Water Resources Research

COMMENTARY

10.1029/2020WR027514

Fill-and-Spill: A Process Description of Runoff Generation at the Scale of the Beholder

Jeffrey J. McDonnell[1,2], Christopher Spence[3], Daniel J. Karran[1], H. J. (Ilja) van Meerveld[4], and Ciaran J. Harman[5,6]

[1]Global Institute for Water Security, School of Environment and Sustainability, University of Saskatchewan, Saskatoon, Canada, [2]School of Geosciences, University of Birmingham, Birmingham, UK, [3]Environment and Climate Change Canada, Saskatoon, Canada, [4]Department of Geography, University of Zurich, Zurich, Switzerland, [5]Department of Environmental Health and Engineering, Johns Hopkins University, Baltimore, MD, USA, [6]Department of Earth and Planetary Science, Johns Hopkins University, Baltimore, MD, USA

Key Points:

- Runoff is at the scale of the beholder. We need to move from a notion of uniqueness of place to uniqueness of scale
- Fill-and-spill, together with its components is common to all event runoff systems
- Fill-and-spill as a framework is perhaps a guide for field hydrologists on what to measure, in what order and why

Correspondence to:

J. J. McDonnell,
jeffrey.mcdonnell@usask.ca

Citation:

McDonnell, J. J., Spence, C., Karran, D. J., van Meerveld, H. J., & Harman, C. J. (2021). Fill-and-spill: A process description of runoff generation at the scale of the beholder. *Water Resources Research*, *57*, e2020WR027514. https://doi.org/10.1029/2020WR027514

Received 15 MAR 2020
Accepted 16 MAR 2021

Abstract Descriptions of runoff generation processes continue to grow, helping to reveal complexities and hydrologic behavior across a wide range of environments and scales. But to date, there has been little grouping of these process facts. Here, we discuss how the "fill-and-spill" concept can provide a framework to group event-based runoff generation processes. The fill-and-spill concept describes where vertical and lateral additions of water to a landscape unit are placed into storage (the fill)—and only when this storage reaches a critical level (the spill), and other storages are filled and become connected, does a previously infeasible (but subsequently important) outflow pathway become activated. We show that fill-and-spill can be observed at a range of scales and propose that future fieldwork should first define the scale of interest and then evaluate what is filling-and-spilling at that scale. Such an approach may be helpful for those instrumenting and modeling new hillslopes or catchments because it provides a structured way to develop perceptual models for runoff generation and to group behaviors at different sites and scales.

1. Introduction

The famous French mathematician Henri Poincaré said that "Science is built up of facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house" (Poincaré, 1902). Our perceptual understanding of runoff processes sometimes appears as a heap of facts (or observations), with many of them dating back to the First International Hydrological Decade and the following years, when key observations were made regarding translatory flow (Hewlett & Hibbert, 1967), saturation excess overland flow (Dunne & Black, 1970), saturated wedges (Weyman, 1973), old water effusion (Martinec, 1975), the importance of topographic convergence (Anderson & Burt, 1978) and preferential flow (Mosley, 1979). Since then, field hydrologists have continued to document more and more runoff responses in more and more areas.

What we have learned from these myriad studies is that (i) runoff processes are scale dependent, making the "up-scaling" of field-based process observations very difficult to nearly impossible, (ii) runoff concepts developed for humid areas (e.g., the variable source area concept) often do not apply in semi-arid and arid areas, and (iii) runoff responses are non-linear (thresholds, hysteresis, etc.) because regardless of whether the dominant flowpaths occur at the surface as overland flow or as subsurface stormflow within the soil or subsoil or saprolite or bedrock, some storage needs to be filled before runoff occurs.

Our heap of facts about runoff generation obtained by process observations have continued to grow, but there has been little construction of these facts into a body of knowledge. This seems to be a crucial next step in catchment hydrology, as it has been in other areas of science where observations gain significance when one uses reason to group them in meaningful ways (Beveridge, 1950). Indeed, "grouping" is a theme in the development of many fields; notably Darwin's (1859) observation that "science consists of grouping facts so that general laws or conclusions may be drawn from them" and Hughlings-Jackson's (1882) advice that we have multitudes of facts but "we require, as they accumulate, organizations of them into higher knowledge; we require generalizations and working hypotheses."

The importance and need for such grouping has been made previously. Dooge (1986) pushed us to do this in his "searching for hydrologic laws," the epic paper aimed at "searching for regularities" in runoff hydrology. Others have commented on the need for catchment comparisons (Barthold & Woods, 2015; Buttle, 2006; Falkenmark & Chapman, 1989; McDonnell & Woods, 2004; Wagener et al., 2007) or the development of similarity measures (e.g., Knoben et al., 2018) and the need for more general frameworks for hydrologic processes (Clark et al., 2017). Dooge (1986) noted the importance of searching for analogies as a fruitful way to get to the construction of models and new theories. He noted the problem of parameterizing the effects of microscale processes at the macroscale.

We know that catchments are not simply a linear superposition of soil core scale processes (McDonnell, 2003), thus making the use of simple additive perceptual models of runoff generation often futile (Beven, 2002) and scaling difficult (Blöschl & Sivapalan, 1995). The recognition that runoff generation is extremely scale dependent has created an impasse for the incorporation of rudimentary runoff generation concepts into land surface models (Fan et al., 2019).

Here, we assert that event-based "fill-and-spill" appears to group together runoff generation process behavior across all scales. Meaning, that all runoff processes have in common that vertical and lateral additions of water to a landscape unit of interest are first placed into storage and only when this storage reaches a critical level (the fill), is a previously infeasible (but subsequently important) outflow pathway activated (the spill).

In the broadest hydrologic sense, fill-and-spill is simple vernacular for "storage excess." Independent of scale, there are always a series of stores within soil columns, hillslopes, and catchments that require filling before water can flow to down-gradient areas. Heterogeneity in the storage capacity within each scale results in a spatially and temporally diverse number of stores. The macroscale response that is observed at any moment in time depends on how well connected these diverse stores are to an observed outlet. Fill-and-spill describes the process of these various (local) storage areas becoming progressively filled and connected as a prerequisite to producing a (macroscale) runoff response.

Fill-and-spill has been evoked to explain hydrological connectivity and runoff generation in many environments, ranging from hillslopes in temperate climates (Du et al., 2016; Tromp-van Meerveld et al., 2006), alpine areas (Muir et al., 2011), the high arctic tundra (Helbig et al., 2013; Rushlow & Godsey, 2017; Young et al., 2010) to tropical savannah (Farrick & Branfireun, 2014) and rainforests (Howard et al., 2010; Zimmermann et al., 2014). It has been used to explain runoff generation on frozen ground (Coles & McDonnell, 2018), surface water dynamics in the prairie pothole region (Leibowitz et al., 2016) and between small lakes (Mielko & Woo, 2006). It has been applied to urban infrastructure (Orlowski et al., 2014), perched groundwater tables (Ali et al., 2011), forested swamps (Martin, 2011), and interconnected bogs (Connon et al., 2015). Meltwater between lakes on the Greenland ice sheet have shown fill-and-spill behavior (Arnold et al., 2014) along with giant lakes on Mars (Balme et al., 2011; Ghatan et al., 2005; Warner et al., 2013).

Here we argue that if beauty is in the eye of the beholder, then fill-and-spill is at the scale of the beholder. In other words, we think that to search for scale invariant processes is futile; rather it is more useful to first define our scale of interest and to then evaluate how and what is filling and spilling at that scale. This commentary is for field hydrologists and those instrumenting and interrogating hillslopes and catchments because we think that fill-and-spill could be a way to group what we measure at current and future field sites. We build on recent work that has pointed at aspects of fill-and-spill as a way forward (Ali et al., 2013; McDonnell, 2013; Spence, 2010) and define fill-and-spill, and its manifestations at the plot, hillslope and catchment scales. We end the commentary with a vision for how the use of the fill-and-spill concept might help guide the development of perceptual models for runoff generation at different scales and propose some specific tests that could be helpful to move these ideas forward.

2. Fill-and-Spill Description

Hints of event-based fill-and-spill processes date back to the late 1960s and early 1970s with the early model formulations of overland flow and subsurface stormflow (see Figure 1 in DeBoer & Johnson, 1971; the PhD thesis work by Haan, 1967; and Figure 10 in Hewlett & Nutter, 1970). The probability-distribution principle for runoff production at plot- and basin scales of Moore (1985) also points in this direction where he noted

10.1029/2020WR027514

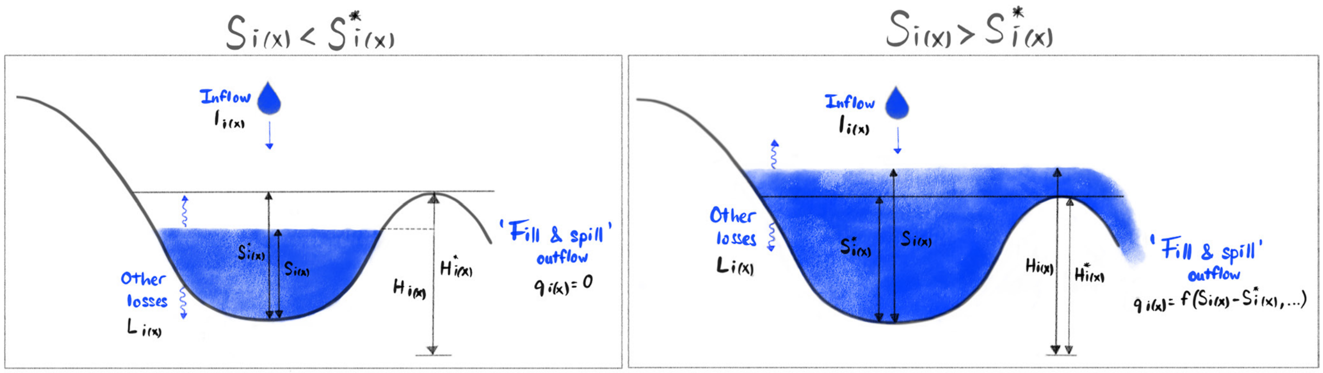

Figure 1. Our simple perceptual model of storage excess at a single scale.

that "the volume of direct runoff" is "a result of storage capacities of increasing depth being progressively replenished and starting to spill."

The term "fill-and-spill" as a rainfall-runoff process was first introduced by Spence and Woo (2003) who used it to explain runoff generation on the Canadian Shield. They showed that instead of runoff being generated from a rising water table in the bottom of the valley, runoff was generated first at upslope locations with smaller storage capacities, which subsequently spilled downslope to fill storage in lower locations. The filling of these heterogeneous storage areas was in large contrast to the concepts of saturation overland flow and variable source area hydrology for humid catchments. While not called fill-and-spill per se, Leibowitz and Vining (2003) proposed essentially the same mechanism for temporal connectivity of a prairie pothole complexes. Tromp-van Meerveld and McDonnell (2006a, 2006b) found that such a concept could explain the threshold nature of subsurface stormflow; runoff generation at the hillslope scale required connectivity of "saturated subsurface patches" that formed at the soil-bedrock interface. Since then, the term "fill-and-spill" has been used in a 100+ papers in the hydrology literature to describe processes across scales from the plot, to hillslope, catchment and beyond.

As noted already, fill-and-spill, in the broadest hydrologic sense, is "storage excess." It is the notion that all vertical and lateral additions of water to a landscape unit are first placed into storage, and outflow occurs only when this storage is filled. This can be represented mathematically by:

$$\begin{aligned} &\textit{if } S_{i(x)} \le S^*_{i(x)} \textit{ then } q_{i(x)} = 0; \\ &\textit{if } S_{i(x)} > S^*_{i(x)} \textit{ then } q_{i(x)} = \frac{1}{T}\int_0^T f_{i(x)}\left(S_{i(x)} - S^*_{i(x)}, \ldots\right)\, dt \end{aligned} \tag{1}$$

where $q_{i(x)}$ is the outflow via the activated pathway from landscape unit i at scale x over a period T, S is the storage held within that landscape unit, S^* is the critical threshold storage, and $f_{i(x)}$ is some (possibly hysteretic) function that describes the rate of release via this activated pathway (Figure 1).

The storage itself varies in time:

$$\frac{dS_{i(x)}}{dt} = I_{i(x)} - q_{i(x)} - L_{i(x)} \tag{2}$$

where I, includes all vertical and lateral inputs into landscape unit i, and L includes all vertical and lateral outputs (or losses) from i other than q.

Taken more broadly, fill-and-spill can be understood as a hydrologic example of an activation barrier or potential barrier, which is a widely used concept in physics (Dyre, 1988; Menzinger & Wolfgang, 1969; Razavy, 2003). This can be more than analogy if the potential energy of the system rises as the storage fills. The storage threshold can then be seen as the potential energy threshold at which the new pathway is activated.

This is familiar territory in hydrology, as potential energy in hydrologic landscape units is commonly linked to storage (e.g., moisture content vs. soil water potential; or water surface elevations vs. hydraulic head under hydrostatic conditions).

Thus if we denote H the energy at a given storage, the expression above can be represented by:

$$\begin{aligned} &if\ H_{i(x)} \le H^*_{i(x)}\ then\ q_{i(x)} = 0 \\ &if\ H_{i(x)} > H^*_{i(x)}\ then\ q_{i(x)} = \frac{1}{T}\int_0^T f_{i(x)}\left(H_{i(x)} - H^*_{i(x)}, \ldots\right)\, dt \end{aligned} \tag{3}$$

Fill-and-spill applies where a threshold activation dominates the runoff generation dynamics. This is not to say that it precludes other streamflow generation mechanisms. Indeed, there must be other pathways for the water (and potential energy) to leave the system (represented above by the losses L) so that the storage can be less than the threshold prior to the next event. This threshold-like fill-and-spill behavior can occur in different places (i.e., at the surface, near the surface or at great depths in the soil or saprolite) and at different landscape units or scales.

The rapid, non-linear change in behavior resulting from event-driven fill-and-spill activation is—as we will show in the next section—commonly the "biggest game in town," that is, the magnitude and direction of the flux that occurs above the threshold is substantially different from the fluxes that occur below it. Even where the threshold is somewhat "fuzzy," when looked at closely, the excursions above and below it are the emergent behavior. Thus the main message here is that runoff at any scale is best considered as a catastrophic failure of storage (Ward et al., 2018). As we will discuss in the next section, activation thresholds can dominate the movement of water at scales ranging from pores to headwaters and beyond. We argue that where it does occur, understanding this threshold is the key to understanding system behavior.

The main difference between fill-and-spill and "simple" storage excess (as described in Equations 1–3 above) is that there isn't a single (average) storage capacity S^* that needs to be filled. Rather, multiple storages are filled and flow from one storage area is used to fill the storage in the receiving area, often resulting in a cascading type of flow path. In other words, the larger scale storage threshold is the emergent behavior that results from multiple smaller storage areas filling and spilling and becoming connected. This combination of storages that need to be filled and the connectivity of these storages may cause a sharper runoff threshold and a larger total storage capacity S^* than would perhaps otherwise be expected. It also leads to spatially variable storage, water potentials and losses—which affect the interpretation of point scale observations and may have important implications for ecological processes, weathering, solute transport, etc.

The function $f_{i(x)}$ that describes the rate of release once the threshold is reached, is often approximated by a power-law permeability-saturation relationship (e.g., Mualem, 1976) for soils, and storage-discharge relations at the catchment scale (e.g., Gray, 1961; Wittenberg, 1999). But this may be more accurately represented by a relationship that is hysteretic, that is, one that depends on additional time-varying factors (cf. Beven, 2006). If only the total quantity of water released q is of interest, then the details of $f_{i(x)}$ may not matter, and it may be sufficient to view the system through the lens of runoff *events* (what the probability theory literature calls a *point process*). This is most appropriate for the Goldilocks conditions when the inputs are intermittent (i.e., precipitation events), the time taken to return to the threshold condition (T in the equation above) is not too large, and the other losses L during that time period are not too big. Through this lens, the losses L might vary gradually in time, but the inputs I and "spill" q might be treated as brief pulses (or even effectively instantaneous). This framework has been explored extensively in ecohydrology to represent soil moisture dynamics (Botter et al., 2007; McGrath et al., 2007; Rodriguez-Iturbe & Porporato, 2005), where the threshold represents a field capacity integrated over the root zone depth. This work has highlighted how the probability distribution describing q is determined by the magnitude and frequency of input events I and the antecedent moisture conditions, which are controlled by L.

3. Fill-and-Spill Manifestations From Micro to Macro Scales

3.1. Pore Scale

Analogies to fill-and-spill at the pore-scale can be found in the abrupt, threshold-based displacement of water and air during infiltration and drainage. This involves the filling of one fluid phase and the spilling of the other once a potential threshold is exceeded (DiCarlo, 2013). A "hold-back pile-up" conceptual model (Eliassi & Glass, 2001, 2003) has been used to represent the temporary "hold-back" of the initial wetting front during infiltration. Once the accumulated water potential exceeds the pore entry pressure threshold, the water fills the pores and advances downwards. The conceptual model for drainage is largely based on an "invasion percolation process" (Blunt & Scher, 1995), where atmospheric air and air trapped in ganglia progressively invade smaller and smaller pore openings via a displacement process called the "Haines jump" (Haines, 1930). Haines jumps occur when the capillary pressure across the fluid interface exceeds the pressure threshold required to invade an opening (Herring et al., 2018). Within milliseconds, air rushes in and displaces water in the invaded pores, often simultaneously in avalanche-like cascades (Moebius & Or, 2012). These instabilities observed during two-phase immiscible fluid displacement are strikingly similar to lateral flow at the larger plot-and hillslope scales described below.

3.2. Plot and Hillslope Scale

3.2.1. Surface Runoff

Surface micro-topography has a major effect on surface storage, connectivity and the initiation of overland flow (Darboux et al., 2001; Frei et al., 2010, 2012). The effect of the filling and connection of surface depressions on overland flow generation has largely been studied in agricultural fields (e.g., Peñuela et al., 2016) and more recently also on frozen agricultural hillslopes (Coles & McDonnell, 2018). The "puddle to puddle model" of fill-and-spill by Chu et al. (2015) is a classic example. Others, such as Appels et al. (2011) and Peñuela et al. (2016) have used a similar approach. Plot scale fill-and-spill behavior has also been shown to create hot spots for biogeochemical activity (Frei et al., 2012) and surface-subsurface water exchange in the riparian zone (Frei et al., 2010).

3.2.2. Subsurface Flow

Fill-and-spill was introduced as a subsurface hillslope-scale runoff generation process by Tromp van Meerveld and McDonnell (2006b). They showed that rainfall fills up depressions in the bedrock and that once the storage capacity of these depressions was filled, the water spilled over microtopographic relief, connecting subsurface saturated areas to lower slope sections in a cascading way. When connectivity was achieved, the instantaneous subsurface stormflow rate increased more than fivefold and total subsurface stormflow was more than 75 times larger. This fill-and-spill behavior observed at the Panola research slope, a forested, humid, continental, subtropical watershed, was also later observed at Maimai, a steep, wet, headwater catchment, (Graham et al., 2010) and at the H. J. Andrews catchment (old-growth mountainous region in the Pacific Northwest) (McGuire & McDonnell, 2010), as well as the low angle coastal plain hillslopes at the Savannah River Site (Du et al., 2016; Jackson et al., 2016).

3.3. Catchment Scale

Spence and Woo (2006) conceptualized the importance of filling a series of heterogeneous stores in a catchment before significant flow could occur at the outlet. The size, location and connectivity of the storage areas on the landscape dictated the volume of water that was directed to downslope or downstream locations in the catchment; and controlled the overall storage state that determined flow at the outlet (Spence & Woo, 2006). Filling of storage areas across a catchment does not occur as a continuum, but follows a series of discontinuous thresholds from smaller to larger landscape units. Others have shown for the ~100 km^2 watershed scale how the smaller (1–5 km^2) catchments that comprise the larger watershed fill-and-spill along abrupt non-linear transitions controlled by sub-basin slope gradient and geomorphic characteristics—and these processes ultimately drive the activation of the larger watershed response (Sayama et al., 2011). Often, upscaling these thresholds to the catchment scale and superimposing the different efficiencies with which runoff is transmitted for the different areas produces a very non-linear relationship between inputs and runoff (Ali et al., 2015). Field studies have demonstrated that the nature of this non-linearity is related to the geometric and topological structure of the catchment that determines which areas fill, coalesce and connect (spill) along pathways to the catchment outlet (Devito et al., 2005; Jencso & McGlynn, 2011; Oswald et al., 2011; Quinton et al., 2003; Rains, 2011; Spence, 2007).

The Prairie pothole landscape, in the northern Great Plains of North America, provides a classic example of fill-and-spill at the catchment scale. The numerous surface depressions (i.e., the potholes) dominate the poorly integrated drainage network. They receive snowmelt and rainfall until they fill and permit periodic surface hydrological connections to other depressions and eventually to higher order streams. The resulting dynamic hydrological connectivity and streamflow response is highly threshold mediated (e.g., Brooks et al., 2018; Haque et al., 2018; Shaw et al., 2012). Understanding the influence of fill-and-spill on the periodicity of hydrological isolation and connection is crucial to explain how these depressions control floods, buffer nutrient transport and enhance biodiversity (Golden et al., 2016; Leibowitz et al., 2018; Rains et al., 2016).

One of the first reports of fill-and-spill in a large catchment comes from stochastic flood frequency models for the Lake Warden catchment in Australia. There, runoff generation thresholds resulted in a break in the slope of the flood frequency curves (Kusumastuti et al., 2007) because antecedent storage was a dominant control on lake-overflow events (Kusumastuti et al., 2008). Bowling and Lettenmaier (2010) used a macroscale hydrological model to assess the effects of lakes and wetlands in two large arctic river basins and found that the system was consistent with the element threshold concept of Spence and Woo (2006), particularly for runoff generation in summer. A similar result was found by Gibson et al. (2016), who used stable isotope-based partitioning, in the Athabasca River Basin of Canada, to show that most rain and snowmelt events resulted in the indirect displacement of pre-event water by fill-and-spill mechanisms. Incorporating fill-and-spill behavior into catchment scale models has also proven beneficial, as demonstrated by Mekonnen et al. (2014) who showed that their fill-and-spill algorithm greatly enhanced the performance of existing hydrological models for the Assiniboine River Basin in Canada.

4. Heterogeneity as the Key to Process Understanding

Spatial heterogeneity in the landscape has been a source of frustration for hydrologic science (McDonnell et al., 2007). It is particularly frustrating if we want to use models based on the Darcy Law and the Richards equation to understand and predict hydrologic dynamics. It seems like we need to make measurements everywhere at fine enough spatial scale to capture the heterogeneity—a practical impossibility. Given that impossibility, what should we do instead? Often the heterogeneity is averaged-over in some sense, and replaced with spatially uniform parameters. But critically, neglecting the heterogeneity while retaining the governing equations means that the threshold-like shifts in flow magnitude and direction are missed. And these are often the most important control on the hydrologic response; and cannot be replicated by the same governing equation with "effective" (averaged) parameters.

The power of heterogeneity can be seen in an example at the hillslope scale and the use of Darcy-Richards models at the Panola experimental hillslope (see work by Camporese et al., 2019 using CATHY; Ameli et al., 2015 using Hydrogeopshere; James et al., 2010 using TOUGH2; Hopp & McDonnell, 2009 using HYDRUS 3D). In each of these modeling applications, these sophisticated models did not capture the fill-and-spill runoff dynamics observed when the soil/regolith boundary was represented as a smooth surface, even when the measured average soil depth was used, or if the bedrock was considered to be impermeable ($L = 0$). When the observed irregularity of that boundary (i.e., the storage distribution) was incorporated into the model, then each of these models was able to represent the magnitude and timing of the hillslope hydrograph. Most interesting from this example is that when a stochastic representation of the spatial variation in hillslope storage depths was attempted, even a simple conceptual model was able to accurately represent slope-scale fill-and-spill (Tromp van Meerveld & Weiler, 2008 using Hill-Vi).

It is important to note here that we do not view the Panola research site as representative of all hillslopes (or even all hillslopes within the Panola catchment). Rather it has been extremely useful as a testbed for understanding runoff generation process, because the boundary conditions have been well mapped and these boundary conditions appear to control the hillslope response. At very different trenched and instrumented experimental field slopes in South Carolina (Jackson et al., 2016), New Zealand (Graham et al., 2010), Oregon (McGuire & McDonnell, 2010) and Saskatchewan (Coles & McDonnell, 2018) we have seen very similar hold-back behavior linked to fill-and-spill and whole slope connectivity. The message across these very different systems has been the same: when the key heterogeneity of the system—at the scale of interest—is known, then a Darcy-Richards model *can* provide a physically reasonable and useful representation of fill-and-spill behavior. But when that heterogeneity is not described, such a model fails to reproduce the observed flow dynamics.

It can be difficult to obtain data at the desired resolution to adequately represent this heterogeneity in an explicit manner. However, what is important is that the heterogeneity be represented well enough that its control on runoff response and behavior can be captured. We do not need to know the exact configuration of the storage areas in order to be able to simulate their filling and spilling. There is evidence that there may be ways to represent heterogeneity implicitly and capture functional hydrological connectivity dynamics and runoff behavior at the scale of the beholder. And the example given earlier that a geospatial representation of the bedrock topography can be sufficient to correctly simulate the observed thresholds, as well as the observed spatial variability in the thickness of the transient saturated layer is encouraging (Tromp van Meerveld & Weiler, 2008).

Recognizing the role of surface storage filling and connectivity on the triggering of plot-scale surface runoff, Antoine et al. (2009) proposed a "relative surface connectivity function." This function provides a quantitative link between surface roughness and overland flow (Peñuela et al., 2016) and can be extrapolated to other scales (Peñuela et al., 2013). There is a large body of literature that has been built on the use of probability distributions of hydraulic conductivity, rainfall intensity, and infiltration capacity (Dunne et al., 1991; Smith & Freeze, 1979; Smith & Hebbert, 1979) to represent how sub-scale conditions influence the field or catchment scale runoff response. Similar attempts have been made with storage capacity (Mekonnen et al., 2014; Moore, 1985), but because of the potential for parallel, series, and nested relationships between storage areas, it is perhaps more difficult for such approaches to succeed since the topological relationships need to be accounted for as well (Buttle, 2006). Furthermore, relationships between storage and runoff can be highly non-linear and hysteretic (O'Kane & Flynn, 2007), which can make it difficult to develop parsimonious means of identifying effective threshold parameters.

A (renewed) focus on fill-and-spill process description posits that *heterogeneity is the key to understanding and simulating outflow dynamics*. But not all heterogeneity is of equal importance. Only the spatial structures that determine the threshold are of first-order importance for characterizing controls on runoff generation. Looked at this way, fill-and-spill offers us some hope for dealing with the challenge of heterogeneity; we don't need to measure everything everywhere. We can instead focus on measuring or finding ways to estimate the key thresholds at the scale of the problem we are interested in. For the person heading into the field and wondering what to measure, it is the knowledge that the larger-scale behavior is not controlled by the mean field, but by critical "choke points" or "energy barriers" that must be exceeded. These choke points are the critical features to be found and identified.

5. Uniqueness of Scale

So which threshold matters for runoff generation? The answer to this question depends inherently on the scale of interest; again our mantra: Fill-and-spill is at the scale of the beholder. Figure 2 illustrates how the thresholds that matter at a given scale may not be meaningful at smaller scales, and may be obscured by other thresholds at larger scales. This is aligned with the view that these thresholds are often expressions of the heterogeneity of the landscape, and that this heterogeneity exists at all scales. Furthermore, we should not expect there to be simple relationships between the critical storage thresholds at smaller and larger scales.

We argue that upscaling process laws by averaging small-scale physical laws will not be adequate where a fill-and-spill type mechanism dominates. This will be true even where the upscaling incorporates higher-order corrections to account for heterogeneity. The reason for this, we suggest, is that the larger-scale behavior is not controlled by the mean field, but by the aforementioned critical "choke points," "energy barriers" or "gate keepers" (Phillips et al., 2011) that must be exceeded. So we are not presenting a formal mathematical framework that is substantially distinct from bucket-type models. However, the perceptual framework (the

10.1029/2020WR027514

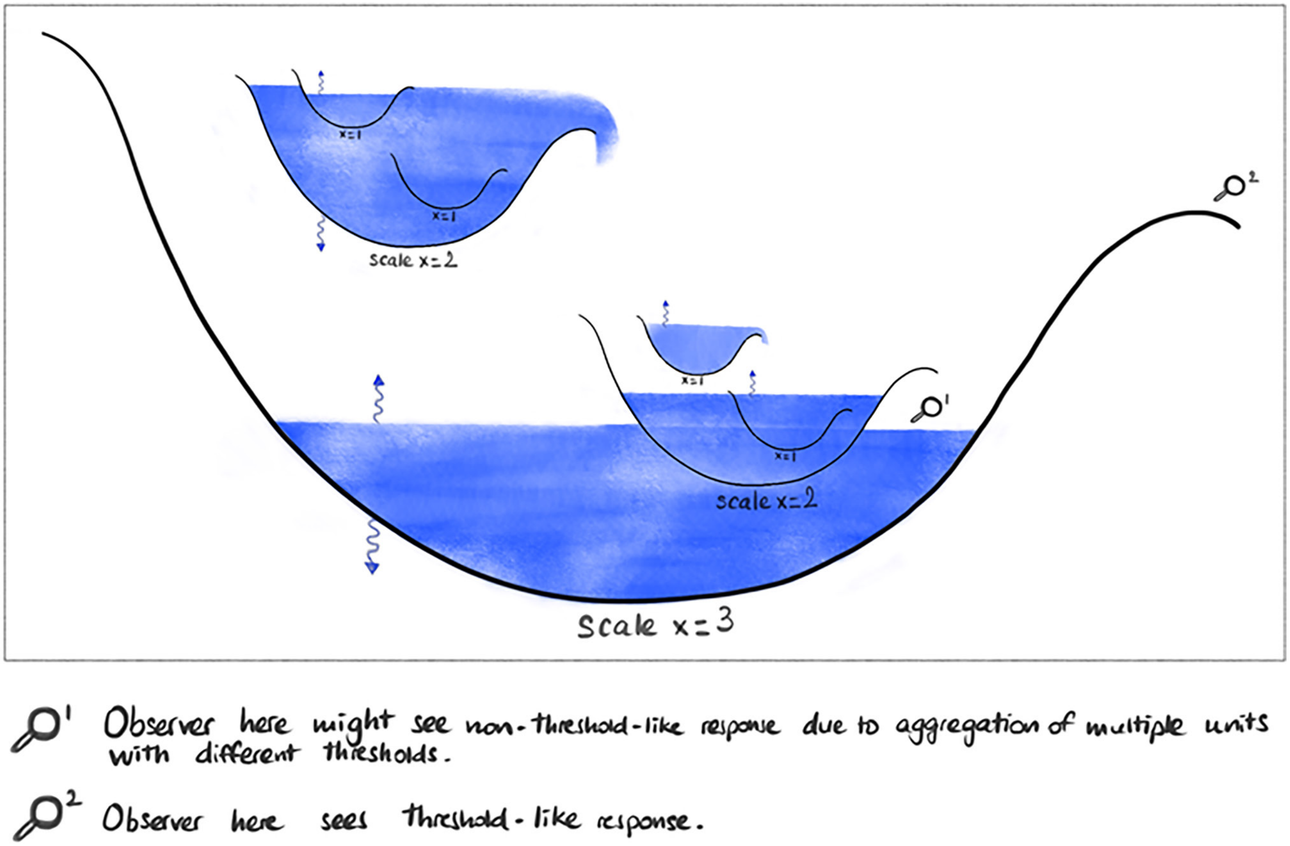

Figure 2. Fill and spill across scales. The thresholds that matter at scale $x = 1$ do not aggregate in simple ways to the thresholds that matter at scale $x = 2$, and similarly at scale $x = 3$. At larger scales, structures and processes can be important that did not matter at smaller scales. Note that at any given scale, the threshold-like activation may also be obscured in observations of aggregate runoff from several landscape units, each with a slightly different threshold, or where the threshold changes in time (Figure 2).

physical reality that the mathematics approximates in some sense) motivates an approach in the future that looks for the relevant choke points or storage areas.

Figure 2 shows conceptually how, if one is observing flow generated by a collection of landscape units i at scale x (e.g., fields) the thresholds $S^*_{i(x)}$ determine the response. The thresholds $S^*_{j(x-1)}$ of units j at a smaller nested scale x–1 that make up the x-scale units (e.g., soil pedons within each field) can arise from quite distinct mechanisms. Consequently, the storage threshold at the larger scale unit may be larger than the sum of the storage thresholds of its parts.

Consider an agricultural field where soil has become saturated down to a plow-pan, and ponding has begun. If the ponded water flows toward a topographic depression within the field, runoff from the field as a whole may not be appreciable until that larger depression has filled to some threshold level. Even though the thresholds have been exceeded for all the landscape units within the area, the runoff that is produced fills an additional storage that prevents flow from leaving the area (even if it doesn't prevent water from being redistributed internally). See an example of this for a frozen agricultural field in Coles and McDonnell (2018).

6. How Might Fill-and-Spill get us Closer to Grouping Runoff Processes?

Ecologists realized early on that in open environmental systems—such as plots, hillslopes, catchments and large watersheds—all we can do is to search for patterns and model parameters (Eberhardt & Thomas, 1991). The fill-and-spill "pattern" links the threshold for runoff generation to the filling of storage, and acknowledges the heterogeneity in storage.

Going into a research catchment with the fill-and-spill mechanism as a guide can help change the goal from an amorphous "characterization" to "measurements against a framework" with questions like: How does the system fill? What is the threshold for spill? How much loss occurs along the connected flowpath? What physiographic and climatic controls dictate how the fill-and-spill processes at one scale influence the redistribution of water that is important for the threshold response at larger scales? How does vegetation's removal of water enhance or diminish fill-and-spill in different vegetation-climate-soil settings? The answers to these questions can also help to group catchments or runoff processes.

We acknowledge that there are still instances where other threshold processes are conflated with fill-and-spill. Question going forward include: How can we develop a sharper definition of fill-and-spill? We know that spatial heterogeneity occurs everywhere and at all scales and that there are a number of threshold-like processes in hydrology; how can theory be guided by this? Should such discussions be driven by inductive inference (as done in this commentary) and how can a deductive approach help? Should discussion of fill-and-spill be a discussion of fill-*store*-spill as outlined by Bogaard and Greco (2016) as linked to their focus on landslide initiation? Important new work has begun to critically examine fill-and-spill at the hillslope scale in light of preferential flow at the plot scale (Nyquist et al., 2018), critically assessing the relative amounts of fill versus spill at the hillslope scale in low angled terrain where this ratio is different than for upland sites (Du et al., 2016), or describing fill-and-spill linked to shrink swell clays (Stewart et al., 2015).

One specific test that we think could be useful to include, is to explore how threshold activation "scales." Cammeraat (2004) has an excellent example of this for a semi-arid system where the threshold rainfall necessary to induce flow was inversely related to the area (from plot to hillslope to headwater catchment and larger watersheds). He also found that the variance of this threshold collapsed with scale. Humid and arid systems might have different scaling behavior at the headwater catchment scale, as the water table and riparian zone are always in contact with the stream in humid systems, but perhaps not in arid systems. The threshold rainfall to activate a response at the headwater scale may be smaller in humid systems, and reached more quickly at plot-, hillslope- and larger watershed scales. Such threshold scaling intercomparisons across hydroclimate and hydrogeological environments could be highly revealing.

Lastly, we know that catchments store and release water. The simple equations presented herein are not new—in many ways, they are the essence of the early Stanford Watershed Model (Crawford & Burges, 2004) and the PDM model of Moore (1985) that underlies virtually all of the most widely used catchment scale conceptual models. However, what we have tried to show with the examples provided above is the commonality of fill-and-spill behavior across a wide range of scales; and that the process of event-based filling, sub-grid spilling, connectivity and loss along the flowpath (at the plot, hillslope or catchment scale) leads to a threshold response in systems at all scales.

This commentary is a call for recognizing the many manifestations of fill-and-spill reported in the literature in the past 15 years. What we distill from this is that fill-and-spill occurs *at the scale of the beholder*. It is emergent behavior *at the scale of interest*. It is our pathway to the grouping of facts, so that, as noted at the outset "general laws or conclusions may be drawn from them." Our message here is thus a pragmatic one: do not strive for a new theory that is an assembly of a sequence of scale, but rather something analogous to a "leap of scales" (Beven, 1983)—or what Klemeš (1983) called "the levels of scale at which a meaningful conceptualization of physical processes is possible." In other words, we argue that it is important to focus on learning (1) where the dominant storage areas are that connect and disconnect, (2) the types of losses along the flowpath (as described by Jackson et al., 2014 and Klaus & Jackson, 2018) and, (3) how this leads to a non-linear runoff response.

7. Summary and Outlook

This is a WRR AGU-100 Centennial Commentary and as such, we write in a provocative style to spark debate. There have been many excellent research—and synthesis papers published on hydrological scaling in the past decade that offer better (theoretical) insights. We do not claim to compete with them. Similarly, others have written reviews and comments on connectivity (Rinderer et al., 2018) and threshold behaviors (Ali et al., 2013). We note that the description of runoff processes continues to become more complex. What we offer here is the very simple idea that runoff is at the scale of the beholder and that recognition of fill-and-spill, together with its components of connectivity, thresholds and losses along the flowpath, is a thing for field hydrologists to look for in their quest to understand how their systems work during events. And perhaps guide analysis of how different systems (either at different locations or different scales) group and compare.

We are motivated by the desire to group and arrange behavior in such a way that perhaps others will see connections and help build higher knowledge. This may help slice the Gordian Knot of scale dependence (and often, event-size and intensity dependence) of runoff generation and free ourselves to recognize that runoff is not additive. We argue that "fill-and-spill" manifests at all scales and places and that there is indeed commonality across different systems and scales (cf. McDonnell, 2013). That all watersheds are not unique—but rather fill-and-spill in some way—where the relative dominance of individual fluxes dictates whether it is most of the time filling or spilling—and that searching for these tipping points can be fruitful in perceptual model development. It perhaps provides a guide for field hydrologists on what to measure, in what order and why; with an explicit eye to defining *a priori* the scale of interest and then working to define that scale's key fill-and-spill behavior and controls. Lastly, while Poincaré said that "*Les faits ne parlent pas*" (Facts do not speak), we think that the grouping of fill-and-spill facts speaks volumes.

Data Availability Statement

No new data are presented in this commentary therefore there are no data to include in a repository.

Acknowledgments
The authors thank many colleagues and former students for fruitful discussions of fill-and-spill over the past many years: Luisa Hopp, Anna Coles, Genevieve Ali and Scott Leibowitz. The authors thank A. Ireson, J. Pomeroy, K. Shook, A. Ameli, J. Craig, T. Bogaard, and A. Ward for recent useful discussions. But any nonsense is ours alone. J. J. McDonnel thanks the Luxembourg Institute for Science and Technology and Laurent Pfister for sabbatical support during the writing of this paper. Kim Janzen is thanked for her edits. The authors thank our 3 reviewers and the editor, Charlie Luce, for very useful critical feedback that helped improve the final version and alerting us to a few additional key references. The authors especially thank one critical reviewer who in the end, disagreed with parts of our message but was magnanimous to support publication of opinions contrary to their own for the sake of scientific debate. The authors thank Fabrizio Fenicia for reminding us of the PDM model of Moore; the authors thank also attendees of the December 2019 Berkeley Catchment Science Symposium where an earlier version of this work was presented and marathon question and answer session following that talk surfaced a number of critical questions that helped sharpen our ideas. Finally, the authors thank Kim Janzen, Pedro Hervé-Fernández, Cody Millar, Veva McDonnell and Magali Nehemy for useful final discussions. Finally, we thank Louise Arnal for her help with the final figures.

References

Ali, G., Oswald, C. J., Spence, C., Cammeraat, E. L. H., Mcguire, K. J., Meixner, T., & Reaney, S. M. (2013). Towards a unified threshold-based hydrological theory: Necessary components and recurring challenges. *Hydrological Processes*, *27*(2), 313–318. https://doi.org/10.1002/hyp.9560

Ali, G., Tetzlaff, D., Mcdonnell, J. J., Soulsby, C., Carey, S., Laudon, H., et al. (2015). Comparison of threshold hydrologic response across northern catchments. *Hydrological Processes*, *29*(16), 3575–3591. https://doi.org/10.1002/hyp.10527

Ali, G. A., L'Heureux, C., Roy, A. G., Turmel, M.-C., & Courchesne, F. (2011). Linking spatial patterns of perched groundwater storage and stormflow generation processes in a headwater forested catchment. *Hydrological Processes*, *25*(25), 3843–3857. https://doi.org/10.1002/hyp.8238

Ameli, A. A., Craig, J. R., & McDonnell, J. J. (2015). Are all runoff processes the same? Numerical experiments comparing a Darcy-Richards solver to an overland flow-based approach for subsurface storm runoff simulation. *Water Resources Research*, *51*, 10008–10028. https://doi.org/10.1002/2015WR017199

Anderson, M. G., & Burt, T. P. (1978). Toward more detailed field monitoring of variable source areas. *Water Resources Research*, *14*(6), 1123–1131. https://doi.org/10.1029/WR014i006p01123

Antoine, M., Javaux, M., & Bielders, C. (2009). What indicators can capture runoff-relevant connectivity properties of the micro-topography at the plot scale?. *Advances in Water Resources*, *32*(8), 1297–1310. https://doi.org/10.1016/j.advwatres.2009.05.006

Appels, W. M., Bogaart, P. W., & van der Zee, S. E. A. T. M. (2011). Influence of spatial variations of microtopography and infiltration on surface runoff and field scale hydrological connectivity. *Advances in Water Resources*, *34*(2), 303–313. https://doi.org/10.1016/j.advwatres.2010.12.003

Arnold, N. S., Banwell, A. F., & Willis, I. C. (2014). High-resolution modelling of the seasonal evolution of surface water storage on the Greenland Ice Sheet. *The Cryosphere*, *8*(4), 1149–1160. https://doi.org/10.5194/tc-8-1149-2014

Balme, M. R., Gallagher, C. J., Gupta, S., & Murray, J. B. (2011). Fill and spill in Lethe Vallis: A recent flood-routing system in Elysium Planitia, Mars. *Geological Society, London, Special Publications*, *356*(1), 203–227. https://doi.org/10.1144/SP356.11

Barthold, F. K., & Woods, R. A. (2015). Stormflow generation: A meta-analysis of field evidence from small, forested catchments. *Water Resources Research*, *51*(5), 3730–3753. https://doi.org/10.1002/2014WR016221

Beven, K. (1983). Surface water hydrology-runoff generation and basin structure. *Reviews of Geophysics*, *21*(3), 721–730. https://doi.org/10.1029/RG021i003p00721

Beven, K. (2002). Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrological Processes*, *16*(2), 189–206.

Beven, K. (2006). Searching for the Holy Grail of scientific hydrology: $Q_t = (S, R, \Delta t)A$ as closure. *Hydrology and Earth System Sciences*, *10*(5), 609–618. https://doi.org/10.5194/hess-10-609-2006

Beveridge, W. I. (1950). *The art of scientific investigation*. Blackburn Press.

Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling: A review. *Hydrological Processes*, *9*(3–4), 251–290. https://doi.org/10.1002/hyp.3360090305

Blunt, M. J., & Scher, H. (1995). Pore-level modeling of wetting. *Physical Review E*, *52*(6), 6387–6403. https://doi.org/10.1103/PhysRevE.52.6387

Bogaard, T. A., & Greco, R. (2016). Landslide hydrology: From hydrology to pore pressure. *WIREs Water*, *3*(3), 439–459. https://doi.org/10.1002/wat2.1126

Botter, G., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2007). Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response. *Water Resources Research*, *43*(2), 1–14. https://doi.org/10.1029/2006WR005043

Bowling, L. C., & Lettenmaier, D. P. (2010). Modeling the effects of lakes and wetlands on the water balance of arctic environments. *Journal of Hydrometeorology*, *11*(2), 276–295. https://doi.org/10.1175/2009JHM1084.1

Brooks, J. R., Mushet, D. M., Vanderhoof, M. K., Leibowitz, S. G., Christensen, J. R., Neff, B. P., et al. (2018). Estimating wetland connectivity to streams in the prairie pothole region: An isotopic and remote sensing approach. *Water Resources Research*, *54*(2), 955–977. https://doi.org/10.1002/2017WR021016

Buttle, J. (2006). Mapping first-order controls on streamflow from drainage basins: The T3 template. *Hydrological Processes*, *20*(15), 3415–3422. https://doi.org/10.1002/hyp.6519

Cammeraat, E. L. H. (2004). Scale dependent thresholds in hydrological and erosion response of a semi-arid catchment in southeast Spain. *Agriculture, Ecosystems & Environment*, *104*(2), 317–332. https://doi.org/10.1016/j.agee.2004.01.032

Camporese, M., Paniconi, C., Putti, M., & McDonnell, J. J. (2019). Fill and spill hillslope runoff representation with a Richards equation-based model. *Water Resources Research*, *55*(11), 8445–8462. https://doi.org/10.1029/2019WR025726

Chu, X., Padmanabhan, G., & Bogart, D. (2015). Microrelief-controlled overland flow generation: Laboratory and field experiments. *Applied and Environmental Soil Science*, *2015*, 1–11. https://doi.org/10.1155/2015/642952

Clark, M. P., Bierkens, M. F. P., Samaniego, L., Woods, R. A., Uijlenhoet, R., Bennett, K. E., et al. (2017). The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism. *Hydrology and Earth System Sciences*, *21*(7), 3427–3440. https://doi.org/10.5194/hess-21-3427-2017

Coles, A. E., & McDonnell, J. J. (2018). Fill and spill drives runoff connectivity over frozen ground. *Journal of Hydrology*, *558*, 115–128. https://doi.org/10.1016/j.jhydrol.2018.01.016

Connon, R. F., Quinton, W. L., Craig, J. R., Hanisch, J., & Sonnentag, O. (2015). The hydrology of interconnected bog complexes in discontinuous permafrost terrains. *Hydrological Processes*, *29*(18), 3831–3847. https://doi.org/10.1002/hyp.10604

Crawford, N. H., & Burges, S. J. (2004). The history of the Stanford watershed model. *Water Resources IMPACT*, *6*(2), 3–5.

Darboux, F., Davy, P., Gascuel-Odoux, C., & Huang, C. (2001). Evolution of soil surface roughness and flowpath connectivity in overland flow experiments. *Catena*, *46*, 125–139. https://doi.org/10.1016/S0341-8162(01)00162-X

Darwin, C. (1859). *The origin of species*. New York, NY: PF Collier & son.

DeBoer, D. W., & Johnson, P. (1971). Simulation of runoff from depression characterized watersheds. *Transactions of the ASAE*, *14*(4), 615–620.

Devito, K., Creed, I., Gan, T., Mendoza, C., Petrone, R., Silins, U., & Smerdon, B. (2005). A framework for broad-scale classification of hydrologic response units on the boreal plain: Is topography the last thing to consider?. *Hydrological Processes*, *19*(8), 1705–1714. https://doi.org/10.1002/hyp.5881

DiCarlo, D. A. (2013). Stability of gravity-driven multiphase flow in porous media: 40 years of advancements. *Water Resources Research*, *49*(8), 4531–4544. https://doi.org/10.1002/wrcr.20359

Dooge, J. C. I. (1986). Looking for hydrologic laws. *Water Resources Research*, *22*(9 S), 46S–58S. https://doi.org/10.1029/WR022i09Sp0046S

Du, E., Rhett Jackson, C., Klaus, J., McDonnell, J. J., Griffiths, N. A., Williamson, M. F., et al. (2016). Interflow dynamics on a low relief forested hillslope: Lots of fill, little spill. *Journal of Hydrology*, *534*, 648–658. https://doi.org/10.1016/j.jhydrol.2016.01.039

Dunne, T., & Black, R. D. (1970). Partial area contributions to storm runoff in a small New England watershed. *Water Resources Research*, *6*(5), 1296–1311. https://doi.org/10.1029/wr006i005p01296

Dunne, T., Zhang, W., & Aubry, B. F. (1991). Effects of rainfall, vegetation, and microtopography on infiltration and runoff. *Water Resources Research*, *27*(9), 2271–2285. https://doi.org/10.1029/91WR01585

Dyre, J. C. (1988). The random free-energy barrier model for ac conduction in disordered solids. *Journal of Applied Physics*, *64*(5), 2456–2468. https://doi.org/10.1063/1.341681

Eberhardt, L. L., & Thomas, J. M. (1991). Designing environmental field studies. *Ecological Monographs*, *61*(1), 53–73. https://doi.org/10.2307/1942999

Eliassi, M., & Glass, R. J. (2001). On the continuum-scale modeling of gravity-driven fingers in unsaturated porous media: The inadequacy of the Richards equation with standard monotonic constitutive relations and hysteretic equations of state. *Water Resources Research*, *37*(8), 2019–2035. https://doi.org/10.1029/2000WR900403

Eliassi, M., & Glass, R. J. (2003). On the porous continuum-scale modeling of gravity-driven fingers in unsaturated materials: Numerical solution of a hypodiffusive governing equation that incorporates a hold-back-pile-up effect. *Water Resources Research*, *39*(6), 1–11. https://doi.org/10.1029/2002WR001535

Falkenmark, M., & Chapman, T. (1989). *Comparative hydrology: An ecological approach to land and water resources*. UNESCO.

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., et al. (2019). Hillslope hydrology in global change research and earth system modeling. *Water Resources Research*, *55*(2), 1737–1772. https://doi.org/10.1029/2018WR023903

Farrick, K. K., & Branfireun, B. A. (2014). Soil water storage, rainfall and runoff relationships in a tropical dry forest catchment. *Water Resources Research*, *50*(12), 9236–9250. https://doi.org/10.1002/2014WR016045

Frei, S., Knorr, K. H., Peiffer, S., & Fleckenstein, J. H. (2012). Surface micro-topography causes hot spots of biogeochemical activity in wetland systems: A virtual modeling experiment. *Journal of Geophysical Research: Biogeosciences*, *117*(4), 1–18. https://doi.org/10.1029/2012JG002012

Frei, S., Lischeid, G., & Fleckenstein, J. H. (2010). Effects of micro-topography on surface-subsurface exchange and runoff generation in a virtual riparian wetland—A modeling study. *Advances in Water Resources*, *33*(11), 1388–1401. https://doi.org/10.1016/j.advwatres.2010.07.006

Ghatan, G. J., Head, J. W., & Wilson, L. (2005). Mangala Valles, Mars: Assessment of early stages of flooding and downstream flood evolution. *Earth Moon Planet*, *96*(1–2), 1–57. https://doi.org/10.1007/s11038-005-9009-y

Gibson, J. J., Yi, Y., & Birks, S. J. (2016). Isotope-based partitioning of streamflow in the oil sands region, northern Alberta: Towards a monitoring strategy for assessing flow sources and water quality controls. *Journal of Hydrology: Regional Studies*, *5*, 131–148. https://doi.org/10.1016/j.ejrh.2015.12.062

Golden, H. E., Sander, H. A., Lane, C. R., Zhao, C., Price, K., D'Amico, E., & Christensen, J. R. (2016). Relative effects of geographically isolated wetlands on streamflow: A watershed-scale analysis. *Ecohydrology*, *9*(1), 21–38. https://doi.org/10.1002/eco.1608

Graham, C. B., Woods, R. A., & McDonnell, J. J. (2010). Hillslope threshold response to rainfall: (1) A field based forensic approach. *Journal of Hydrology*, *393*(1–2), 65–76. https://doi.org/10.1016/j.jhydrol.2009.12.015

Gray, D. M. (1961). Interrelationships of watershed characteristics. *Journal of Geophysical Research*, *66*(4), 1215–1223. https://doi.org/10.1029/JZ066i004p01215

Haan, C. T. (1967). Hydraulics of watersheds characterized by depressional storage (p. 197, PhD thesis). University of Iowa. Retrieved from https://lib.dr.iastate.edu/rtd/3934

Haines, W. B. (1930). Studies in the physical properties of soil. V. The hysteresis effect in capillary properties, and the modes of moisture distribution associated therewith. *Journal of Agricultural Sciences*, *20*(1), 97–116. https://doi.org/10.1017/s002185960008864x

Haque, A., Ali, G., & Badiou, P. (2018). Hydrological dynamics of prairie pothole wetlands: Dominant processes and landscape controls under contrasted conditions. *Hydrological Processes*, *32*(15), 2405–2422. https://doi.org/10.1002/hyp.13173

Helbig, M., Boike, J., Langer, M., Schreiber, P., Runkle, B. R. K., & Kutzbach, L. (2013). Spatial and seasonal variability of polygonal tundra water balance: Lena River Delta, northern Siberia (Russia). *Hydrogeology Journal*, *21*(1), 133–147. https://doi.org/10.1007/s10040-012-0933-4

Herring, A. L., Gilby, F. J., Li, Z., McClure, J. E., Turner, M., Veldkamp, J. P., et al. (2018). Observations of nonwetting phase snap-off during drainage. *Advances in Water Resources*, *121*(March), 32–43. https://doi.org/10.1016/j.advwatres.2018.07.016

Hewlett, J. D., & Hibbert, A. R. (1967). Factors affecting the response of small watersheds to precipitation in humid areas. *Forest Hydrology*, *1*, 275–290. https://doi.org/10.1177/0309133309338118

Hewlett, J. D., & Nutter, W. L. (1970). The varying source area of streamflow from upland basins. In *Proceedings of the Symposium on interdisciplinary aspects of watershed management, Montana State University (pp. 65–83)*. Bozeman Montana.

Hopp, L., & McDonnell, J. J. (2009). Connectivity at the hillslope scale: Identifying interactions between storm size, bedrock permeability, slope angle and soil depth. *Journal of Hydrology*, *376*(3–4), 378–391. https://doi.org/10.1016/j.jhydrol.2009.07.047

Howard, A. J., Bonell, M., Gilmour, D., & Cassells, D. (2010). Is rainfall intensity significant in the rainfall-runoff process within tropical rainforests of northeast Queensland? The Hewlett regression analyses revisited. *Hydrological Processes*, *24*(18), 2520–2537. https://doi.org/10.1002/hyp.7694

Hughlings-Jackson, J. (1882). Localised convulsions from tumour of the brain. *Brain*,*5*(3), 364–374.

Jackson, C. R., Bitew, M., & Du, E. (2014). When interflow also percolates: Downslope travel distances and hillslope process zones. *Hydrological Processes*, *28*(7), 3195–3200. https://doi.org/10.1002/hyp.10158

Jackson, C. R., Du, E., Klaus, J., Griffiths, N. A., Bitew, M., & McDonnell, J. J. (2016). Interactions among hydraulic conductivity distributions, subsurface topography, and transport thresholds revealed by a multitracer hillslope irrigation experiment. *Water Resources Research*, *52*(8), 6186–6206. https://doi.org/10.1002/2015WR018364

James, A. L., McDonnell, J. J., Tromp-Van Meerveld, I., & Peters, N. E. (2010). Gypsies in the palace: Experimentalist's view on the use of 3-D physics-based simulation of hillslope hydrological response. *Hydrological Processes*, *24*(26), 3878–893. https://doi.org/10.1002/hyp.7819

Jencso, K. G., & McGlynn, B. L. (2011). Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation. *Water Resources Research*, *47*(11), 1–16. https://doi.org/10.1029/2011WR010666

Klaus, J., & Jackson, C. R. (2018). Interflow is not binary: A continuous shallow perched layer does not imply continuous connectivity. *Water Resources Research*, *54*(9), 5921–5932. https://doi.org/10.1029/2018WR022920

Klemeš, V. (1983). Conceptualization and scale in hydrology. *Journal of Hydrology*, *65*(1–3), 1–23. https://doi.org/10.1016/0022-1694(83)90208-1

Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A quantitative hydrological climate classification evaluated with independent streamflow data. *Water Resources Research*, *54*(7), 5088–5109. https://doi.org/10.1029/2018WR022913

Kusumastuti, D. I., Sivapalan, M., Struthers, I., Reynolds, D. A., Murray, K., & Turlach, B. A. (2008). Thresholds in the storm response of a catchment-lake system and the occurrence and magnitude of lake overflows: Implications for flood frequency. *Water Resources Research*, *44*(2), 1–15. https://doi.org/10.1029/2006WR005628

Kusumastuti, D. I., Struthers, I., Sivapalan, M., & Reynolds, D. A. (2007). Threshold effects in catchment storm response and the occurrence and magnitude of flood events: Implications for flood frequency. *Hydrology and Earth System Sciences*, *11*(4), 1515–1528. https://doi.org/10.5194/hess-11-1515-2007

Leibowitz, S. G., Mushet, D. M., & Newton, W. E. (2016). Intermittent surface water connectivity: Fill and spill vs. fill and merge dynamics. *Wetlands*, *36*, 323–342. https://doi.org/10.1007/s13157-016-0830-z

Leibowitz, S. G., & Vining, K. C. (2003). Temporal connectivity in a prairie pothole complex. *Wetlands*, *23*(1), 13–25. https://doi.org/10.1672/0277-5212(2003)023[0013:TCIAPP]2.0.CO;2

Leibowitz, S. G., Wigington, P. J., Schofield, K. A., Alexander, L. C., Vanderhoof, M. K., & Golden, H. E. (2018). Connectivity of streams and wetlands to downstream waters: An integrated systems framework. *Journal of the American Water Resources Association*, *54*(2), 298–322. https://doi.org/10.1111/1752-1688.12631

Martin, J. E. (2011). Hydrologic response of a small forested swamp complex, Seymour Valley, British Columbia. *Canadian Geographer*, *55*(4), 457–469. https://doi.org/10.1111/j.1541-0064.2011.00385.x

Martinec, J. (1975). Subsurface flow from snowmelt traced by tritium. *Water Resources Research*, *11*(3), 496–498. https://doi.org/10.1029/WR011i003p00496

McDonnell, J., & Woods, R. (2004). On the need for catchment classification. *Journal of Hydrology*, *299*(1–2), 2–3. https://doi.org/10.1016/S0022-1694(04)00421-4

McDonnell, J. J. (2003). Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response. *Hydrological Processes*, *17*(9), 1869–1875. https://doi.org/10.1002/hyp.5132

McDonnell, J. J. (2013). Are all runoff processes the same?. *Hydrological Processes*, *27*(26), 4103–4111. https://doi.org/10.1002/hyp.10076

McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., et al. (2007). Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research*, *43*(7). https://doi.org/10.1029/2006WR005467

McGrath, G. S., Hinz, C., Sivapalan, M., Mcgrath, G. S., Hinz, C., & Temporal, M. S. (2007). Temporal dynamics of hydrological threshold events. *Hydrology and Earth System Sciences*, *11*(2), 923–938. https://doi.org/10.5194/hess-11-923-2007

McGuire, K. J., & McDonnell, J. J. (2010). Hydrological connectivity of hillslopes and streams: Characteristic time scales and nonlinearities. *Water Resources Research*, *46*(10), 1–17. https://doi.org/10.1029/2010WR009341

Mekonnen, M. A., Wheater, H. S., Ireson, A. M., Spence, C., Davison, B., & Pietroniro, A. (2014). Towards an improved land surface scheme for prairie landscapes. *Journal of Hydrology*, *511*, 105–116. https://doi.org/10.1016/j.jhydrol.2014.01.020

Menzinger, M., & Wolfgang, R. (1969). The meaning and use of the Arrhenius activation energy. *Angewandte Chemie International Edition in English*, *8*(6), 438–444. https://doi.org/10.1002/anie.196904381

Mielko, C., & Woo, M.-k. (2006). Snowmelt runoff processes in a headwater lake and its catchment, subarctic Canadian shield. *Hydrological Processes*, *20*(4), 987–1000. https://doi.org/10.1002/hyp.6117

Moebius, F., & Or, D. (2012). Interfacial jumps and pressure bursts during fluid displacement in interacting irregular capillaries. *Journal of Colloid and Interface Science*, *377*(1), 406–415. https://doi.org/10.1016/j.jcis.2012.03.070

Moore, R. J. (1985). The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal*, *30*(2), 273–297. https://doi.org/10.1080/02626668509490989

Mosley, M. P. (1979). Streamflow generation in a forested watershed, New Zealand. *Water Resources Research*, *15*(4), 795–806. https://doi.org/10.1029/WR015i004p00795

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, *12*(3), 513–522. https://doi.org/10.1029/WR012i003p00513

Muir, D. L., Hayashi, M., & Mcclymont, A. F. (2011). Hydrological storage and transmission characteristics of an alpine talus. *Hydrological Processes*, *25*(19), 2954–2966. https://doi.org/10.1002/hyp.8060

Nyquist, J. E., Toran, L., Pitman, L., Guo, L., & Lin, H. (2018). Testing the fill-and-spill model of subsurface lateral flow using ground-penetrating radar and dye tracing. *Vadose Zone Journal*, *17*(1), 1–13. https://doi.org/10.2136/vzj2017.07.0142

O'Kane, J. P., & Flynn, D. (2007). Thresholds, switches and hysteresis in hydrology from the pedon to the catchment scale: A non-linear systems theory. *Hydrology and Earth System Sciences Discussions*, *11*(1), 443–459.

Orlowski, N., Lauer, F., Kraft, P., Frede, H.-G., & Breuer, L. (2014). Linking spatial patterns of groundwater table dynamics and streamflow generation processes in a small developed catchment. *Water*, *6*(10), 3085–3117. https://doi.org/10.3390/w6103085

Oswald, C. J., Richardson, M. C., & Branfireun, B. A. (2011). Water storage dynamics and runoff response of a boreal shield headwater catchment. *Hydrological Processes*, *25*(19), 3042–3060. https://doi.org/10.1002/hyp.8036

Peñuela, A., Darboux, F., Javaux, M., & Bielders, C. L. (2016). Evolution of overland flow connectivity in bare agricultural plots. *Earth Surface Processes and Landforms*, *41*(11), 1595–1613. https://doi.org/10.1002/esp.3938

Peñuela, A., Javaux, M., & Bielders, C. L. (2013). Scale effect on overland flow connectivity at the plot scale. *Hydrology and Earth System Sciences*, *17*(1), 87–101. https://doi.org/10.5194/hess-17-87-2013

Phillips, R. W., Spence, C., & Pomeroy, J. W. (2011). Connectivity and runoff dynamics in heterogeneous basins. *Hydrological Processes*. *25*(19), 3061–3075. https://doi.org/10.1002/hyp.8123

Poincaré, H. (1902). *La Science et l'Hypothèse*. Flammarion.

Quinton, W. L., Hayashi, M., & Pietroniro, A. (2003). Connectivity and storage functions of channel fens and flat bogs in northern basins. *Hydrological Processes*, *17*(18), 3665–3684. https://doi.org/10.1002/hyp.1369

Rains, M. C. (2011). Water sources and hydrodynamics of closed-basin depressions, cook inlet region, Alaska. *Wetlands*, *31*(2), 377–387. https://doi.org/10.1007/s13157-011-0147-x

Rains, M. C., Leibowitz, S. G., Cohen, M. J., Creed, I. F., Golden, H. E., Jawitz, J. W., et al. (2016). Geographically isolated wetlands are part of the hydrological landscape. *Hydrological Processes*, *30*(1), 153–160. https://doi.org/10.1002/hyp.10610

Razavy, M. (2003). *Quantum Theory of Tunneling*. World Scientific. Retrieved from https://doi.org/10.1142/4984

Rinderer, M., Ali, G., & Larsen, L. G. (2018). Assessing structural, functional and effective hydrologic connectivity with brain neuroscience methods: State-of-the-art and research directions. *Earth-Science Reviews*, *178*, 29–47. https://doi.org/10.1016/j.earscirev.2018.01.009

Rodriguez-Iturbe, I., & Porporato, A. (2005). *Ecohydrology of water-controlled ecosystems: Soil moisture and plant dynamics. Ecohydrology of water-controlled ecosystems: Soil moisture and plant dynamics (Vol. 9780521819)*. Cambridge University Press. Retrieved from https://doi.org/10.1017/CBO9780511535727

Rushlow, C. R., & Godsey, S. E. (2017). Rainfall-runoff responses on Arctic hillslopes underlain by continuous permafrost, North Slope, Alaska, USA. *Hydrological Processes*, *31*(23), 4092–4106. https://doi.org/10.1002/hyp.11294

Sayama, T., Mcdonnell, J. J., Dhakal, A., & Sullivan, K. (2011). How much water can a watershed store?. *Hydrological Processes*, *25*(25), 3899–3908. https://doi.org/10.1002/hyp.8288

Shaw, D. A., Vanderkamp, G., Conly, F. M., Pietroniro, A., & Martz, L. (2012). The fill-spill hydrology of prairie wetland complexes during drought and deluge. *Hydrological Processes*, *26*(20), 3147–3156. https://doi.org/10.1002/hyp.8390

Smith, L., & Freeze, R. A. (1979). Stochastic analysis of steady state groundwater flow in a bounded domain: 2. Two-dimensional simulations. *Water Resources Research*, *15*(6), 1543–1559. https://doi.org/10.1029/WR015i006p01543

Smith, R. E., & Hebbert, R. H. B. (1979). A Monte Carlo analysis of the hydrologic effects of spatial variability of infiltration. *Water Resources Research*, *15*(2), 419–429. https://doi.org/10.1029/WR015i002p00419

Spence, C. (2007). On the relation between dynamic storage and runoff: A discussion on thresholds, efficiency, and function. *Water Resources Research*, *43*(12), 1–11. https://doi.org/10.1029/2006WR005645

Spence, C. (2010). A paradigm shift in hydrology: Storage thresholds across scales influence catchment runoff generation. *Geography Compass*, *4*(7), 819–833. https://doi.org/10.1111/j.1749-8198.2010.00341.x

Spence, C., & Woo, M.-K. (2003). Hydrology of subarctic Canadian shield: Soil-filled valleys. *Journal of Hydrology*, *279*(1–4), 151–166. https://doi.org/10.1016/S0022-1694(03)00175-6

Spence, C., & Woo, M.-K. (2006). Hydrology of subarctic Canadian shield: Heterogeneous headwater basins. *Journal of Hydrology*, *317*(1–2), 138–154. https://doi.org/10.1016/j.jhydrol.2005.05.014

Stewart, R. D., Abou Najm, M. R., Rupp, D. E., Lane, J. W., Uribe, H. C., Arumí, J. L., & Selker, J. S. (2015). Hillslope run-off thresholds with shrink-swell clay soils. *Hydrological Processes*, *29*(4), 557–571. https://doi.org/10.1002/hyp.10165

Tromp-Van Meerveld, H. J., & McDonnell, J. J. (2006a). Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope. *Water Resources Research*, *42*(2), 1–11. https://doi.org/10.1029/2004WR003778

Tromp-Van Meerveld, H. J., & McDonnell, J. J. (2006b). Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water Resources Research*, *42*(2), 1–11. https://doi.org/10.1029/2004WR003800

Tromp-van Meerveld, I., & Weiler, M. (2008). Hillslope dynamics modeled with increasing complexity. *Journal of Hydrology*, *361*(1–2), 24–40. https://doi.org/10.1016/j.jhydrol.2008.07.019

Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment classification and hydrologic similarity. *Geography Compass*, *1*(4), 901–931. https://doi.org/10.1111/j.1749-8198.2007.00039.x

Ward, A. S., Schmadel, N. M., & Wondzell, S. M. (2018). Simulation of dynamic expansion, contraction, and connectivity in a mountain stream network. *Advances in Water Resources*, *114*, 64–82. https://doi.org/10.1016/j.advwatres.2018.01.018

Warner, N. H., Sowe, M., Gupta, S., Dumke, A., & Goddard, K. (2013). Fill and spill of giant lakes in the eastern valles marineris region of mars. *Geology*, *41*(6), 675–678. https://doi.org/10.1130/G34172.1

Weyman, D. R. (1973). Measurements of the downslope flow of water in a soil. *Journal of Hydrology*, *20*(3), 267–288. https://doi.org/10.1016/0022-1694(73)90065-6

Wittenberg, H. (1999). Baseflow recession and recharge as nonlinear storage processes. *Hydrological Processes*, *13*(5), 715–726. https://doi.org/10.1002/(SICI)1099-1085(19990415)13:5<715::AID-HYP775>3.0.CO;2-N

Young, K. L., Assini, J., Abnizova, A., & De Miranda, N. (2010). Hydrology of hillslope-wetland streams, Polar Bear Pass, Nunavut, Canada. *Hydrological Processes*, *24*(23), 3345–3358. https://doi.org/10.1002/hyp.7751

Zimmermann, B., Zimmermann, A., Turner, B. L., Francke, T., & Elsenbeer, H. (2014). Connectivity of overland flow by drainage network expansion in a rain forest catchment. *Water Resources Research*, *50*(2), 1457–1473. https://doi.org/10.1002/2012WR012660