A Paradigm Shift in Hydrology: Storage Thresholds Across Scales Influence Catchment Runoff Generation

Christopher Spence*
Environment Canada, Canada

Abstract
A paradigm shift is occurring in the science surrounding runoff generation processes. Results of recent field investigations in landscapes and during periods previously unobservable are shaping new ideas on how runoff is generated and transferred from the hillslope to the catchment outlet. The previous paradigm saw runoff generation and contributing area variability as a continuum. The new paradigm is based not on continual storage satisfaction and runoff generation but threshold-mediated, connectivity-controlled processes dictated by heterogeneity in the catchment. This review focuses on the body of literature summarizing research on storage, storage thresholds and runoff generation, particularly over the last several years during which this paradigm shift has occurred. Storage thresholds that control the release of water exist at scales as small as the soil matrix and as large as the catchment. Hysteresis in storage–runoff relationships at all scales manifest because of these thresholds. Because storage thresholds at a range of scales have now been recognized as important, connectivity has become an important concept crucial to understanding how water is transferred through a catchment. This new paradigm requires basins to be instrumented within the context of a water budget investigation, with measurements taken within key catchment units, in order to be successful. New model approaches that incorporate connectivity are required to address the findings of field hydrologists. These steps are crucial if our community wishes to adopt the holistic view of the catchment necessary to answer the questions posed to us by the society.

Introduction
The amount of storage in a catchment has long been recognized as important for runoff generation. When investigating storage–runoff relationships, hydrologists of all types remain strongly influenced by ideas like the partial area concept of Betson (1964) and the variable source area concept of Hewlett and Hibbert (1967). Both concepts view runoff generation as a function of the volume or rate of inputs relative to the ability of landscape components to receive them. Hewlett and Hibbert (1967) eminently stated one way in which storage influences runoff response. They observed:

The yielding proportion of the watershed shrinks and expands depending on the rainfall amount and the antecedent wetness of the soil. When subsurface flow of water from upslope exceeds the capacity of the soil profile to transmit it, the water will come to the surface and the channel length will grow. (Hewlett and Hibbert 1967)

This statement has shaped hydrological research for 40 years, and was on the vanguard of a massive effort to understand and model runoff generation. Following Hewlett and Hibbert (1967) were seminal field studies that demonstrated the spatial and temporal dynamics of variable source areas (e.g. Dunne and Black 1970) and modelling studies that related these areas to topography (e.g. Beven and Kirkby 1979). Runoff from a
catchment was envisioned as a function of continual shrinking and expanding saturated hillslope source areas around the stream channel. Difficulty in applying models based on these ideas in some environments has lead hydrologists to conduct field studies over the last 10–15 years (e.g. Allan and Roulet 1994; Branfireun and Roulet 1998; Quinton et al. 2003; Spence and Woo 2003; Tromp van Meerveld and McDonnell 2006). Results have revealed that there are nuances about the spatial distribution of storage and storage thresholds. Changes to areas contributing to runoff are not continuous or contiguous, contributing areas are not necessarily centred around the stream, and the stream network is not always synonymous with the topographic drainage network. Furthermore, a key recent paper by Zehe et al. (2005) provided model results that showed that it is not the mere quantity of water stored in a catchment that is influential. A change in storage can have a disproportionate effect on runoff response depending on how close parts of the basin are to meeting saturation thresholds. They demonstrated that a change in the spatial pattern in which storage capacity is met changes the spatial pattern of contributing areas. This results in temporal differences in concentration of flow and in turn a change in the hydrograph shape, all with the same rainfall input and average antecedent conditions.

The idea that storage thresholds impact runoff response is not new, but thoughts are substantially changing. Research of runoff generation across a range of scales is supporting a shift to a new paradigm that views runoff generation not as the function of continual storage accumulation or depletion but as a threshold-mediated, connectivity-controlled process dictated by storage heterogeneity in soils, hillslopes and catchments. It is the literature that has lead to the paradigm shift that is the focus of this review. The nature of this focus means that much attention will be on situations where storage effects are most pronounced, but attempts have been made throughout the manuscript to show the reader there are a few general themes that make this new paradigm applicable across a range of scales and landscapes.

Storage and storage thresholds across spatial scales

UNSATURATED SOIL MATRIX

For water to be stored in soil, surface tension must retain water against the force of gravity. Smaller void radii can support longer columns of water. If the water table drops, voids now in the unsaturated zone need to empty until a suitably smaller controlling radius can be found. As water drains from the soil, a larger number of voids become empty, leaving a more circuitous route for the remaining water (Knapp 1978), resulting in asymptotic drainage. Furthermore, void heterogeneity dictates that saturation is not a complete reversal of drainage. The capillary pressure needed to drain pores is controlled by the minimum radii, and the capillary pressure to fill empty pores is controlled by the maximum radii. The result is the non-linearity and hysteresis between the pressure head, water content and soil hydraulic conductivity (Rawls et al. 1992) common to analytically described moisture characteristic curves (e.g. Brooks and Carey 1964; van Genuchten 1980). The hysteresis is more often than not disregarded in many practical, and academic, applications because of its complicating nature. However, its importance has been noted by the likes of Torres et al. (1998) and Quinton et al. (2008) who demonstrated the impact of soil column scale hysteresis on hillslope runoff response.
MACROPORES

There remains no consensus on the definition of a macropore. Macropores can be classified by size or origin or both (Beven and Germann 1982). In this review, macropores are defined to include all pores that are large enough to conduct water a significant distance downslope at rates higher than the surrounding soil matrix. This definition incorporates the role that an increased pore diameter has on velocities and the applicability of Darcy’s Law (Freeze and Cherry 1979). It also recognizes the importance of macropore connectivity in promoting pipeflow (Beven and Germann 1982; Bouma 1981; Jones 1971). Just as capillary pressure and gravity thresholds dictate hydraulic conductivity in the unsaturated soil matrix, the frequency at which macropores conduct pipeflow is governed by macropore storage thresholds and the ability of the soil matrix to receive and conduct water. The rate of transfer from the soil matrix, rainfall or snowmelt along the macropore needs to equal or exceed infiltration losses from the macropore to the soil matrix in order to fill the macropore and initiate and sustain pipeflow. Unless there is a source of upslope water from exposed bedrock (Peters et al. 1995) or another macropore (Carey and Woo 2001), this may require the water table to be above the elevation of the macropore (McCaig 1983; Roberge and Plamondon 1987). It may require the soil matrix infiltration capacity or hydraulic conductivity to be exceeded. When the rate at which water is applied to the macropore is higher than the surrounding soil matrix hydraulic conductivity, the soil matrix relinquishes control on the initiation of pipeflow and rainfall intensity becomes important (McDonnell 1990). Pipeflow occurs only when one type of these thresholds are exceeded.

In documenting differences in lag times between the initiation of rainfall and pipeflow during the same storm, Jones (1987) implied that different macropores have different thresholds of response. The differences in macropore response are due to two controls. The first is the nature of the area supplying water to the macropore (i.e. the possible topographic catchment area contributing to the macropore and/or the location of the macropore relative to source water). The second is the nature of the macropore network (i.e. the number, size and connectivity of macropores) (McDonnell 1990). For example, macropores with perennial pipeflow are obviously well connected and have better access to groundwater stores than macropores with ephemeral flow. As the former are already saturated and do not require filling by the soil matrix or upslope water, they react more quickly to rainfall events. For those macropores that are not perennial, the not uncommon requirement for the water table to be above the elevation of the macropore means the moisture state of the surrounding soil matrix and its moisture characteristic curves dictate how, when and where the soil and macropore network become saturated in response to a given input of water.

HILLSLOPES

The discussion above described how subsurface flow, by either matrix flow or pipeflow, has inherent storage thresholds that must be necessarily breached to initiate the transfer of runoff downslope. Once an entire soil column becomes saturated, hydraulic conductivity across that section of the hillslope increases exponentially and a different suite of processes becomes important. This suite of processes that transmits water from saturated hillslopes includes Hortonian flow, subsurface flow, saturation overland flow and fill-and-spill runoff. There have been numerous excellent reviews, summaries and commentaries (e.g. Anderson and Burt 1990; Beven 2007) discussing the relative importance of each in
different landscapes to which the reader is directed for a general overview. The focus of this discussion will continue to be on the role of storage and storage thresholds on the presence of these hillslope processes.

Each runoff generation process is dictated by its own inherent thresholds. The spatial distribution of Hortonian runoff is controlled by where infiltration capacity is exceeded (Betson 1964). Saturation overland flow is generated at the topographic surface when the rate at which water is supplied to the soil column is higher than the hydraulic conductivity (Hewlett and Hibbert 1967), which tends to be at a maximum near the soil surface (e.g. Quinton and Marsh 1999). In one of the original contributions discussing saturation overland flow, Dunne and Black (1970) noted that 1 in 25 year storms were needed to generate saturation overland flow from New England hillslopes. The correct combination of soil thickness, soil type, slope, possible contributing area and local hydrometeorology are necessary. Dunne and Black’s (1970) classic comparison of convex and concave slopes and Devito et al.’s (1996) contribution highlighting the importance of soil thickness and contributing area show the importance of soil and topography on where saturation overland flow can occur. Central to the fill-and-spill concept of runoff generation (Spence and Woo 2003) is that storage is spatially variable and key stores across the hillslope need to be satisfied before water is transmitted by this process (Tromp van Meerveld and McDonnell 2006). Fill-and-spill runoff requires a certain combination of hillslope topology and topography that would permit a mechanism by which water can coalesce in upslope locations. All these mechanisms result in dynamic contributing areas on hillslopes over space and time. Some of the best examples of dynamic contributing areas on hillslopes produced by a variety of mechanisms are shown in Dunne (1978), Tromp van Meerveld and McDonnell (2006) and McNamara et al. (2005).

Different runoff mechanisms have been documented on the same hillslopes under different wetness conditions (Montgomery and Dietrich 1995). The potential exists for any process to occur on any hillslope, but it is the interplay between inputs and storage thresholds on each hillslope that defines when and how frequent each process’ thresholds will be crossed. Figure 1 illustrates a hypothetical hillslope on which unsaturated flow, pipeflow, saturated matrix flow, saturation overland flow, direct precipitation on saturated areas, and Hortonian flow all occur. Increasing rainfall thresholds and different efficiencies are assumed for each process, and a uniform increase in saturated area with saturation overland flow. Smaller events only produce saturated matrix flow, as thresholds initiating other processes have not been exceeded. The ‘zone of higher contribution’ that has been documented by Jones (1987) in the field for pipeflow is apparent in Figure 1. There is another storage threshold associated with pipeflow upon which the relative influence of pipeflow declines, if not the volume. This is when macropores are full, but not overflowing, such that neither the soil matrix nor macropores can conduct all the water entering the hillslope. This triggers saturation excess overland flow. Only the largest events produce Hortonian flow, but as this process results in ponding at the surface, a portion of the runoff must come from direct precipitation on saturated areas. This may be complicated on more heterogeneous hillslopes where the physiography of the hillslope is also important. Allan and Roulet (1994) noted the simultaneous presence of both saturation overland and Hortonian overland flow on Canadian Shield hillslopes. The relative contribution of each depended on when thresholds were breached in landscapes patches with different storage thresholds.

Dunne and Black (1970) observed that small portions of a hillslope (7.5%) can produce disproportionate amounts of runoff (50%). This field study confirmed that some parts of a hillslope can have higher efficiencies than others. Jones (1987) conceptualized that
pipeflow and saturated matrix flow each have different contributing areas, but these areas may not necessarily be mutually exclusive. Pipeflow contributing areas sometimes contain a saturated soil matrix that provides some water downslope. Runoff generation processes may not be mutually exclusive but juggle predominance because of differences in efficiencies, which may be expressed through recession coefficients (Montgomery and Dietrich 1995). It is also expressed in hydraulic conductivities. Dunne and Black (1970) note that once saturation overland flow occurs, the overland flow has velocities 500 times

![Figure 1](image-url)
faster than the fastest subsurface stormflow, which is still continuing. The contributing area has not changed with the introduction of saturation overland flow, but the proportion of water in contributing areas with a higher efficiency increases. Tromp van Meerveld and McDonnell (2006) demonstrated how the initiation of fill-and-spill runoff increased transmission velocities from the hillslope by an order of magnitude. The operation of saturation overland flow also increases the amount of area eligible for the highly efficient runoff process of direct precipitation on saturated areas.

As with macropores and a saturated soil matrix, field studies of hillslope runoff in a diversity of landscapes imply hydrologic connectivity, defined as saturated conditions adjacent to one another along a flowpath to the outlet, is a necessary condition for the transmission of widespread lateral surface flow. In their investigation of a peatland in the Canadian Shield, Branfireun and Roulet (1998) showed that significant runoff from the hillslope only occurred once water tables in key locations reached threshold levels. This phenomenon is highly non-linear to the point where hillslope storage and stream discharge can be hysteretic (Branfireun and Roulet 1998) in that water tables can be higher during the rising limb for the same level of runoff during recession. McNamara et al. (2005) showed that mid slope capacities must be filled in order for upslope subsurface flows to connect to downslope positions, expanding the contributing area. The occurrence of hillslope runoff processes are not static, do not follow steady state principles and are threshold-mediated (McGlynn and McDonnell 2003). Assuming otherwise has not improved our understanding of how different landscape units interact over time to produce a catchment runoff signal.

CATCHMENTS

The upscaling of hillslope runoff response to the catchment has perplexed many a hydrologist. This is an important conceptual and practical problem to be solved because almost all water resource management problems occur at the catchment scale. It is often accepted that as catchment size increases, complex local patterns become attenuated. In reality, it is not always so easy. Most catchment scale studies have involved some degree of modelling because of the logistical difficulties of instrumenting large areas. Assumptions had to be made on the nature of how hillslope processes upscale to the catchment, which may have provided us with fortuitous answers to our questions.

There has been no study that has definitively shown how small scale storage thresholds influence catchment scale runoff response, but there have been several that have demonstrated the impact of hillslope scale features. Hewlett and Hibbert (1967) and especially Dunne and Black (1970) illustrated how different storage thresholds associated with different topographic slopes (i.e. straight, concave or convex) dictate catchment runoff response. In a period when the interaction between basin geomorphology and hydrology was of great interest, Boyd et al. (1979) and Rodriguez-Iturbe and Valdés (1979) both took a phenomenological and probabilistic approach to the storm hydrograph concepts of Nash (1957), Dooge (1959) and Wooding (1965) and the idea that topography dictates runoff response to develop the concepts commonly referred to as the geomorphic instantaneous unit hydrograph (GIUH). The GIUH methodology includes the assumption that each stream reach in a catchment is comparable to an independent reservoir reacting to effective rainfall. In addition, the efficiency with which each reach transmits runoff exhibits a statistical function that is tied to basin form which is expressed through the Horton laws of stream order, length and area (Horton 1945). Robinson et al. (1995) built upon the GIUH by permitting efficiency to change with the extent of the stream network.
This was an important step as field studies since have demonstrated that spatially variable water stores across hillslopes (Tromp van Meerveld and McDonnell 2006) and catchments (Bowling et al. 2003; Spence 2006) are crucial to the lateral transfer of water in higher efficiency contributing areas increases.

At the same time as the GIUH methodology was being introduced, Beven and Kirkby (1979) and O’Loughlin (1981) presented predictive methodologies able to account for the different observed thresholds of saturation overland flow that occur in different geomorphological and soil conditions. This approach assumed locations where saturation deficits could be exceeded and saturation overland flow generated could be based primarily on topography; from (Beven and Kirkby 1979):

$$\ln \left( \frac{a}{\tan \beta} \right) > \frac{S_T}{m} - \frac{S}{m} + \lambda$$

where \(a\) is contour length, \(\beta\) is local slope angle, \(S_T\) local maximum storage, \(m\) is a constant related to recession and \(S\) and \(\lambda\) are functions representing the amount of available storage. These topographic-index models have been widely applied, more often than not successfully (Kirkby 1997), but as with anything that becomes ubiquitous and popular (e.g. Disney), they have been exposed to their share of criticism. In recent years, two key findings have come to light that imply limits to the applicability of Eq. 1. The first is recognition that just as with macropores and an unsaturated soil matrix, runoff efficiency is not spatially uniform (e.g. McNamara et al. 2005). The second is the recognition that a broader definition is needed of where saturation deficits are generally exceeded in a catchment. Field studies in more heterogeneous catchments have suggested that topography is but one influence on where saturation thresholds are exceeded (e.g. Allan and Roulet 1994). Wet climates can encourage topographically controlled connectivity, but in dry regions topography may not always act as a surrogate when there are other primary influences on soil moisture distribution (Grayson et al. 1997). Typology of sub-catchment units within a heterogeneous landscape controls storage thresholds and potential evapotranspiration losses. Topology, or relative location of these landscape units, dictates the availability of water directed to any specific location in the catchment, controlling storage states (Spence and Woo 2006). In recognizing the importance of topology and typology in addition to topography, Buttle (2006) proposed that the first order controls on catchment runoff generation follow a T³ (typology, topology and topography) classification. The reaction to Buttle’s proposal has been mixed, but it has generated much needed discussion questioning the dominance of topography as a predictor for runoff generation.

Topography is now recognized as only one possible control of storage thresholds across catchments. The focus is now shifting away from using a surrogate for storage (i.e. topography) to storage itself as the predictor for runoff response. Both Spence (2007) and Kirchner (2009) demonstrated that much can be learned about runoff response by simply placing catchments within a theoretical construct where discharge is a function of storage. A new paradigm is emerging that recognizes it is the different attributes of the storage across the catchment; where it is; when it occurs and how it becomes accessible, that are proving important to catchment runoff response. Furthermore, as the discussions earlier on the nature of storage at smaller scales imply, satisfaction of storage across a catchment does not occur as a continuum, but following a series of discontinuous thresholds. Often, upscaling these thresholds to the catchment scale and superimposing the different runoff processes in Figure 1 together produces a non-linear relationship between inputs and runoff. However, these thresholds can be so high in places, that runoff can be completely
intercepted, stored, evapotranspired and prevented from reaching the basin outlet (Mielko and Woo 2006; Stichling and Blackwell 1957). The most profound effect at the catchment scale of this hillslope scale threshold behaviour is a diverse contributing area. Roulet and Woo (1988), Quinton et al. (2003), Laudon et al. (2007) and Spence (2006) all showed how the distribution of storage thresholds at the hillslope scale dictate the organization of the stream network, contributing areas overtime and runoff response at the catchment outlet. In very different catchments, both McGlynn et al. (2004) and Spence and Woo (2006) documented how noting the location and topology of key landscape units within a catchment provides insight into where each of the classical runoff mechanisms occur, and the relative importance of each on catchment runoff response. It is important to note how different are the research catchments of McGlynn et al. (2004) and Spence and Woo (2006) in order to demonstrate the wide applicability of their findings. McGlynn et al. worked at the wet steep Maimai catchment in New Zealand. Spence and Woo worked on gentle Canadian Shield terrain in a boreal climate with one-tenth the rainfall.

The existence of smaller scale thresholds that control catchment runoff manifests in the form of non-linear response functions to changes in storage, which may be distinct to individual catchments, or types of catchments. McGlynn et al. (2004) and Spence et al. (2009) have demonstrated that these functions are hysteretic, and are comparable to patterns between storage and outflow documented at the smaller scales discussed earlier; including soil columns (Brooks and Carey 1964), hillslopes (Dunne 1978) and stream channels (Chow et al. 1988) (Figure 2). Spence et al. (2009) were able to relate portions of these hysteretic curves to different catchment functions. There are a host of problems

Fig. 2. Examples from the literature demonstrating hysteresis between storage and water flux at a range of scales including a soil matrix (a), hillslope (b) stream channel (c) and catchment (d). Adapted from van Genuchten (1980), Branfireun and Roulet (1998), Chow et al. (1988) and Spence et al. (2009).
and issues with deriving and applying storage–discharge curves at the catchment scale. Nonetheless, recognizing their existence is useful as they permit a numerical representation of the physical relationship between storage in dynamic contributing areas and streamflow response at the catchment outlet. Compilation of curves from a diversity of catchments may aid in classification (McDonnell and Woods 2004). The appearance of connectivity and hysteresis functions at the catchment scale because of traits and processes occurring at smaller scales is what Sivapalan (2005) has termed emergent behaviour. Analysis of the signatures of this emergent behaviour (i.e. degree of connectivity and magnitude and direction of hysteresis) in catchment storage-discharge curves would permit basins to be classed not only by their streamflow response (e.g. Church 1974), but also incorporate hydrological function and could infer the first order controls responsible for catchment behaviour (Buttle 2006).

What does this mean for how we should investigate catchment streamflow response?

IMPLICATIONS FOR FIELD METHODS

Hydrologists have sometimes derided themselves for not having an overarching scientific foundation. This may be because so many of the historical roots of the field are in water engineering and management. One applicable concept across a multitude of spatial and temporal time scales is the water budget, which generally can be expressed as:

\[ \Delta S = I - O \]  

where \( \Delta S \) is change in storage, \( I \) is inputs and \( O \) is outputs. The subdivision of \( I \) and \( O \) into specific fluxes on the right hand side of the equation changes among catchments and time. For instance, snowmelt has to be included when investigating cold catchments. If studies take place over a short time period, certain slow processes such as deep percolation may be ignored. The energy budget, which drives the entire hydrological system, is closely associated with the water budget through an evapotranspiration/latent heat term, \( E \), if included as part of \( O \) in Eq. 2:

\[ R_n = H + E + G \]  

where \( R_n \) is net radiation, \( H \) is sensible heat and \( G \) is heat storage.

Not enough catchment field studies report water budget estimates, as too many focus only on precipitation/snowmelt and runoff. Although most hillslope studies have plethora of wells, piezometers, soil moisture probes and tensiometers, only some calculate storage or cumulative change in storage. Although having a direct measurement of antecedent moisture at their fingertips, many still use individual wells or antecedent precipitation as an index. Storm and water year precipitation/runoff ratios are often less than 50%. Where does the majority of the water go? Strangely, evapotranspiration is often an afterthought, or assumed to be non-existent or inconsequential. The benefit of nested instrumented watersheds in improving understanding of streamflow production has been well established (Pomeroy et al. 2005; Prowse et al. 2007; Sivapalan et al. 2003). A case can also be made for instrumenting these hillslopes and watersheds for full water budget evaluation. Doing so provides insight on water fluxes, and how they draw upon basin storage, that is possible no other way. Some may argue that basin heterogeneity and measurement error are too high to calculate a reliable catchment water budget. Hydrologists should always be suspicious when their water budget calculations come too close to balancing.
However, evidence shows that individual fluxes are often large enough that the error bounds do not prevent fair conclusions to be made. It is important to attempt water budget evaluations as part of this new paradigm. A better understanding will be gained of where and how storage thresholds are being breached across the catchments if attempts are made to measure all fluxes.

The application of isotopes and other hydrochemical approaches to catchment studies has provided remarkable insight into runoff generation processes. Chemistry provides evidence of the dynamics of source areas that hydrometric approaches cannot. A key limitation of water chemistry methodologies is that they do not permit the determination by which processes the water gets to the stream. This has to be inferred (McGlynn and McDonnell 2003). In turn, it becomes difficult to make conclusions about which storage thresholds are being crossed and how efficient are different parts of the catchment. The use of a multi-tracer approach (e.g. isotopes with silica or chloride) may alleviate this problem. Further to this point, the author wishes to reiterate the opinion of Buttle and Peters (1997) who demonstrated why hydrochemical approaches should always be used in conjunction with hydrometric methods. This helps narrow the different possibilities hydrologists have to contend with when results from one method are inconclusive as to which source areas, pathways and runoff processes are predominant at a given time.

McGlynn et al. (2004) suggested that rather than investigate specific runoff mechanisms, a more appropriate course of action for understanding watersheds may be to focus on key catchment units that collect and convey water. It may be possible to combine the recommendations of McGlynn et al. (2004) and findings of Spence et al. (2009) and derive storage–streamflow curves for research catchments and their key component parts. This could provide a qualitative and quantitative means to develop upscaling relationships useful for predicting streamflow. Furthermore, complementary full water budget measurements within these key catchment units would provide insight on how catchments and their parts behave and function.

**IMPLICATIONS FOR MODELLING AND PREDICTION**

Troch et al. (2008) recently commented that hydrology is moving towards a more holistic view that recognizes the importance of landscape heterogeneity. Crucial to a successful migration to this new paradigm is recognition of the role of storage and its thresholds across heterogenous landscapes. A common approach to modelling storage thresholds under the old paradigm is epitomized by Sugawara (1967) in which stores on the landscape are conceptualized by tanks of different capacities. What has not been fully recognized in such models is the importance of topology and connectivity. This is relatively easy in distributed models as key linkages can be explicitly represented. It is also surmountable at the hillslope scale or smaller where we are confident in the location of important intermittent linkages where, again, they can be explicitly represented in models. For example, Reaney et al. (2006) developed the Connectivity of Runoff Model (CRUM) which explicitly accounts for surface depressional storage; parameterized using surface roughness and slope. Lehmann et al. (2007) applied an explicit model of an instrumented hillslope at Panola, Georgia, based upon the principles of percolation theory that performed very well.

Semi-distributed models require more elegant approaches where sub-grid connectivity needs to be parameterized. Many lateral transfer schemes in semi-distributed models used at the sub-grid scale are rudimentary. For example, the Canadian Land Surface Scheme (CLASS), uses a simple additive approach that does not permit any interaction among
sub-grid landscape units. Topography-based models represent connectivity based on slope and antecedent wetness. Deriving the probability of saturation overland flow across a catchment using Eq. 1 is a fair approach as in many landscapes, the saturated areas are in downslope locations aligned with drainage networks (Buttle et al. 2004). However, as noted earlier, we are discovering that in many landscapes, saturation overland flow occurs in areas unrelated to topography. New sub-grid parameterization schemes will need to account for the dynamic distributions of unit topology, morphology and storage transience, as these dictate the spectrum of landscape unit function. This is particularly important if demands are placed on modellers to properly reproduce sediment, carbon or nutrient fluxes in addition to water.

Meerkerk et al. (2009) are correct; hydrologists need to follow the lead of landscape ecologists who have an accepted universal definition of connectivity. This is required before we can provide consistent quantitative measures of connectivity. Ecologists have accepted a definition that recognizes that as long as habitat is contiguous, it provides connection (Li and Reynolds 1993). Contiguous habitat is not necessarily quality habitat that will provide a connected pathway for all individuals of a species, but field observations show that this is generally acceptable. Western et al. (2001) defined hydrological connectivity in a similar manner as this landscape ecology definition; where locations of similar moisture state are adjacent, they are connected. Physically, this does not imply hydrological connectivity, in that a drop of water can actually move between adjacent locations. Usually slope dictates unidirectional connectivity, and drainage divides sever hydrological connectivity. Bracken and Croke (2007) propose a good definition adapted from fluvial geomorphology; connectivity occurs when there is the potential for a specific particle to move through the system and it is transferred from one zone or location to another. They also propose a means by which to measure it by looking at the relative values of applied water and ability to transfer downslope.

Two primary research gaps for modellers need to be addressed. First, we need to discern a better way to parameterize sub-grid or sub-catchment units with appropriate time constants relating storage and discharge (perhaps from characteristic curves) so that we can properly simulate how well they can sustain downslope flow and maintain connections. Second, a key unknown to be addressed is how to better parameterize sub-grid or sub-catchment topology. The transfer functions used in unit hydrograph theory and topographic indices have previously been estimated with assumptions of static stream networks, effective precipitation, homogenous runoff responses and assumed probability distributions, all of which can be to be severely limiting in some landscapes. By even recognizing connectivity as a viable concept nullifies these assumptions and necessitates parameterizing topology among disparate sub-catchment units.

Concluding remarks

A paradigm is defined by the questions posed within it, the phenomena of interest, the methodologies employed and the nature with which data is interpreted. A new paradigm in hydrology is emerging (Figure 3) because recent observations of runoff generation could not be explained readily by existing theories. Field studies have begun to reveal that runoff is not only non-linear but also a hysteretic, threshold-mediated, connectivity-controlled process. The thresholds that control runoff at the catchment scale exist at a range of subordinate scales. Thresholds in the soil matrix produce hysteresis between storage and the rate at which soil water can move. Still within the soil, but in macropores, the rate of inputs relative to the ability of the macropore to conduct water determines
when macropores are important for downslope water movement. As such, the threshold at which macropores become active is intrinsically linked to the hysteretic nature of soil moisture–conductivity curves. In addition to pipeflow, there is a suite of runoff generation processes that occurs on hillslopes. Each is dictated by its own thresholds. The physio-climatic situations on each component of the hillslope determine when and where any process will tend to occur. Differences in efficiencies with each process cause different areas to provide disproportionate amounts of runoff. In some landscapes, this will create patterns indicative of the variable source area concept. In others, intermittent connectivity and non-contiguous variable source areas are the norm. Our community is learning that in these landscapes, understanding hillslope runoff patterns requires ideas respectful of heterogeneity, connectivity, hysteresis and non-steady state conditions. Continual storage satisfaction and runoff generation is the exception, rather than the rule.

The second reason for this paradigm shift is that questions were being posed to the hydrological research community on the hydrological, chemical and ecological behaviour and function of catchments that it could not answer. Answering these questions will require more holistic investigative techniques than the previous reductionist ones. To improve our understanding of catchment response, nested basins should be instrumented within the context of a water budget investigation, with measurements taken within key catchment units. New model approaches that incorporate connectivity are required to address the findings of field hydrologists and permit constituent and runoff pathways to be adequately simulated. This is particularly crucial if our community wishes to answer the questions posed to us by society.

Acknowledgement

The author would like to acknowledge the support of Environment Canada, and several funding sources over the years, including the Natural Sciences and Engineering Research Council and the Canadian Foundation for Climate and Atmospheric Sciences. Some
colleagues and students whom I would like to thank for fruitful discussions about what I think is an exciting paradigm shift in hydrology include Ming-ko Woo, Jim Buttle, Brian Branfireun, Sean Carey, Jeff McDonnell, Murugesu Sivapalan, Xiu Juan Guan and Ross Phillips, among others.

Short Biography

Christopher Spence holds a B.A. (Hons.) and M.Sc. from the University of Regina, Regina, Canada and a Ph.D. from McMaster University, Hamilton, Canada. He currently is a research scientist for Environment Canada in Saskatoon, Canada. He was appointed an adjunct professor in the Department of Geography at the University of Saskatchewan in 2006. His research focuses on hydrometeorological processes in cold regions. He has a particular interest in the hydrological behaviour and function of complex landscapes such as the Canadian Shield and Prairie. He has a beautiful wife with whom he has two daughters (7 and 9 years). He is getting a bit tired of Disney entertainment.

Note

* Corresponding address: C. Spence, Environment Canada, 11 Innovation Boulevard, Saskatoon, SK, Canada S7N 3H5. E-mail: chris.spence@ec.gc.ca

References


