will dominate topographic variability but Wolock rejected only four catchments as having little or no relief in the DEM data base.

Figure 4 presents some patterns for soil transmissivity, rainfall and storm responses for Kings Creek catchment, a catchment of 11.7 km² that is part of the Konza Prairie Natural Reserve near Manhattan, Kansas. The patterns are overlaid on the topographic DEM for the catchment. Figure 4b gives the precipitation pattern for a storm which occurred on August 4, 1987, while Figures 4c and 4d give the model-predicted catchment response showing areas of storm runoff due to infiltration excess and saturation excess processes as a function of initial catchment dryness. The purpose of Figure 4 is to illustrate how the predicted patterns vary within the catchment and with initial dryness. Looking at Figure 4, one can appreciate the variability in the catchment response at the hillslope scale. In fact, if one divides up the catchment into smaller subcatchments, as is shown in Figure 5 for Kings Creek, one can investigate this variability over different scale. With increased scale, the increased sampling of hillslopes should lead to a decrease in the difference between subcatchment response. At some scale, the variance between storm responses for catchments of the same area should reach a minimum.

Wood et al. [1988a] suggest that this threshold scale represents a "Representative Elementary Area" which is proposed to be the fundamental building block for catchment modeling. The REA is a critical scale at which implicit continuum assumptions can be used without knowledge of the actual patterns of topographic, soil, or rainfall fields, although it would be necessary to account

![Figure 5. Natural division of Kings Creek. Division was carried out by the techniques of Band and Wood [1988].](image-url)
for the underlying variability of these parameters through distributional functions.

To investigate this concept, the storm responses were modeled for the four rainstorms shown in Figure 6. Using the natural subcatchments of Figure 5, the storm runoff was calculated for each subcatchment. Since the four storms resulted in different rainfall depths, the storm runoff was rescaled by the average rainfall depth over the catchment. This ratio of storm runoff to rainfall depth is plotted in Figure 7 against the subcatchment area as represented in pixels. Here one pixel is 900 m$^2$, and 1100 pixels is approximately 1 km$^2$. By visual inspection we suggest that the size of the representative elementary area is approximately 1 km$^2$. It is interesting to note that in modeling a different catchment with steeper topography and using simulated rainfall patterns, Wood et al. [1988a] also found an REA of the order of 1 km$^2$.

The results of Figure 7 suggest that at scales larger than the REA it should be possible to simplify the representation of catchment responses, while still retaining the important effects of heterogeneity in the hydrological processes. For catchment areas much larger than the correlation lengths of soil and rainfall variability (assumptions consistent with our REA theory) one needs not consider the actual patterns in rainfall, topography, and soil but can consider them statistically through their means and variances.

Using the statistical distribution of the topographic-soil index, one can determine the fraction of the catchment that will be saturated due to the local soil storage being full. These areas will generate saturation excess runoff at the rate $\bar{p}$, the mean rainfall rate. For that portion of the catchment where infiltration occurs, the local expected infiltration rate, $m_s$, can be calculated as the difference between the mean rainfall rate, $\bar{p}$, and the local expected infiltration rate, $m_s$. This difference can be expressed as

$$m_s [\ln (aT_i/T; \tan \beta)] = \bar{p} - m_s [\ln (aT_i/T; \tan \beta)]$$

(7)

where both the local expected runoff rate and the local expected infiltration rate are (probabilistically) conditioned on the topographic-soil index, $\ln (aT_i/T; \tan \beta)$. The runoff production from the catchment is found by integrating, usually numerically, the conditional rate over the statistical

---

**Figure 6.** Rainfall intensity pattern for August 4, 1987, at four different times: (a) 0815, (b) 0845, (c) 0900, and (d) 0930.
distribution of the topographic-soil index. Figure 8 gives results for the storm patterns presented in Figure 6. The increase observed over the time intervals is indicative of an increase in the runoff ratio during the storm.

For the simplified case where only infiltration excess is considered, it is possible to derive analytically the macroscale model [Sivapalan, 1986; Wood et al., 1988a]. Figure 9 gives results for one set of simulations in which the macroscale model is compared to the averaged microscale model.

4. CATCHMENT VARIABILITY AND HYDROLOGIC SIMILARITY

The results from the REA analysis suggest that progress has been made in understanding the transition from microscale to macroscale parameterization as we increase scale in the presence of spatial variability. This does not address the question as to how catchments may be hydrologically similar. Definitions of similarity used in the past have been based on physioclimatic characterizations without explicit recognition of the environmental controls on runoff generation. The work of Wood and Hebson [1986] and Sivapalan et al. [1987, 1989] redefine this to consider similarity relationships for catchment runoff responses that are independent of basin scale.

The development of a scaled model of storm response requires the specification of the appropriate scaling parameters. This requires scaling the topographic-soil index function, the soil characteristic relationships, and the rainfall intensity and duration. The scaling must be consistent in that the underlying equations are not altered.

Earlier in the paper it was argued that variability in topographic-soil index played a major role in determining both the patterns and magnitudes of generated runoff. It is quite reasonable that this index be described in a probability distributional form, and in doing so it allows one to define a convenient and appropriate scaling parameter. Let $x = \ln \left( \frac{a}{\tan \beta} \right)$ and let the areal distribution of $x$ be
Figure 8. Comparison of storm runoff generated from the distributed model and from the macroscale model at four time intervals on August 4, 1987: (a) 0815, (b) 0845, (c) 0900, and (d) 0930.

(a) Mean sub-catchment area (900 M pixels)

(b) Mean sub-catchment area (900 M pixels)

(c) Mean sub-catchment area (900 M pixels)

(d) Mean sub-catchment area (900 M pixels)
Figure 9. Comparisons of simulation and analytical results for the case of spatially variable soils and rainfall and runoff production by infiltration excess: (a) runoff rate and (b) cumulative infiltration volume, $\lambda_r = 625$ m [from Wood et al., 1988a].

Described by a three-parameter gamma probability distribution with mean $\lambda = \chi \phi + \mu$, variance $\chi^2 \phi$, and coefficient of variation $C_r = \chi \sqrt{\phi} / \lambda$. Figures 10a and 10b show, as an illustration, the topographic index fitted for two catchments, where $A_1/A$ represents that fraction of the catchment having an index value greater than the abscissa value. Wolock's [1988] analysis of 145 catchments found that the three-parameter gamma probability distribution fitted the calculated topographic index quite well.

Field evidence presented by Nielsen et al. [1973], Russo and Bresler [1981], Beven [1983a] and others show that saturated hydraulic conductivity of soil tends to follow a lognormal probability distribution. In developing our scaled model, we assume that variability in hydraulic conductivity represents the major influence in soil variability. This is justified with the argument that variability in runoff production will be less sensitive to other soil parameters. We assume that saturated hydraulic conductivity is lognormally distributed, which we further assume will also hold for hydraulic transmissivity. In general, we will have little or no information about the actual pattern of conductivity or transmissivity with respect to topography. If we are concerned at modeling scales larger than the REA scale, such patterns may not be important. This is supported by the results of Beven [1983b] who found through Monte Carlo simulation that the effect of soil patterns is small when its spatial correlation length scale is small relative to the catchment scale.

To a first approximation, we assume that the soil transmissivity, $T$, is independent of the topographic index and that the probability distribution for $\ln (T/T)$ is normal with zero mean and variance $\sigma_T^2$. By assuming for the topographic index, $\ln (a / \tan \beta)$, a three-parameter gamma probability distribution with parameters $\mu$, $\phi$, and $\chi$, and by
defining \( x^* = \ln (aT/L \tan \beta) \), the topographic-soil index, then \( x^* \) is approximately distributed as a gamma \((\mu^*, \phi^*, \chi^*)\), where

\[
\begin{align*}
\mu^* &= \mu \\
\frac{1}{\phi^*} &= \frac{1}{\phi} + \frac{\sigma_T^2}{\chi^2} \\
\chi^* &= \chi + \frac{\sigma_T^2}{\chi^2}
\end{align*}
\]

It can be shown that the scaled parameter \( v = x^* - \mu/\lambda - \mu \) follows a one-parameter gamma distribution,

\[
f_v(v) = \frac{\phi^*}{\Gamma(\phi^*)} (\phi^*v)^{\phi^*-1} \exp(-\phi^*v)
\]

with shape parameter \( \phi^* \), scale parameter \( 1/\phi^* \), mean 1, and variance \( 1/\phi^* \). The variable \( v \) is scale independent, and \( \phi^* \) reflects the variability in soil and topography and is independent of the scale of the catchment. Thus \( \phi^* \) captures in a scaled manner the net hydrological effects of the convergence, divergence, concavity, convexity, and soil transmissivity characteristics of the hillslopes constituting the catchment. Within our theory of hydrologic similarity, the topographic-soil effects of two catchments having the same \( \phi^* \) would be the same. Figure 10b shows the cumulative distributions of the scaled topographic-soil index for two catchments. Notice that their scaled topographic indices are almost identical, implying that the effect of topography on “scaled” storm surface runoff production would be similar for both catchments. Figure 10c shows the cumulative distribution function \( f_v(v|\phi^*) \) for different values of the parameter \( \phi^* \).

Based on the scaling which defined \( v \), a consistent scaling for the water table depth would be \( f(\lambda - \mu) \). For the saturation excess runoff generation, using (5), (3) can be written as

\[
z^* = Q^* + (1 - v)
\]

where \( Q^* = -\ln [Q(0)/Q_0]/(\lambda - \mu) \), the scaled average water table depth which represents the catchment scaled wetness at the start of the rain event.

To scale the infiltration excess portion of the storm response, different scaling relationships are needed. Intuitively, we will require scaling parameters that consider soil characteristics and storm duration, since the amount of the rainfall which infiltrates depends on these two factors. We propose for the time dimension the scaling parameter \( t^* \), the storm duration, and for the length dimension \( \psi_2(\theta, - \theta) \), parameters of the Brooks-Corey soil moisture characteristic relationship [Brooks and Corey, 1964]. The Brooks-Corey

Figure 10. (a) Cumulative \( \ln (a/tan \beta) \) distributions of the Coweeta Experimental Basin and a subcatchment of the Davidson River catchment (unconnected points) and their gamma approximations (solid curves), (b) cumulative distributions of the scaled topographic index \( v \) for the two basins, and (c) cumulative distribution of \( v \) for different values of \( \phi^* \) [from Sivapalan et al., 1989].
soil moisture relationship is a functional relationship that relates suction pressure, \( \psi \), to the soil moisture, \( \theta \).

\[
\theta(\psi) = \theta_s + \left( \theta_r - \theta_s \right) \left( \frac{\psi}{\psi_c} \right)^B
\]

(11)

Here \( \psi_c \) is the depth of the capillary fringe, \( \theta_s \) is the saturation soil moisture content (which is equal to the soil porosity), and \( \theta_r \) is the residual soil moisture content, the value of \( \theta \) when the soil dries and \( -\psi \) becomes very large. The above scaling leads to the following scaled parameters:

\[
\kappa^* = \frac{t_i}{t_{id}}
\]

(12a)

\[
p^* = \frac{p_{i*}}{\psi_c (\theta_r - \theta_s)}
\]

(12b)

\[
g^*(t^*) = g(t) \frac{t_{id}}{\psi_c (\theta_r - \theta_s)}
\]

(12c)

Equation (6) still holds but in the scaled parameter space. Sivapalan et al. [1987] provide the details for the scaled parameters. Similarly, (7) would also hold in the scaled parameter domain. Through scaling the runoff production process equations, the explicit consideration of spatial and temporal scales has been removed without sacrificing the modeling of the hydrological processes. Scaling of the model equations has identified five dimensionless catchment similarity parameters and three dimensionless auxiliary conditions which govern the scaled storm response. The similarity parameters include two scaled soil hydraulic conductivity parameters, the mean scaled hydraulic conductivity, \( K^*_a \), and the coefficient of variation for \( K^*_a \); two soil moisture characteristics, the scaled size of the capillary fringe, \( \psi_c^* \), and a Brooks-Corey parameter \( B^* \); and a scaled topographic-soil index parameter \( \phi^* \). Two catchments can be said to be hydrologically similar in terms of surface runoff production within the assumptions of our model if they are identical in these five parameters, regardless of scale. The two auxiliary conditions include a scaled initial flow \( Q^*_i \) and two rainfall distribution parameters, the ratio of the mean rainfall rate to the mean soil hydraulic conductivity and the coefficient of variation of the rainfall rate.

One purpose of such scaling is to investigate the relationship between environmental controls (topography, soils, and climate) and runoff production without the compounding effects of scale. Sivapalan et al. [1987] carried out sensitivity analysis for a number of parameters, and the reader is referred there for details. Some conclusions of those runs are worth discussing here.

As expected, topography plays a major role in defining the dominate runoff mechanism. In fact, varying the soil heterogeneity, keeping the mean hydraulic conductivity constant, had very little effect on the saturated excess runoff production. Finer grained soils tended to increase runoff by producing saturated areas more quickly and increasing their extent. Increasing (spatially) mean rainfall rates relative to mean soil conductivity has a profound effect on runoff, as one would expect: strongly increasing both saturated areas and infiltration excess runoff. On the other hand, increasing rainfall variability (e.g., convective storms or a large catchment when compared to the storm size) had little impact on saturation excess production.

It is doubtful that there would be any two catchments that satisfy jointly all the dimensionless parameters. On the other hand, there could be a range of values for each of the parameters for which similarity could be established between catchments. In addition, sensitivity analysis may indicate the most probable hydrologic regimes that would occur under given topographic, soil, and climatic combinations. Finally, it is hoped that these results would help in guiding field experiments and in interpreting data from such studies.

5. FLOOD FREQUENCY, SIMILARITY, AND SCALE

The determination of the peak runoff from a rain event requires that the generated runoff be routed down the channel network to the basin outlet or gaging point. The analysis of this process, which includes the hillslope response and the channel response, by channel routing using simplified equations of fluid flow has been investigated by Kirkby [1986] and Wood et al., [1988b]. The review of this work is beyond the scope of this paper. A more simplified approach that captures the aggregated response of the basin is to model the channel response as a simple linear response function, i.e., an instantaneous unit hydrograph (IUH). The IUH is the response to a unit input occurring over a time \( \tau \), as \( \tau \rightarrow 0 \).

Henderson [1963] showed that if the response function was simplified to be triangular in shape, then this "triangular unit hydrograph" was sufficient for the prediction of the channel response. This simplification required that the peak and time to peak be preserved between the actual response function and the simplified response function. Wood and Hebson [1986] derived the peak discharge \( Q_p \) from a catchment due to storm runoff generated at a constant rate \( q \) with duration \( t_q \) by utilizing the simplified triangular IUH having a peak \( h^*_q \) and time to peak \( t_p^* \). The resulting expression for \( Q_p \) is as follows:

\[
Q_p = q \frac{h^*_q t_q}{4} \left( 1 - \frac{h^*_q t_q}{4} \right) \quad h^*_q t_q \leq 2
\]

(13a)

\[
Q_p = q \frac{h^*_q t_q}{2} \quad h^*_q t_q > 2
\]

(13b)

The parameterization of (13) could be based on any number of expressions for the IUH peak and time to peak. A particularly interesting one is to use the results from Rodrigues-Iturbe and Valdes [1979] where equations for the peak and time to peak are developed in terms of Horton's geomorphological laws [Horton, 1945]. Channel networks were first studied quantitatively by Horton [1945] who defined an ordering system for individual links of the network based on the network structure. Horton
showed that the expected values for the number of streams, the length of streams, and the stream slope were significantly different for each order and suggested that the structure of individual networks could be represented by a small number of descriptive parameters.

By using these relationships of Horton to describe the structure of the channel network within a catchment and subsequently the response function of the catchment (which they call the geomorphologic unit hydrograph), Rodriguez-Iiturbe and Valdes [1979] represent similarity in basin responses for basins having the same Horton ratios and scale. They found that the unit hydrograph peak could be expressed as

\[ h_p = 0.364 R_L^{0.43} \frac{v}{L} \]

(14)

where \( R_L \) is Horton’s length ratio, \( v \) is the peak velocity, and \( L \) is the length of the highest-order stream. The peak \( h_p \) is the discharge per unit area per unit depth of runoff (effective rainfall), resulting in units \( \text{T}^{-1} \).

Wood and Hebson [1986] defined a scaling time \( \tau = L/\nu \) which they define as the characteristic basin response time that they used to define a dimensionless geomorphic unit hydrograph (DGUH). This led to a scaled DGUH peak discharge of \( h_p = 0.364 R_L^{0.43} \) and a scaled excess storm duration, \( t_d = t_d/\tau \). The scaled peak runoff function, due to scaled generated runoff \( q^* \), is found by substituting these scaled variables into (13).

The results from Wood and Hebson [1986] only considered temporally constant excess rainfall which allowed an analytic solution for the scaled flood frequency curve. Further, they had presupposed the probability distribution of the excess rainfall and did not consider more complex mechanisms for runoff generation. Sivapalan et al. [1989] relaxed these considerations and found the flood frequency curves by numerical simulation. Their runoff generation model is that described earlier in this paper which considers both infiltration excess and saturation excess (but not subsurface storm flow) mechanisms. Flood frequency curves usually plot the largest observed peak discharge during any year (annual peak discharge) against the probability that it will be equalled or exceeded in any year. The exceedance probability is usually expressed as its inverse and denoted as the “return period” for that annual peak discharge.

Only one set of results from Sivapalan et al. [1989] will be discussed here. The shape of the flood frequency curve appears to result from a complex interaction of topography, soil, and climate. Figure 11 shows three scaled flood frequency curves for a catchment with fixed topography; they differ in the mean rainfall when compared to the mean soil conductivity. The curves are the average curves from 25 repetitions of 4000 storms; i.e., 4000(25) storm simulations. The top two curves represent an environment where the flood peaks are dominated by infiltration excess runoff; thus the distributional characteristics of the rainfall are passed through to the flood frequency curve. The bottom curve, which represents a low-rainfall climate, is dominated at low return periods by saturation excess runoff and at high return periods by infiltration excess runoff. Sensitivity analysis shows that the saturation excess flood curve is extremely stable while the infiltration excess curve is more variable (see Figure 12). At this time we believe that the above modeling has produced an important result, that for some catchments the flood frequency curve at low return periods is defined by catchment characteristics (through saturated excess storm runoff) while at high return periods it is defined by rainfall characteristics (through infiltration excess storm runoff).

This finding appears to be supported by data. Figure 13 presents annual flood peak data for three catchments, and their shapes resemble the shapes of Figure 12. More work is needed to verify this hypothesis, but the modeling results provide some guidance for this work.

The purpose of this example is to demonstrate how similarity-based analyses can be used to study basic hydrological processes over a range of scales and to reveal characteristics and interactions among processes previously overlooked. Research is ongoing to gather data over a range of catchments representing different climates and scales with which this theory can be further tested and applied.
6. CONCLUSIONS

This review has summarized research by the authors in two important interrelated areas of catchment hydrology: the effects of scale on hydrologic parameterization and the concept of hydrologic similarity. Our research into scale problems in hydrology attempts to understand the effect of spatial variability on the parameterization of hydrological processes and to seek consistent models across a range of scales. Field studies over the last 30 years have shown that two main mechanisms generate surface runoff and that these may both occur within a catchment during a rain event. These mechanisms are the infiltration excess runoff of Horton and the saturation excess runoff of Dunne. Their spatially variable controls (topography, soil, rainfall, and vegetation) lead to patterns of runoff generation across the catchment.

In the work by Wood et al. [1988a] the concept of a representative elementary area was proposed as the fundamental scale for catchment modeling. It is argued that for accurate surface runoff modeling at scales smaller than the REA, these patterns of the spatially variable controls must be modeled exactly. (Such an argument is supported by the lack of success in the literature toward accurately modeling small-scale experimental catchment areas.) For areas larger than the REA, only the statistical representations of these controls need to be considered. Further, at these larger scales, simplified models, based on these statistics, should suffice as models of the catchment.

Figure 12. The envelope, over 25 repetitions, of the flood frequency distribution for the sensitivity run, $\frac{P^*}{K_0} = 0.5$ [from Sivapalan et al., 1989].

Figure 13. Flood frequency distributions for three watersheds: (a) North Nashua River, (b) Davidson River, and (c) Santa Paula Creek [after Moughamian et al., 1987].
response. For one set of simulations, Wood et al. [1988a] found an REA of the order of 1 km. Simplified models consistent with the REA theory were successful in simulating the large-scale behavior.

The research in similarity was motivated by the desire to understand the influence of environmental controls on catchment storm responses independent of basin scale. Unlike previous research that related similarity to physioclimatic regimes our definition of similarity is based on similarity in runoff generation and catchment response. It was shown that consistent scaling parameters could be defined and a scaled storm response model developed. Such a scaled model was used by Sivapalan et al. [1987] to investigate catchment responses and by Wood and Hebson [1986] and Sivapalan et al. [1989] to study flood frequency characteristics.

The research reported here is a first step in clarifying the issues involved in understanding the parameterization of hydrological processes across a range of scales. There remains a vast amount of research to which the ideas presented here can contribute. Let us suggest only two. In the United States the recommended approach to flood frequency studies uses regionalized values for the coefficient of skewness statistic [Interagency Advisory Committee on Water Data, 1982]. The regionalization is based on geographical regions without thought or consideration to differing catchment response characteristics. The concept of similarity as presented here can help in understanding the relationship between the controls within the catchment and its flood frequency characteristics and can help to form the basis for grouping catchments for statistical flood analyses.

The second area where the research presented here can be extended to is the parameterization of the land portion of global climate models. The REA results give confidence that simpler macroscale models will perform well at large grid scales which have significant subgrid variability, thus questioning the wisdom of detailed land parameterization which ignores process heterogeneity. Continuing research is comparing the macroscale models developed here with current GCM parameterizations of land surface hydrology. In addition, the concepts of similarity can help in parameterizing the macroscale models by utilizing data from a number of sites and transferring data to other areas around the Earth. This effectively expands the available data base into regions where data may be sparse or lacking.

ACKNOWLEDGMENTS. The work on which this paper is based was supported in part by U.S. NASA grants NAG-5-491 and NAG-1392, U.S. Department of the Interior (USGS) grant 14-08-0001-G1138, and U.K. NERC grant GR3/6264. This support is gratefully acknowledged.

Garrison Sposito was the editor in charge of this paper. He thanks William E. Dietrich and Vijai K. Gupta for their assistance in evaluating the technical content of the paper and an anonymous associate editor for serving as cross-disciplinary referee.

REFERENCES


