Upscaling of lateral groundwater flow processes in watershed models

by

Mark Ranjram

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Civil Engineering (Water)

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Peter A. Troch Full Professor, Department of Hydrology and Atmospheric Sciences, University of Arizona
Supervisor	James R. Craig Associate Professor, Department of Civil and Environmental Engineering, University of Waterloo
Internal Member	Nandita Basu Professor, Department of Civil and Environmental Engineering, University of Waterloo
Internal Member	Bryan A. Tolson Professor, Department of Civil and Environmental Engineering, University of Waterloo
Internal-external Member	David L. Rudolph Professor, Department of Earth and Environmental Sciences, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Mark Ranjram was the sole author for Chapters 1 and 5 which were written under the supervision of Dr. James R. Craig and were not written for publication. Chapters 2 and 4 were written with the intent to publish, with Chapter 2 containing contributions by co-authors. Chapter 3 is a published manuscript.

All work in this thesis was supported by the Natural Sciences and Engineering Research Council of Canada (PGSD3-518790-2018). As lead author of all thesis chapters, I was primarily responsible for study design, data collection and synthesis, and drafting manuscripts. Exceptions to sole authorship of material are as follows.

Research presented in Chapter 2

This research was conducted at the University of Waterloo by Mark Ranjram under the supervision of Dr. James R. Craig. Dr. Craig, Hannah Burdett, Leland Scantlebury, and Mahkameh Taheri contributed to reference collection and provided text for the draft manuscript. Dr. Craig contributed significantly to Section 2.2 in this manuscript. Hannah Burdett is the co-author of Section 2.3.4. Leland Scantlebury is the co-author of Section 2.3.5 and developed the finalized version of Figure 2.1. Makhameh Taheri is the co-author of Section 2.3.1.4. Mark Ranjram contributed significant editorial and intellectual input on all the co-authored sections and wrote all other portions of the manuscript.

Research presented in Chapter 3

This research was conducted at the University of Waterloo by Mark Ranjram under the supervision of Dr. James R. Craig. Dr. Craig contributed to the design of the study and provided iterative feedback during the data collection and analysis phase. Mark Ranjram conducted all other aspects of the study. Mark Ranjram wrote the draft manuscript, which Dr. Craig contributed intellectual input on.

This work was published in *Water Resources Research* on March 14, 2021. The published version of this manuscript is included in this thesis for deposit in the University of Waterloo's thesis repository following a six-month embargo which ended on September 14, 2021.

Citation:

> Ranjram, M., & Craig, J. R. (2021). Use of an efficient proxy solution for the hillslope-storage Boussinesq problem in upscaling of subsurface stormflow. Water Resources Research, 57, 2020WR029105. https://doi.org/10.1029/2020WR029105

Abstract

Hydrologic models are essential tools for hydrologists and water resources engineers. However, there is a critical gap between our knowledge of the appropriate model representation of physical processes at the point scale and at the watershed scale. Due to the ubiquitous presence of heterogeneity, this scale disconnect can have significant implications for practical watershed modelling. Closing this gap requires the identification and implementation of appropriate upscaling approaches. Upscaling refers to the derivation of relationships which translate small scale process descriptions into constitutive relationships that are applicable at larger scales. While several approaches in the literature have been successful in generating such relations for vertical flow (i.e., laterally independent) processes, such as infiltration or evapotranspiration, attempts to address lateral flow processes have been limited.

This work establishes upscaling relationships and identifies critical watershed-scale landform controls for lateral subsurface stormflow, being the lateral transfer of water through the saturated subsurface of a hillslope into a surface water network during a recharge event. This work also establishes a novel and comprehensive literature review synthesizing decades of applied upscaling work in the hydrology literature. This review is intended to provide a practical starting point for future research into upscaling approaches in computational hydrology.

Subsurface stormflow is here characterized by the hillslope-storage Boussinesq (hsB) equation at the hillslope scale, and manifests as an aggregate recession curve at the basin sale. A surrogate model developed herein, the hsB Proxy, rapidly provides high-quality approximate solutions to this non-linear governing equation, and thus enables the application of the equation across 50 basins composed of hundreds of hillslopes. Upscaling relationships are generated from the aggregate recession behaviour in these basins in response to a recharge time series. This work describes the development of the hsB proxy, and the associated development of novel upscaling relationships that capably reproduce the aggregate recession behaviour of a basin in response to a recharge time series, using only the distribution of hillslope-scale properties as input. The insights thus generated provide a new, explicit connection between the topographic characteristics of a basin, the history and magnitude of recharge, and the large-scale recession response. These relationships allow hydrologic models to include the insights of detailed hillslope drainage physics without the associated computational cost.

Acknowledgements

I would like to thank Professor James Craig for his supervision and support over the last four and a half years. Professor Craig is, simply, a legend, and I join the humble trail of graduate students who affirm it to the world: your guidance, insight, humour, and critical thought has made me a deeper thinker and a sharper mind, and made this Ph. D. a joy. It was an absolute honour.

I would like to thank my thesis committee, Professor Nandita Basu, Professor Dave Rudolph, Professor Bryan Tolson, and Professor Peter Troch for their insights and comments on this work.

I would like to thank the Hydrology Research Group in the Civil and Environmental Engineering department at the University of Waterloo for their support and comradery. I would especially like to thank Hannah Burdett, Leland Scantlebury, and Mahkameh Taheri for their gracious contributions to the upscaling review paper.

The financial support of NSERC and the University of Waterloo is gratefully acknowledged.

Dedication

This thesis is dedicated to Rachele, who's grace and love has been my baseline for seven years. Here's to the four-hour bus rides, the three-hour commutes, the home we finally got to make together, and the home we are going to make together. It's been a wild four years and you've been my anchor through it all. I couldn't ask for a better partner. You deserve an emotional co-authorship on everything in this work.

This thesis is dedicated to Maman Joon, who faced a new world with dignity and grit, and gave her twin boys a perfect childhood. Who kicked cancer's *** and worked two jobs to put rent in the bank and food on the table. Thank you for all you've done.

This thesis is dedicated to Laleh Joon, who's worked through tougher problems than any you'll find in this work, and who taught me the value of laughter and joy. You deserve a thousand degrees for all the **** you've been through.

This thesis is dedicated to Hazhir, who paved the way forward, and taught me how to raise my head and live with fierce pride and confidence. Hazhir, with every year that passes, I check in on the responsibilities you were carrying when you were that age, and it boggles my mind. Any achievements I've had are nothing compared to the challenges you've taken on. I wouldn't be who I am without you.

This thesis is dedicated to Kavian, who reflects the best in all of us, and makes us shine so bright. There is literally no better recipe for a companion in this world than spending every waking hour together from -0.75 to 18 years, and then embarking on the exact same professional pathway. It makes me smile so big that you made it, Professor.

And finally, this thesis is dedicated to Reza, who never knew he raised two doctors.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Scale issues are inherent in many aspects of hydrology, arising from a simple constraint: humanity can only observe the environment over a fixed range of spatial extent. In the field, hydrologists can obtain detailed characterizations of dozens of square metres, but to extend that characterization to a regional or watershed scale requires assumptions about how the small-scale behaviour applies beyond the scale of observation. In the office, hydrologists are tasked with building, calibrating, and utilizing hydrological models to answer critical water management questions at the watershed scale, making use of governing constitutive relations for the suite of relevant hydrological processes in their model that were derived at, or informed by, the observation scale. In both cases, the hydrologist has small-scale understanding in hand and must translate that understanding to a larger scale. This is the essence of upscaling. Unfortunately, upscaling is not trivial: it requires a comprehensive understanding of small-scale heterogeneity, and/or defensible simplifications of this heterogeneity. And yet, upscaling has been successfully applied throughout the hydrological literature to generate large-scale descriptions of small-scale processes. This notion of upscaling small-scale governing constitutive relationships is the focus of this thesis. Although several landmark publications exist cataloguing scale issues in hydrology generally (e.g., Klemeš, 1983; Blöschl and Sivapalan, 1995); and several publications exist cataloguing scale issues as an aspect of an in-depth review of specific processes (e.g., Vereecken *et al.*, 2007, 2019); this thesis contains a novel, comprehensive literature review combining these two approaches, providing a detailed assessment of upscaling strategies and conclusions across a suite of hydrological processes, including runoff, infiltration, evapotranspiration, and snow processes. In lieu of a classical literature review providing background on the specific lateral groundwater flow problem considered in the title of this work, the second chapter of this thesis presents a larger-scope review of computational upscaling methods in hydrology in which the relevant runoff upscaling literature is a subset (Section 2.3.1)

The question of upscaling is investigated more directly in the third and fourth chapters of this thesis, which explore the link between the recession behaviour of a basin and the lateral subsurface flow through hillslopes in response to recharge events (i.e., subsurface stormflow). That there is a connection between the recession of the hydrograph and subsurface flow has long been accepted, but more and more complexity has been considered in this problem over time. The original works in this field, looking to parameterize a single subsurface reservoir using observed recession behaviour, have been extended to interpret the influence of a network of independent, heterogeneous subsurface reservoirs on the recession rather than a single reservoir. Independent of the complexity of the conceptualization, the translation of small-scale subsurface flow mechanics into a large-scale recession relationship is of clear utility in hydrology, where, for computational and conceptual reasons, it is difficult to include detailed solutions to groundwater flow equations. Together, the third and fourth chapters of this thesis demonstrate that the classical approach of conceptualizing the subsurface as an

abstracted single reservoir releasing water according to a power-law recession relationship can be justified by aggregating physically-based descriptions of subsurface stormflow. That is, rather than a single reservoir (e.g., Brutsaert and Nieber, 1977), these works utilize hundreds of independent heterogeneous hillslopes and that are derived directly from topography and characterized by equations of groundwater flow as opposed to conceptual storage-outflow relationships (e.g., Clark *et al.*, 2009; Harman *et al.*, 2009). In this way, these works serve both to upscale as well as to translate expertise across hydrological silos: by minimizing the computational expense of a robust groundwater flow equation (Chapter 3), the results of detailed numerical modelling of hundreds of hillslopes can be rapidly aggregated to provide a physically-based understanding of the power-law recession behaviour at the basin scale. Then, by applying this groundwater flow equation to hundreds of hillslopes across dozens of basins, generalized upscaling relationships can be extracted (Chapter 4).

In this way, this thesis intends to provide a baseline for future upscaling research in hydrology, while also presenting a novel upscaling analysis.

1.1 Objectives

This thesis can thus be categorized as having three objectives, as follows:

1. Provide an in-depth review of computational upscaling approaches in hydrology. This critical review provides a synthesis of the methods and insights of upscaling analyses across a suite of hydrological processes, organizes these methods into a consistent ontology, and provides generalized conclusions about the challenges and successes of upscaling methods across these processes (Chapter 2).
2. Demonstrate the efficacy of a physically-based conceptualization of subsurface flow mechanics through hillslopes in generating the convenient power-law recession behaviour at the basin scale, and create a tool that allows these mechanics to be rapidly applied to thousands of hillslopes, removing the burden of computational cost from the upscaling exercise (Chapter 3).
3. Utilize this proof-of-concept to analyze the aggregate hillslope subsurface stormflow behaviour across dozens of basins, to generate upscaling relationships that can directly estimate the recession behaviour of a basin from topographic characteristics which inform the subsurface flow equations. In this way, the hillslope-scale understanding of the subsurface physics is upscaled directly to the basin scale (Chapter 4).

Chapter 2

A review of computational upscaling approaches in surface water hydrology

2.1 Introduction

Water moves through the environment by several distinct flow pathways. Understanding these fluxes of water, including lateral overland flow, shallow subsurface runoff, vertical infiltration, evapotranspiration, and various snow processes such as melt, blowing snow transport, and interception, is essential to the science of hydrology. Through rigorous observations of nature, hydrologists have developed governing constitutive relationships which, embedded within computational algorithms, enable practical estimation of water fluxes at the scale of observation: given the state of the system (e.g., soil saturation) we may calculate fluxes (e.g., percolation) using constitutive relations (e.g., Darcy's law). Hydrologic models combine and apply these algorithms in an attempt to comprehensively describe the movement of water through a catchment. However, fundamental to the application of these algorithms is a question of scale: if a flux is approximated by a relationship derived from observations at a local scale in a homogeneous system (over several metres), application of that same relationship at much larger scales in an inevitably heterogenous catchment requires care. In hydrologic modelling, this problem is partially alleviated by the utility of distributed and semi-distributed discretization schemes, which (ideally) subdivide the landscape into scale-appropriate, relatively homogeneous units. When the governing constitutive relationships are applied at a scale larger than the scale at which they were derived, or when a new relationship is derived to express the net impact of heterogeneous small-scale fluxes at a larger scale, upscaling strategies must be used. In this review paper, the term "upscaling" describes this act of generating a large-scale representation of a hydrologic process that accounts for small-scale heterogeneity. A computational upscaling method is thus a set of rules that performs a many-to-one conversion: the many heterogeneous fluxes within a domain (resulting from heterogeneity in parameters and states) are converted into a single homogeneous flux that adequately represents the small-scale heterogeneity. In this way, we make a distinction with alternative definitions of "upscaling" in the hydrological literature which relate to the translation of small-scale data to larger scales, such as aggregating rasterized data or interpolating point data for large-scale applications.

This review presents an historical overview and ontology of computational upscaling strategies in the watershed modelling literature. Specifically, this work catalogues various methods used to upscale governing constitutive relationships which characterize hydrologic fluxes at the local scale into constitutive relationships applied in hydrologic models at the Hydrologic Response Unit (HRU) scale or catchment scale. The local scale is here typically confined to be on the order of 1-10 m^2, representing typical scales of observation, while the HRU or catchment scale is on the order of 0.1-100 km^2; we do

not discuss upscaling from the microscale (<1 m^2) to the local scale. This work does not aim to catalogue scale issues, which are observations of the occurrence of scaling or observations of correlations between hydrologic variables with increasing observation scale (e.g., Blöschl and Sivapalan, 1995; Blöschl, 2001), but rather focuses on the computational treatment of these scale issues. Additionally, the focus of this work is on surface and shallow subsurface water fluxes, and so we exclude any review of transport phenomena upscaling or upscaling in aquifer subsurface flow (for rigorous discussions of these, see e.g., Cushman et al., 2002; Rubin, 2003; Frippiat and Holeyman, 2008; Yang et al., 2021). Finally, this work does not include temporal upscaling approaches: the focus is exclusively on spatial upscaling.

To provide necessary context for and organization of the computational upscaling approaches catalogued in this work, shared definitions and an accompanying mathematical treatment of upscaling are first provided in Sections 2.1.1 and 2.2. Section 2.3 presents a catalogue of key computational upscaling approaches historically used in surface water hydrology, categorized by process: runoff, infiltration, evapotranspiration, and snow-related processes. Section 2.3 also describes how these methods have been integrated into hydrological modelling platforms and land surface schemes, and the issues arising when doing so. Finally, Section 2.4 provides a discussion of current challenges in computational upscaling which hopefully serves as a useful reference point for future research into upscaling problems.

Please note that the "power-law recession of the hydrograph" subsection of the runoff section (Section 2.3.1) is quite detailed compared to all other processes catalogued. When this manuscript is submitted for publication, an alternative, condensed version of this subsection will be substituted. The expanded version of this subsection is included in this thesis to support Chapters 3 and 4, and serves as the distilled literature review relevant to these chapters, reflecting the lack of a traditional literature review chapter in this thesis.

2.1.1 Definitions

Upscaling generally refers to the translation of small-scale information to a representative form at a larger scale (Blöschl and Sivapalan, 1995). Here, we focus solely on spatial upscaling and refine this definition to focus specifically on the upscaling of hydrological processes; that is, the conversion of a spatially heterogeneous flux response across an HRU or catchment (where that heterogeneity arises due to heterogeneity in parameters and states) into a single equivalent homogenous form at the HRU or catchment scale. This is inherently a simplifying exercise, requiring some loss of information about small-scale heterogeneity. This conversion can be achieved by several different techniques. It is worthwhile to provide definitions of three critical terms in these techniques: states, parameters, and fluxes. States are typically the quantification of the storage elements of a mass balance, which in the context of hydrological processes are measures of water quantity (the depth of a snowbank or the hydraulic head in an aquifer are examples of states, though other intensive properties such as density or temperature may represent system state). Parameters are the environmental controls of the rate of

change of a mass balance, and may be physically based or abstracted (hydraulic conductivity and the power-law coefficients of baseflow recession are examples of physically based and abstracted parameters, respectively). Fluxes are the rates of change of a mass balance, representing the movement of mass into and out of storage elements (snowmelt and runoff are examples of fluxes). Additionally, we use the term "constitutive relationship" to describe the mathematical equations which relate hydrologic fluxes as a function of parameters and states.

In this refined definition, "process upscaling" represents the highest tier of a hierarchical classification of upscaling methodologies (depicted in Figure 2.1). First, we define all "process upscaling" as the act of deriving a functional relationship for a large-scale constitutive relationship from information about the process and the heterogeneity in the process at the small scale. For illustration, a simple form of upscaling is spatial averaging, in which a distribution of fluxes is represented by a mean value and other local spatial information is discarded. Each new functional relationship can be derived from an analysis of smaller-scale information (the "bottom-up" approach), or it can be derived from an analysis of large-scale information (the "top-down" approach) (Klemeš, 1983). In the bottom-up approach, process upscaling is achieved by explicitly translating some knowledge of small-scale heterogeneity into a useful large-scale constitutive relation. Bottom-up approaches include: Exact averaging (B-1); Effective parameterization (B-2); Distribution-based upscaling (B-3); and Stochastic differential equation approaches (SDEs, B-4). In the top-down approach, process upscaling is achieved by inferring constitutive relations from observed aggregate system behaviour, e.g., by backing out apparent relations from spatially integrated observations such as streamflow. In top-down approaches, a consideration of large-scale information produces insight into how the smaller-scale components can generate observed emergent behaviour. Top-down approaches include Empiricism (T-1) and Inverse modelling (T-2). A third approach, here termed "naïve upscaling," does not technically qualify as a process upscaling methodology in that it uses small-scale constitutive relations at the larger scale with no acknowledgement of important heterogeneity below the computational scale. It is included visually as a component of the hierarchy (Figure 2.1), but this approach is not catalogued in this review. Several of these classifications, including exact averaging, effective parameterization, distribution-based modelling, and inverse modelling roughly map upon those previously proposed by Bierkens et al. (2000). However, whereas Bierkens et al. (2000) considers the classification of upscaling methodologies from a perspective of specifying when certain methods should be applied to a problem based on the information available to the investigator, here we use such classifications to establish a historical accounting of the manner in which hydrological processes have been upscaled. That is, Bierkins et al. (2000) provides a detailed catalogue of various algorithmic approaches, whereas this review provides a more general treatment of the algorithmic approaches with a focus on cataloguing the context in which the upscaling methodologies have been historically applied in hydrology.

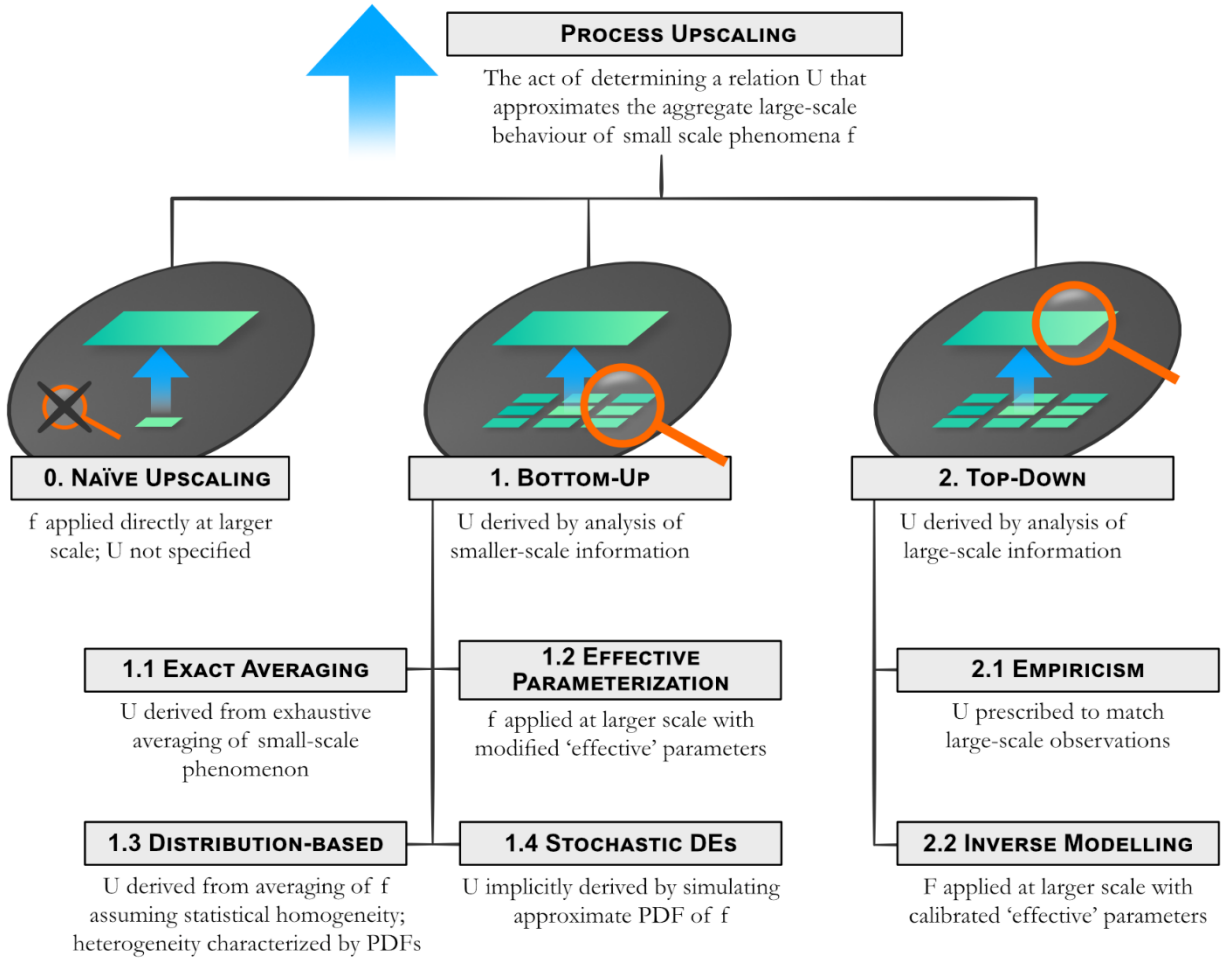

Figure 2.1. Hierarchy of process upscaling.

2.1.2 Bottom-up approaches

Here, we classify bottom-up process upscaling approaches into four distinct categories: Exact averaging; Effective parameterization; Distribution-based techniques; and Stochastic Differential Equations (SDEs).

1. **B-1. Exact averaging**. Exact averaging is the integration of exhaustively known small-scale model fluxes (and/or the states and parameters that inform the constitutive relationship producing such fluxes) over a larger domain. This approach is fundamental to the notion of bottom-up upscaling: the large-scale flux is inherently known if all small-scale fluxes are known. As such, exact averaging can provide a useful benchmark for other upscaling approaches which attempt to simplify local heterogeneity. Exact averaging is a mathematically trivial (but computationally expensive and parameter-intensive) form of upscaling, requiring no assumptions. The large-scale information is simply an aggregation of small-scale information in hand.

2. **B-2. Effective parameterization**. Effective parameterization is the act of applying known small-scale constitutive relations at the large scale with modified ('effective') parameter values that have been altered to allow the scale-inappropriate constitutive relationship to reasonably reproduce the expected large-scale flux. Although small-scale constitutive relationships are not necessarily valid at the large scale, effective parameterization approaches assume that the bulk parameterization of small-scale variability can reasonably replicate the large-scale flux that would be derived by exact averaging. Strategies for determining effective parameters at a given scale of interest vary widely. Effective parameters can also be derived by the inverse modelling top-down upscaling approach (T-2), and, as such, is likely the most common form of upscaling used in practice.
3. **B-3. Distribution-based upscaling**. Distribution-based upscaling is the generation of a large-scale constitutive relationship which uses closed-form probability distributions (PDFs) of small-scale states and/or parameters as inputs. This approach is distinct from exact averaging in that the small-scale fluxes are not exhaustively known: the spatial distribution of small-scale fluxes is not explicitly characterized, but rather prescribed. The form of the PDF is often selected by mathematical convenience and/or justified by field observations. In either case, the assumed PDF is making a strong assumption of the heterogeneity in the domain of interest and so it must be defensible to ensure relevant upscaling results. As in effective parameterization, the resultant large-scale constitutive relationship can simply be the small-scale constitutive relationship, assumed to be valid at the larger scale.
4. **B-4. Stochastic Differential Equations (SDEs)**. SDEs attempt to simulate small-scale hydrologic fluxes and/or states as transient PDFs. As with distribution-based upscaling approaches, SDEs do not account for the spatial distribution of fluxes and presume that the small-scale constitutive laws are known. Once solved, SDEs may be used to produce ensemble mean fluxes and states by integrating the solved PDFs (e.g., Kavvas, 2003). The required simplifications in these methods are primarily in how the (potentially very complicated) covariance structure between states, parameters, and fluxes is characterised; typically, only key lower-order correlations are retained.

2.1.3 Top-down approaches

Two subsets of top-down process upscaling approaches are identified: Empiricism and Inverse modelling.

1. **T-1. Empiricism.** Empiricism refers to the act of directly assigning a large-scale conceptual model that describes the aggregate behaviour of the small-scale process of interest. This large-scale model is usually inferred by experiment and an analysis of observed behaviour. The key feature of this approach is the generation of new differential equations and/or constitutive relations. A popular example is the use of simple non-linear storage reservoirs to represent baseflow or interflow: the constitutive relation determining baseflow from storage is inferred

from the observed recession characteristics of the hydrograph, and represents a simplified conceptualization of a much more complex set of processes in a heterogeneous system.

2. **T-2. Inverse modelling.** Inverse modelling refers to the act of deriving effective parameters (B-2) via calibration for a process description at larger scales. As a calibration exercise, this approach requires numerical modelling results generated from point-scale inputs and process descriptions. As in exact averaging (B-1), the numerical model produces a large-scale flux or state by explicitly resolving small-scale heterogeneity as defined by the model user. In inverse modelling, this aggregate result is used to back-calculate an effective parameterization that optimally reproduces the large-scale flux (e.g., as known from the hydrograph) when substituted into the point-scale relationship. In this case, the upscaling is typically implicit.

2.1.4 Naïve upscaling

Naïve upscaling refers to the use of a small-scale process description and small-scale parameters at larger scales without recognition of scaling effects. This approach implicitly assumes homogeneity of processes and parameters at the computational scale, and sees use, for example, in many land surface schemes which use pedotransfer functions. Given the definition above, this is not an upscaling approach at all, and is included in Sections 2.1 and 2.2 to highlight by contrast what this review considers process upscaling.

2.2 Mathematical Treatment

In this section, a more rigorous mathematical treatment is provided to reinforce the upscaling definitions of Section 2.1. Here we start with a general expression for an instantaneous flux of water [L/T] at an arbitrary location x in 3D space,

$$q(x) = f(x, \theta(x), \theta_x(x), P(x)) \tag{2.1}$$

where $\theta(x)$ is the spatially distributed state variable vector, $P(x)$ is the spatially distributed parameter vector, and $\theta_x(x)$ denotes the spatial derivatives (of any order) of the state variables. The function f is a point-scale relation that is often assumed to be known. It is important to note that these point-scale relations are not necessarily describing detailed point-scale physics; rather, they are an accepted form used to characterize the process at the point-scale. For example, the Green-Ampt characterization of infiltration is a simplification of the more detailed physics of the Richards' equation, but may still act as an appropriate point-scale description.

As an example, the instantaneous flux of water in partially saturated porous media according to Darcy-Buckingham law can be given as:

$$q(x) = -k(\psi(x), k_s(x)) \cdot \left(\frac{\partial \psi(x) + z}{\partial x} \right) \tag{2.2}$$

where $k_s \in P$, $\psi \in \theta$, $\frac{\partial \psi}{\partial x} \in \theta_x, z \in \theta$, and k is a fixed characteristic relationship. The entire functional form of this constitutive law defines the function f. With this general expression for $q(x)$ and a general expression for the total volume of water V of fixed density in a control volume, we obtain via mass balance

$$\frac{d}{dt}V(\theta, P) = -\int_{\Gamma} q(x) \cdot \eta \, d\Gamma \tag{2.3}$$

where η is the outward normal vector along the control volume surface Γ. The estimation of the total water volume (i.e., calculation of V) in a control volume is usually straightforward. However, the potentially complex and non-linear dependency of the function f upon heterogeneous parameters and states means that calculation of the integral on the right for simulating this mass balance can become problematic, particularly when we have insufficient information about the heterogeneity of both parameters and states along Γ.

While this overview can readily be extended to three dimensions, we are going to present the mathematical explanation of the various upscaling approaches in terms of a vertical flux in a two-dimensional plane, as this readily applies to how we often conceptualize infiltration, evaporation, snowmelt, and other hydrological processes on the land surface, and should help to simplify the discussion somewhat. We will also assume for simplicity that all parameters are temporally invariant.

2.2.1 Exact averaging

Spatial upscaling algorithms are essentially strategies for spatial averaging with limited knowledge about subscale states and parameters, and different strategies can be evaluated with relation to exact averaging (B-1) when all microstates are known. For instance, to exactly calculate the mean instantaneous flux over an arbitrary area A, we can integrate the vertical flux q_z over the area:

$$\overline{q_z} = \frac{1}{A}\int_{A} q_z(x, y; \theta(x, y), P(x, y))dA \tag{2.4}$$

This is our upscaling process: converting a heterogeneous distributed flux - $q_z(x, y)$ - into a homogeneous equivalent - $\overline{q_z}$ - over some domain A. This integration incurs no errors on its own: if we explicitly know our detailed states θ and parameters P and the constitutive relation f everywhere in the domain, we can readily perform this integration numerically. However, lack of knowledge regarding the spatial distribution of states, parameters, the flux relation, or all three (not to mention computational costs) often precludes using this operation, and we must make assumptions about these distributions. First, with a few exceptions, it is typical in hydrological models to only store information about mean state variables in an HRU, e.g., mean water content of a soil layer or mean snow water

equivalent on the land surface. Therefore, we need to be able to approximate this integral knowing only the mean state vector $\underline{\theta}$, and (perhaps) limited information m_θ regarding the distribution of states such as the standard deviation or probability that a state exceeds a given value (e.g., percent snow covered area tells us the percentage of land covered with non-zero snow). Thus, the upscaling problem amounts to finding the function U which approximates this (exact) average with imperfect knowledge of $\theta(x,y)$ and $P(x,y)$:

$$\overline{q_z} = \frac{1}{A}\int_A q_z\big(x,y;\theta(x,y),P(x,y)\big)dA \cong U\big(\underline{\theta}, m_\theta, P'\big) \tag{2.5}$$

where P' is a vector of scale-appropriate parameters that may be tightly/explicitly or loosely/implicitly related to the original distributed parameter vector. Bottom-up approaches generate our approximating function U by making simplifying assumptions about the distribution of inputs, whereas top-down approaches estimate U via observation of relations between bulk fluxes and states. A third approach, naïve upscaling, assumes that the small-scale values are identical to the large-scale values.

2.2.2 Naïve upscaling

The most straightforward approach for approximating the large-scale form is just to assume the parameter, spatial field, and relevant constitutive relation are uniform, i.e., $P \cong \underline{P}$, $\theta \cong \underline{\theta}$, and $f(x,y,\theta,P) \cong f(\theta,P)$:

$$\overline{q_z} \cong \frac{1}{A}\int_A q_z\big(x,y;\underline{\theta},\underline{P}\big)dA \cong U\big(\underline{\theta},\underline{P}\big) = q_z\big(\underline{\theta},\underline{P}\big) \tag{2.6}$$

This is the standard assumption used by most land surface schemes parameterized using pedotransfer functions or other relationships derived from point-scale physical observations. The equivalency is now an approximation, reflecting the inevitable error in the assumption that the heterogeneous domain may be treated as homogeneous. Equation 2.6 is only mathematically an equivalency in the case where (1) the domain states and parameters are completely uniform or (2) the constitutive law is a linear function of one state θ_j and P is uniform or (3) the constitutive law is a linear function of one parameter P_j and the state is uniform. These latter two cases are rare. Minor violations of uniformity may lead to inconsequential errors in this equivalency – indeed, much of computational hydrology is predicated on the assumption that this is the case. However, even simple correlations between P and θ or between individual states θ_j can cause problems. For instance, high conductivity soils tend to be dryer than low conductivity soils, leading to a local dependency of moisture state (θ_j) upon local soil parameter (P_j). Direct use of mean properties and mean states for simulating this variably wet landscape – the most rudimentary process upscaling approach - will lead to error. In addition, even the proper averaging operation used to determine $\underline{\theta}$ or $\underline{P}$ may be

questionable – for instance, proper averaging of hydraulic conductivity may entail using either the geometric mean, arithmetic mean, or something in-between (Journel, 1986) depending upon the (often unknown) direction of flow.

2.2.3 Inverse modelling and effective parameterization

In practice, we often calibrate our large-scale models to observations, modifying the mean domain parameters in such a way that may compensate for the scaling issues inherent in the naïve upscaling approach:

$$\overline{q_z} \cong q_z\left(\underline{\theta}, \underline{P}^*\right) \tag{2.7}$$

where $\underline{P}^*$ is the vector of "effective parameters" (B-2). Notably, use of effective parameterization implicitly assumes that the constitutive law f is unchanged at larger scales and that either (1) the state vector is uniform in space or (2) that the constitutive law f is unchanged when using the spatial mean state. In many cases, this assumption is warranted, and has repeatedly been shown to be successful (e.g., Brutsaert and Nieber, 1977; Binley et al., 1989a; Feddes et al., 1993b; Kirchner, 2009; Deng et al., 2021). However, there is ample literature to suggest that this approximation can be very problematic in cases where the constitutive law is non-linearly or otherwise variably dependent upon state (e.g., Zhu and Sun, 2009), includes thresholds (e.g., Craig et al., 2010), or when the parameters and states are uncorrelated (e.g., Zhu and Mohanty, 2002, 2003). Despite this, a great deal of the spatial upscaling literature in hydrology has focused upon general approaches for effective parameterization of models (e.g., Feddes et al., 1993a; Samaniego et al., 2010). Similarly, a great deal of effort has been expended to calibrate models predicated on point-scale physical constitutive relationships. This calibration is a watershed-specific attempt at upscaling via effective parameterization. When effective parameters are informed by large-scale fluxes or states, as in a calibration exercise, the upscaling processes is inherently top-down, and is given the label "inverse modelling" in this review (T-2). This is to distinguish calibration-type exercises from bottom-up approaches which generate effective parameters from a simplification of known or assumed small-scale heterogeneity.

2.2.4 Distribution-based upscaling

The next level of complexity in upscaling approaches explicitly recognizes the sub-HRU-scale heterogeneity in parameters and states, and tries to characterize some information about this heterogeneity, often in the form of a probability distribution of parameters, states, or both (B-3). In general, this uses a different averaging identity:

$$\overline{q_z} \cong \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\theta P}(\theta, P) q_z(\theta, P) d\theta dP \tag{2.8}$$

where $f_{\theta P}(\theta, P)$ is the joint probability distribution of parameter and state in the domain of interest. Note with comparison to Equation 2.4 that the explicit representation of space is removed. The approximate nature of this expression is due to this loss of spatial dependence – it assumes that each differential element is independent, i.e., two individual locations in space with identical state and parameter vectors will respond identically regardless of the state of adjacent parcels. Such an assumption impedes simulation of processes where spatial adjacency is critical such as infiltration run-on (e.g., Corradini et al., 1998), blowing snow (e.g., Essery et al., 1999) or shallow lateral groundwater flow. In general, this joint probability distribution will be rather complex to estimate and is often approximated using one or more critical distributions of state or parameter or both, with the remaining parameters and states treated as uniform. For instance, if we consider that only the heterogeneity of one parameter P_j is important, this expression may be simplified to

$$\overline{q_z} \cong \int_{-\infty}^{\infty} f_{P_j}(P_j) \cdot q_z(\underline{\theta}, \underline{P}_{\neq j}, P_j) dP_j \tag{2.9}$$

i.e., the mean flux may be realized by convolving the flux over the distribution of all parameter values P_j. This approach has been utilized extensively in runoff algorithms predicated on fill-and-spill of soil or wetland stores (e.g., Moore, 1985, 2007; Mekonnen et al., 2014, 2016, 2017; Ahmed et al., 2020). When the expression cannot be simplified, or when the integral of the probability distribution is difficult to solve, Monte Carlo approaches may be employed. Alternatively, when the expression is simple enough to have an analytical solution that is readily inverted, a derived distribution approach may be used. In the derived distribution approach, the probability distribution of fluxes may be directly solved from the probability distribution of parameters or states (Zhao et al., 1980; e.g., Wood et al., 1986; Zhao, 1992)

2.2.5 Stochastic Differential Equations

The final, and generally most complicated means of upscaling hydrologic phenomena is via solution of stochastic differential equations (SDEs, B-4). In these approaches, instead of tracking just the mean fluxes and states within a landscape the evolving PDF of states (e.g., soil moisture) within a computational element is explicitly simulated, subject to simplifying assumptions about the parameter distribution and covariance structure defining the relation between states and parameters. As with distribution-based upscaling, these approaches generally must assume spatial independence and statistical homogeneity in space. In surface water hydrology, the work of M. Levent Kavvas and collaborators has dominated the literature in this area (e.g., Kavvas, 2003; Yoon and Kavvas, 2003; Ohara et al., 2008, 2014). Many of these approaches are predicated on the use of constructs similar to the 'master key conservation equation' of Kavvas (2003), a restatement of the ensemble conservation equations in a computational element. Using this approach, a simple form of the resultant SDE for the case of a single state variable θ (a Fokker-Planck equation) is generated:

$$\frac{\partial f_\theta(\theta,t)}{\partial t} = -\frac{\partial}{\partial \theta}[f_\theta(\theta,t)\cdot G(\theta,P)] + \frac{\partial^2}{\partial \theta^2}[f_\theta(\theta,t)\cdot H(\theta,P)] \quad (2.10)$$

where $G(\theta,P)$ and $H(\theta,P)$ are non-trivial terms which include time integrals of covariance functions (and their derivatives) determined from the form of the point-scale constitutive equation $q_z(\theta,P)$. The solution to this equation, subject to appropriate initial conditions, is the transient probability distribution of changing states within the computational element, $f_\theta(\theta,t)$. Note that here, the constitutive relationships defining the state-flux relation are built into the SDE, and that this equation is an implicit integro-differential nonlinear equation; such equations are not trivial to solve without making simplifying assumptions about the correlations between states and parameters. These necessarily lead to additional approximations beyond that of statistical homogeneity. Alternative mathematical approaches are available to generate SDEs from point-scale constitutive relations, but they still typically yield partial differential equations for PDFs of system states over time.

As most applications of SDEs yield a time-varying PDF of states and not fluxes, it may not necessarily be straightforward to generate the upscaled fluxes $\overline{q_z}$ without (typically numerical) integration of an expression similar to Equation 2.8. While the mathematical challenges associated with this class of methods are significant, the application of SDEs has provided much insight into the scaling behaviour of hydrological processes (Haltas and Kavvas, 2011; He and Ohara, 2019) and has been used within some practical modelling frameworks (Kavvas et al., 2004, 2013).

2.2.6 Empiricism

Empirical approaches (Γ-1) are not conveniently summarized by mathematical representations, as the upscaling function U in Equation 2.5 is directly assumed by an assigned conceptual model. That is, the large-scale process representation is not derived by an aggregation of the small-scale heterogeneity, but from a careful analysis of experimental and/or observed aggregate behaviour that suggests a constitutive relationship. For example, many subsurface runoff upscaling studies connect the small-scale drainage of the shallow subsurface to the recession of the hydrograph, such that this power-law relationship may be used as an upscaled, emergent representation of aggregate small-scale runoff processes. As a general example, high resolution numerical simulations produce large-scale results that may be summarized by a new conceptual model, as opposed to effective parameters assigned to smaller-scale process descriptions. Many popular conceptual hydrological models (e.g., GR4J (Perrin et al., 2003) or HBV (Bergstrom, 1995)) may implicitly treat scale issues via empirical upscaling, even if not intentionally doing so, as development of such models is in part informed by their performance in simulating hydrographs from heterogeneous basins.

2.2.7 Discretization

It is relevant to distinguish discretization approaches from upscaling approaches as defined in this work. Spatial discretization partitions a watershed into M HRUs (or equivalently, gridded response

units (GRUs) or representative elementary watersheds (REWs)), each intended to divide the landscape into a number of hydrologically homogeneous (or sufficiently homogeneous) regions, i.e., supporting a higher fidelity averaging:

$$\overline{q_z} \cong \frac{1}{A}\sum_{k=1}^{M} \bar{\mathrm{q}}_z^k\left(\underline{\theta}_k, \underline{P}_k\right) \cdot A_k \tag{2.11}$$

where the index k denotes belonging in a given quasi-homogeneous element of area A_k. In the extreme case, as M gets very large, we again approach the limiting case of Equation 2.4. This limit may be considered analogous to the Freeze-Harlan blueprint for hydrological modelling (Freeze and Harlan, 1969), the basis of many integrated watershed models which represent the world as a finely discretized grid or mesh at which the constitutive laws, parameters, and instantaneous states are assumed correctly known and scale-appropriate. The success of these ostensibly upscaling-free models is constrained by our ability to parameterize both state and parameters, and in the integrity of our constitutive laws at these scales. In this paper, since no attempt is made to directly address the likely disconnect between scales of simulation and the scales of the physical processes, discretization alone is not considered upscaling, though it helps models to respect the presence of heterogeneity and may be used in conjunction with upscaling at the HRU scale.

2.3 Upscaling hydrologic processes

With definitions and mathematical context in hand, we now have a fixed ontological framework through which we can view how different computational upscaling approaches have been applied in hydrology. Here, we review these applications in the context of four distinct components of the water cycle: runoff (2.3.1), infiltration (2.3.2), evapotranspiration (2.3.3), and snow-related processes (2.3.4). In each section, we highlight key upscaling advances and insights, while also providing insight into missing or under-investigated areas of upscaling research. We argue that several popular hydrological simulation strategies and algorithms are, in fact, upscaling approaches, even though they have typically not been classified as such. A final section (2.3.5) discusses the practical application of upscaling in hydrological modelling platforms.

2.3.1 Runoff magnitude and timing

This section catalogues upscaling methodologies applied to runoff processes, including: using a large-scale power-law relationship to upscale shallow subsurface flow (2.3.1.1); relating the geomorphological structure of the basin to its large-scale runoff characteristics (2.3.1.2); the topographic index approach for upscaling variable contributing area from an irregular terrain (2.3.1.3); approaches for handling landscapes characterized by filling-and-spilling soil or wetland elements (2.3.1.4); and the stochastic treatment of the kinematic wave equation (2.3.1.5).

2.3.1.1 Power-law recession of the hydrograph

There is a rich history of upscaling literature attempting to connect the recessional limb of the hydrograph to subsurface flow processes. These studies utilize the observed behaviour of recession to inform a power-law conceptual model of aggregate subsurface flow at the basin scale. In this top-down empirical approach (T-1), the basin subsurface is conceptualized as a reservoir that reproduces the observed recessional limb of the hydrograph. In the seminal work of this approach, Brutsaert and Nieber (1977) made two foundational contributions. First, they found that the recession behaviour of streamflow can be characterized as independent of time by plotting flow versus its derivative in time, resulting in a power-law representation of recession response:

$$-\frac{dQ}{dt} = c_1 Q^{c_2} \tag{2.12}$$

where Q represents flow [L^3/T], t represents time [T], and c_1 [$1/T$] and c_2 [-] are coefficients of the straight-line fit between flow and its derivative in log-log space. Second, they found that the coefficients of this power-law are controlled by subsurface flow. Specifically, the authors utilized analytical solutions of the Boussinesq equation (i.e., one-dimensional, unconfined, homogeneous, saturated flow, draining into a surface network from a specified initial condition) that take on a power-law form to replicate the observed recession behaviour in six real-world basins. In their analysis, they found that the observed long-term and short-term recession behaviour could be replicated using a single power-law reservoir, and thus were able to use observed recession behaviour to derive effective subsurface parameters (T-2) in each basin.

Thus, Brutsaert and Nieber (1977) made an important upscaling insight: that the observed recession behaviour of a heterogeneous basin can be replicated by a single effective subsurface reservoir. A large number of studies iterated on these ideas by considering various degrees of subsurface complexity, and the efficacy of the approach was underscored by numerous successful comparisons to real-world basin behaviour (e.g., Zecharias and Brutsaert, 1988; Vogel and Kroll, 1992; Troch et al., 1993; Brutsaert and Lopez, 1998; Szilagyi et al., 1998; Huyck et al., 2005). A flat, linear reservoir conceptualization was applied to twenty-three basins (4.4-388 km^2) in Massachusetts (Vogel and Kroll, 1992) and twenty-two basins (1-533 km^2) in Oklahoma (Brutsaert and Lopez, 1998). A linear reservoir with a sloping bed was applied to nineteen basins (~200 km^2) in the Allegheny mountains and Appalachian plateau (Vogel and Kroll, 1992), and a basin in East Flanders, Belgium (144 km^2) (Troch et al., 1993). Huyck et al. (2005) applied a non-linear sloping reservoir with exponentially-varying planar width to the Zwalm catchment, Belgium (114 km^2). Szilagyi et al. (1998) tested the approach of Brutsaert and Nieber (1977) via numerical modelling of synthetic catchments (0.04 and 0.06 km^2) with varying degrees of areal heterogeneity in conductivity and depth, and found that single effective parameters of Brutsaert and Nieber (1977) provided acceptable estimates under the specified heterogeneity. In a general assessment of simple hydrological modelling approaches, Kirchner (2009) reinforced the feasibility of representing the storage-discharge response of certain catchments using reservoir relationships derived from streamflow recession behaviour.

Although the single reservoir approach was successfully applied in real-world basins, the use of a single effective subsurface parametrization, solved under specific boundary and initial conditions enabling analytical solutions, was found to be restrictive (Tallaksen, 1995). Indeed, Rupp et al. (2006) demonstrated that analytical solutions derived from a linearization of the Boussinesq equation are a poor facsimile of the expected non-linear Boussinesq behaviour when the reservoir is sloped, or when vertical heterogeneity in conductivity exists. Bogaart et al. (2013) further challenged the validity of results derived from analytical solutions to the linearized Boussinesq equation.

In their original work, Brutsaert and Nieber (1977) suggested that the superposition of many parallel linear reservoirs can replicate the observed recession behaviour. Moore (1997) independently expanded on this notion through an investigation of five conceptual reservoir models applied to a catchment in British Columbia, Canada (0.17 km^2), concluding that a model of two parallel reservoirs was able to best replicate observed recession behaviour. Clark et al. (2009) demonstrated that the superposition of the drainage response of three hypothetical linear reservoirs, parameterized by an abstracted characteristic timescale representing the speed at which the reservoirs release water, was able to successfully emulate the observed non-linear recession response in the Panola Mountain Research Watershed, Georgia (0.41 km^2). Harman et al. (2009) provided added insight in this basin by investigating the aggregate behaviour of many linear reservoirs, parameterized by hypothetical distributions of characteristic timescales, and demonstrated that many sub-watershed units shedding water at different rates can replicate observed recession behaviour. Similarly, Gao et al. (2017) was able to replicate the behaviour of thirty individual recession events in the Lantang Watershed, China (1080 km^2) using two parallel linear reservoirs. Significantly, these studies demonstrate that the upscaled recession behaviour of a basin can be produced by heterogeneity in sub-watershed response, independent of the subsurface flow mechanics. Thus, the utility of single effective parameters derived from empirical single-reservoir analysis is brought into doubt.

Adding complexity to the physically-based approach, Harman and Sivapalan (2009) investigated the controls of periodic recharge and areal heterogeneity in conductivity on one-dimensional linearized Boussinesq flow through many reservoirs, now characterized as hillslopes and handled via numerical grid cells with unique slope and subsurface parameters. Their analysis demonstrated that the recession power-law coefficients derived under varying recharge conditions are distinct (as opposed to the previously documented studies, which exclusively evaluate the drainage response from some fixed initial condition). This insight is reinforced by studies which suggest that recession coefficients may vary with event and season (Shaw and Riha, 2012; Thomas et al., 2015; Chen and Krajewski, 2016; Karlsen et al., 2019). Ali et al. (2014) investigated the non-linear reservoir response of hillslopes characterized via Richards' equation for steady-state variably-saturated subsurface flow and accounting for vertical heterogeneity in conductivity, although the aggregation of such hillslopes in fifty US catchments (66.6-9062 km^2) was unable to replicate observed recession behaviour. Ranjram and Craig (2021) (Chapter 3 of this thesis) demonstrated the efficacy of reproducing the observed aggregate recession response in the Panola Mountain Research Watershed

by simulating a network of physically-based hillslopes characterized by the non-linear hillslope-storage Boussinesq equation for saturated, unconfined subsurface flow through sloping hillslopes with wedge-shaped areas, both using transient and steady-state recharge inputs. More generally, Snowdon and Craig (2016) used saturated and variably-saturated groundwater flow simulations to derive upscaling relationships capable of predicting mean surface discharge from mean subsurface head across a basin.

Three techniques of evolving complexity have thus been applied in the literature to upscale subsurface flow through the landscape into the power-law recession of the hydrograph at the watershed-scale. In the first approach, the functional power-law form of the analytical solution for flow through a single subsurface reservoir justifies an equivalency to the power-law recession behaviour of the hydrograph (T-1) (Brutsaert and Nieber, 1977; Zecharias and Brutsaert, 1988; Vogel and Kroll, 1992; Troch et al., 1993; Brutsaert and Lopez, 1998; Szilagyi et al., 1998; Huyck et al., 2005; Kirchner, 2009). The subsurface parameters of the analytical solution are then back-calculated to produce power-law coefficients that match the observed recession behaviour (T-2). In the second approach, the physical justification for the power-law behaviour of subsurface reservoirs is set aside, and abstracted reservoirs, explicitly characterized as linear reservoirs with varying timescales of response, are aggregated to reproduce large-scale behaviour (Moore, 1997; Clark et al., 2009; Harman et al., 2009; Gao et al., 2017). These studies demonstrated that the observed power-law behaviour of recessions at the watershed scale can arise from the heterogeneity of sub-watershed units, irrespective of the subsurface flow mechanics used to justify the power-law solution of a single reservoir. In the final approach, the aggregate behaviour of many parallel subsurface reservoirs (now more directly conceptualized as hillslopes) is considered (Harman and Sivapalan, 2009; Ali et al., 2013; Snowdon and Craig, 2016; Ranjram and Craig, 2021). These reservoirs are characterized by physical descriptions of subsurface flow, and each reservoir is assigned unique subsurface parameters, thus enabling a physically-based description of heterogeneity in the basin, as well as accounting for transience in recharge input for the first time. These techniques utilize bottom-up exact averaging (B-1) to develop an aggregate flow response from explicitly simulated sub-basin-scale elements.

2.3.1.2 Geomorphology

The aggregate flow response of a basin has been linked directly to geomorphology. In this approach, the properties of the overland flow network are used to produce a simplified basin-scale description of the distributed, heterogeneous small-scale rainfall-runoff response. The insight of generalizing the runoff response of a basin according to its drainage network is first credited to Rodríguez-Iturbe and Valdés (1979). In their work, an instantaneous unit hydrograph (IUH) was analytically derived by the convolution of two matrices, one characterizing overland flow pathways and the other characterizing the amount of time a parcel of water will spend in different portions of the landscape. Thus, the small-scale production and routing of surface runoff is substituted by these larger-scale matrices, providing a direct physical basis for the often-used unit hydrograph regularly employed by engineering hydrologists. A simplified mathematical description of this relationship is provided by (Rigon et al., 2016) as,

$$IUH(t) = \sum_{\gamma \in \Gamma} p_\gamma \left(p_{\gamma 1} \cdot p_{\gamma \Omega} \right) \tag{2.13}$$

where γ represents a portion of the landscape; Γ represents all portions of the landscape; p_γ represents the probability of rainfall entering γ; and $(p_{\gamma 1} \cdot p_{\gamma \Omega})$ represents the convolution of the pathways between each γ and the time spent in each γ. More directly accounting for geomorphology, Rodríguez-Iturbe and Valdés (1979) derived functions relating the peak and timing of the IUH to five flow network parameters, although contemporary literature finds these relationships unfavourable (Rigon et al., 2016). Nonetheless, this geomorphological IUH approach has proved popular, with subsequent developments authoritatively reviewed by Rigon et al. (2016). These developments include a more flexible definition of the overland routing pathways given advances in digital elevation data (e.g., Rinaldo et al., 1991; Naden, 1992; Snell and Sivapalan, 1994) as well as an allowance for fast and slow runoff routing from channels and hillslopes, respectively (e.g., van der Tak and Bras, 1990).

Combining the insights of the power-law recession approach and the IUH approach, Biswal and Marani (2010) considered the hypothesis that the recession behaviour of a basin is controlled by its geomorphological characteristics, rather than the response of subsurface reservoirs. That is, Biswal and Marani (2010) conceptualized a time-varying Active Drainage Network (ADN) that drains the subsurface at a constant rate and thus controls the recession response. The ADN is characterized by two geomorphological parameters, which are related by a power-law function, as follows,

$$N(t) = c_{g1} G(t)^{c_{g2}} \tag{2.14}$$

where $N(t)$ represents the time-varying number of upstream sources (the first geomorphological parameter); $G(t)$ represents the time-varying length of the network [L] (the second parameter), and c_{g1} [1/L] and c_{g2} [-] are the coefficients of this relationship derived from observation. Using observed data, Biswal and Kumar (2013) demonstrated that this "geomorphic recession curve" is directly related to the power-law form of observed recessions (Equation 2.12) when basins are steep (in terms of topographic relief). A subsequent study of thirty-four steep basins reinforced that the ADN approach is an appropriate representation of observed recession behaviour (Biswal and Kumar, 2014). Biswal and Kumar (2013) and Mutzner et al. (2013) independently relaxed the assumptions of the ADN approach to account for variable drainage of the subsurface across the basin. Biswal and Kumar (2013) also derived a relationship generalizing the c_{g2} coefficient of the geomorphic recession curve as a function of the structure of the overland flow network. Additional studies have generalized the two coefficients of the geomorphic recession curve as functions of initial subsurface storage (Biswal and Marani, 2014) and recharge and past streamflow (Biswal and Kumar, 2014).

In the geomorphological IUH approach, the small-scale release of runoff across different landscape units is generalized as a travel time distribution, representing in a bulk sense how long it takes for a parcel of water to move through the overland flow network before reaching the basin outlet

(Rodríguez-Iturbe and Valdés, 1979; Rigon et al., 2016). The resultant unit hydrograph, produced by a convolution of probability distributions, thus represents a distributed upscaling approach (B-3), wherein the basin-scale runoff is characterized using the small-scale variability of runoff travel times (although this distribution takes on an assumed distribution and so small-scale physics are not utilized). In the ADN approach, the small-scale release of runoff is treated as a subsurface discharge that drains into every channel of the overland flow network, with the numbers and lengths of these channels changing over the course of a rainfall-runoff event (Biswal and Marani, 2010). The basin-scale recession of runoff is then characterized as a function of the numbers and lengths of the channels making up this evolving overland flow network. The resultant geomorphic recession curve thus represents an empirical approach (T-1), wherein a new power-law relationship governing basin-scale runoff has been derived by substituting the aggregate subsurface release across the overland flow network with a description of the time-varying evolution of the ADN. The shared utility of both approaches is the relative ease of acquiring and analyzing overland flow pathways from topographic data. However, in both cases the small-scale physics are substituted for assumed forms, either an assigned distribution or an assigned rate of subsurface runoff release over time.

2.3.1.3 Topographic index approaches

A popular treatment of subsurface runoff was first introduced by Beven and Kirkby (1979), which conceptualized subsurface flow as follows (e.g., Beven et al., 2021)

$$q = T_o \tan(\beta)\, e^{-D/m} \tag{2.15}$$

where q is flow through the subsurface [L²/T], T_o is transmissivity when the subsurface is saturated [L²/T], $\tan(\beta)$ is the hydraulic gradient equivalent to the topographic slope [-], D is the depth of water in the subsurface [L], and m is a parameter controlling the exponential decrease in transmissivity as saturation decreases [L]. This conceptualization is a reconfiguration of Darcy flow in an unconfined saturated medium driven solely by topography, and was used by Beven and Kirkby (1979) to derive the "topographic index," a metric which indicates when a portion of the landscape will be saturated in response to steady rainfall,

$$\lambda = \ln\left(\frac{\alpha}{\tan(\beta)}\right) \tag{2.16}$$

where λ is the topographic index at any location where α, the upslope contributing area [L²], and $tan(\beta)$ are known. The topographic index is a reconceptualization of point-scale topography-driven Darcy flow that can be applied at any scale at which gridded (originally contoured) terrain elevation data is available. In application, the index serves to highlight regions of a basin that will reach similar levels of high or low surface saturation in response to a steady rainfall rate (Ambroise et al., 1996; Beven, 1997). The topographic index is an upscaled representation of point-scale subsurface flow derived by top-down empiricism (T-1): the results of point-scale subsurface flow modelling were

compared against known topographic information to generate a new, larger-scale conceptual model summarizing the effects of rainfall-induced, topography-driven subsurface flow and contributing area over a landscape.

The topographic index is limited by the assumptions of the point-scale subsurface flow model, specifically those of a steady subsurface response to a steady rainfall event; the equivalence of water table gradients to topographic gradients; and the exponential behaviour of subsurface conductivity (Beven, 1997; Beven et al., 2021). However, the subsurface flow mechanics have been relaxed or improved to allow for transience (Beven and Freer, 2001); water table gradients (Quinn et al., 1991); additional conductivity-depth relationships (Ambroise et al., 1996); and heterogeneous transmissivity (Beven, 1986; Woods et al., 1997). The topographic index has seen wide use through its inclusion in the TOPMODEL hydrologic modelling platform and its variants (as discussed in Section 2.3.5).

2.3.1.4 Filling surface depressions/heterogenous soil storage

In some landscapes, runoff to the outlet is intercepted by surface depressions. These systems exhibit unique "fill-and-spill" behaviour, wherein water is stored on the landscape until a storage threshold is reached, triggering the overflow (or spilling) of these depressions and a subsequent hydrologic connection to the outlet (Spence and Woo, 2003). As such, the generation of runoff in these landscapes is dependent on whether the volume of precipitation satisfies a known deficit representing this overflow threshold. Fill-and-spill systems are composed of dozens or hundreds of depressions (Spence and Woo, 2006; Shaw et al., 2012, 2013; McDonnell et al., 2021), making them computationally inefficient modelling problems (Shook and Pomeroy, 2011; Shook et al., 2013). This fill-and-spill conceptual model has also been widely applied to represent runoff from heterogeneously saturated soils (Zhao et al., 1980; Moore, 1985). Translating the heterogeneity of small-scale storage deficits and the associated variability of surface runoff fluxes across a landscape into a representative large-scale flux is fundamentally an upscaling problem.

The algorithmic treatment of runoff generation across a landscape resulting from the exceedance of a storage threshold is first credited to the Xinanjiang model (Zhao et al., 1980; Zhao, 1992). The Xinanjiang model is a conceptual semi-distributed framework of runoff simulation that calculates overland flow based on saturation excess runoff, using a soil moisture storage parameter as the spill threshold. To account for landscape heterogeneity, this model utilizes the nonuniform distribution of soil storage capacity to calculate the proportion of the landscape that will shed water as follows,

$$\frac{f}{F} = 1 - \left(1 - \frac{W'_M}{W'_{MM}}\right)^B \tag{2.17}$$

where f/F is the fractional pervious area [-], W'_M is the variable soil moisture capacity [L], W'_{MM} is the maximum soil moisture capacity [L], and B is a shape factor [-]. In this conceptualization, saturation-excess runoff is generated when precipitation (less evaporation) exceeds the water storage

capacity of the catchment as characterized by the heterogenous soil moisture capacity. This heterogeneity is defined by a probability distribution, and so this technique utilizes a bottom-up distribution-based approach (B-3), and more specifically, a derived distribution approach, in which the PDF of soil heterogeneity is explicitly translated into a PDF of the saturated area. Several extensions have been made to the Xinanjiang approach, including a more realistic handling of soil capacity-saturation relationships in storage units (Jayawardena and Zhou, 2000). The Variable Infiltration Capacity (VIC) model (Wood et al., 1992) is a well-known extension which incorporates infiltration-excess runoff. The VIC model itself has seen several extensions, allowing for more soil layers (Liang et al., 1994); frozen soils (Liang and Xie, 2001, 2003); and subsurface percolation losses (Todini, 1996).

The Xinanjiang approach of analytically characterizing the heterogeneity of storage units was further refined by the Probability Distributed Model (PDM) (Moore, 1985, 2007), which utilizes a PDF of storage capacity to characterize the landscape. In the PDM conceptualization, storage units are spatially heterogenous and the entire system exhibits a transient critical storage capacity, such that the total runoff generated from a rainfall event is determined as,

$$V(t+\Delta t) = \int_{C^*(t)}^{C^*(t+\Delta t)} F(c)\, dc \tag{2.18}$$

where V is the generated runoff [L], t is time [T], Δt is the rainfall interval [T], c^* is a critical storage capacity which evolves over time [L], and $F(c)$ is the cumulative distribution function (CDF) of the storage capacity of soil elements. The PDM approach represents a bottom-up distribution-based upscaling approach (B-3) which converts the unknown exact heterogeneity of storage units into an assumed closed-form distribution. The approach is inherently flexible, in that the form of $F(c)$ and c^* can be modified to suit the unique fill-and-spill properties of different landscapes. For example, the PDM Runoff generation model (PDMROF) is a modification applied to wetland-dominated landscapes, which exhibit a unique behaviour wherein the contributing area of the storage unit is dynamically linked to its level of saturation (Mekonnen et al., 2014). In PDMROF, the heterogeneity of wetland capacity is defined by a Pareto distribution function, and this assumption has been successful in application to basins in the Canadian Prairies (Mekonnen et al., 2014; Ahmed et al., 2020). Mekonnen et al. (2016) conducted an analysis of fine-scale topographic information and demonstrated that an exponential distribution and pareto distribution were both able to replicate the observed distribution of storage capacity, with the exponential distribution being a parsimonious choice due to its single distribution parameter. The exponential distribution was also employed by (Zeng et al., 2020) in their puddle-based unit model (PBU-PDM), an extension of PDM that can account for the dynamic behaviour of surface depressions with variable contributing area by a discretization of the landscape into ponding units and channel-based units with distinct runoff behaviour.

Accurately characterizing the location, volume, contributing area, and dynamic fill behaviour of surface depressions is not trivial (Spence and Woo, 2003; Hayashi et al., 2016). The use of assumed

PDF descriptions of these storage units is thus popular, and once characterized in this way, the fill-and-spill problem is naturally suited to bottom-up distribution-based upscaling approaches. The validity of assumed distributions is demonstrated by experimental application in several basins (Mekonnen et al., 2014, 2016, 2017; Ahmed et al., 2020), although distribution-based approaches are inherently lacking in their inability to account for spatial structure, such as cascading surface depression features (Shook and Pomeroy, 2011; Shaw et al., 2013; Connon et al., 2015) or dominant "gatekeeper" storage units which may be disproportionately responsible for the spill behaviour of a basin (Phillips et al., 2011).

2.3.1.5 Kinematic Wave Equation

In hydrologic applications, the small-scale physics of overland flow are often characterized using the kinematic wave equation (KWE), an approximation of the Saint Venant equations governing the change in overland flow depth in time and space via conservation relationships of momentum and mass (Kavvas and Govindaraju, 1991). The KWE thus represents an explicit accounting of overland flow as follows,

$$\frac{\partial y(x,t)}{\partial t} = -\frac{\partial}{\partial x}[\alpha(x) y(x,t)^m] + q(x,t) \tag{2.19}$$

where y represents flow depth [L]; α represents properties of the overland flow surface (e.g., bed slope and surface roughness) [-]; m is a coefficient dependent on flow state (laminar or turbulent) [-]; and q represents the lateral inflow of water [L/T] (Yoon and Kavvas, 2003).

To apply the KWE at larger (hillslope) scales, (Kavvas and Govindaraju, 1991) mathematically converted the deterministic differential equation into a Fokker-Plank equation characterizing the evolution of the probability distribution of flow depths across a hillslope with stochastic inflows and surface roughness. Solutions to this stochastic differential equation are obtained by mathematical remapping of the SDEs as Partial DEs using cumulant expansion theory, and by specifying the time-dependent mean, covariance, and cross-covariance of the uncertain input variables. A special case, considering deterministic surface roughness, leads to a solution dependent only on the mean and covariance of the stochastic inflows. Kavvas (2003) provided a general framework for the conversion of the KWE into its stochastic form, as well as general solutions for the mean overland flow depth when both inflows and overland flow properties are heterogeneous. In an accompanying work, Yoon and Kavvas (2003) evaluated the simplifications required to utilize the stochastic form of the KWE at the hillslope scale to produce an average overland flow depth. These simplifications assume a deterministic surface roughness and focus on the efficacy of approximating the diffusive term of the SDE by making assumptions about the autocovariance of stochastic rainfall input. The authors successfully replicated the evolution of the PDF of overland flow depth by assuming that overland flow depths reach an immediate steady-state after the application of rainfall, which simplifies the autocovariance function. In a related set of studies, the stochastic KWE equation was reformulated in

application to rills (overland channels across hillslopes) and inter-rill areas (overland surfaces that drain downslope and into rills) (Tayfur and Kavvas, 1994, 1998).

The stochastic overland flow approach is a stochastic differential equations upscaling approach (B-4) that substitutes the known small-scale physics of overland flow with a mean response at larger scales. The approach respects the known distribution of water inputs and surface properties, and is mathematically rigorous, explicitly tracking the evolution of the resultant probability density function of overland flow depths. However, the numerical and analytical solutions of the stochastic overland flow description are not trivial: especially challenging is the need to account for input and parameter covariances which are difficult to assign and often require simplifying assumptions (e.g., Yoon and Kavvas, 2003). In general, applications of stochastic KWE approaches are limited by the necessity of simplifying assumptions, although the underlying mathematics are robust.

Harman et al. (2010) introduced a "subordinated" KWE which utilizes stochastic subordination techniques to transform the KWE into a form suitable for hillslopes with moderate subsurface heterogeneity in conductivity and hence variable subsurface velocities. Zhang et al. (2017) extended this approach to allow for preferential flow paths, although this extension is difficult to parameterize. The subordinated KWE approach is a distribution-based upscaling approach (B-3) wherein the heterogeneity in conductivity is mapped directly to the flow response in steep, straight hillslopes which produce piston-type flow responses under homogeneous or mildly heterogeneous conditions.

2.3.2 Infiltration

Spatial variability in infiltration across a basin occurs primarily due to the spatial heterogeneity of soil hydraulic properties (SHPs) and precipitation (Vereecken et al., 2019). These heterogeneous controls, and a desire to evaluate infiltration at the basin scale, have resulted in a rich history of infiltration upscaling literature. Although there are seminal works handling the spatial heterogeneity of SHPs and their control on the groundwater flow equation (Yeh et al., 1985a, 1985b, 1985c; Binley et al., 1989a, 1989b; Yeh, 1989), the focus here is on the influence of such heterogeneity on infiltration processes specifically (i.e., vertical shallow subsurface flow). Further, the focus of this work is spatial upscaling, and so studies that have investigated the effects of the temporal heterogeneity of rainfall inputs on infiltration are not included (Eagleson, 1978a, 1978b; Cordova and Bras, 1981; Russo and Bresler, 1982). Similarly, studies of subsurface similarity or regionalization are not included, although a substantial discussion of these topics is provided by Vereecken et al. (2019). Although infiltration and runoff processes are inherently linked, infiltration processes in this section are distinct from the runoff processes catalogued in Section 2.3.1 with regards to the level of detail considered in the associated studies. In Section 2.3.1, runoff processes are characterized explicitly from the perspective of hillslope drainage, fill-and-spill behaviour, and variable contributing areas, and infiltration is generalized either as a prescribed boundary condition or a fixed sink. Conversely, in this section

infiltration is explicitly resolved by governing constitutive relationships applied over heterogeneous soils and runoff results are generalized as the rate of water that fails to infiltrate.

The influence of spatial heterogeneity in SHPs on large-scale infiltration has been demonstrated in several studies focused on characterizing the influence of assumed heterogeneity on upscaled infiltration. In a foundational study, Smith and Hebbert (1979) used a Monte Carlo analysis to demonstrate that changing the empirical SHP properties across a basin changed the upscaled infiltration rate derived by an analytical solution to Richards' equation (Smith and Parlange, 1978). Andersson and Shapiro (1983) and Ünlü et al (1990) conducted similar analyses demonstrating the influence of heterogeneity on upscaled behaviour. Although these bottom-up exact averaging approaches illustrate the value of accounting for heterogeneity, they make no attempt to generate upscaling relationships or conclusions that can be used to simplify the heterogeneity of SHPs at the larger scale. However, many studies have attempted to translate the point-scale heterogeneity of SHPs and infiltration fluxes to the basin scale via upscaling techniques. The infiltration upscaling literature is distinguished by the choice of infiltration model (i.e., the full Richards equation or the Green-Ampt approximation (GA)) and variably-saturated SHP model (e.g., Gardner, 1958; Brooks and Corey, 1964; Clapp and Hornberger, 1978; van Genuchten, 1980). Here, the upscaling approaches applied to the infiltration process are distinguished by their use of Richards' equation or the Green-Ampt approximation.

2.3.2.1 Richards' equation

Richards' equation characterizes the vertical movement of water across a porous medium with variable saturation as follows,

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K\left(\frac{\partial h}{\partial z} + 1\right)\right] \tag{2.20}$$

where θ is water content [-], t is time [T], $\partial/\partial z$ is the vertical gradient [1/L], K is the water-content-dependent hydraulic conductivity [L/T], and h is hydraulic head [L].

The efficacy of inverse modeling (T-2) for deriving effective SHP parameters for Richards' infiltration was demonstrated generally by Feddes et al. (1993a, 1993b) in a series of proof-of-concept exercises using remote sensing and finely discretized field data. Kabat et al. (1997) identified inverse modelling as the most promising technique for incorporating transient van Genuchten and Clapp-Hornberger SHP models into large-scale applications of Richards' equation. However, Smith and Diekkrüger (1996), in an analysis of field data, concluded that large-scale SHP functions for unsaturated Richards's flow could not be verified by observation. As such, these insights suggest that although effective parameterizations have been derived in the literature and proven valuable for large scale modelling, they are theoretical conveniences that cannot be directly verified. This limitation notwithstanding, several studies provide useful conclusions about the utility and validity of effective parameters (B-2) for Richards' equation, as follows.

Smith and Diekkrüger (1996) conducted a Monte Carlo analysis of a combined van Genuchten-Brooks-Corey SHP model coupled to an algebraic solution of the transient Richards equation to a ponded boundary condition (Smith and Parlange, 1978). The SHPs were fit to lognormal and normal distributions based on observed data. The authors determined that effective values of saturated conductivity (lognormal), residual soil moisture (normal), and saturated soil moisture (normal) were successful in replicating expected upscaled flux behaviour. However, no effective value existed for the shape parameter controlling the change in conductivity with saturation.

Considering the Brooks-Corey model in isolation, Zhu and Mohanty (2003) determined that its piecewise description of the conductivity-saturation function was inappropriate for effective parameterization approaches. However, all other studies utilizing Brooks-Corey in this review ignore this piecewise discontinuity in conductivity. Zhu and Mohanty (2002) conducted an in-depth analysis of the suitability of upscaled Brooks-Corey and Gardner models for steady Richards' modelling under varying degrees of parameter covariance in a bottom-up effective parameterization analysis considering arithmetic and geometric means. The authors identified the arithmetic mean for saturated conductivity and the geometric mean for pore size distribution (α) as the most suitable averaging methods when these parameters are assumed to follow lognormal distributions. The authors demonstrated that the correlation between these SHPs directly impacted the success of the effective parameterization, and that effective parameters were most successful when these SHPs were perfectly correlated. The correlation between these parameters, now including their use in the van Genuchten SHP model, was further investigated by Zhu and Mohanty (2003) which applied an inverse modelling approach (T-2) and concluded that an increased correlation between conductivity and α lead to larger effective conductivities but had no effect on the effective α parameter (again classifying the heterogeneity in these parameters via lognormal distributions). Liu et al. (2016b) used stochastic theory to define an effective Brooks-Corey saturated conductivity that respected known parameter covariances and mean soil moisture. Due to the complexity of this approach, the authors evaluated whether this effective value could be replaced by a mean value, concluding that this was only tenable when the variance of the lognormally distributed saturated conductivity was low.

Zhu et al. (2007) evaluated a novel power averaging approach to generate effective Gardner and van Genuchten SHPs for steady Richards' infiltration. The authors applied this approach to field data to demonstrate that the SHPs may have characteristic power averaging coefficients, although they were careful to state that this may be a bias of the potentially limited heterogeneity of their data sets. Jana and Mohanty (2012) used bottom-up effective parameterization (specifically, a novel averaging approach that accounts for spatial correlation in its weighting) to demonstrate that effective van Genuchten SHP parameters are dependent on subgrid topography when these parameters are applied to regions larger than several hundred metres. Liu and Bodvarsson (2003) derived upscaled unsaturated conductivity curves for porous media domains with large air entry values. These upscaled curves were produced by exact averaging of small-scale curves, requiring the assumption that the

suction head was spatially uniform at the larger scale. The utility of these curves was demonstrated via numerical modelling under various degrees of heterogeneity.

These studies allow us to make several generalizations about effective SHP parameters applied to Richards' equation. First, effective parameters derived via inverse modelling (T-2) are generally successful, whereas bottom-up averaging approaches (B-2) are appropriate under specific conditions of (1) low variance and (2) high spatial correlation. Further, bottom-up approaches aggregating to regions larger than several hundred metres may be sensitive to the distribution of subgrid topography. Strictly, the Brooks-Corey conductivity must be handled piecewise and so should not be tenable to effective parameterization, although this restriction is only considered in one study. Although each SHP model handles saturation differently, the work of Zhu and Mohanty (2003), which provides general conclusions applicable to the van Genuchten, Brooks-Corey, and Gardner models, suggests that the effective behaviour of the SHPs in each model may be similar.

Several studies have produced upscaled constitutive relationships capable of characterizing the mean Richards infiltration response to prescribed heterogeneity. Yeh et al. (1985a) transformed the point-scale Richards equation for 3D flow into a large-scale equation solving for the mean steady-state infiltration rate given an effective unsaturated conductivity relationship, a distribution-based approach (B-3). Although this conceptualization included lateral flow, the authors demonstrated solutions for one-dimensional infiltration. Chen et al. (1994) transformed the Richards' equation into a stochastic differential equation (B-4) characterizing the ensemble average soil moisture across the domain in response to the moments and covariance of three spatially variable SHPs (saturated conductivity, water content, and matric potential). This upscaled equation accounts for spatial dependence in SHPs, but a simplification of the covariance is required, such that the equation was only successfully able to reproduce ensemble mean results when variance in saturated conductivity was small. Liu et al. (2016b) also derived an upscaled Richards equation using stochastic theory (B-3), but utilized a unique approach that solves for small-scale SHP relationships at a mean soil moisture value and then aggregates and offsets these values by correction terms controlled by the covariance and means of the SHPs. The authors used this model to demonstrate that representative effective SHP parameters were only tenable when variance in saturated conductivity was small.

2.3.2.2 Green-Ampt approximation

The GA approximation simplifies the Richards equation into two analytical solutions, distinguished based on the occurrence of ponding at the surface as follows (e.g., Dingman, 2002),

$$F(t) = \begin{cases} wt, & t \le t_p \\ wt_p + |\psi|(\theta_s - \theta_0) \ln\left(\dfrac{F(t) + |\psi|(\theta_s - \theta_0)}{wt_p + |\psi|(\theta_s - \theta_0)}\right) + k_s\left(t - t_p\right), & t > t_p \end{cases} \tag{2.21}$$

where $F(t)$ is the cumulative infiltration volume [L] at time t [T], w is the rainfall rate [L/T], t_p is the time to ponding [T], ψ is suction head at the wetting front [L], θ_s is porosity [-], θ_0 is initial soil moisture [-], and k_s is the saturated hydraulic conductivity [L/T].

The efficacy of effective SHP parameters (B-2) for the GA infiltration model is dependent on the degree of assumed homogeneity and in the handling of the top boundary condition. In the simplest case, Bresler and Dagan (1983) demonstrated that an effective saturated conductivity for the GA model satisfied the results of the full Richards equation when a steady-state rainfall condition existed and when all other SHPs were homogeneous. Several additional studies reinforced this conclusion by successfully defining an effective saturated conductivity for the GA model under steady rainfall conditions and with no variation in other soil parameters (Corradini et al., 2002; Kim et al., 2005; Ojha et al., 2017). Corradini et al. (2002) further derived a semi-empirical saturated conductivity function that could reproduce the infiltration response to both steady and variable rainfall inputs when the top boundary condition was tied to a run-on process (being the saturation excess transfer of water across the landscape to unsaturated downslope regions which then infiltrate this runoff). Kim et al. (2005) demonstrated that when rainfall rates are large, the spatial variability in conductivity dominates the upscaled response, whereas when rainfall rates are low, it is the rainfall magnitude that dominates. Ojha et al. (2017) formalized the connection between effective saturated conductivity and heterogeneity by generating closed-form analytical expressions relating effective conductivity to the lognormal moments of the distribution of conductivity and the given rainfall rate.

A key assumption in these studies is that saturated conductivity is the only heterogeneous property of the soil – all other SHPs are specified as effectively constant, or an implied assumption is made that the variability of saturated conductivity is much larger than the variation in any other SHPs. In contrast, Craig et al. (2010) analyzed the upscaled GA response to a steady input under heterogeneity in both conductivity as well as initial soil moisture, porosity, and suction head, and concluded that no effective parameterization was able to successfully replicate analytically-derived upscaled behaviour when all parameters were heterogeneous. Thus, effective parameters applied to the GA problem require an essential assumption of homogeneity in SHPs, and the resultant infiltration must be produced in response to a steady input when run-on is not considered. Sivapalan and Wood (1986) and Wood et al. (1986), using the Philips infiltration equation, another approximation of the full Richards infiltration response, reached a similar conclusion that an average conductivity biased the expected mean infiltration rate produced by accounting for the heterogeneity across the basin.

The mathematical tractability of the GA approach has resulted in several closed-form descriptions of upscaled infiltration using distribution-based approaches (B-3). Chen et al. (1994) derived an analytical solution to the upscaled GA problem investigated by Dagan and Bresler (1983) by using a spatially independent distribution of saturated conductivity. Govindaraju et al. (2001) derived an analytical solution informed by a spatially correlated lognormal distribution in saturated conductivity. The authors used this solution to demonstrate that the ensemble mean infiltration rate was independent of the covariance of saturated conductivity, rather than demonstrating the

dependence on the mean and variance. Craig et al. (2010) reconsidered the GA problem to derive a novel approximate upscaled analytical solution in response to spatially independent heterogeneity in saturated conductivity and an aggregate SHP accounting for initial moisture content, porosity, and suction head.

Kim et al. (2005) derived a stochastic differential equation form (B-4) of the GA approximation, accounting for spatial variability in both saturated conductivity and rainfall intensity. However, to produce solutions of the derived stochastic differential equation, the distribution of lognormal conductivity could not include spatial correlation, thus eliminating any accounting of covariance. Ojha et al. (2014) utilized a top-down direct substitution approach (T-1) to develop four dimensionless Green-Ampt solutions in response to a single rainfall event under various rainfall and saturation conditions. These upscaled solutions depend on a spatially independent, lognormally distributed conductivity field, and successfully reproduced the soil moisture observed at an experimental field site.

The literature thus contains many variations of upscaled relationships and effective parameters for infiltration processes utilizing the Green-Ampt approximation. Notably, many studies draw upscaling conclusions by focusing on the variability in saturated conductivity alone. Such an agreement on a dominant aspect of SHP heterogeneity simplifies the upscaling process while still enabling relevant conclusions. Second, accounting for the covariance of SHPs can be considered an unnecessary complication when considering GA solutions (and covariance has been noted as negligible in Richards' equation applications where saturated conductivity exhibits small variance). Critically, a fundamental assumption regarding saturated conductivity, that this parameter exhibits a lognormal distribution of heterogeneity in space, is applied in nearly every GA and Richards' application in this section. The validity of this assumption is ultimately traced to studies in the literature examining field data (e.g., Nielsen et al., 1973), and it is evident that the convenience of this accepted heterogeneity has engendered a variety of upscaling conclusions.

2.3.3 Evapotranspiration

Evapotranspiration (ET) is unique in the library of hydrologic processes in that it is inherently coupled to the physics of plant physiology and respiration. In hydrological modelling, ET is represented by historically accepted empirical equations which translate the complex mechanisms of water vapour exchange from plant stomata into a useful flux (Dingman, 2002). However, the broader literature discussing ET upscaling accounts for more fundamental relationships which consider the inherent coupling between plant respiration and gas exchange fluxes including both water vapour and carbon dioxide (e.g., Jarvis and Mcnaughton, 1986; McNaughton and Spriggs, 1986; McNaughton and Jarvis, 1991; Ding et al., 2014; De Kauwe et al., 2017; Luo et al., 2018). Such studies utilize the leaf as the fundamental scale, and are focused on translating exchange fluxes to the canopy scale or regional (climate modelling) scale. The result of this dichotomy of approaches, with the coupling of leaf physiology and carbon-water exchange processes on the one hand and the hydrological empiricism in

modelling on the other, means that the handling of ET as an upscaled process in hydrological applications is limited.

The Penman-Monteith (PM) equation (Monteith, 1965) is widely used to produce ET fluxes in hydrological applications (e.g., Dingman, 2002; Schymanski and Or, 2017). The PM equation is an empirical approximation of ET (T-1) over a vegetated surface, and is effectively an upscaled representation of the exchange of water vapour across plant stomata, as follows (e.g., Dingman, 2002)

$$ET = \frac{\Delta(K + L) + \rho_a c_a C_{at} e_a^*(1 - W_a)}{\rho_w \lambda_v \left(\Delta + \gamma(1 + C_{at}/C_{can})\right)} \tag{2.22}$$

where Δ is the slope of the saturation vapour pressure-temperature curve , K is net shortwave radiation input, L is net longwave radiation input , ρ_a is the density of air, c_a is the heat capacity of air, C_{at} is the atmospheric conductance as a function of wind speed and height, e_a^* is the saturation vapour pressure, W_a is the relative humidity, ρ_w is the density of water, λ_v is the latent heat of vapourization, γ is the psychrometric constant, and C_{can} is the canopy conductance. In the PM conceptualization, the mechanics of ET are analogized to the behaviour of electric circuits (Monteith, 1965; Dingman, 2002) and the influence of the plant is summarized by the conductance parameter. This conductance can be related to a single leaf ("leaf conductance") or an entire canopy ("canopy conductance") (Dingman, 2002; Schymanski and Or, 2017). However, in practice conductance is empirically derived at the canopy scale (e.g., Lindroth, 1985; Stewart, 1988; Komatsu et al., 2012; Song et al., 2020; Deng et al., 2021). The assumption that Penman-Monteith can be applied at the canopy-scale using an effective canopy conductance is known as the "big leaf" scheme, and this assumption is common in hydrological applications (Dingman, 2002; Schymanski and Or, 2017; Luo et al., 2018). In the plant physiology literature, additional "dual leaf" and "dual source leaf" schemes are used, although these effective parameterization approaches are directly coupled with carbon exchange metrics (Ding et al., 2014).

The lack of ET upscaling procedures in the hydrology literature beyond the specification of an effective canopy conductance (strictly this conductance is not aggregated (B-2) or inferred (T-2), but directly measured) highlights some significant gaps in the handling of ET in hydrological applications. First, as previously discussed, hydrological models do not invoke the fundamental physics of plant vapour exchange but rather empirical approximations of these physics that depend on difficult to measure parameters. Indeed, there are a variety of ET approximations that are best suited to different environment and climate settings and the language of hydrology is not in understanding which of these approximations best replicates the fundamental physics, but rather, which is suitable for the practical question at hand (Allen, 1986; Federer et al., 1996). The interest in hydrology for an ET relationship that suitably estimates water fluxes without a deep concern for the physics of leaf vapour exchange is underlined by Schymanski and Or (2017), which repeated Penman's fundamental experiment with modern methodologies and demonstrated that the Penman relationship, foundational to the application of ET processes in hydrology, is missing critical terms for two-sided

leaf vapour exchange and leaves with distinct stomatal behaviour. Without explicitly accounting for the fundamental physics, upscaling investigations are inherently limited. The empirically derived approximations of ET (T-1) are parameterized with scale-specific effective parameters, and thus the point scale of ET in hydrology is the scale at which the empirical parameters are derived. This contrasts with the plant physiology literature which explicitly treats the leaf as the point scale. In general, ET upscaling efforts in hydrology are made redundant due to the convenient overlap between the canopy scale and the basin scale. Because the canopy scale is large relative to the basin scale, the heterogeneity of canopy types in a basin is appropriately characterized as discrete lumps rather than continuous distributions (i.e., if there are four types of vegetation in a basin, there will be four canopy conductance values, and explicitly accounting for this heterogeneity will not incur a significant computational expense).

The difficulty in assessing the fundamental heterogeneity of ET fluxes at the sub-canopy scale is connected to the scales at which ET processes are observed (further explored in Section 2.4). Eddy-covariance towers, for example, provide an estimate of ET not over a single leaf but rather over a, "spatially homogeneous and structurally uniform vegetative canopy" (Munger et al., 2012), while sap flow techniques provide an estimate at a single tree (e.g., Mackay et al., 2010; Song et al., 2020; Deng et al., 2021). At the other end of the scale spectrum, empirical ET relationships have been derived to estimate ET from satellite imagery via energy balance models (e.g., Raupach and Finnigan, 1995; Allen et al., 2005; Liu et al., 2016a; Singh and Senay, 2016). However, with observations limited to the remote sensing pixel scale, these ET relationships are unable to invoke sub-pixel heterogeneity and thus are more tenable to global circulation models than catchment-scale hydrological models. The lack of available point-scale observation data in determining large scale ET behaviour is reflected in the use of the term "upscaling" in some ET literature to mean the interpolation of point scale data to produce a heterogeneous grid at the large scale (e.g., Xu et al., 2018), rather than the generation of large-scale fluxes.

2.3.4 Snow

Characterizing the melt, blowing transport, and interception of snow is essential for hydrological investigations of basins at high latitude and/or high altitude. Remote sensing tools allow hydrologists to readily classify the snow cover across a basin, but understanding how snow cover evolves requires an upscaled understanding of point-scale processes and related heterogeneity.

2.3.4.1 Snowmelt

Snowmelt occurs when atmospheric and radiative forcings trigger a melt threshold which converts snow cover into liquid water. In hydrological applications, snow cover is represented using a snow-covered area (SCA) metric, indicating the fraction of the landscape covered by the snow pack. The liquid water produced by melt is represented by a snow water equivalence (SWE), which accounts for the lack of one-to-one correlation between snow depth and water depth due to the porosity of the snowpack.

Snow depletion curves (SCDs) characterize the melt relationship between SCA and SWE, and are useful for understanding the impact of landscape heterogeneity on melt. SCDs are dependent on physiographic basin properties (e.g., land cover, topography) and prevailing local weather conditions. As such, these curves are upscaled descriptions of snowmelt developed from field data (T-1). Due to the variability of field conditions, many SCDs have been reported in the literature. Each curve relates the SCA to a model-simulated variable, such as the cumulative production of snowmelt (Luce et al., 1999) or the current state of the SWE in the snowpack (e.g., Dunne and Leopold, 1978). In practice, the conceptual form of the SCD is assumed correct and upscaling insights are generated by explicit consideration of the statistical heterogeneity of SWE, a distribution-based upscaling exercise (B-3). Under homogeneous melting conditions, the SCD is simplified by relating the SCA directly to a shift in the PDF of accumulated SWE at a known initial state (Donald et al., 1995; Pomeroy et al., 1998; Liston, 1999; Luce et al., 1999; Essery and Pomeroy, 2004; Luce and Tarboton, 2004), as follows (Luce et al., 1999)

$$SCA = \int_{SWE_{acc}}^{\infty} f_{SWE}(SWE) \tag{2.23}$$

where SWE_{acc} is the accumulated initial SWE, and f_{swe} is the known PDF of SWE that does not change throughout the melt period. The PDF of initial SWE thus dominates the handling of the melt process. A number of PDFs have been assigned to this initial state, including the lognormal distribution (Donald et al., 1995; Liston, 2004), gamma distribution (Godio and Rege, 2016), and the normal distribution (Marchand and Killingtveit, 2005). Helbig et al. (2015) determined that the gamma distribution appropriately characterized melt for a small area (less than 500 square meters), whereas normal distributions were appropriate when considering a larger area in the same domain. Liston (2004) found success characterizing initial SWE as a lognormal distribution at regional and global scales. When the landscape is entirely covered in snow (SCA equals landscape area), a Gaussian distribution has been successfully applied to the initial distribution of SWE (Dunne and Leopold, 1978; Ferguson, 1984; Buttle and Mcdonnell, 1987; Luce et al., 1999; He et al., 2019). However, melt rates are frequently variable across a landscape due to terrain and vegetation effects, as well as differences in energy state for different depths across the SWE distribution (Donald et al., 1995; DeBeer, 2012). The spatial distributions skew to a non-Gaussian distribution when portions of the underlying ground surface begin to appear (He et al., 2019). To capture non-homogeneous snowmelt conditions, Liston (1999) applied the SCD to subregions of a model domain over which melt can be assumed to be uniform. Many studies have used the approach of dividing a basin into elevation zones; however, if there are very fine scale variations in melt the approach may fail (DeBeer, 2012). For distributed modelling applications, Luce et al. (1999) uniquely characterized the snow distribution in each grid cell using a cell-specific empirical SCD. Despite the fact that this technique appears to be quite effective (Luce and Tarboton, 2004), defining the form of the empirical SCD relationship is not trivial (Ohara et al., 2008).

Snow melt is thus upscaled by top-down empiricism (T-1), wherein the complex physics of melt are converted into a simple emergent large-scale relationship (the SCD) which converts bulk snow cover to bulk snow water equivalent. The form of the SCD is controlled by an initial distribution of SWE across the basin approximated from observed data or fit to an assumed distribution, which are variably successful depending on the size and homogeneity of the domain. In this way, bottom-up distribution-based approaches are also utilized (B-3). The SCD approach has been successfully applied at the sub-basin and distributed modelling grid scale to account for the heterogeneity in the melt process resulting from unique local climate and topography.

An alternative approach for upscaling snowmelt explicitly incorporates the stochasticity affecting snow distribution in a finite area due to uneven snowmelt and redistribution (Ohara et al., 2008; He and Ohara, 2019). This is a stochastic differential equations-based approach (B-4) that requires the parameterization of input PDFs and covariances. The approach was shown to be useful for describing snowpack dynamics in terms of evolving PDFs, but the conclusions were not validated with field data (Ohara et al., 2008; He and Ohara, 2019). He et al. (2019) further suggested that study with field data is necessary for better understanding the evolution of snow spatial distribution at the sub-grid scale. Alternatively, Beaton et al. (2019) evaluated the efficacy of applying a single effective snow depth at numerical grid scales produced via aggregation of point-scale snow depth data, concluding that the upscaled mean value was tenable at reproducing mean SWE except during spring freshet.

2.3.4.2 Blowing snow transport and interception

The physics of point-scale blowing snow transport and interception have been upscaled to larger, sub-basin scales through rigorous field experiments and related empiricism (e.g., Hedstrom and Pomeroy, 1998; Pomeroy and Li, 2000). These processes are grouped together because of the similar approaches use to derive the constitutive relationships and the more restrictive penalties of their empiricism.

The snow pack is constantly evolving due to blowing snow transport, which redistributes snow and so produces a spatially evolving distribution of depth across a basin. Observations of intermittent and spatially variable snow transport have suggested that transport is controlled by small-scale variations in snow cover properties and boundary-layer flow; however, blowing snow models often do not capture small-scale variations due to fixed assumptions of constant wind direction and speed (Pomeroy and Li, 2000). As such, blowing snow is characterized as a top-down empirical upscaled relationship (T-1) describing the aggregate point-scale transport of snow across an area [L/T] with uniform wind direction and speed as follows (e.g., Pomeroy and Li, 2000),

$$transport = p[-\nabla F - E_B] \tag{2.24}$$

where p is the probability that blowing snow occurs across this domain, F is the downwind transport rate [L/T] resulting from empirical descriptions of erosional processes related to wind speed (Pomeroy et al., 1993), and E_B is the sublimation flux [L/T] of the blowing snow (representing a sink on the subsequent volume of redistributed snow). Basin-scale transport may then be realized using an area-weighted sum of fluxes from each portion of the landscape. The probability of occurrence of blowing snow in a given sub-domain has been characterized by a Gaussian CDF of wind speeds, with mean and variance controlled by empirical functions related to temperature and snow age (Li and Pomeroy, 1997; Pomeroy and Li, 2000). This conceptualization of upscaled blowing snow was successfully incorporated into snow mass balance models in the Canadian Prairies and arctic (Pomeroy and Li, 2000); however, the empirical nature of the CDF of wind speed occurrence limits the application of the model to areas where relevant observation data has been collected.

The interception of snow by vegetation canopy results in a critical store of water in cold regions that is difficult to model (Mazzotti et al., 2021). The heterogeneity of the canopy controls the distribution of interception across a basin, and so characterizing the large-scale interception storage is an upscaling exercise in which the interception of snow at the branch scale is translated to the canopy scale. Numerous studies have catalogued observed interception at the branch and stand scale (e.g., Satterlund and Haupt, 1967; Schmidt and Gluns, 1991; Hedstrom and Pomeroy, 1998). Satterlund and Haupt (1967) introduced a top-down empirical upscaled snowfall interception model (T-1) at the canopy scale, informed by observations at the branch and stand scale, as follows (modified from Moeser et al., 2015),

$$I = I_{max}(1 + e^{-kP})^{-1} \tag{2.25}$$

where I is canopy interception [L], I_{max} is the maximum interception storage [L], P is the snowfall measured in an open clearing (with no canopy interception) [L], and k is a proportionality constant controlled by species-specific canopy metrics including Leaf Area Index, canopy coverage, canopy height, and mean forested fetch length (Hedstrom and Pomeroy, 1998) [1/L]. Hedstrom and Pomeroy (1998) extended this model by developing an empirical relationship for I_{max} as a function of snow density, and parameterized the interception relationship for two canopy species (black spruce and jack pine) with good comparison to observed data. However, the authors noted that specific observations of varied canopy species are required to apply the model generally and suggested that further validation and model improvement may be necessary. Moeser et al. (2015) reinforced the difficulty of improving and parameterizing this model by cataloguing the various methods available to collect interception data, including the standard direct method (i.e., cutting and hanging a tree on a tower scale as in Hedstrom and Pomeroy (1998)); time lapse photography; and gamma ray instrumentation. The authors explicitly noted that the expense and difficulty of these observation techniques limited the collection of large-scale canopy data. Moeser et al. (2015) also conducted a statistical analysis of highly resolved LIDAR measurements of a forest canopy and developed regression models to more accurately predict I_{max} by accounting for the greater canopy topography (the proportion of the canopy open to the sky, controlled by stand height and the width of the canopy) and the spatial heterogeneity of the canopy structure. This LIDAR-derived I_{max} produced improved estimates of

interception using Equation 2.25 than the empirical relationship of Hedstrom and Pomeroy (1998). Mazzotti et al. (2021) further demonstrated the benefit of finely resolved remote sensing data in producing a more detailed accounting of canopy and snowpack variability across a domain, although like Moeser et al. (2015) the analysis was based on finely-resolved gridded data as opposed to a new constitutive relationship for canopy interception.

Blowing snow transport and interception are thus similarly characterized by top-down empirical upscaling relationships (T-1) which require detailed site-specific information. However, this dependence on rigorously assigned empirical parameters limits their application in general hydrological modelling. Further, these processes are constrained by the scale at which observations are tenable and hence the scale at which the governing constitutive relationship is defined, which in both cases is not the fundamental point scale. That is, blowing snow transport is characterized at the scale at which a portion of the landscape exhibits uniform wind and transport characteristics, as opposed to the snow particle scale; while interception is characterized at the canopy scale as opposed to the branch scale. As such, although the governing constitutive relationships are themselves upscaled forms of point-scale physics, the relationships are not tenable to further upscaling analysis, in a manner identical to the limitations of upscaling ET processes (Section 2.3.3). As an added complicating factor, the relevant large-scale information parameterizing blowing snow and interception equations is difficult to collect, for example, requiring the identification of a uniform parcel of land and the measurement of transport gradients in blowing snow transport (Hedstrom and Pomeroy, 1998), or more severely, the destructive requirement of cutting a tree and attaching it to a built tower scale for interception measurements (Moeser et al., 2015). These processes are contrasted with melt, which is similarly defined by an empirical relationship, but one that is parameterized by a single distribution, the spatial heterogeneity of initial melt, and hence more generally applicable.

2.3.5 Computational upscaling approaches in modelling tools

Upscaling insights are inherently useful for applied hydrologic modelling: if basin-scale fluxes and states can be determined from small-scale variability without the computational burden of solving a finely discretized system of small-scale inputs (or collecting such input data), valid hydrological conclusions can be created at a fraction of the computational cost. In this section, we highlight two modelling platforms in which the application of upscaling ideas is essential to their novelty: the Watershed Environment Hydrology Model (WEHY; Kavvas et al., 2004) and the mesoscale Hydrological Model (mHM; Samaniego et al., 2010). We single out these two platforms because they focus on a large suite of hydrological processes, and so attempt to comprehensively apply upscaling ideas across a number of hydrologic fluxes. However, upscaling methods for a single hydrologic flux have been incorporated into specific hydrological modelling platforms, and we here include a discussion of two relevant examples: the handling of fill-and-spill runoff in the Xinanjiang model (Zhao et al., 1980; Zhao, 1992) and its extensions VIC and PDMROF; as well as the handling of subsurface runoff in the popular TOPMODEL platform (Beven and Kirkby, 1979).

The Watershed Environmental Hydrology Model (WEHY) (Kavvas et al., 2004) is a physically-based hydrological model built to accommodate spatial heterogeneity through ensemble-averaged conservation equations (B-4) (Gelhar and Axness, 1983; Kavvas and Karakas, 1996; Kavvas, 2003). The hydrologic module includes stochastic differential equation-based upscaled representations of: vertical unsaturated flow (Section 2.3.2.2), subsurface flow, overland flow (Section 2.3.1.5), and snow accumulation (Section 2.3.3.1). Watersheds are discretized into model computational units (MCUs), which are rectangular hillslopes or first-order watersheds containing interrill areas and rills. Delineation of MCUs and their estimated geomorphologic and soil statistical moment parameters is covered in Chen et al. (2004b). Discharges calculated from MCUs are routed into a stream network, represented using a 1D diffusive wave approximation, and a regional groundwater flow component, represented using an upscaled version of the 2D Boussinesq equation. Additional sediment/contaminant transport processes are included in an environmental module. The inputs to the model are atmospheric forcings (e.g., precipitation, temperature) and the distribution parameters of the point-scale processes, limiting the complexity of the parameter estimation processes (Kavvas et al., 2004). The WEHY model has also been extended to include atmospheric processes through coupling to an atmospheric circulation model (WEHY-HCM; Kavvas et al., 2013). The WEHY and WEHY-CHM models have been deployed on a mountainous, forest-covered watershed in Japan (Chen et al., 2004a), various watersheds in California (e.g., Kure et al., 2013; Ohara et al., 2014; Jang et al., 2017; Trinh et al., 2017), and monsoon-influenced watersheds in Peninsular Malaysia (Amin et al., 2017).

The mesoscale Hydrologic Model (mHM) is a spatially distributed conceptual model that incorporates upscaling through multiscale parameter regionalization (MPR) (Samaniego et al., 2010). Rather than incorporating parameter heterogeneity with non-physical or lumped input parameters, MPR represents a robust and effective procedure for estimating these values from measured data. Linear and non-linear transfer functions map relationships between model parameters and basin data at the measurement scale. These relationships are controlled by global parameters that parsimoniously adjust these transfer functions. Upscaling to the model (meso) scale is achieved through effective parameters (B-2) calculated from the measurement-scale parameters using various upscaling operators (e.g., arithmetic mean, geometric mean, maximum difference). The mHM model, combined with the MPR technique, has been tested on various watersheds in Germany (Samaniego et al., 2010), the U.S. (Kumar et al., 2013), the greater Pan-European region (Samaniego et al., 2017), and numerous other European watersheds (Rakovec et al., 2016a, 2016b). The ability of MPR to transfer global parameters between watersheds across different climates with minimal or no additional calibration (Kumar et al., 2013) is a powerful element of mHM.

The TOPMODEL hydrologic modelling platform is a popular modelling tool (Beven et al., 2021), that utilizes the topographic index (Section 2.3.1.3) in its algorithmic handling of subsurface runoff. The model, first introduced by Beven and Kirkby (1979), provided immediate computational benefit in its ability to convert contoured topographic information from physical maps into useful hydrologic insight, providing an understanding of what portions of the landscape would be similarly

saturated by a rainfall event. A robust contemporary review of TOPMODEL, including its history and its evolution, is available in Beven et al. (2021). TOPMODEL is unique in its flexible definition of the "large scale", which is an arbitrary grid cell size at which the modeller is comfortable assuming homogeneity of the topographic index. TOPMODEL has been successfully applied throughout its history to a large number of diverse basins (e.g., Beven, 1997; Beven et al., 2021).

The Xinanjiang model (Zhao et al., 1980; Zhao, 1992) and its extensions (including VIC, PDM, and others, Section 2.3.1.4) are modelling tools for solving runoff in landscapes with threshold-based, fill-and-spill runoff behaviour. Traditional rainfall-runoff approaches are not applicable in such systems, which require an explicit handling of the disconnect between precipitation inputs and runoff outputs. These models were designed to rapidly handle the heterogeneity of storage components across these landscapes using probability distributions, and so are examples of distribution-based upscaling approaches (B-3). The rapid handling of heterogeneity stands in contrast to resolving each storage unit explicitly in a distributed modelling approach, which has been shown to be computationally impractical (Shook et al., 2013). Pareto or exponential distributions are accepted representations of the distribution of storage capacities throughout the landscape in these models (Mekonnen et al., 2014, 2016; Ahmed et al., 2020; Zeng et al., 2020)

2.4 Challenges and open questions

2.4.1 The limits of observation

The challenges of computational upscaling in hydrology have been directly linked to the scale at which we observe hydrological processes: the inconsistent overlap between the scale of these processes and the scale of convenient observation creates difficulty in our ability to fully characterize hydrological processes at the basin scale (Klemeš, 1983; Blöschl and Sivapalan, 1995). Because hydrological processes can occur at the edges of convenient observation scales, it is difficult to obtain complete large-scale descriptions of hydrological physics. For example, we can observe an infiltration flux over a few square metres using a ring infiltrometer, but it is infeasible to observe a mean infiltration flux over an entire basin using current technology. Small-scale constitutive relationships are thus often limited to the scale of lab or field measurements, where we can observe parameters, states, and fluxes at the same scale, thus providing a spatial boundary to these relationships that enables closure. At catchment or even HRU scales, we may be able to resolve bulk parameters and states to some degree, but apart from streamflow, are nearly always lacking observations of bulk fluxes.

This mismatch between observation scale and process scale is a fundamental limitation. Advances in remote sensing have provided access to hydrological information beyond small-scale observation scales, but converting satellite data to useful hydrological information is not trivial (e.g., Bahrami et al., 2021), and what information is available is often too coarsely resolved to minimize the discrepancy between process and observation scales (i.e., the observation scales are often much larger than the hillslope or catchment scales). The paucity of upscaling algorithms in the evapotranspiration section of this review (Section 2.3.3) provides a valuable example of this mismatch: the physics of ET

are resolved at the leaf scale, and the literature has introduced an upscaled representation utilizing effective parameters to convert these physics to the canopy or basin scale. However, remote sensing of ET-related parameters is resolved at scales much larger than a canopy or a leaf (e.g., Wang and Liang, 2008). Without scale-appropriate observation data, top-down upscaling investigations are inherently limited, and bottom-up approaches lose valuable validation information. Thus, progress in upscaling ET algorithms to the basin scale has been fundamentally limited. These conclusions are also applicable to interception and blowing snow transport algorithms (Section 2.3.4.2), which similarly are represented by empirical approximations of the fundamental physics at convenient observation scales. As a contrasting example, we can consider the great diversity of upscaling approaches in runoff processes (Section 2.3.1), which benefit significantly from the information provided by the basin hydrograph. The hydrograph is an integration of the spatially variable runoff processes across a basin and so represents an upscaled observation that can be obtained at the point scale. This unique configuration allows for the validation of bottom-up upscaling approaches and has informed top-down approaches, enabling robust investigations into the upscaling of runoff processes.

It is important to recognize the role technology has played in providing information beyond convenient observation scales. As the satellite, microscope, and telescope have provided access to new frontiers of information, future hydrologists may have access to tools which enable rapid, simple basin-scale assessment of heterogeneity and fluxes (McCabe et al., 2017). Once the spatial extents of convenient observation scales have been extended, new insights into the connection between sub-basin heterogeneity and basin-scale conclusions can be generated, pushing the science of hydrology onward.

2.4.2 Numerical modelling and point-scale physics

Bottom-up upscaling approaches attempt to derive upscaled process descriptions by aggregating the point-scale physics of a process. The aggregation of point-scale physics is a fundamental aspect of hydrological modelling and is explicit in every distributed model. Advances in computational processing power may thus suggest that distributed modelling should replace the need to produce upscaling relationships, as the point-scale physics can be immediately resolved within a finely discretized distributed model grid cell. Although this approach may be tenable for some applications, short-circuiting the conclusions of a distributed model with a fully resolved upscaling relationship provides both computational advantages as well as a deeper understanding of the aggregate process. By upscaling hydrological processes, we attempt to deal with the fundamental issue of closure (Beven, 2006), to generate descriptions of processes that are valid within identifiable large-scale boundaries. The notion of a Representative Elementary Watershed (Reggiani et al., 1998) arises out of this idea and reflects an understanding that a system is relatively well behaved at a certain scale (the scale at which there is an identifiable boundary to the process and at which the heterogeneity in the basin is ergodic). Even if computational power achieved a potency such that any researcher could comprehensively simulate all processes in any basin, upscaling investigations would still produce necessary contributions to the science of hydrology, moving the science forward and ensuring that

our models produce results that we can understand. As modelling gets cheaper and easier, the intersection between bottom-up and top-down approaches should grow fuzzy: instead of developing conceptual models that justify observed aggregate behaviour (in the top-down approach), the small-scale physics should be directly aggregated to understand what aspects of the basin (measurable at our convenient observation scales) control the aggregate behaviour of a basin. Increased computational faculties should inform upscaling studies, rather than ignore them.

A related question about the utility of upscaling investigations is the relative convenience of calibration in producing effective basin-scale parameters. Model calibration has proven itself to be useful as both a project-specific top-down upscaling approach as well as a traditional and historically accepted tool to ensure our models produce useful results. Calibration is practical, but it should be the objective of hydrologic research to understand what controls the calibrated coefficients and parameters in a basin, succinctly described by Kirchner (2006) as, "getting the right answers for the right reasons." As computational efficiency increases, there is more room to understand what aspects of a basin are responsible for calibrated effective parameters.

It is important to acknowledge that the point-scale process descriptions available in hydrology may be imperfect. Ultimately, point-scale physics reflect an upscaling of molecular-scale physics, and although the desire to close the link between the laws of thermodynamics and basin scale physics is intriguing, it may be beyond practical reason to produce upscaling relationships from the very-small scale to the point scale. As such, the library of extant and traditionally applied point-scale hydrological relationships, often empirically derived as opposed to representing an exact averaging of sub-point-scale information, may be a fundamental lower-bound on the physics of hydrology.

2.4.3 Assumed heterogeneity, dominant heterogeneity

If we accept that we have a reliable understanding of the physics at the point scale, and that these physics can be aggregated to produce an appropriate description at the large scale, we are faced with the difficult task of understanding the distribution of heterogeneity across our scale of interest. Correctly characterizing the heterogeneity across a basin is not trivial. However, repeated observation of a property may indicate that the property exhibits a convenient statistical structure that can be characterized by a known probability distribution function. The catalogue of infiltration studies in this review (Section 2.3.2) demonstrates the utility of a widely accepted statistical distribution in an upscaling investigation. The understanding that hydraulic conductivity follows a lognormal distribution in space means that infiltration studies can specify a reasonable approximation of observed heterogeneity using only a mean and variance, as opposed to the more difficult task of specifying individual values across a basin. In contrast, the lack of accepted statistical distributions for key parameters in the snow literature (Section 2.3.4) is a clear limitation on upscaling efforts. Indeed, efforts in the snow literature to demonstrate and justify the efficacy of gamma, normal, or lognormal distributions of snow depth across a domain provide an example of how the science must first be familiar with convenient statistical distributions of a parameter before any reasonable upscaling efforts

can take place. However, it is important to distinguish between the utility of convenient statistical distributions for upscaling studies, and the utility of these distributions generally. Fogg and Zhang (2016) considered the application of conductivity distributions in stochastic subsurface hydrology and suggested that an over-dependence on these assumed distributions has resulted in analyses that utilize biased estimates of conductivity when the given domains exhibit structure untenable to description by a distribution function. That is, assumed distributions of a parameter may not be fully representative in all cases, and indeed, the structure of heterogeneity rather than the distribution alone may be required to upscale properly. In the context of upscaling studies, experimental control should dictate that an assumed parameter distribution is justified, such that the upscaling conclusions are defensible.

The challenge of characterizing basin heterogeneity is complicated by the challenge of understanding the dominant heterogeneity in a basin. Dominant heterogeneity may be classified into two aspects. First, in the context of the infiltration example above, we note that most infiltration upscaling studies ignore the spatial variability of other soil parameters because they are accepted to be much less variable than hydraulic conductivity. As such, an assumption of homogeneity is defended and these parameters may be assigned a single representative value while still enabling rich upscaling conclusions. Certainly, these studies would be improved by evaluating the variability of all parameters, but such an expectation is impractical given the difficulty of collecting information at the appropriate scale. The first aspect of dominant heterogeneity simplifies upscaling analyses: if a parameter exhibits a much larger degree of heterogeneity *and* influence relative to other parameters, these other parameters may be defensibly classified as homogenous.

The second aspect of dominant heterogeneity involves problems that are significantly controlled by a single element that dictates the upscaled response. For example, in cascading fill-and-spill systems, a single storage element may be responsible for triggering a connection between the landscape and the surface water network (e.g., Phillips et al., 2011). Similarly, in shallow subsurface flow, the existence of a preferential flow pathway may produce a distinct conductivity which does not respect the conductivity of the bulk porous medium (e.g., Nimmo et al., 2021). Failing to characterize such "gatekeeper" elements would result in a dramatic misinterpretation of the behaviour of the landscape. These gatekeeper elements may require detailed observation data that can be physically unrealistic to collect, as is the case with the delineation of macropores in the subsurface (e.g., Rinaldo et al., 2011). However, systems dominated by a few elements may be a boon to upscaling studies, as the identification of a dominant heterogeneity implies that all other elements can be ignored in favour of a detailed characterization of a few elements.

2.5 Conclusion

Computational upscaling approaches allow hydrologists to account for the small-scale heterogeneity of parameters, fluxes, and states at basin scales. Investigations into upscaling approaches provide two critical benefits. First, the derivation of an upscaled constitutive relationship provides a fundamental scientific insight into the emergent behaviour of hydrologic processes at larger scales.

Second, because upscaled relationships provide basin-scale insights without explicitly simulating small-scale heterogeneity, upscaled constitutive relationships can provide computational benefits to hydrologic modelling. In this review, we have generated a formal classification of the methodologies that have been used to generate upscaling insights in hydrology, providing a concise set of definitions (Section 2.1) and a rigorous mathematical treatment of these ideas (Section 2.2). The utility and challenges of these approaches is entrenched in the numerous studies that have generated upscaling insights in the literature for runoff, infiltration, evapotranspiration, and snow-related fluxes, and these studies are comprehensively documented in this review for the first time (Section 2.3). The insights of these studies were then synthesized to produce a distilled summary of the gaps in upscaling approaches and the associated open questions for research into computational upscaling approaches (Section 2.4).

2.6 References

Ahmed MI, Elshorbagy A, Pietroniro A. 2020. Toward Simple Modeling Practices in the Complex Canadian Prairie Watersheds. *Journal of Hydrologic Engineering* 25 (6): 04020024 DOI: 10.1061/(ASCE)HE.1943-5584.0001922

Ali M, Fiori A, Bellotti G. 2013. Analysis of the nonlinear storage–discharge relation for hillslopes through 2D numerical modelling. *Hydrological Processes* 27: 2683–2690 DOI: 10.1002/HYP.9397

Ali M, Ye S, Li H yi, Huang M, Leung LR, Fiori A, Sivapalan M. 2014. Regionalization of subsurface stormflow parameters of hydrologic models: Up-scaling from physically based numerical simulations at hillslope scale. *Journal of Hydrology* 519 (PA): 683–698 DOI: 10.1016/j.jhydrol.2014.07.018

Allen RG. 1986. A Penman for All Seasons. *Journal of Irrigation and Drainage Engineering* 112 (4): 348–368 DOI: 10.1061/(ASCE)0733-9437(1986)112:4(348)

Allen RG, Tasumi M, Morse A, Trezza R. 2005. A landsat-based energy balance and evapotranspiration model in Western US water rights regulation and planning. *Irrigation and Drainage Systems* 19 (3–4): 251–268 DOI: 10.1007/s10795-005-5187-z

Ambroise B, Beven K, Freer J. 1996. Toward a generalization of the TOPMODEL concepts: Topographic indices of hydrological similarity. *Water Resources Research* 32 (7): 2135–2145 DOI: 10.1029/95WR03716

Amin MZM, Shaaban AJ, Ercan A, Ishida K, Kavvas ML, Chen ZQ, Jang S. 2017. Future climate change impact assessment of watershed scale hydrologic processes in Peninsular Malaysia by a regional climate model coupled with a physically-based hydrology model. *Science of The Total Environment* 575: 12–22 DOI: 10.1016/J.SCITOTENV.2016.10.009

Andersson J, Shapiro AM. 1983. Stochastic analysis of one-dimensional steady state unsaturated flow: A Comparison of Monte Carlo and Perturbation Methods. *Water Resources Research* 19 (1): 121–133 DOI: 10.1029/WR019i001p00121

Bahrami A, Goïta K, Magagi R, Davison B, Razavi S, Elshamy M, Princz D. 2021. Data assimilation of satellite-based terrestrial water storage changes into a hydrology land-surface model. *Journal of Hydrology*. DOI: 10.1016/j.jhydrol.2020.125744

Beaton AD, Metcalfe RA, Buttle JM, Franklin SE. 2019. Investigating snowpack across scale in the northern Great Lakes–St. Lawrence forest region of Central Ontario, Canada. *Hydrological Processes* 33 (26): 3310–3329 DOI: 10.1002/hyp.13558

Bergstrom, S. 1995. Computer models of watershed hydrology. Water Resources Publications, Highlands Ranch, Colorado, Ch. The HBV Model, 443-476.

Beven K. 1986. Runoff Production and Flood Frequency in Catchments of Order n: An Alternative Approach. Scale Problems in Hydrology, Springer, Dordrecht; 107–131. DOI: 10.1007/978-94-009-4678-1_6

Beven K. 1997. TOPMODEL: A critique. *Hydrological Processes* 11 (9): 1069–1085 DOI: 10.1002/(SICI)1099-1085(199707)11:9<1069::AID-HYP545>3.0.CO;2-O

Beven K. 2006. Searching for the Holy Grail of scientific hydrology: Qt=(S, R, Δt)A as closure. *Hydrology and Earth System Sciences* 10 (5): 609–618 DOI: 10.5194/hess-10-609-2006

Beven K, Freer J. 2001. A dynamic topmodel. *Hydrological Processes* 15 (10): 1993–2011 DOI: 10.1002/hyp.252

Beven K, Kirkby M, E. Freer J, Lamb R. 2021. A history of TOPMODEL. *Hydrology and Earth System Sciences* 25 (2): 527–549 DOI: 10.5194/hess-25-527-2021

Beven KJ, Kirkby MJ. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin* 24 (1): 43–69 DOI: 10.1080/02626667909491834

Bierkens MFP, Finke PA, Willigen P de. 2000. Upscaling and downscaling methods for environmental research. Kluwer Academic Publishers, Dordrecht

Binley A, Beven K, Elgy J. 1989a. A physically based model of heterogeneous hillslopes: 2. Effective hydraulic conductivities. *Water Resources Research* 25 (6): 1227–1233 DOI: 10.1029/WR025i006p01227

Binley A, Elgy J, Beven K. 1989b. A physically based model of heterogeneous hillslopes: 1. Runoff production. *Water Resources Research* 25 (6): 1219–1226 DOI: 10.1029/WR025i006p01219

Biswal B, Kumar DN. 2013. A general geomorphological recession flow model for river basins. *Water Resources Research* 49 (8): 4900–4906 DOI: 10.1002/wrcr.20379

Biswal B, Kumar DN. 2014. Study of dynamic behaviour of recession curves. *Hydrological Processes* 28: 784–792 DOI: 10.1002/HYP.9604

Biswal B, Marani M. 2010. Geomorphological origin of recession curves. *Geophysical Research Letters* 37 DOI: 10.1029/2010GL045415

Biswal B, Marani M. 2014. 'Universal' recession curves and their geomorphological interpretation. *Advances in Water Resources* 65: 34–42 DOI: 10.1016/J.ADVWATRES.2014.01.004

Blöschl G. 2001. Scaling in hydrology. *Hydrological Processes* 15 (4): 709–711 DOI: 10.1002/hyp.432

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

Bogaart P, Rupp D, Selker J, Velde YVD. 2013. Late-time drainage from a sloping Boussinesq aquifer. *Water Resources Research* 49: 7498–7507 DOI: 10.1002/2013WR013780

Bresler E, Dagan G. 1983. Unsaturated flow in spatially variable fields: 2. Application of water flow models to various fields. *Water Resources Research* 19 (2): 421–428 DOI: 10.1029/WR019i002p00421

Brooks RH, Corey AT. 1964. HYDRAULIC PROPERTIES OF POROUS MEDIA. Fort Collins, Colorado.

Brutsaert W, Lopez JP. 1998. Basin-scale geohydrologic drought flow features of riparian aquifers in the Southern Great Plains. *Water Resources Research* 34: 233–240 DOI: 10.1029/97WR03068

Brutsaert W, Nieber J. 1977. Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resources Research* 13 (3): 637-643 DOI: 10.1029/WR013i003p00637

Buttle JM, Mcdonnell JJ. 1987. Modelling the areal depletion of snowcover in a forested catchment. *Journal of Hydrology* 90: 43–430

Chen B, Krajewski W. 2016. Analysing individual recession events: sensitivity of parameter determination to the calculation procedure. *Hydrological Sciences Journal* 61: 2887–2901 DOI: 10.1080/02626667.2016.1170940

Chen Z, Govindaraju RS, Kavvas ML. 1994. Spatial averaging of unsaturated flow equations under infiltration conditions over areally heterogeneous fields: 1. Development of models. *Water Resources Research* 30 (2): 523–533 DOI: 10.1029/93WR02885

Chen ZQ, Kavvas ML, Fukami K, Yoshitani J, Matsuura T. 2004a. Watershed Environmental Hydrology (WEHY) Model: Model Application. *Journal of Hydrologic Engineering* 9 (6): 480–490 DOI: 10.1061/(ASCE)1084-0699(2004)9:6(480)

Chen ZQ, Kavvas ML, Yoon JY, Dogrul EC, Fukami K, Yoshitani J, Matsuura T. 2004b. Geomorphologic and Soil Hydraulic Parameters for Watershed Environmental Hydrology (WEHY) Model. *Journal of Hydrologic Engineering* 9 (6): 465–479 DOI: 10.1061/(ASCE)1084-0699(2004)9:6(465)

Clapp RB, Hornberger GM. 1978. Empirical equations for some soil hydraulic properties. *Water Resources Research* 14 (4): 601–604 DOI: 10.1029/WR014i004p00601

Clark MP, Rupp DE, Woods RA, Tromp-van Meerveld HJ, Peters NE, Freer JE. 2009. Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale. *Hydrological Processes* 23 (2): 311–319 DOI: 10.1002/hyp.7154

Connon RF, Quinton WL, Craig JR, Hanisch J, Sonnentag O. 2015. The hydrology of interconnected bog complexes in discontinuous permafrost terrains. *Hydrological Processes* 29 (18): 3831–3847 DOI: 10.1002/HYP.10604

Cordova JR, Bras RL. 1981. Physically based probabilistic models of infiltration, soil moisture, and actual evapotranspiration. *Water Resources Research* 17 (1): 93–106 DOI: 10.1029/WR017i001p00093

Corradini C, Govindaraju RS, Morbidelli R. 2002. Simplified modelling of areal average infiltration at the hillslope scale. *Hydrological Processes* 16 (9): 1757–1770 DOI: 10.1002/hyp.394

Corradini C, Morbidelli R, Melone F. 1998. On the interaction between infiltration and Hortonian runoff. *Journal of Hydrology* 204 (1–4): 52–67 DOI: 10.1016/S0022-1694(97)00100-5

Craig JR, Liu G, Soulis ED. 2010. Runoff-infiltration partitioning using an upscaled Green-Ampt solution. *Hydrological Processes* 24 (16): 2328–2334 DOI: 10.1002/hyp.7601

Cushman JH, Bennethum LS, Hu BX. 2002. A primer on upscaling tools for porous media. *Advances in Water Resources* 25 (8–12): 1043–1067 DOI: 10.1016/S0309-1708(02)00047-7

Dagan G, Bresler E. 1983. Unsaturated flow in spatially variable fields: 1. Derivation of models of infiltration and redistribution. *Water Resources Research* 19 (2): 413–420 DOI: 10.1029/WR019i002p00413

DeBeer C. 2012. Simulating Areal Snowcover Depletion and Snowmelt Runoff in Alpine Terrain. University of Saskatchewan. Available at: https://harvest.usask.ca/handle/10388/ETD-2012-04-432 [Accessed 14 September 2021]

Deng J, Yao J, Zheng X, Gao G. 2021. Transpiration and canopy stomatal conductance dynamics of Mongolian pine plantations in semiarid deserts, Northern China. *Agricultural Water Management* 249: 106806 DOI: 10.1016/j.agwat.2021.106806

Ding R, Kang S, Du T, Hao X, Zhang Y. 2014. Scaling Up Stomatal Conductance from Leaf to Canopy Using a Dual-Leaf Model for Estimating Crop Evapotranspiration. *PLoS ONE* 9 (4): e95584 DOI: 10.1371/journal.pone.0095584

Dingman SL. 2002. Physical hydrology. Prentice Hall.

Donald JR, Soulis ED, Kouwen N, Pietroniro A. 1995. A Land Cover-Based Snow Cover Representation for Distributed Hydrologic Models. *Water Resources Research* 31 (4): 995–1009 DOI: 10.1029/94WR02973

Dunne T, Leopold LB. 1978. Water in environmental planning. W. H. Freeman: San Francisco.

Eagleson PS. 1978a. Climate, soil, and vegetation: 5. A derived distribution of storm surface runoff. *Water Resources Research* 14 (5): 741–748 DOI: 10.1029/WR014i005p00741

Eagleson PS. 1978b. Climate, soil, and vegetation: 3. A simplified model of soil moisture movement in the liquid phase. *Water Resources Research* 14 (5): 722–730 DOI: 10.1029/WR014i005p00722

Essery R, Pomeroy J. 2004. Implications of spatial distributions of snow mass and melt rate for snow-cover depletion: Theoretical considerations. *Annals of Glaciology* 38: 261–265 DOI: 10.3189/172756404781815275

Essery R, Li L, Pomeroy J. 1999. A distributed model of blowing snow over complex terrain. *Hydrological Processes* 13 (14–15): 2423–2438 DOI: https://doi.org/10.1002/(SICI)1099-1085(199910)13:14/15<2423::AID-HYP853>3.0.CO;2-U

Feddes RA, Menenti M, Kabat P, Bastiaanssen WGM. 1993a. Is large-scale inverse modelling of unsaturated flow with areal average evaporation and surface soil moisture as estimated from remote sensing feasible? *Journal of Hydrology* 143 (1–2): 125–152 DOI: 10.1016/0022-1694(93)90092-N

Feddes RA, de Rooij GH, van Dam JC, Kabat P, Droogers P, Stricker JNM. 1993b. Estimation of Regional Effective Soil Hydraulic Parameters by Inverse Modeling. Water Flow and Solute Transport in Soils, Springer, Berlin, Heidelberg; 211–231. DOI: 10.1007/978-3-642-77947-3_14

Federer CA, Vörösmarty C, Fekete B. 1996. Intercomparison of Methods for Calculating Potential Evaporation in Regional and Global Water Balance Models. *Water Resources Research* 32 (7): 2315–2321 DOI: 10.1029/96WR00801

Ferguson RI. 1984. Magnitude and modelling of snowmelt runoff in the Cairngorm mountains. *Hydrological Sciences Journal* 29 (1): 49–62 DOI: 10.1080/02626668409490921

Fogg GE, Zhang Y. 2016. Debates—Stochastic subsurface hydrology from theory to practice: A geologic perspective. *Water Resources Research* 52 (12): 9235–9245 DOI: 10.1002/2016WR019699

Freeze RA, Harlan RL. 1969. Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology* 9 (3): 237–258 DOI: 10.1016/0022-1694(69)90020-1

Frippiat CC, Holeyman AE. 2008. A comparative review of upscaling methods for solute transport in heterogeneous porous media. *Journal of Hydrology* 362 (1–2): 150–176 DOI: 10.1016/j.jhydrol.2008.08.015

Gao M, Chen X, Liu J, Zhang Z, Cheng Q. 2017. Using Two Parallel Linear Reservoirs to Express Multiple Relations of Power-Law Recession Curves. *Journal of Hydrologic Engineering* 22: 4017013 DOI: 10.1061/(ASCE)HE.1943-5584.0001518

Gardner WR. 1958. Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Science* **85** (4): 228–232

Gelhar LW, Axness CL. 1983. Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research* 19 (1): 161–180 DOI: 10.1029/WR019i001p00161

van Genuchten MT. 1980. A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Science Society of America Journal* 44 (5): 892–898 DOI: 10.2136/sssaj1980.03615995004400050002x

Godio A, Rege RB. 2016. Analysis of georadar data to estimate the snow depth distribution. *Journal of Applied Geophysics* 129: 92–100 DOI: 10.1016/J.JAPPGEO.2016.03.036

Govindaraju RS, Morbidelli R, Corradini C. 2001. Areal Infiltration Modeling over Soils with Spatially Correlated Hydraulic Conductivities. *Journal of Hydrologic Engineering* 6 (2): 150–158 DOI: 10.1061/(ASCE)1084-0699(2001)6:2(150)

Haltas I, Kavvas ML. 2011. Scale Invariance and Self-Similarity in Hydrologic Processes in Space and Time. *Journal of Hydrologic Engineering* 16 (1): 51–63 DOI: 10.1061/(ASCE)HE.1943-5584.0000289

Harman C, Sivapalan M. 2009. Effects of hydraulic conductivity variability on hillslope-scale shallow subsurface flow response and storage-discharge relations. *Water Resources Research* 45 (1) DOI: 10.1029/2008WR007228

Harman CJ, Sivapalan M, Kumar P. 2009. Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resources Research* **45** (9) DOI: 10.1029/2008WR007392

Harman CJ, Reeves DM, Baeumer B, Sivapalan M. 2010. A subordinated kinematic wave equation for heavy-tailed flow responses from heterogeneous hillslopes. *Journal of Geophysical Research* 115 (1): F00A08 DOI: 10.1029/2009JF001273

Hayashi M, van der Kamp G, Rosenberry DO. 2016. Hydrology of Prairie Wetlands: Understanding the Integrated Surface-Water and Groundwater Processes. *Wetlands* 36 (S2): 237–254 DOI: 10.1007/s13157-016-0797-9

He S, Ohara N. 2019. Modeling Subgrid Variability of Snow Depth Using the Fokker-Planck Equation Approach. *Water Resources Research* 55 (4): 3137–3155 DOI: 10.1029/2017WR022017

He S, Ohara N, Miller SN. 2019. Understanding subgrid variability of snow depth at 1-km scale using Lidar measurements. *Hydrological Processes* 33 (11): 1525–1537 DOI: 10.1002/hyp.13415

Hedstrom NR, Pomeroy JW. 1998. Measurements and modelling of snow interception in the boreal forest. *Hydrological Processes* DOI: 10.1002/(SICI)1099-1085(199808/09)12:10/11

Helbig N, van Herwijnen A, Magnusson J, Jonas T, Helbig N, van Herwijnen A, Magnusson J, Jonas T. 2015. Fractional snow-covered area parameterization over complex topography. *Hydrology and Earth System Sciences 19* (3): 1339–1351 DOI: 10.5194/HESS-19-1339-2015

Huyck AAO, Pauwels VRN, Verhoest NEC. 2005. A base flow separation algorithm based on the linearized Boussinesq equation for complex hillslopes. *Water Resources Research* 41 (8): 1–18 DOI: 10.1029/2004WR003789

Jana RB, Mohanty BP. 2012. On topographic controls of soil hydraulic parameter scaling at hillslope scales. *Water Resources Research* 48 (2) DOI: 10.1029/2011WR011204

Jang S, Kure S, Ohara N, Kavvas ML, Chen ZQ, Carr KJ, Anderson ML. 2017. Application of WEHY-HCM for Modeling Interactive Atmospheric-Hydrologic Processes at Watershed Scale to a Sparsely Gauged Watershed. *Sustainability* 2017, Vol. 9, Page 1554 9 (9): 1554 DOI: 10.3390/SU9091554

Jarvis PG, Mcnaughton KG. 1986. Stomatal Control of Transpiration: Scaling Up from Leaf to Region. *Advances in Ecological Research* 15 (C): 1–49 DOI: 10.1016/S0065-2504(08)60119-1

Jayawardena AW, Zhou MC. 2000. A modified spatial soil moisture storage capacity distribution curve for the Xinanjiang model. *Journal of Hydrology* 227 (1–4): 93–113 DOI: 10.1016/S0022-1694(99)00173-0

Journel AG. 1986. Geostatistics: Models and tools for the earth sciences. *Mathematical Geology* 1986 18:1 18 (1): 119–140 DOI: 10.1007/BF00897658

Kabat P, Hutjes RWA, Feddes RA. 1997. The scaling characteristics of soil parameters: From plot scale heterogeneity to subgrid parameterization. *Journal of Hydrology* 190 (3–4): 363–396 DOI: 10.1016/S0022-1694(96)03134-4

Karlsen RH, Bishop K, Grabs T, Ottosson-Löfvenius M, Laudon H, Seibert J. 2019. The role of landscape properties, storage and evapotranspiration on variability in streamflow recessions in a boreal catchment. *Journal of Hydrology* 570: 315–328 DOI: 10.1016/J.JHYDROL.2018.12.065

De Kauwe MG, Medlyn BE, Knauer J, Williams CA. 2017. Ideas and perspectives: how coupled is the vegetation to the boundary layer? *Biogeosciences* 14: 4435–4453 DOI: 10.5194/bg-14-4435-2017

Kavvas ML, Govindaraju RS. 1991. Stochastic overland flows - Part 1: Physics-based evolutionary probability distributions. *Stochastic Hydrology and Hydraulics* 5 (2): 89–104 DOI: 10.1007/BF01543052

Kavvas ML, Karakas A. 1996. On the stochastic theory of solute transport by unsteady and steady groundwater flow in heterogeneous aquifers. *Journal of Hydrology* 179 (1–4): 321–351 DOI: 10.1016/0022-1694(95)02835-8

Kavvas ML, Chen ZQ, Dogrul C, Yoon JY, Ohara N, Liang L, Aksoy H, Anderson ML, Yoshitani J, Fukami K, et al. 2004. Watershed Environmental Hydrology (WEHY) Model Based on Upscaled Conservation Equations: Hydrologic Module. *Journal of Hydrologic Engineering* 9 (6): 450–464 DOI: 10.1061/(ASCE)1084-0699(2004)9:6(450)

Kavvas ML, Kure S, Chen ZQ, Ohara N, Jang S. 2013. WEHY-HCM for Modeling Interactive Atmospheric-Hydrologic Processes at Watershed Scale. I: Model Description. *Journal of Hydrologic Engineering* 18 (10): 1262–1271 DOI: 10.1061/(ASCE)HE.1943-5584.0000724

Kavvas MLL. 2003. Nonlinear Hydrologic Processes: Conservation Equations for Determining Their Means and Probability Distributions. *Journal of Hydrologic Engineering* 8 (2): 44–53 DOI: 10.1061/(ASCE)1084-0699(2003)8:2(44)

Kim S, Kavvas ML, Yoon J. 2005. Upscaling of Vertical Unsaturated Flow Model under Infiltration Condition. *Journal of Hydrologic Engineering* 10 (2): 151–159 DOI: 10.1061/(ASCE)1084-0699(2005)10:2(151)

Kirchner JW. 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research* 42 (3) DOI: 10.1029/2005WR004362

Kirchner JW. 2009. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research* 45 (2) DOI: 10.1029/2008WR006912

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

Komatsu H, Onozawa Y, Kume T, Tsuruta K, Shinohara Y, Otsuki K. 2012. Canopy conductance for a Moso bamboo (Phyllostachys pubescens) forest in western Japan. *Agricultural and Forest Meteorology* 156: 111–120 DOI: 10.1016/j.agrformet.2012.01.004

Kumar R, Livneh B, Samaniego L. 2013. Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme. *Water Resources Research* 49 (9): 5700–5714 DOI: 10.1002/wrcr.20431

Kure S, Jang S, Ohara N, Kavvas ML, Chen ZQ. 2013. WEHY-HCM for Modeling Interactive Atmospheric-Hydrologic Processes at Watershed Scale. II: Model Application to Ungauged and Sparsely Gauged Watersheds. *Journal of Hydrologic Engineering* 18 (10): 1272–1281 DOI: 10.1061/(ASCE)HE.1943-5584.0000701

Li L, Pomeroy JW. 1997. Probability of occurrence of blowing snow. *Journal of Geophysical Research*: Atmospheres 102 (D18): 21955–21964 DOI: 10.1029/97JD01522

Liang X, Xie Z. 2001. A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models. *Advances in Water Resources* 24 (9–10): 1173–1193 DOI: 10.1016/S0309-1708(01)00032-X

Liang X, Xie Z. 2003. Important factors in land–atmosphere interactions: surface runoff generations and interactions between surface and groundwater. *Global and Planetary Change* 38 (1–2): 101–114 DOI: 10.1016/S0921-8181(03)00012-2

Liang X, Lettenmaier DP, Wood EF, Burges SJ. 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*: Atmospheres 99 (D7): 14415–14428 DOI: 10.1029/94JD00483

Lindroth A. 1985. Canopy Conductance of Coniferous Forests Related to Climate. *Water Resources Research* 21 (3): 297–304 DOI: 10.1029/WR021i003p00297

Liston GE. 1999. Interrelationships among Snow Distribution, Snowmelt, and Snow Cover Depletion: Implications for Atmospheric, Hydrologic, and Ecologic Modeling. *Journal of Applied Meteorology* 38 (10): 1474–1487 DOI: 10.1175/1520-0450(1999)038<1474:IASDSA>2.0.CO;2

Liston GE. 2004. Representing Subgrid Snow Cover Heterogeneities in Regional and Global Models. *Journal of Climate* 17 (6): 1381–1397 DOI: 10.1175/1520 0442(2004)017<1381:RSSCHI>2.0.CO

Liu HH, Bodvarsson GS. 2003. Upscaling of constitutive relations in unsaturated heterogeneous tuff matrix. *Journal of Hydrology* 276 (1–4): 198–209 DOI: 10.1016/S0022-1694(03)00071-4

Liu S, Xu Z, Song L, Zhao Q, Ge Y, Xu T, Ma Y, Zhu Z, Jia Z, Zhang F. 2016a. Upscaling evapotranspiration measurements from multi-site to the satellite pixel scale over heterogeneous land surfaces. *Agricultural and Forest Meteorology* 230–231: 97–113 DOI: 10.1016/j.agrformet.2016.04.008

Liu Z, Zha Y, Yang W, Kuo Y-M, Yang J. 2016b. Large-Scale Modeling of Unsaturated Flow by a Stochastic Perturbation Approach. *Vadose Zone Journal* 15 (3): vzj2015.07.0103 DOI: 10.2136/vzj2015.07.0103

Luce CH, Tarboton DG. 2004. The application of depletion curves for parameterization of subgrid variability of snow. *Hydrological Processes* 18 (8): 1409–1422 DOI: 10.1002/hyp.1420

Luce CH, Tarboton DG, Cooley KR. 1999. Sub-grid parameterization of snow distribution for an energy and mass balance snow cover model. *Hydrological Processes* 13 (12–13): 1921–1933 DOI: https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1921::AID-HYP867>3.0.CO;2-S

Luo X, Chen JM, Liu J, Black TA, Croft H, Staebler R, He L, Arain MA, Chen B, Mo G, et al. 2018. Comparison of Big-Leaf, Two-Big-Leaf, and Two-Leaf Upscaling Schemes for Evapotranspiration Estimation Using Coupled Carbon-Water Modeling. *Journal of Geophysical Research: Biogeosciences* 123 (1): 207–225 DOI: 10.1002/2017JG003978

Mackay DS, Ewers BE, Loranty MM, Kruger EL. 2010. On the representativeness of plot size and location for scaling transpiration from trees to a stand. *Journal of Geophysical Research*: Biogeosciences 115 (G2): n/a-n/a DOI: 10.1029/2009jg001092

Marchand W-D, Killingtveit Å. 2005. Statistical probability distribution of snow depth at the model sub-grid cell spatial scale. *Hydrological Processes* 19 (2): 355–369 DOI: 10.1002/hyp.5543

Mazzotti G, Webster C, Essery R, Jonas T. 2021. Increasing the Physical Representation of Forest-Snow Processes in Coarse-Resolution Models: Lessons Learned From Upscaling Hyper-Resolution Simulations. *Water Resources Research* 57 (5): DOI: 10.1029/2020WR029064

McCabe MF, Rodell M, Alsdorf DE, Miralles DG, Uijlenhoet R, Wagner W, Lucieer A, Houborg R, Verhoest NEC, Franz TE, et al. 2017. The future of Earth observation in hydrology. *Hydrology and Earth System Sciences* 21 (7): 3879–3914 DOI: 10.5194/HESS-21-3879-2017

McDonnell JJ, Spence C, Karran DJ, Meerveld HJ (Ilja) van, Harman CJ. 2021. Fill-and-Spill: A Process Description of Runoff Generation at the Scale of the Beholder. *Water Resources Research* 57 (5): e2020WR027514 DOI: 10.1029/2020WR027514

McNaughton KG, Jarvis PG. 1991. Effects of spatial scale on stomatal control of transpiration. *Agricultural and Forest Meteorology* 54 (2–4): 279–302 DOI: 10.1016/0168-1923(91)90010-N

McNaughton KG, Spriggs TW. 1986. A mixed-layer model for regional evaporation. *Boundary-Layer Meteorology* 34 (3): 243–262 DOI: 10.1007/BF00122381

Mekonnen BA, Mazurek KA, Putz G. 2016. Incorporating landscape depression heterogeneity into the Soil and Water Assessment Tool (SWAT) using a probability distribution. *Hydrological Processes* 30 (13): 2373–2389 DOI: 10.1002/HYP.10800

Mekonnen BA, Mazurek KA, Putz G. 2017. Modeling of nutrient export and effects of management practices in a cold-climate prairie watershed: Assiniboine River watershed, Canada. A*gricultural Water Management* 180: 235–251 DOI: 10.1016/J.AGWAT.2016.06.023

Mekonnen MA, Wheater HS, Ireson AM, Spence C, Davison B, Pietroniro A. 2014. Towards an improved land surface scheme for prairie landscapes. *Journal of Hydrology* 511: 105–116 DOI: 10.1016/J.JHYDROL.2014.01.020

Moeser D, Stähli M, Jonas T. 2015. Improved snow interception modeling using canopy parameters derived from airborne LiDAR data. *Water Resources Research* 51 (7): 5041–5059 DOI: 10.1002/2014WR016724

Monteith JL. 1965. Evaporation and environment. *Symposia of the Society for Experimental Biology* (19): 205–234 Available at: https://repository.rothamsted.ac.uk/item/8v5v7/evaporation-and-environment [Accessed 27 October 2020]

Moore RD. 1997. Storage-outflow modelling of streamflow recessions, with application to a shallow-soil forested catchment. *Journal of Hydrology* 198 (1–4): 260–270 DOI: 10.1016/S0022-1694(96)03287-8

Moore RJ. 1985. The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal* 30 (2): 273–297 DOI: 10.1080/02626668509490989

Moore RJ. 2007. The PDM rainfall-runoff model. *Hydrology and Earth System Sciences* 11 (1): 483–499 DOI: 10.5194/HESS-11-483-2007

Munger JW, Loescher HW, Luo H. 2012. Measurement, Tower, and Site Design Considerations. In Eddy Covariance, Springer Netherlands; 21–58. DOI: 10.1007/978-94-007-2351-1_2

Mutzner R, Bertuzzo E, Tarolli P, Weijs S, Nicotina L, Ceola S, Ceola S, Tomasic N, Rodriguez-Iturbe I, Parlange M, Rinaldo, A. 2013. Geomorphic signatures on brutsaert base flow recession analysis. *Water Resources Research* 49: 5462–5472 DOI: 10.1002/WRCR.20417

Naden PS. 1992. Spatial variability in flood estimation for large catchments: The exploitation of channel network structure. *Hydrological Sciences Journal* 37 (1): 53–71 DOI: 10.1080/02626669209492561

Nielsen DR, Biggar JW, Erh KT. 1973. Spatial variability of field-measured soil-water properties. *Hilgardia* 42 (7): 215–259 DOI: 10.3733/hilg.v42n07p215

Nimmo JR, Perkins KS, Plampin MR, Walvoord MA, Ebel BA, Mirus BB. 2021. Rapid-Response Unsaturated Zone Hydrology: Small-Scale Data, Small-Scale Theory, Big Problems. *Frontiers in Earth Science* 9: 123 DOI: 10.3389/feart.2021.613564

Ohara N, Kavvas ML, Chen ZQ. 2008. Stochastic Upscaling for Snow Accumulation and Melt Processes with PDF Approach. *Journal of Hydrologic Engineering* 13 (12): 1103–1118 DOI: 10.1061/(ASCE)1084-0699(2008)13:12(1103)

Ohara N, Kavvas ML, Chen ZQ, Liang L, Anderson M, Wilcox J, Mink L. 2014. Modelling atmospheric and hydrologic processes for assessment of meadow restoration impact on flow

and sediment in a sparsely gauged California watershed. *Hydrological Processes* 28 (7): 3053–3066 DOI: 10.1002/HYP.9821

Ojha R, Corradini C, Morbidelli R, Govindaraju RS, Ojha R, Corradini C, Morbidelli R, Govindaraju RS. 2017. Effective Saturated Hydraulic Conductivity for Representing Field-Scale Infiltration and Surface Soil Moisture in Heterogeneous Unsaturated Soils Subjected to Rainfall Events. *Water* 9 (2): 134 DOI: 10.3390/w9020134

Ojha R, Morbidelli R, Saltalippi C, Flammini A, Govindaraju RS. 2014. Scaling of surface soil moisture over heterogeneous fields subjected to a single rainfall event. *Journal of Hydrology* 516: 21–36 DOI: 10.1016/j.jhydrol.2014.01.057

Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology* 279 (1-4): 275-289 DOI: 10.1016/S0022-1694(03)00225-7

Phillips RW, Spence C, Pomeroy JW. 2011. Connectivity and runoff dynamics in heterogeneous basins. *Hydrological Processes* 25 (19) DOI: 10.1002/hyp.8123

Pomeroy JW, Li L. 2000. Prairie and arctic areal snow cover mass balance using a blowing snow model. *Journal of Geophysical Research Atmospheres* 105 (D21): 26619–26634 DOI: 10.1029/2000JD900149

Pomeroy JW, Gray DM, Landine PG. 1993. The Prairie Blowing Snow Model: characteristics, validation, operation. *Journal of Hydrology* 144 (1–4): 165–192 DOI: 10.1016/0022-1694(93)90171-5

Pomeroy JW, Gray DM, Shook KR, Toth B, Essery RLH, Pietroniro A, Hedstrom N. 1998. An evaluation of snow accumulation and ablation processes for land surface modelling. *Hydrological Processes* 12 (15): 2339–2367 DOI: https://doi.org/10.1002/(SICI)1099-1085(199812)12:15<2339::AID-HYP800>3.0.CO;2-L

Quinn P, Beven K, Chevallier P, Planchon O. 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes* 5 (1): 59–79 DOI: 10.1002/hyp.3360050106

Rakovec O, Kumar R, Attinger S, Samaniego L. 2016a. Improving the realism of hydrologic model functioning through multivariate parameter estimation. *Water Resources Research* 52 (10): 7779–7792 DOI: https://doi.org/10.1002/2016WR019430

Rakovec O, Kumar R, Mai J, Cuntz M, Thober S, Zink M, Attinger S, Schäfer D, Schrön M, Samaniego L. 2016b. Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins. *Journal of Hydrometeorology* 17 (1): 287–307 DOI: 10.1175/JHM-D-15-0054.1

Ranjram M, Craig JR. 2021. Use of an Efficient Proxy Solution for the Hillslope-Storage Boussinesq Problem in Upscaling of Subsurface Stormflow. *Water Resources Research* 57 (4) DOI: 10.1029/2020WR029105

Raupach MR, Finnigan JJ. 1995. Scale issues in boundary-layer meteorology: Surface energy balances in heterogeneous terrain. *Hydrological Processes* 9 (5–6): 589–612 DOI: 10.1002/hyp.3360090509

Reggiani P, Sivapalan M, Majid Hassanizadeh S. 1998. A unifying framework for watershed thermodynamics: Balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics. *Advances in Water Resources* 22 (4): 367–398 DOI: 10.1016/S0309-1708(98)00012-8

Rigon R, Bancheri M, Formetta G, de Lavenne A. 2016. The geomorphological unit hydrograph from a historical-critical perspective. *Earth Surface Processes and Landforms* 41 (1): 27–37 DOI: 10.1002/esp.3855

Rinaldo A, Beven KJ, Bertuzzo E, Nicotina L, Davies J, Fiori A, Russo D, Botter G. 2011. Catchment travel time distributions and water flow in soils. *Water Resources Research* 47 (7) DOI: 10.1029/2011WR010478

Rinaldo A, Marani A, Rigon R. 1991. Geomorphological dispersion. *Water Resources Research* 27 (4): 513–525 DOI: 10.1029/90WR02501

Rodríguez-Iturbe I, Valdés JB. 1979. The geomorphologic structure of hydrologic response. *Water Resources Research* 15 (6): 1409–1420 DOI: 10.1029/WR015i006p01409

Rubin Y. 2003. Applied Stochastic Hydrogeology. Oxford University Press. DOI: 10.1093/OSO/9780195138047.001.0001

Rupp D, Selker JS. 2006. Information, artifacts, and noise in dQ/dt − Q recession analysis. *Advances in Water Resources* 29: 154–160 DOI: 10.1016/J.ADVWATRES.2005.03.019

Rupp D, Rupp DE, Selker JS. 2006. On the use of the Boussinesq equation for interpreting recession hydrographs from sloping aquifers. *Water Resources Research* 42 DOI: 10.1029/2006WR005080

Russo D, Bresler E. 1982. A univariate versus a multivariate parameter distribution in a stochastic-conceptual analysis of unsaturated flow. *Water Resources Research* 18 (3): 483–488 DOI: 10.1029/WR018i003p00483

Samaniego L, Kumar R, Attinger S. 2010. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research* 46 (5) DOI: 10.1029/2008WR007327

Samaniego L, Kumar R, Thober S, Rakovec O, Zink M, Wanders N, Eisner S, Müller Schmied H, Sutanudjaja E, Warrach-Sagi K, et al. 2017. Toward seamless hydrologic predictions across spatial scales. *Hydrology and Earth System Sciences* 21 (9): 4323–4346 DOI: 10.5194/HESS-21-4323-2017

Satterlund DR, Haupt HF. 1967. Snow Catch by Conifer Crowns. *Water Resources Research* 3 (4): 1035–1039 DOI: 10.1029/WR003i004p01035

Schmidt RA, Gluns DR. 1991. Snowfall interception on branches of three conifer species. *Canadian Journal of Forest Research* 21 (8): 1262–1269 DOI: 10.1139/x91-176

Schymanski SJ, Or D. 2017. Leaf-scale experiments reveal an important omission in the Penman–Monteith equation. *Hydrology and Earth System Sciences* 21 (2): 685–706 DOI: 10.5194/hess-21-685-2017

Shaw DA, Pietroniro A, Martz LW. 2013. Topographic analysis for the prairie pothole region of Western Canada. *Hydrological Processes* 27 (22): n/a-n/a DOI: 10.1002/hyp.9409

Shaw DA, Vanderkamp G, Conly FM, Pietroniro A, Martz L. 2012. The Fill-Spill Hydrology of Prairie Wetland Complexes during Drought and Deluge. *Hydrological Processes* 26 (20): 3147–3156 DOI: 10.1002/hyp.8390

Shaw SB, Riha S. 2012. Examining individual recession events instead of a data cloud: Using a modified interpretation of dQ/dt-Q streamflow recession in glaciated watersheds to better

inform models of low flow. *Journal of Hydrology* 434: 46–54 DOI: 10.1016/J.JHYDROL.2012.02.034

Shook K, Pomeroy JW, Spence C, Boychuk L. 2013. Storage dynamics simulations in prairie wetland hydrology models: evaluation and parameterization. *Hydrological Processes* 27 (13): 1875–1889 DOI: 10.1002/hyp.9867

Shook KR, Pomeroy JW. 2011. Memory effects of depressional storage in Northern Prairie hydrology. *Hydrological Processes* 25 (25): 3890–3898 DOI: 10.1002/hyp.8381

Singh RK, Senay GB. 2016. Comparison of four different energy balance models for estimating EvapoTranspiration in the Midwestern United States. *Water (Switzerland)* 8 (1): 9 DOI: 10.3390/w8010009

Sivapalan M, Wood EF. 1986. Spatial Heterogeneity and Scale in the Infiltration Response of Catchments. In Scale Problems in Hydrology, Gupta VK, Rodriguez-Iturbe I, Wood EF (eds).Springer, Dordrecht; 81–106. DOI: 10.1007/978-94-009-4678-1_5

Smith RE, Diekkrüger B. 1996. Effective Soil Water Characteristics and Ensemble Soil Water Profiles in Heterogeneous Soils. *Water Resources Research* 32 (7): 1993–2002 DOI: 10.1029/96WR01048

Smith RE, Hebbert RHB. 1979. A Monte Carlo Analysis of the hydrologic effects of spatial variability of infiltration. *Water Resources Research* 15 (2): 419–429 DOI: 10.1029/WR015i002p00419

Smith RE, Parlange J-Y. 1978. A parameter-efficient hydrologic infiltration model. *Water Resources Research* 14 (3): 533–538 DOI: 10.1029/WR014i003p00533

Snell JD, Sivapalan M. 1994. On geomorphological dispersion in natural catchments and the geomorphological unit hydrograph. *Water Resources Research* 30 (7): 2311–2323 DOI: 10.1029/94WR00537

Snowdon AP, Craig JR. 2016. Effective groundwater-surface water exchange at watershed scales. *Hydrological Processes* 30 (12): 1849–1861 DOI: 10.1002/hyp.10759

Song L, Zhu J, Zheng X, Wang K, Lü L, Zhang X, Hao G. 2020. Transpiration and canopy conductance dynamics of Pinus sylvestris var. mongolica in its natural range and in an introduced region in the sandy plains of Northern China. *Agricultural and Forest Meteorology* 281: 107830 DOI: 10.1016/j.agrformet.2019.107830

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

Spence C, Woo MK. 2006. Hydrology of subarctic Canadian Shield: heterogeneous headwater basins. *Journal of Hydrology* 317 (1–2): 138–154 DOI: 10.1016/J.JHYDROL.2005.05.014

Stewart JB. 1988. Modelling surface conductance of pine forest. *Agricultural and Forest Meteorology* 43 (1): 19–35 DOI: 10.1016/0168-1923(88)90003-2

Szilagyi J, Parlange MB, Albertson JD. 1998. Recession flow analysis for aquifer parameter determination. *Water Resources Research* 34 (7): 1851–1857 DOI: 10.1029/98WR01009

van der Tak LD, Bras RL. 1990. Incorporating hillslope effects into the geomorphologic instantaneous unit hydrograph. *Water Resources Research* 26 (10): 2393–2400 DOI: 10.1029/WR026i010p02393

Tallaksen LM. 1995. A review of baseflow recession analysis. *Journal of Hydrology* 165 (1–4): 349–370 DOI: 10.1016/0022-1694(94)02540-R

Tayfur G, Kavvas ML. 1994. Spatially Averaged Conservation Equations for Interacting Rill-Interrill Area Overland Flows. *Journal of Hydraulic Engineering* 120 (12): 1426–1448 DOI: 10.1061/(ASCE)0733-9429(1994)120:12(1426)

Tayfur G, Kavvas ML. 1998. Areally-averaged overland flow equations at hillslope scale. *Hydrological Sciences* 43 (3): 43 DOI: 10.1080/02626669809492132

Thomas BF, Vogel R, Famiglietti J. 2015. Objective hydrograph baseflow recession analysis. *Journal of Hydrology* 525: 102–112 DOI: 10.1016/J.JHYDROL.2015.03.028

Todini E. 1996. The ARNO rainfall—runoff model. *Journal of Hydrology* 175 (1–4): 339–382 DOI: 10.1016/S0022-1694(96)80016-3

Trinh T, Ishida K, Kavvas ML, Ercan A, Carr K. 2017. Assessment of 21st century drought conditions at Shasta Dam based on dynamically projected water supply conditions by a regional climate model coupled with a physically-based hydrology model. *Science of The Total Environment* 586: 197–205 DOI: 10.1016/J.SCITOTENV.2017.01.202

Troch PA, De Troch FP, Brutsaert W. 1993. Effective water table depth to describe initial conditions prior to storm rainfall in humid regions. *Water Resources Research* 29 (2): 427–434 DOI: 10.1029/92WR02087

Ünlü K, Nielsen DR, Biggar JW. 1990. Stochastic analysis of unsaturated flow: One-dimensional Monte Carlo simulations and comparisons with spectral perturbation analysis and field observations. *Water Resources Research* 26 (9): 2207–2218 DOI: 10.1029/WR026i009p02207

Vereecken H, Pachepsky Y, Bogena H, Montzka C. 2019. Upscaling Issues in Ecohydrological Observations. In Observation and Measurement of Ecohydrological Processes, 435–454. DOI: 10.1007/978-3-662-48297-1_14

Vogel RM, Kroll CN. 1992. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resources Research* 28 (9): 2451–2458 DOI: 10.1029/92WR01007

Wang K, Liang S. 2008. An improved method for estimating global evapotranspiration based on satellite determination of surface net radiation, vegetation index, temperature, and soil moisture. *Journal of Hydrometeorology* 9 (4): 712–727 DOI: 10.1175/2007JHM911.1

Wood EF, Lettenmaier DP, Zartarian VG. 1992. A land-surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research*: Atmospheres 97 (D3): 2717–2728 DOI: 10.1029/91JD01786

Wood EF, Sivapalan M, Beven KJ. 1986. Scale effects in infiltration and runoff production. *Proceedings of the Budapest Symposium, July 1986:* 375-387

Woods RA, Sivapalan M, Robinson JS. 1997. Modeling the spatial variability of subsurface runoff using a topographic index. *Water Resources Research* 33 (5): 1061–1073 DOI: 10.1029/97WR00232

Xu T, Guo Z, Liu S, He X, Meng Y, Xu Z, Xia Y, Xiao J, Zhang Y, Ma Y, et al. 2018. Evaluating Different Machine Learning Methods for Upscaling Evapotranspiration from Flux Towers to the Regional Scale. *Journal of Geophysical Research: Atmospheres* 123 (16): 8674–8690 DOI: 10.1029/2018JD028447

Yang X, Sun H, Yang Y, Liu Y, Li X. 2021. Recent progress in multi-scale modeling and simulation of flow and solute transport in porous media. *Wiley Interdisciplinary Reviews: Water* 8 (6): e1561 DOI: 10.1002/WAT2.1561

Yeh J. 1989. One-Dimensional Steady State Infiltration in Heterogeneous Soils. *Water Resources Research* 25 (10): 2149-2158 DOI: 10.1029/WR025i010p02149

Yeh T-CJ, Gelhar LW, Gutjahr AL. 1985a. Stochastic Analysis of Unsaturated Flow in Heterogeneous Soils: 1. Statistically Isotropic Media. *Water Resources Research* 21 (4): 447–456 DOI: 10.1029/WR021i004p00447

Yeh T-CJ, Gelhar LW, Gutjahr AL. 1985b. Stochastic Analysis of Unsaturated Flow in Heterogeneous Soils: 2. Statistically Anisotropic Media With Variable α. *Water Resources Research* 21 (4): 457–464 DOI: 10.1029/WR021i004p00457

Yeh T-CJ, Gelhar LW, Gutjahr AL. 1985c. Stochastic Analysis of Unsaturated Flow in Heterogeneous Soils: 3. Observations and Applications. *Water Resources Research* 21 (4): 465–471 DOI: 10.1029/WR021i004p00465

Yoon J, Kavvas ML. 2003. Probabilistic Solution to Stochastic Overland Flow Equation. *Journal of Hydrologic Engineering* 8 (2): 54–63 DOI: 10.1061/(ASCE)1084-0699(2003)8:2(54)

Zecharias YB, Brutsaert W. 1988. Recession characteristics of groundwater outflow and base flow from mountainous watersheds. *Water Resources Research* 24 (10): 1651–1658 DOI: 10.1029/WR024i010p01651

Zehe E, Lee H, Sivapalan M. 2006. Dynamical process upscaling for deriving catchment scale state variables and constitutive relations for meso-scale process models. *Hydrology and Earth System Sciences* 10 (6): 981–996 DOI: 10.5194/hess-10-981-2006

Zeng L, Shao J, Chu X. 2020. Improved hydrologic modeling for depression-dominated areas. *Journal of Hydrology* 590: 125269 DOI: 10.1016/J.JHYDROL.2020.125269

Zhang Y, Baeumer B, Chen L, Reeves DM, Sun H. 2017. A fully subordinated linear flow model for hillslope subsurface stormflow. *Water Resources Research* 53 (4): 3491–3504 DOI: 10.1002/2016WR020192

Zhao R-J. 1992. The Xinanjiang model applied in China. *Journal of Hydrology* 135 (1–4): 371–381 DOI: 10.1016/0022-1694(92)90096-E

Zhao RJ, Zhang YL, Fang LR, Liu XR, Zhang QS. 1980. The Xinanjiang model. In Hydrological Forecasting Symposium: Proceedings of the Oxford Symposium, 15-18 April, 1980. International Association of Hydrological Sciences: 351–356.

Zhu J, Mohanty BP. 2002. Upscaling of soil hydraulic properties for steady state evaporation and infiltration. *Water Resources Research* 38 (9): 17-1-17–13 DOI: 10.1029/2001wr000704

Zhu J, Mohanty BP. 2003. Effective hydraulic parameters for steady state vertical flow in heterogeneous soils. *Water Resources Research* 39 (8) DOI: 10.1029/2002WR001831

Zhu J, Sun D. 2009. Effective Soil Hydraulic Parameters for Transient Flows in Heterogeneous Soils. *Vadose Zone Journal* 8 (2): 301–309 DOI: 10.2136/vzj2008.0004

Zhu J, Young MH, van Genuchten MT. 2007. Upscaling Schemes and Relationships for the Gardner and van Genuchten Hydraulic Functions for Heterogeneous Soils. *Vadose Zone Journal* 6 (1): 186–195 DOI: 10.2136/vzj2006.0041

Chapter 3

Use of an efficient proxy solution for the hillslope-storage Boussinesq problem in upscaling of subsurface stormflow

3.1 Introduction

It is common in watershed modelling to characterize subsurface storage as a reservoir in which the release of water from the subsurface is determined by a simple, minimally parameterized function such as a power law function (e.g., Brutsaert & Nieber, 1977; Kirchner, 2009). The practical efficacy of this approach for simulating baseflow and subsurface stormflow has been demonstrated repeatedly (e.g., Ali et al., 2014; Harman et al., 2009; Kirchner, 2009); the power law relationship has also been observed via analysis of hydrograph recession curves (e.g., Patnaik et al., 2018; Ye et al., 2014). In examining the power-law form of this relationship, it is natural to ask: what enables the simplification of complex groundwater mechanics into a reservoir formulation with only one dependent variable and two coefficients? Previous studies have demonstrated that breaking down the basin into individual hillslopes and evaluating the aggregate behaviour of subsurface flow mechanics at those hillslopes is a valuable place to start (Ali et al., 2014; Harman et al., 2009).

A common choice for describing subsurface flow at the hillslope scale is the Boussinesq equation, which describes one-dimensional subsurface flow through a cross-section. Troch et al. (2003), following the work of Fan and Bras (1998) and Childs (1971), developed the Hillslope-Storage Boussinesq Equation (hsB), a modification of the Boussinesq formulation to describe the saturated, one-dimensional flow through hillslopes with known planar shape characteristics, captured through a hillslope width function describing the change in width as a function of distance from the downslope stream boundary. Thus, the unique planar shape of hillslopes in a watershed can be respected without resorting to simulating two-dimensional flow. The hsB has been linearized and solved analytically for specific cases of hillslope shape in which the width profile varies according to an exponential function (Dralle et al., 2014; Troch et al., 2003, 2004). The hsB can also be solved numerically to satisfy any hillslope shape (e.g., Hazenberg et al., 2015). The hsB has been a valuable tool for exploring questions of hillslope subsurface flow. As such there have been several developments and applications of the hsB since Troch et al. (2003). Matonse and Kroll (2009) applied the non-linear hsB (now with a variable conductivity as a function of saturation) to investigate the recession behaviour of the Maimai catchment represented using 1, 3, and 10 hillslopes. Carrillo et al. (2011) built a hydrologic model around the hsB and applied this model in a Prediction in Ungauged Basins exercise evaluating hydrologic similarity between 12 MOPEX catchments. Troch et al. (2013) applied the model of Carrillo et al. (2011) to investigate of the controls of landscape and climate on long-term water balances in these basins. Broda et al. (2012) developed a coupling between the hsB and an Analytical Element Method model of regional groundwater flow. Sahoo et al. (2018) altered the hsB to account

for surface ponding, unsaturated zone processes, bedrock leakage, and root-zone water balance for improved prediction of subsurface flow in a real-world catchment. In this work, we present a methodology, the hsB Proxy, that interrogates the results of high-quality numerical solutions of the full, non-linear hsB for wedge-shaped hillslopes with homogenous conductivity (having no analytical solution) and exploits patterns in these results to replicate the results of the hsB at dramatically reduced computational times (Note: herein, the use of "hsB" refers to this specific wedge-shaped, homogeneous conductivity configuration of the hsB). In this way, the conceptual and physical basis of the hsB can be applied to upscaling problems requiring solutions for thousands of hillslopes without resorting to computationally costly numerical methods. Insights of the interrogation of the hsB simulations and the generation of the hsB Proxy include the realization that the results of the hsB under a single recharge event can be linearly scaled by superposition to handle a range of recharge events. Further, a non-dimensionalization exercise demonstrates that results of the hsB can be scaled directly to handle a range of porosity, conductivity, and downstream hillslope width parameters. Finally, the interrogation reveals that the drainage response of a hillslope can be related via a power-law relationship to the bedrock slope of the hillslope. The ability of the Proxy to replicate the numerical solutions of the hsB under a range of hillslope shapes and recharge values is demonstrated. Then, the efficacy of the hsB in evaluating upscaling problems is demonstrated through two applications. First, the Proxy is compared against the upscaled recession behaviour in the Panola Mountain Research Watershed (Clark et al., 2009). These experiments demonstrate the utility of the Proxy and the non-linear hsB in accounting for topographic and transient recharge information in such upscaling problems. Then, the Proxy is used to derive a single effective hillslope representation of a basin composed of 867 individual hillslopes. The hsB Proxy is thus presented as a useful modelling tool that enables the simulation of: (1) the hillslope-scale subsurface flow response to a transient recharge forcing, given (2) hillslopes with unique planar shapes and bedrock slopes, such that (3) a network of heterogeneous hillslopes can be rapidly solved and analyzed. The Proxy therein enables the application of the physical insights of the full non-linear hsB for wedge-shaped hillslopes with transient recharge to problems of hillslope subsurface flow upscaling at significantly reduced computational costs. This work will explore and resolve three specific objectives:

1. Build a Proxy model that rapidly emulates the full non-linear hsB solution for wedge-shaped hillslopes under quite general conditions. The quality of the Proxy model is tested by an error metric indicating the discrepancy between the Proxy and numerical solutions.
2. Demonstrate the computational and conceptual efficacy of the Proxy by comparison to a numerical solver applied to a basin composed of hundreds of hillslopes. The efficacy of the Proxy is tested by evaluating the goodness-of-fit of the aggregate hillslope flow response generated by the Proxy and a numerical solver, and a comparison of the computation time required for each.
3. Utilize the Proxy model to apply the non-linear hsB for wedge-shaped hillslopes to two upscaling problems: (1) the derivation of a basin scale recession response from sub-basin scale reservoirs; and (2) the derivation of a single effective hillslope that best captures the aggregate flow response of a basin. The utility of the Proxy in the first application is demonstrated by

recreating the observed basin scale recession response as the aggregate response from dozens of DEM-derived hillslopes. The approach is compared to a previous conceptual reservoir steady-state analysis of the Panola experimental basin recession response (that of Clark et al. (2009)) and extended to a fully transient analysis enabled by the non-linear hsB. The utility of the Proxy in the second application is demonstrated through the rapid calibration of a single hillslope that reasonably reproduces the aggregate flow behaviour of a numerical solver applied to a basin, as characterized by a goodness-of-fit metric.

3.2 Methodology

The hsB Proxy is derived from the results of thousands of numerical simulations of the hsB. Here, the connection between the Proxy and the hsB is introduced. First, the hsB is defined and non-dimensionalized. Simplifications from this non-dimensionalization are carried into the development of the hsB Proxy, and the procedure and assumptions used to create the Proxy are presented.

3.2.1 The hillslope-storage Boussinesq equation

The hsB is an extension of the Boussinesq equation applied to hillslopes with known planar shape and bedrock slope (Troch et al., 2003). Application of a mass balance along a slice of the hillslope yields,

$$f\frac{\partial h(x,t)}{\partial t} = -\frac{1}{w(x)}\frac{\partial Q(x,t)}{\partial x} + N \tag{3.1}$$

where *f* is porosity [-], h is groundwater head relative to bedrock [L], t is time [T], x is the distance from the downslope end of the hillslope [L], $w(x)$ is the width function which describes the width of the hillslope at location x [L], Q represents volumetric flow rate [L^3/T], and N represents uniform recharge [L/T] applied to the top of the hillslope. The flow rate may be calculated using a modified version of Darcy's law,

$$Q(x,t) = -w(x)h(x,t)k\left(cos\theta\frac{\partial h(x,t)}{\partial x} + sin\theta\right) \tag{3.2}$$

where k is hydraulic conductivity (L/T) and θ is bedrock slope. In this work, the hsB is solved numerically using the discretization scheme of Hazenberg et al. (2015), which replaces the explicit variable accounting of $w(x)$ in Equations 3.1 and 3.2 with a constant representative width at each numerical grid cell, recognizing that a grid cell must be assigned a constant width (at fine discretization levels the numerical discretization respects the smoothly changing $w(x)$ function). To solve Equation 3.1, hillslopes are assigned a no-flow boundary at the upslope interface and the downslope condition is modified to enforce a specified zero-head boundary, consistent with the approach of Troch et al. (2003) and related studies utilizing the hsB (e.g., Broda et al., 2012; Dralle et al., 2014; Matonse & Kroll, 2009; Sahoo et al., 2018). It is significant to note that as an extension of the Boussinesq equation, the hsB will be sensitive to a non-zero boundary condition (Chor and Dias, 2015). As such, the application of the hsB Proxy developed here is currently only applicable to problems which assume a

zero-head downslope boundary condition. To remove the dependency of the Proxy solutions upon hydraulic conductivity and porosity, the non-linear problem is re-cast in terms of dimensionless time,

$$\frac{\partial h(x,\hat{t})}{\partial \hat{t}} = -\frac{1}{w(x)}\frac{\partial}{\partial x}\left[w(x)h(x,\hat{t})\left(cos\theta\frac{\partial h(x,\hat{t})}{\partial x} + sin\theta\right)\right] + \widehat{N} \tag{3.3}$$

where dimensionless time and recharge are defined as $\hat{t} = f/K \cdot t$ and $\widehat{N} = N/k$, respectively. In the case of a constant width function, an exact analytical solution is available (Bartlett and Porporato, 2018). However, we here limit the analysis to a more general linear width function,

$$w(x) = w_m x + w_b \tag{3.4}$$

where w_m represents the change in width per unit length (this can be positive or negative) and w_b represents the width at the $x = 0$ downslope interface. To more succinctly include the w_m and w_b parameters in the Proxy, a non-dimensional parameter X (upslope width fraction) is introduced, which characterizes the upslope width as some fraction of the downslope width,

$$w(L) = X \cdot w_b \tag{3.5}$$

$$X = \frac{w_m L}{w_b} + 1 \tag{3.6}$$

where L is the hillslope length and $w(L)$ is the upslope width. In hsB terminology, hillslopes are convergent when $X > 1$ (upslope width greater than downslope width) and divergent when $X < 1$ (upslope width less than downslope width). Substituting Equation 3.6 into Equations 3.4, 3.3, and 3.2 reveals that the governing equation is not dependent on the w_b parameter, and that the flow exiting the hillslope is linearly scaled by w_b. Thus, solutions to Equation 3.3 for a fixed value of w_b and any given combination of $\widehat{N}$, θ, X, L, and initial condition $h(x,0)$ may be re-scaled for any arbitrary porosity, hydraulic conductivity, and w_b value.

3.2.2 The hsB Proxy

The general strategy for generating the hsB Proxy is to first generate thousands of numerical solutions of the hsB for a wide range of X, L, and θ, with a fixed $h(x,0)$ value and $\widehat{N}$ fixed at zero. We then identify regression relationships between the θ parameter and a finite set of fully explanatory solution characteristics which describe the hillslope drainage $Q(0,\hat{t})$ from the fixed $h(x,0)$ initial condition over a range of X and L values. The Proxy support for the wide range of X, L, and θ characteristics is handled via the descriptive power of the regression relationships. The support for variation in conductivity and porosity are handled via scaling of the time parameter in Equation 3.3; support for variation in w_b and the recharge rate/initial condition is handled through linear scaling of the flow solution. With all of these in hand, the hsB Proxy will be able to act as a replacement for

numerical simulations of the hsB with a significant reduction in computational cost. The details of the numerical solutions used to create the Proxy and the efficacy of the regression relations and recharge rate scaling are evaluated below. A flow chart summarizing the generation of the Proxy is available in Figure 3.1. A flow chart summarizing how the Proxy is applied is available in Figure 3.2.

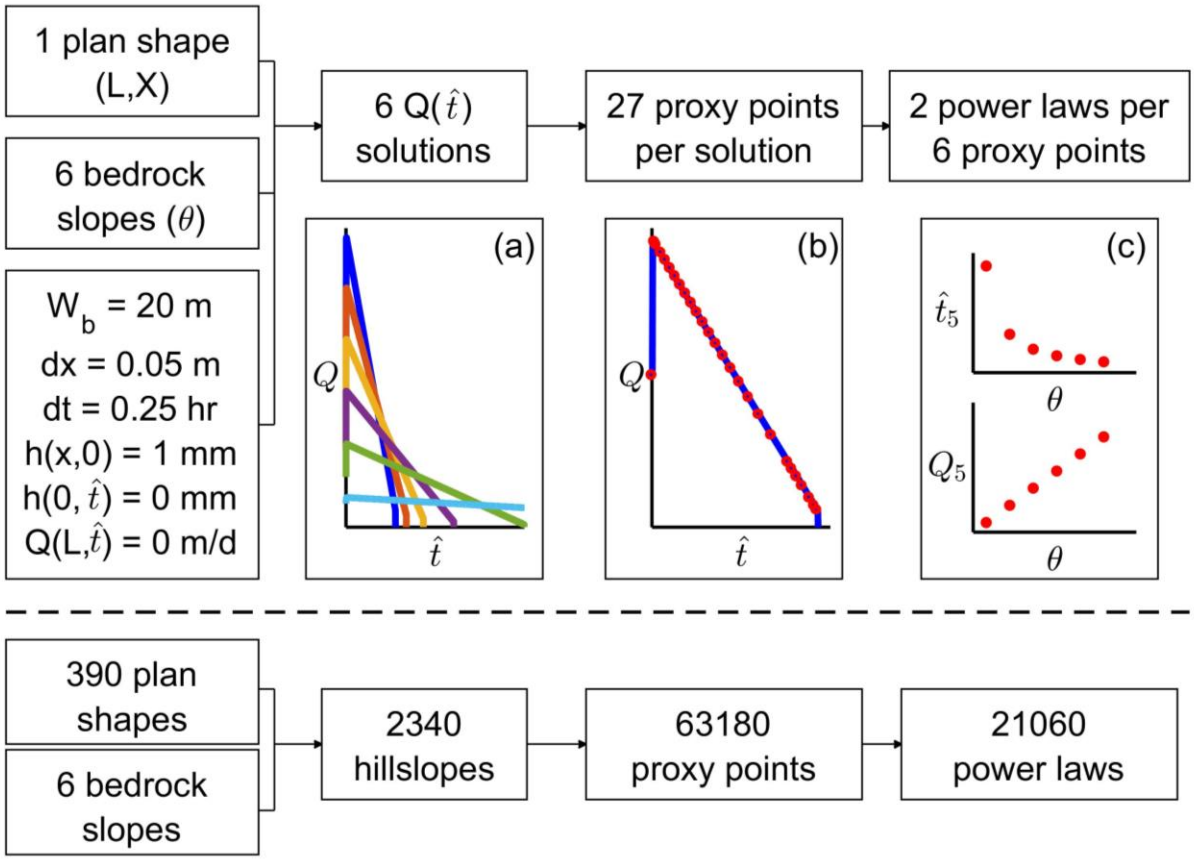

Figure 3.1. Generation of the Proxy. For a given plan hillslope shape (L,X) and six bedrock slope values (θ), 6 numerical Q($\hat{t}$) solutions are derived. Each solution is then converted into 27 proxy points, and each proxy point across the 6 solutions is described by 2 power-law equations relating time and flow to bedrock slope at the given plan shape. The six coloured lines in inset figure (a) illustrate hypothetical Q($\hat{t}$) solutions for six hillslopes, each with identical plan shapes (L,X) but different bedrock slopes (θ). Inset figure (b) illustrates the conversion of one of the Q($\hat{t}$) solutions (blue line) from inset (a) to 27 proxy points (red circles). Each of the Q($\hat{t}$) solutions in (a) is transformed, however only one Q($\hat{t}$) function is illustrated for visual clarity. As such, there are 6 Q($\hat{t}$) values associated with each proxy point. Inset figure (c) illustrates the two power law relationships for a single proxy point (the p = 5% proxy point in this example) across all six solutions. These power laws characterize the increase in flow at the p = 5% proxy point as bedrock slope increases, and the decrease in time to the p = 5% proxy point as bedrock slope increases. Overall, there are 390 plan shapes and 6 bed slopes in the Proxy, which combine to create 2,340 hillslopes and associated numerical solutions. Each of these solutions is translated into 27 proxy points, creating 63,180 total proxy points. For each set of 6 proxy points, there exists two power-law relationships, creating 21,060 power law relationships.

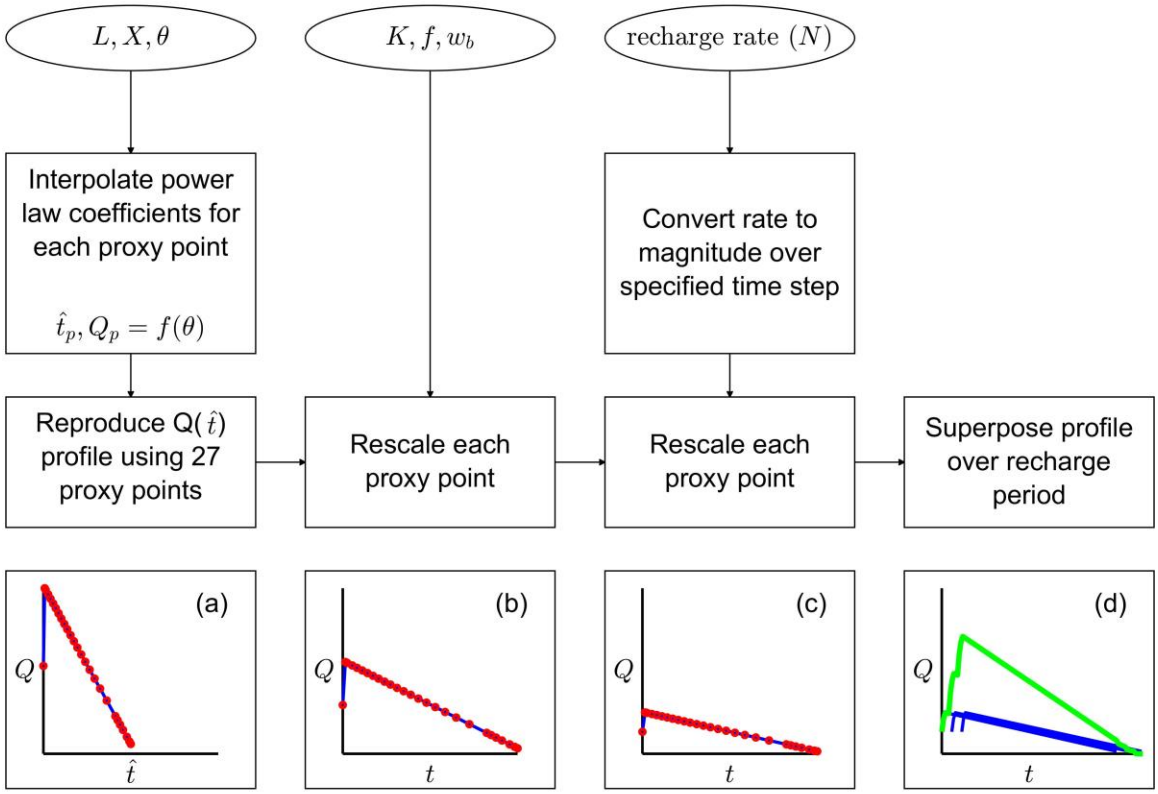

Figure 3.2. Application of the Proxy. Given inputs (ovals), the Proxy interpolates power-law coefficients to generate 27 proxy points representing the flow response to a 1 mm initial condition. These proxy points are then connected with straight lines, scaled, and superposed to generate a facsimile of the drainage response to the input recharge rate. Inset figure (a) illustrates an extracted set of proxy points (red points) and the interpolated Q($\hat{t}$) drainage response through these points (blue line). Inset figure (b) illustrates the rescaling of the drainage response according to the input K,f, and wb values. Note that when the solution is rescaled according to K and f, the solution is transformed out of dimensionless time. Inset figure (c) illustrates the rescaling of the Q(t) drainage response from a 1 mm initial condition to the magnitude specified by the recharge rate and time step. Inset figure (d) illustrates the superposition of the Q(t) response at every timestep of the recharge period: the Q(t) response (blue line) is applied at each timestep and then a superposed solution (green line) is generated describing the drainage response to the recharge rate over the recharge period.

3.2.2.1 Simplified drainage response

A key component of the Proxy is determining how to represent the continuous outflow response from each simulated hillslope in terms of a set of discrete data amenable to regression. This is simplified by taking advantage of the similar structure of outflow solutions for diverging and converging hillslopes. For diverging hillslopes, the outflow solution is generally characterized by a rapid, nearly instantaneous, increase in outflow which peaks and diminishes nearly linearly to almost zero at some time after the peak (Figure 3.3). For converging hillslopes, the outflow slowly increases (again, nearly linearly) to a peak flow rate after which the outflow quickly declines to zero. Because this continuous outflow hydrograph is nearly triangular, it may be described with relatively few degrees of freedom. Here, the continuous drainage response $Q(0, \hat{t})$ of each numerically simulated hillslope is simplified into a finite set of individual $(Q_p, \hat{t}_p)$ pairs, where the subscript p indicates an individual

pair. These points are associated with a percentage of the initial recharge volume remaining in the hillslope by splitting the drainage response (i.e., the outflow hydrograph) into nineteen pairs associated with 5% changes in storage (p = 95%, 90%, …, 10%,5%), and eight additional points providing added fidelity at the edges of the response (p = 97%, 96%, 4%, 3%, 2%, 1%, 0.5%, 0.1%), when the drainage response changes rapidly. That is, the Proxy represents the continuous hsB drainage response as twenty-seven individual pairs of flow magnitudes and times that are can then be linearly interpolated to emulate ("proxy") the continuous outflow. Note that the initial flow value (associated with p = 100%) can be directly calculated, and so is not included as a Proxy point. After the p = 0.1% Proxy point, flow is assumed to be zero.

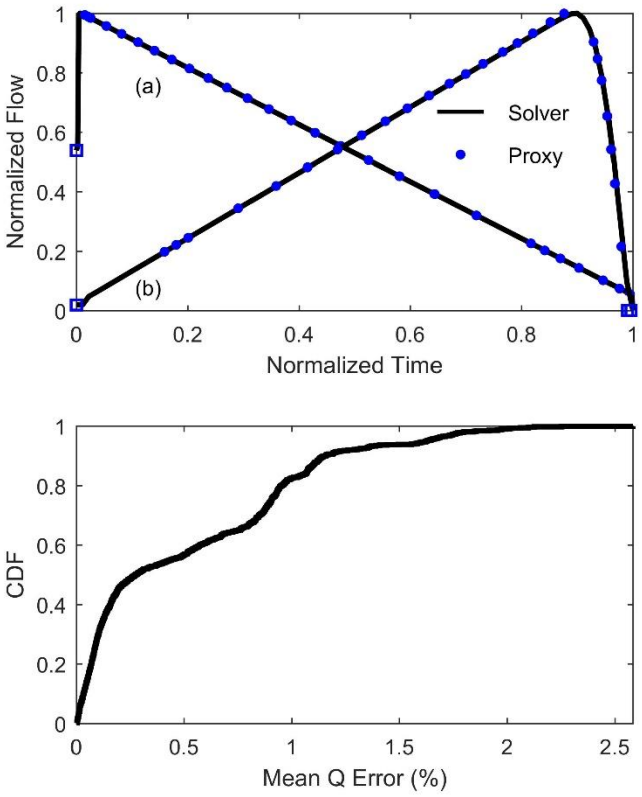

Figure 3.3. The hsB Proxy. (top) Example drainage profiles for (a) divergent and (b) convergent slopes as simulated by the numerical solver and the Proxy. Closed circles indicate proxy points. Squares indicate the initial flow value, which can be solved directly, and the assumption of flow = 0 after the final proxy point. (bottom) Distribution of mean flow error for the 2,340 hillslopes in the Proxy.

3.2.2.2 Numerical solutions

The Proxy is generated from high quality numerical solutions of the hsB. For each of the numerical simulations, all hillslopes are assigned identical boundary conditions and the hsB is initialized at a uniform 1 mm head. That is, the hsB Proxy describes the drainage of a hillslope in response to a recharge event that instantaneously applies 1 mm of head everywhere along the hillslope. The hsB is solved at a fine 0.05 m spatial discretization and 0.25 hr temporal discretization via the discretization scheme of Hazenberg et al. (2015), which generates numerical solutions much more

rapidly than other tested PDE solvers. The scheme is an implicit finite difference formulation that includes a flexible internal time step – this flexible internal time step allows for the generation of solutions at a specified time step while ensuring that numerical tolerance criteria are met when a finer temporal resolution is required. These discretization parameters are selected to ensure stability and to produce a Proxy that benefits from the accuracy derived from such a detailed discretization scheme.

The final hsB Proxy developed here is informed by the numerical solutions of 2,340 hillslopes representing unique combinations of the three hillslope shape parameters: Length (L); upslope width fraction (X); and bedrock slope (θ), solved at a fixed downslope width (w_b) equal to 20 m. The specific hillslope shape parameters used in the development of the proxy are available in Table 3.1. 1,404 of the hillslopes are convergent (X >1) and 936 are divergent (X<1). It is important to note that there are maximum values of hillslope shape parameters that limit the application of the Proxy generated here (Table 3.1). However, there is no reason the regression relationships here cannot be extended in range by further model sampling.

Table 3.1 – Hillslope parameter values in the proxy.

Hillslope Class	**L (m)**	**θ (°)**	**X**
Convergent	20, 44, 69, 93, 118, 142, 167, 191, 216, 240, 265, 290, 315, 340, 365, 390, 415, 440, 465, 490, 515, 540, 565, 775, 1000, 1500	2 5.6 9.6 12.8 16.4 20	1.05, 2.84, 4.63, 6.42, 8.21, 10, 15, 20, 30
Divergent			0.01, 0.198, 0.386, 0.574, 0.762, 0.95

3.2.2.3 Regression relationships

The Proxy simplifies the results of the twenty-seven pairs of discharge-time values derived for each of the 2,340 numerical solutions by exploiting a valuable underlying structure exhibited by the results of the hsB: each of the twenty-seven individual Proxy points exhibit a power-law relationship to the bedrock slope of the hillslope when all other parameters are held constant. That is, for a unique combination of L and X, two power-law relationships exist for each Proxy point, describing how the Proxy point flow and timing varies with bedrock slope (Figure 3.1). The functional form of these regressions are,

$$\hat{t}_p(L,X) = c_{p,\hat{t}}(L,X)\theta^{d_{p,\hat{t}}(L,X)} \tag{3.7}$$

$$Q_p(L,X) = c_{p,Q}(L,X)\theta^{d_{p,Q}(L,X)} \tag{3.8}$$

where p indicates one of the twenty-seven Proxy points; and c and d represent the coefficient and exponent of the power-law relationship, which are unique to each Proxy point and combination of L and X. The Proxy then behaves as a lookup table wherein the two coefficients (c and d) of these power-law relationships are interpolated between known L and X values. Although an underlying

solution structure is being exploited here, any attempts to fit a closed-form relationship through these coefficients was unsuccessful. The validity of this lookup table interpolation approach will be examined in a following section. The lack of an explicit function characterizing the relationship between the coefficients and the remaining hillslope shape parameters is the reason for the maximum values specified in Table 3.1 – beyond these values there is no way to interpolate an appropriate value, and extrapolation is not justified.

3.2.2.4 Linear scaling of recharge event responses

Testing of the Proxy revealed that, despite the non-linearity of the hsB equation, the response of a hillslope is adequately handled via linear scaling of the instantaneous 1 mm recharge event response. Thus, the hsB solution from a 1 mm initial condition can be applied over a range of recharge inputs. The key to this simplification lies in recognizing that a recharge rate at a given timescale is easily scaled to a more tenable magnitude at a finer timescale. For example, a 12 mm/d recharge event applies 0.5 mm of water over a single hour. Converting this to a head via porosity, we can recognize that the 1 mm Proxy initial condition is close enough to this hourly recharge magnitude that the non-linearity of the hsB will not produce a significantly different response when all other equation parameters are held constant (indeed, this is what testing reveals). Critically, testing also reveals that superposing these hourly drainage responses produces an aggregate drainage response that matches numerical results. That is, the drainage response of a hillslope to a 12 mm/d recharge event over a single day can be successfully replicated by superposition of twenty-four 0.5 mm drainage responses. This linear scaling approach thus enables the handling of transient recharge inputs by superposition.

The validity of this linear scaling approach is demonstrated by example. Figure 3.4 illustrates a daily rainfall time series from the Naramata climate station, a daily Environment Canada climate station in British Columbia, Canada (Environment Canada, 2020), with precipitation values ranging from 0.2 to 11.4 mm/d. For the sake of example, this rainfall time series is treated as a recharge time series (if overestimating magnitude, it provides a realistic temporal distribution of daily recharge events). Figure 3.4 also presents the hsB solution of a single hillslope (L = 100 m, W_b = 60 m, X = 0.1, θ = 10°) with this transient recharge input as well as the accompanying results based upon linear scaling of the 1 mm Proxy solution (NSE = 0.999). The good agreement between the scaled Proxy and the numerical solution demonstrates the validity of the linear scaling approach under these real-world recharge rates. The upper limit of acceptable daily recharge rates is evaluated below.

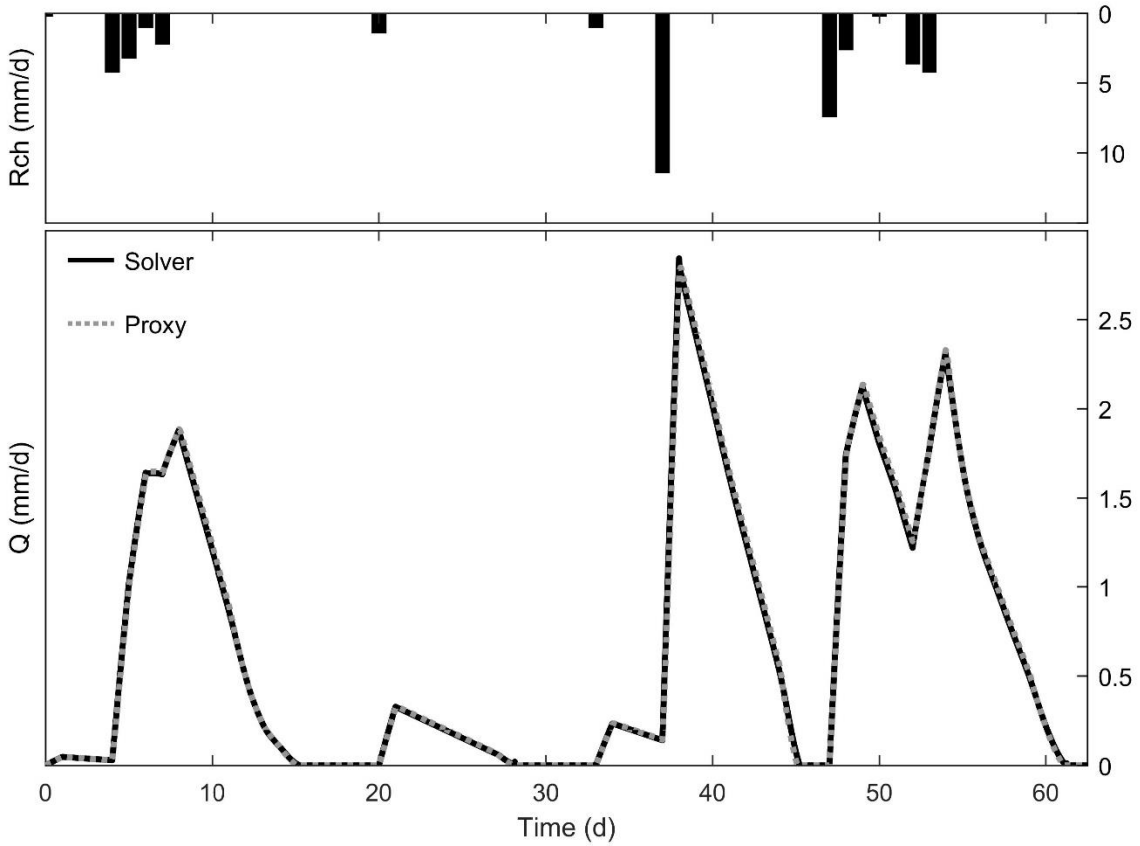

Figure 3.4. A real precipitation (recharge) daily time series (top) and the hillslope subsurface flow response (bottom, NSE = 0.999) for a hypothetical hillslope solved using the numerical solver and the Proxy.

3.3 Results

Example results produced by the Proxy are presented in Figure 3.3 for a convergent and a divergent hillslope, with each exhibiting unique drainage patterns (convergent slopes generally display an increase to a maximum value while divergent slopes consistently decrease). These examples illustrate the excellent handling of most of the drainage profile, with only slight discrepancies at the initial and final periods, when the response of the profile changes rapidly. To characterize the error in the Proxy, a mean flow error metric is calculated for each hillslope as follows,

$$Q_{err} = \frac{1}{n_t} \sum_{n=1}^{n_t} \frac{\left| Q_{sim}^{n} - Q_{prox}^{n} \right|}{\max(Q_{sim})} \times 100\% \tag{3.9}$$

where Q_{err} represents the mean flow error associated with a hillslope, n is the time index, Q_{sim}^{n} represents the flow produced by the numerical solver at time index n, Q_{prox}^{n} represents the flow produced by the Proxy at time index n, $\max(Q_{sim})$ represents the maximum simulated flow value associated with the hillslope, and n_t represents the final time index when the hillslope flow ends. The simulated flow and Proxy flow values are compared at each time step (Δt = 0.25 hours). This error metric indicates the mean distance between the numerical solution which produced the Proxy and the

Proxy itself, normalized into a percentage with respect to the largest simulated flow value. The distribution of mean flow error in the Proxy is presented in Figure 3.3. This error distribution demonstrates that the Proxy is not a perfect facsimile of the results of the hsB but is on average correct within 2.5% of the maximum flow value for a single hillslope. The Proxy performs well under tested recharge rates up to 50 mm/d over a single day, with 96% of the Proxy hillslopes exhibiting less than 10% error and 50% of hillslopes exhibiting less than 2.5% error. In general, this error decreases as recharge rate decreases, and longer, steeper hillslopes exhibit smaller errors than shorter, flatter hillslopes as recharge rates increase. However, the efficacy of the Proxy is more appropriately demonstrated in application to a landscape composed of hundreds of hillslopes, where the computational benefits of the Proxy are evident.

3.3.1 Application of hsB Proxy to derived hillslopes

The practical efficacy of the Proxy approach is demonstrated by comparing the results of the Proxy to the numerical solutions of hundreds of hillslopes derived from a real basin. The comparison is evaluated with respect to (a) computational cost and (b) the difference between the discharge hydrographs of the computationally expensive numerical simulations and the Proxy. The selected basin is upstream of Dennis Creek in British Columbia, Canada, illustrated in Figure 3.5a. The basin is a forested catchment (AAFC, 2020) with a subsurface composition of sandy loam soils (MECSS, 2021) over a bedrock of granite and granodiorite (MEM, 2020), receiving an average of 220 mm of rainfall per year (Environment Canada, 2020). The previously illustrated rainfall time series (Figure 3.4) is again applied as a recharge time series. A hydraulic conductivity of 1 m/hr and a porosity equal to 0.3 are assumed.

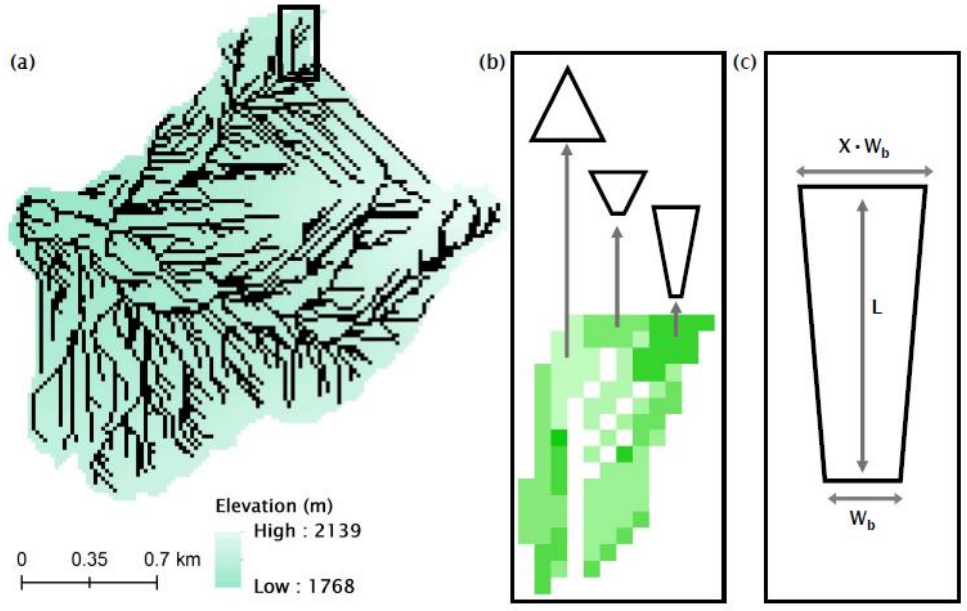

Figure 3.5. (a) The Dennis Creek watershed and stream network. (b) fourteen derived hillslopes in the upstream subset highlighted in (a); each shaded area represents a hillslope. Three wedge-shaped hillslopes derived from these areas are shown. (c) Each of the 867 derived hillslopes in the watershed is converted to a representative planar wedge-shape.

Hillslopes are derived from a 20 m resolution DEM of the watershed (CDEM, 2019) using a modification of the approach of Liu et al. (2012), itself a modification of the approach of Fan and Bras (1998). In this approach, stream network tributaries are assigned a left bank and right bank hillslope, and the most upstream raster cells are assigned a single hillslope (these upstream hillslopes represent headwater hillslopes). The algorithm identifies the contributing DEM cells to each stream bank/upstream cell and traces the surface flow pathways between the cells. The cells are then sorted by flow distance to determine a $w(x)$ profile (related to the number of cells existing at a specified flow distance), as well as the maximum flow distance, which defines the hillslope length. In our modified algorithm, this length (L), the total stream bank cell length (w_b), and the derived contributing area are used to calculate the X parameter of the linear width function (Equation 3.5). The bedrock slope is derived using the tangent of the change in elevation along the longest flowpath divided by the length of this flowpath (L). The assumption that bedrock slope is identical to topographic gradient is made to produce some estimate of bedrock gradient with readily available data.

It is known that the choice of derived stream network has a direct impact on the shape and size of derived hillslopes, and past research efforts have explored the feasibility of various derived networks (Montgomery and Dietrich, 1992; Lin et al., 2006) as well as the more fundamental difficulty of specifying precisely where a channel begins (Wohl, 2018). In this work, we select a stream network associated with a Strahler stream ordering number greater than one, which results in a dense network. The justification is that, given the uncertainty of actual flow pathways, we err on the side of producing as many flow pathways as is tenable. Conceptually, this suggests that the subsurface discharge from hillslopes behaves as an ephemeral stream network that can extend beyond the perennial network. Note that surface routing considerations are ignored as, given the timescale of subsurface flows, surface flow pathways are relatively instantaneous (experiments, conducted but not shown, demonstrated that the discrepancy was negligible).

Figure 3.5a illustrates the derived stream network in the Dennis Creek watershed. As an example of the application of the hillslope derivation algorithm, Figure 3.5b illustrates fourteen derived hillslopes in an upstream subset of the network. The wedge-shaped conceptualization of each hillslope is presented in Figure 3.5c. In this basin, 867 hillslopes are derived (Figure 3.5b depicts only fourteen for the sake of visual clarity). Histograms of the derived hillslope shape properties (L, X, w_b, and θ) are presented in Figure 3.6. These histograms illustrate the mean shape properties (L = 110 m, w_b = 36 m, X = 0.85, θ = 9.8°) and provide a visual representation of the structure of hillslopes within the basin. The distributions of these physically-based hillslope shape parameters will control the aggregate hillslope subsurface flow response.

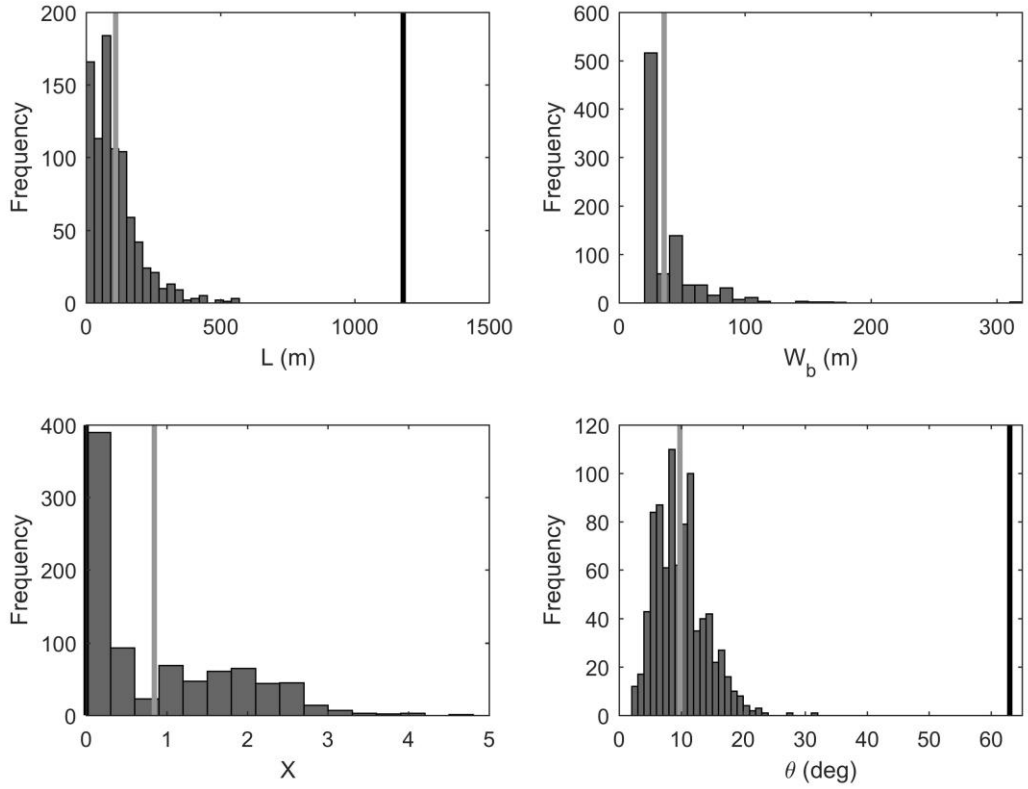

Figure 3.6. Distributions of the hillslope shape properties of the 867 hillslopes derived from the Dennis Creek watershed. Grey vertical lines indicate mean values of L= 110 m, Wb = 36 m, X = 0.85, θ = 9.8°. Black vertical lines indicate single effective slope values (Figure 3.7b) of L = 1180 m, X = 0.01, θ = 63°. The effective single slope value for Wb (= 4430 m) is excluded for visual clarity.

Each of the 867 hillslopes is solved using a numerical solver (i.e., the scheme of Hazenberg et al. (2015)) as well as the Proxy. Figure 3.7a illustrates the aggregate response of all hillslopes (i.e., the sum of each hillslope's subsurface flow contribution at each time step) as calculated by the solver and the Proxy. As a comparison of two hydrographs, the quarter-hourly NSE value of 0.999 indicates that the Proxy is exceedingly successful in emulating the results from the numerical solver. The key distinction, however, lies in the comparison of computational time: 55 hours 32 minutes and 25 seconds for the numerical solution, versus just 7 seconds for the Proxy on the same machine (Intel Core i5-6300 CPU @ 2.3 GHz, 8GB RAM). The significant reduction in computational cost enables support for calibration, uncertainty analysis, sensitivity analysis, ensemble simulation or any other computationally intensive modelling exercise.

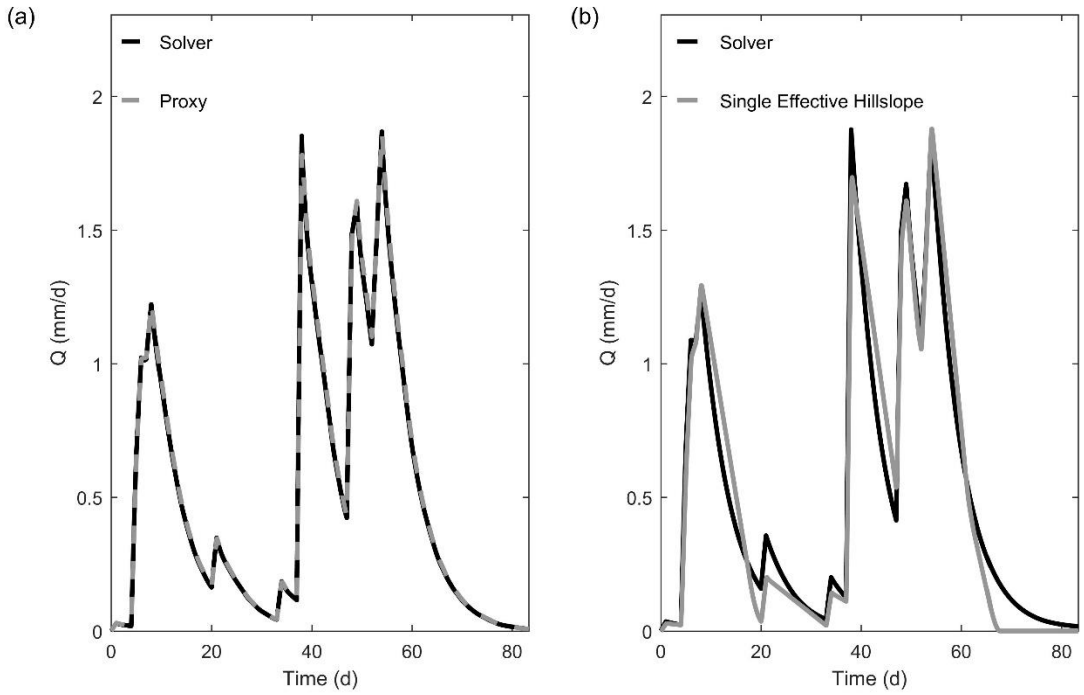

Figure 3.7. Aggregate hillslope subsurface flow results in the Dennis Creek watershed example. (a) The aggregate response as solved by the numerical solver and the Proxy (NSE = 0.999). (b) The aggregate response as solved by the numerical solver compared to the best-fitting single effective hillslope response derived from the Proxy via a Monte Carlo exercise (NSE = 0.956)

3.4 Application to upscaling problems

With the computational benefits of the Proxy approach demonstrated, it is necessary to demonstrate the utility of the hsB and its wedge-shaped hillslopes in the context of known upscaling behaviour. The work of Clark et al. (2009) in the Panola Mountain Research Watershed (PMRW) provides a valuable case study in this regard. Then, the application of the hsB Proxy to an upscaling problem asking whether a network of hillslopes can be replicated by a single effective hillslope is demonstrated.

3.4.1 Upscaling Problem 1: Application to PMRW

Clark et al. (2009) used observed data from the PMRW to evaluate the unique recession behaviour of flow ($-dQ\ /\ dt = f(Q)$) from a single hillslope and from the entire basin. They then developed a modelling framework wherein three linear reservoirs, representing portions of the landscape with similar timescales of response, were aggregated to successfully simulate the overall basin response. The efficacy of this approach – breaking down the basin-scale response into the aggregate response of smaller-scale reservoirs – was also demonstrated in this basin by Harman et al. (2009), which used a more theoretical approach wherein the distribution of timescales across the basin is explicitly characterized by a probability density function. Thus, we can compare how the hsB Proxy

and its network of physically-based wedge-shaped hillslopes compares against these hypothetical linear reservoirs in simulating basin-scale recession behaviour. Field estimates of porosity (=0.5) and hydraulic conductivity (=0.64 m/hr) required by the Proxy are obtained from McIntosh et al. (1999), and 117 hillslopes were derived from a DEM of the basin (U.S. Geological Survey, 2018). Following the methodology of Clark et al. (2009), the hillslopes are first brought to a steady-state in response to the constant application of a 25 mm/d recharge rate, and then the following recession behaviour is analyzed. Figure 3.8a demonstrates the basin-scale recession data in Clark et al. (2009) as well as the Proxy response assuming a homogenous distribution of subsurface properties. The lack of fit between these responses suggests that the conductivity estimate is too low (the estimate of McIntosh et al. (1999) is from a core sample) and that known heterogeneity in the timescales of the reservoirs (hillslopes) is unaccounted. While a range of homogeneous conductivity values were tested, it was only by introducing variability in conductivity that we were able to adequately replicate the observed response. To resolve these limitations, the Proxy is applied to this problem a second time, now allowing for heterogeneity in hillslope conductivity by sampling a normal distribution of log conductivity with mean 10 m/hr and standard deviation 0.7 in log space (allowing for sampling within two orders of magnitude from the mean). These mean and standard deviation values were selected from a manual calibration, with this mean value being an order of magnitude larger than the reported core sample, an implicit reflection of the changing preferential flow pathways at scales larger than a core sample. Taking advantage of the computational speed of the Proxy, this distribution is sampled 1000 times to produce unique distributions of hillslope conductivity across the 117 basin hillslopes, therein ensuring appropriate characterization of the heterogeneity introduced by this distribution. To reinforce the necessity of the Proxy in this experiment, note that the 117,000 hillslope solutions required here would be computationally untenable with a numerical solver. The mean response of the ensemble of 1000 samples is presented in Figure 3.8a. The single best member of the ensemble is also presented in Figure 3.8a, and the associated conductivity distribution is presented in Figure 3.8b. The agreement between this response and the Clark model demonstrates that the hsB Proxy and its 117 physically-based hillslope reservoirs are a viable hypothesis for simulating the aggregate subsurface stormflow behaviour in a basin. This analysis underlines the critical role of heterogeneity in governing subsurface stormflow response and reflects the motivating justification for the creation of the Proxy, which here enabled the rapid analysis of a thousand basin configurations using the hsB and its physically-based reservoir and subsurface characteristics, as opposed to the more abstract linear reservoir conceptualization. It is important to note that although the hsB approach is superior in its physical conceptualization, linear reservoir models are conceptually simpler, and so may be preferred depending on the specific modelling question at hand. For example, in this example 117 physically-based hillslopes are generated, whereas the linear reservoir approach only utilizes three hillslopes. Although the Proxy enables the rapid generation of solutions for these 117 hillslopes, three hillslopes are inherently simpler than 117 hillslopes. In contrast, the hsB approach is more physically justified as well as more parsimonious, requiring only two parameters (mean conductivity and variance) to be calibrated across the entire basin while all other properties (including the number of hillslopes) are automatically derived from topography. This is contrasted to the reservoir approach, which requires the specification of two abstracted parameters per reservoir to control the hillslope flow response.

However, the hsB approach is inherently superior to the linear reservoirs approach when the modelling question at hand depends on transient recharge conditions.

It is worthwhile now to consider whether the observed recession behaviour and associated upscaled model of Clark et al. (2009) can be replicated under transient conditions, which can be handled efficiently by the Proxy. To evaluate this, a daily time series of precipitation from August 1995 to December 1996 in the PMRW (Aulenbach, 2017), matching the observed recession time period in Clark et al. (2009), is obtained and fed into the Proxy as a recharge time series. In light of the significant reduction in recharge pulse magnitudes associated with a precipitation time series, as compared to the unrealistic constant value of 25 mm/d used to produce the steady-state results, a new normal distribution of log conductivity with a manually calibrated mean of 0.5 m/hr and standard deviation of 0.5 in log space (allowing for sampling over 1.5 orders of magnitude from the mean) is sampled. The change in the conductivity distribution between the previous steady-state experiment and this transient experiment occurs because the conductivity distribution calibrated from the steady input does not respect the transient application of water in the basin. As such, the steady-state analysis requires a conductivity distribution that is compensating for the unrealistic constant 25 mm/d input in order to produce recession curves that match the observed recession response. The transient recharge investigated in this experiment is more realistic, and this updated conductivity distribution is now accounting for both the observed recession response as well as this more realistic recharge input. In contrast to the steady-state case, the transient recharge pulses produce more than one recession event. Following Clark et al. (2009), a minimum recession volume of 400 m^3 is required for a recession to be valid for analysis. Once a recession meets this requirement, a unique recession function ($-dQ/dt = f(Q)$) is derived via linear regression. The analysis of Clark et al. (2009) demonstrates a multi-phase response of the recession function, wherein the slope of the recession function changes as different parts of the landscape contribute flow. To capture this multi-phase behaviour, each derived recession function is allowed to have two phases, such that the recession can be characterized by a piece-wise function allowing for a change in dQ/dt as Q values decrease. Once all recession events have been converted into a representative piece-wise recession function, a single representative function is derived as the mean of these responses. Sampling the conductivity distribution 1000 times produces the ensemble mean and single best member of the ensemble presented in Figure 3.8c. The conductivity distribution associated with the single best member is presented in Figure 3.8d. The transient response compares well again to the steady-state models, but is superior in comparison to the observed data, especially in its handling of the apparent concave form of the observed data. As in the steady-state experiment, this experiment demonstrates that observed recession behaviour can be captured by a physically-based discretization of the landscape, as well as with realistic transient recharge impulses as opposed to an idealized steady-state recharge. This reinforces the utility of the Proxy as a tool for exploring subsurface stormflow upscaling problems with minimal conceptual abstraction and improved physical basis.

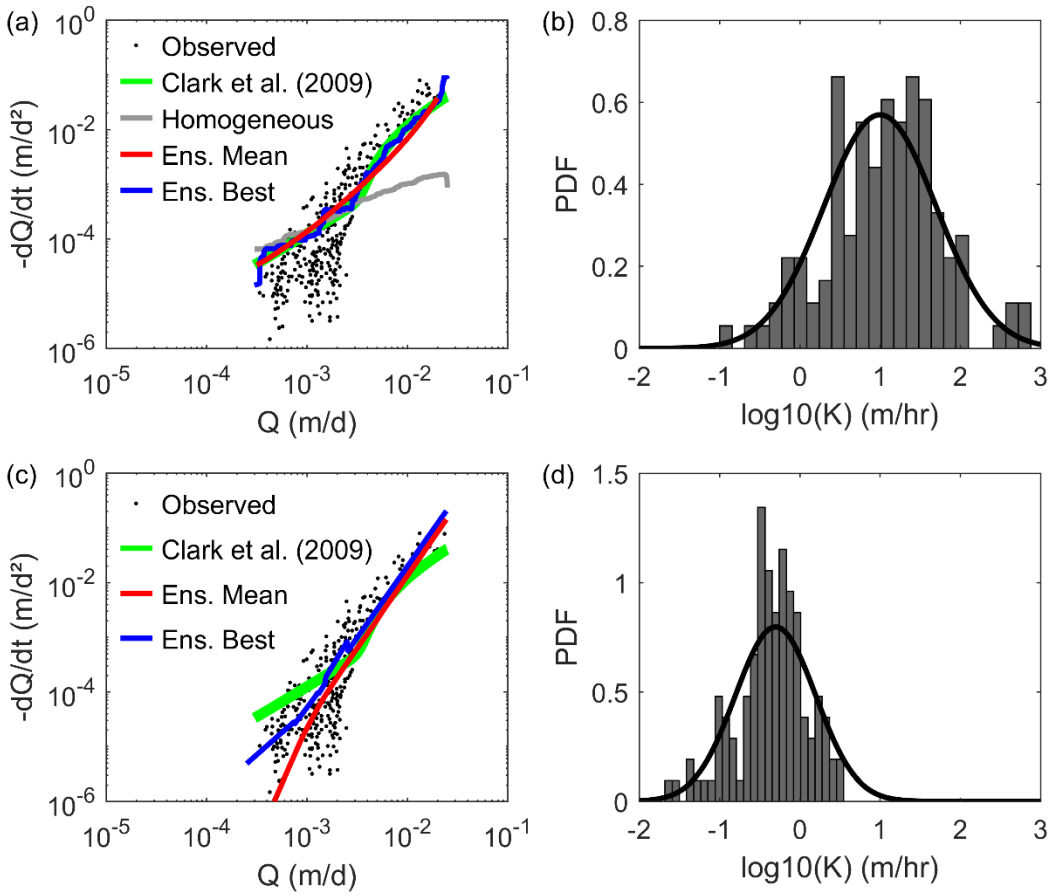

Figure 3.8. Panola Mountain Research Watershed upscaling problem. (a) Observed recession behaviour from Clark et al. (2009), compared to: the Clark et al. (2009) three reservoir model and the Proxy under steady-state recharge. The Proxy results are presented with a single homogenous conductivity ("Homogeneous") and a heterogenous conductivity distribution ("Ens.") – ensemble mean demonstrates the mean response of 1000 samples of the heterogenous conductivity distribution, while ensemble best reflects the single best comparison to the Clark et al. model. (b) Associated conductivity distribution and histogram of sampled conductivity associated with the ensemble best. (c) Comparison between observed behaviour, Clark et al. (2009) model, and Proxy solved under transient recharge conditions. (d) Associated conductivity distribution and ensemble best histogram.

3.4.2 Upscaling Problem 2: calibrating a single effective hillslope

An additional upscaling question that is enabled by the Proxy but untenable with the numerical solver is whether it is possible to derive a single effective hsB hillslope representation of the aggregate flow behaviour in the Dennis Creek Watershed illustrated in Figure 3.7a via calibration. Such a calibration exercise requires solving the hsB hundreds of times for various hillslope characteristics, meaning tens of hours of computational time. However, the Proxy can handle this question directly in minutes, even taking on a more exhaustive Monte Carlo approach where a large, structured parameter space is meticulously sampled. Undertaking such an exercise, using the same recharge time series, conductivity, and porosity utilized to produce the hydrograph in Figure 3.7a, reveals that the most effective single slope representation (NSE = 0.956, Figure 3.7b) of this system is a hillslope with a shape approximate to L = 1180 m, w_b = 4430 m, θ = 63 degrees, and X = 0.01. A comparison of these parameter values with respect to the means and histograms of derived hillslopes in this basin are

presented in Figure 3.6. This analysis presents several interesting discussion items. First, it demonstrates that, while not an obvious facsimile, a single hillslope can be reasonably descriptive of the aggregate response. Second, that the obvious discrepancy is in the ability of a single hillslope to capture the recession behaviour of the aggregate response, which is a consequence of ignoring the variability of drainage response timescales across the basin. As a complement to the previous upscaling experiments reinforcing the influence of heterogeneity in conductivity, this experiment demonstrates that the heterogeneity in hillslope shape and slope is necessary to produce the variable timescales of flow response that result in a smoothly receding recession response given a homogenous conductivity across the basin. Third, this analysis demonstrates that the ability to quantify the conditions in which the basin-scale heterogeneity is reasonably captured by a single effective hillslope would be a valuable upscaling exercise, in that it would set out rules under which hillslope heterogeneity can be folded into a single quantifiable response. The investigation of the relationship between the bulk characteristics of the watershed (e.g., distributions of hillslope shape) and a single effective hillslope response is a key component of future work using the Proxy.

3.5 Conclusion

The hsB Proxy is a tool for the modeller interested in utilizing the hillslope-storage Boussinesq equation (hsB) describing hillslopes with wedge-shaped planar shape and homogenous conductivity under transient recharge. The hsB Proxy reproduces the results of the hsB with minimal loss of fidelity and a significant reduction in computational time, and is useful in its ability to handle a wide range of hillslope shapes as well as its demonstrated efficacy in superposing individual recharge events. The hsB Proxy is simple to use, only requiring the modeller to interpolate a lookup table of power-law coefficients for twenty-seven individual time and flow points which are then linked together to produce a facsimile of the drainage response. Further, the philosophy behind the Proxy is such that the lookup table is highly extendable. That is, the current number of 2,340 hillslopes (and the associated maximum values of hillslope shape parameters) is not static: any modeller solving hillslopes with parameters beyond the current range of Proxy parameters can produce numerical results that can be analyzed and plugged into the Proxy.

The application of the Proxy to two upscaling problems has been demonstrated. In the first problem, the Proxy was used to rapidly generate alternative hypotheses of hillslope-scale flow response to generate a set of hillslopes that produced a viable, physically-based description of observed basin-scale recession response. Then, the Proxy was used to generate an alternative hypothesis of hillslope-scale flow response under non-steady state conditions, using a recharge time series derived from real world precipitation magnitudes. In the second upscaling problem, the Proxy was used to test thousands of single hillslope responses to select the single best effective hillslope to represent the aggregate flow response of a basin. The following objectives have thus been achieved:

1. A Proxy model has been developed that emulates the full non-linear hsB solution for wedge shaped hillslopes within 2.5% error and with minimal computational cost.

2. The efficacy of the Proxy has been demonstrated in application to a basin composed of hundreds of hillslopes. The Proxy is able to reproduce the aggregate flow response of the basin as derived from a numerical solver (NSE = 0.999) at a fraction of the computational cost (over 10,000 times more computationally efficient).
3. The non-linear hsB for wedge-shaped hillslopes has been applied via the Proxy to two computationally demanding upscaling problems. In the first, the Proxy was able to generate a viable hypothesis of the distribution of hillslopes and their properties (size, shape, and hydraulic conductivity) which justified the recession behaviour at the PMRW as compared to a recession model from the literature parameterized at steady-state. Then, transient precipitation data was used to generate an alternate hypothesis justifying the observed recession behaviour without requiring the steady depletion curve assumption. In the second problem, the Proxy enabled the rapid calibration of a single effective hsB hillslope that reasonably reproduces the aggregate behaviour of the Dennis Creek basin. The successful application to these problems demonstrates the utility of the Proxy and the non-linear hsB in investigating such upscaling problems, both from a computational savings perspective by supporting more physically-based conceptualizations in the analysis.

In the future, we intend to apply the Proxy to a large number of basins to identify scale-appropriate relationships between distributed basin characteristics and emergent runoff response at the basin scale. In addition, future work will include the generation of updates to the Proxy, including solutions utilizing alternative downstream boundary conditions and new limits of hillslopes shapes properties.

3.6 References

AAFC. (2020). Agriculture and Agri-Food Canada - Land Use 2010 - Open Government Licence - Canada. Retrieved 13 February 2021, from https://open.canada.ca/data/en/dataset/9e1efe92-e5a3-4f70-b313-68fb1283eadf#wb-auto-6

Ali, M., Ye, S., Li, H. yi, Huang, M., Leung, L. R., Fiori, A., & Sivapalan, M. (2014). Regionalization of subsurface stormflow parameters of hydrologic models: Up-scaling from physically based numerical simulations at hillslope scale. *Journal of Hydrology*, 519(PA), 683–698. https://doi.org/10.1016/j.jhydrol.2014.07.018

Aulenbach, B. T. (2017). Data for estimating monthly water budgets at Panola Mountain Research Watershed, Stockbridge, Ga., water years 1986–2015. Retrieved 31 August 2020, from https://doi.org/10.5066/F7XS5SNV

Bartlett, M. S., & Porporato, A. (2018). A Class of Exact Solutions of the Boussinesq Equation for Horizontal and Sloping Aquifers. *Water Resources Research*, 54(2), 767–778. https://doi.org/10.1002/2017WR022056

Broda, S., Larocque, M., Paniconi, C., & Haitjema, H. (2012). A low-dimensional hillslope-based catchment model for layered groundwater flow. *Hydrological Processes*, 26(18), 2814–2826. https://doi.org/10.1002/hyp.8319

Brutsaert, W., & Nieber, J. (1977). Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resources Research* 13 (3): 637-643 DOI: 10.1029/WR013i003p00637

Carrillo, G., Troch, P. A., Sivapalan, M., Wagener, T., Harman, C., & Sawicz, K. (2011). Catchment classification: hydrological analysis of catchment behavior through process-based modeling

along a climate gradient. *Hydrology and Earth System Sciences*, 15(11), 3411–3430. https://doi.org/10.5194/hess-15-3411-2011

CDEM. (2019). Canadian Digital Elevation Model, 1945-2011 - Open Government Portal. Retrieved 21 April 2020, from https://open.canada.ca/data/en/dataset/7f245e4d-76c2-4caa-951a-45d1d2051333

Childs, E. C. (1971). Drainage of Groundwater Resting on a Sloping Bed. *Water Resources Research*, 7(5), 1256–1263. https://doi.org/10.1029/WR007i005p01256

Chor, T. L., & Dias, N. L. (2015). Technical Note: A simple generalization of the Brutsaert and Nieber analysis. *Hydrology and Earth System Sciences*, 19(6), 2755–2761. https://doi.org/10.5194/hess-19-2755-2015

Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-van Meerveld, H. J., Peters, N. E., & Freer, J. E. (2009). Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale. *Hydrological Processes,* 23(2), 311–319. https://doi.org/10.1002/hyp.7154

Dralle, D. N., Boisramé, G. F. S., & Thompson, S. E. (2014). Spatially variable water table recharge and the hillslope hydrologic response: Analytical solutions to the linearized hillslope Boussinesq equation. *Water Resources Research*, 50(11), 8515–8530. https://doi.org/10.1002/2013WR015144

Environment Canada. (2020, September 17). Daily Data Report for September 1986 - Climate - Environment and Climate Change Canada. Retrieved 29 October 2020, from https://climate.weather.gc.ca/climate_data/daily_data_e.html?hlyRange=%7C&dlyRange=1971-04-01%7C2004-02-29&mlyRange=1971-01-01%7C2004-0201&StationID=1026&Prov=BC&urlExtension=_e.html&searchType=stnProv&optLimit=yearRange&StartYear=1990&EndYear=2016&selRowPerPage=100&Line=515&lstProvince=BC&timeframe=2&Day=16&Year=1986&Month=9#

Fan, Y., & Bras, R. L. (1998). Analytical solutions to hillslope subsurface storm flow and saturation overland flow. *Water Resources Research*, 34(4), 921–927. https://doi.org/10.1029/97WR03516

Harman, C. J., Sivapalan, M., & Kumar, P. (2009). Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resources Research*, 45(9). https://doi.org/10.1029/2008WR007392

Hazenberg, P., Fang, Y., Broxton, P., Gochis, D., Niu, G. -Y., Pelletier, J. D., et al. (2015). A hybrid-3D hillslope hydrological model for use in <scp>E</scp> arth system models. *Water Resources Research*, 51(10), 8218–8239. https://doi.org/10.1002/2014WR016842

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, 45(2). https://doi.org/10.1029/2008WR006912

Lin, W.-T., Chou, W.-C., Lin, C.-Y., Huang, P.-H., & Tsai, J.-S. (2006). Automated suitable drainage network extraction from digital elevation models in Taiwan's upstream watersheds. *Hydrological Processes*, 20(2), 289–306. https://doi.org/10.1002/hyp.5911

Liu, J., Chen, X., Zhang, X., & Hoagland, K. D. (2012). Grid digital elevation model based algorithms for determination of hillslope width functions through flow distance transforms. *Water Resources Research,* 48(4). https://doi.org/10.1029/2011WR011395

Matonse, A. H., & Kroll, C. (2009). Simulating low streamflows with hillslope storage models. *Water Resources Research*, 45(1). https://doi.org/10.1029/2007WR006529

McIntosh, J., McDonnell, J. J., & Peters, N. E. (1999). Tracer and hydrometric study of preferential flow in large undisturbed soil cores from the Georgia Piedmont, USA. *Hydrological Processes*,

13(2), 139–155. https://doi.org/10.1002/(SICI)1099-1085(19990215)13:2<139::AID-HYP703>3.0.CO;2-E

MECSS. (2021). Ministry of Environment and Climate Change Strategy - Soil Survey Spatial View - Licensed under Open Government Licence - British Columbia. Retrieved 13 February 2021, from https://catalogue.data.gov.bc.ca/dataset/soil-survey-spatial-view

MEM. (2020). Ministry of Energy, Mines and Low Carbon Innovation - Bedrock Geology - Licensed under Open Government Licence - British Columbia. Retrieved 13 February 2021, from https://catalogue.data.gov.bc.ca/dataset/bedrock-geology

Montgomery, D. R., & Dietrich, W. E. (1992). Channel initiation and the problem of landscape scale. *Science*, 255(5046), 826–830. https://doi.org/10.1126/science.255.5046.826

Patnaik, S., Biswal, B., Nagesh Kumar, D., & Sivakumar, B. (2018). Regional variation of recession flow power-law exponent. *Hydrological Processes*, 32(7), 866–872. https://doi.org/10.1002/hyp.11441

Sahoo, S., Sahoo, B., & Panda, S. N. (2018). Hillslope-storage Boussinesq model for simulating subsurface water storage dynamics in scantily-gauged catchments. *Advances in Water Resources*, 121, 219–234. https://doi.org/10.1016/j.advwatres.2018.08.016

Troch, P. A., Paniconi, C., & Emiel van Loon, E. (2003). Hillslope-storage Boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. Formulation and characteristic response. *Water Resources Research*, 39(11). DOI: 10.1029/2002WR001728

Troch, P. A., van Loon, A. H., & Hilberts, A. G. J. (2004). Analytical solution of the linearized hillslope-storage Boussinesq equation for exponential hillslope width functions. *Water Resources Research*, 40(8). DOI: 10.1029/2003WR002850

Troch, P. A., Berne, A., Bogaart, P., Harman, C., Hilberts, A. G. J., Lyon, S. W., et al. (2013). The importance of hydraulic groundwater theory in catchment hydrology: The legacy of Wilfried Brutsaert and Jean-Yves Parlange. *Water Resources Research*, 49(9), 5099–5116. https://doi.org/10.1002/wrcr.20407

U.S. Geological Survey. (2018). USGS NED 1/3 arc-second n34w085 1 x 1 degree ArcGrid 2018. Reston, VA: U.S. Geological Survey. Retrieved from http://nationalmap.gov/viewer.html %0D

Wohl, E. (2018, October 1). The challenges of channel heads. *Earth-Science Reviews*. Elsevier B.V. https://doi.org/10.1016/j.earscirev.2018.07.008

Ye, S., Li, H. Y., Huang, M., Ali, M., Leng, G., Leung, L. R., et al. (2014). Regionalization of subsurface stormflow parameters of hydrologic models: Derivation from regional analysis of streamflow recession curves. *Journal of Hydrology*, 519(PA), 670–682. https://doi.org/10.1016/j.jhydrol.2014.07.017

Chapter 4

Upscaling hillslope-scale subsurface flow to inform catchment-scale recession behaviour

4.1 Introduction

Hydrologic models are vital tools for understanding the complex interaction of water fluxes throughout a basin. In a hydrologic model, each hydrologic process must be characterized at a level of detail that makes it part of a holistic whole, with the objective of generating a cohesive understanding of the sources, sinks, fluxes, and stores of water within the basin. In this work, we focus on the characterization of subsurface flow in hydrologic models. The fundamental mass balance algorithms of subsurface flow depend strongly on the heterogeneity of the subsurface, and yet, integrated at the basin-scale, the subsurface has been successfully characterized as a lumped reservoir that releases water according to an abstracted, prescribed relationship, such as a power-law relationship (Tallaksen, 1995). This simplification of complex, heterogeneous subsurface physics into a convenient abstracted lumped reservoir deserves an investigation: Can the physics of subsurface flow be used to inform the simplified subsurface flow description in a hydrological model with minimal computational cost? In contrast to the robust (but computationally expensive) practice of coupling hydrologic and hydrogeologic algorithms, this work is motivated to close the gap between subsurface physics and semi-distributed hydrological models by other means. That is, we conceptualize this problem as an upscaling problem, in which careful consideration of the aggregate behaviour of finely resolved hillslope physics can be used to inform the basin-scale power-law representation of hillslope drainage. Here, we use the term 'upscaling' to refer to the translation of the hillslope-scale physics of subsurface flow to an aggregate basin-scale relationship. The end result of such an upscaling exercise is a set of relationships that can predict the coefficients of the basin-scale power-law relationship directly from the parameters controlling hillslope-scale physics, which in this work are readily derived from topographic information. As such, the upscaling exercise provides value to hydrologic models by providing an estimate of the coefficients of the subsurface power-law reservoir without calibration, while also providing a capacity to estimate recession characteristics in ungauged basins where topographic metrics are the few readily available parameters.

The treatment of subsurface flow as a reservoir parameterized by a few coefficients is well-justified in the literature. Brutsaert and Nieber (1977) were the first to demonstrate that the observed recession behaviour in a hydrograph could be reasonably approximated by a power-law relationship:

$$-\frac{dQ}{dt} = c_1 Q^{c_2} \tag{4.1}$$

where Q represents flow [L/T], t represents time [T], and c_1 [1/T] and c_2 [-] represent the power-law coefficients. Critically, Brutsaert and Nieber (1977) linked this large-scale reservoir to analytical solutions of the Boussinesq equation for one-dimensional, unconfined, saturated groundwater flow through a homogenous medium. In this foundational approach, the conductivity of the single subsurface reservoir could be derived by calibrating the power-law solution of this equation to the observed power-law coefficients derived from the basin hydrograph. Many studies demonstrated the value of this single-reservoir approach by repeated application in real-world basins (Zecharias and Brutsaert, 1988; Vogel and Kroll, 1992; Brutsaert and Lopez, 1998; Huyck et al., 2005). However, several studies raised concerns with the simplified mathematics used in this approach (Tallaksen, 1995; Rupp et al., 2006; Bogaart et al., 2013). Others demonstrated that the large-scale response of a basin can be reproduced by the aggregate response of a number of independent conceptual subsurface reservoirs acting in parallel (Moore, 1997; Clark et al., 2009; Harman et al., 2009; Gao et al., 2017). More recent studies have attempted to synthesize these two ideas by analyzing the response of a larger number of independent subsurface reservoirs utilizing more robust subsurface physics (Ali et al., 2013; Ranjram and Craig, 2021). This work follows Ranjram and Craig (2021), which demonstrated that the hillslope-storage Boussinesq (hsB) equation (Troch et al., 2003), characterizing the unconfined, saturated flow through a hillslope with variable width, could be applied to a basin discretized into individual hillslope reservoirs, and these responses could be aggregated to successfully replicate the observed recession behaviour in a basin. Critically, the authors demonstrated that the observed behaviour could be reproduced from a time series of recharge, as opposed to assuming that the recession behaviour of subsurface units was consistent with the steady-state response, a common assumption in previous approaches. The flexibility of generating subsurface flow responses to a recharge time series reflects the relatively recent insights in the literature that a single-recession response derived through the entire time history of the basin may ignore critical transience in recession behaviour (Shaw and Riha, 2012; Thomas et al., 2015; Chen and Krajewski, 2016; Karlsen et al., 2019). However, the notion of recharge-dependent recession events is a significant departure from the original conceptualization of basin-scale recession that is difficult to incorporate into a predictive modelling framework: individual recession events can be extracted from a time series of streamflow, but predicting recession behaviour generated from a time series of recharge is a more challenging problem.

In this work, we provide a novel bottom-up upscaling justification for the transient power-law recession behaviour of basins. Basins are conceptualized as a network of independent hillslope units producing saturated, unconfined subsurface flow according to the hsB equation. Thirty CAMELS basins (Addor et al., 2017) are used to inform upscaling relationships which enable the prediction of the aggregate recession response of a basin as a function of distributed hillslope metrics. The efficacy of these predictive relationships in representing the transient recession response of a basin is then demonstrated. Twenty additional CAMELS basins are used to verify the utility of the upscaling relationships. A single basin is examined in more detail to demonstrate the efficacy of these conclusions. This work thus considers three specific objectives:

1. Derive upscaling relationships that can estimate transient subsurface-flow-induced recession behaviour using only topographic characteristics of a basin and information about the time series of recharge.
2. Generate rules for scaling the aggregate power-law response of a basin under different homogeneous, effective basin-scale conductivity values.
3. Demonstrate the efficacy of the upscaled recession model via comparison to observed transient recession behaviour in a single basin.

The utility of these upscaling relationships is in their capacity to estimate the transient recession behaviour without requiring an analysis of streamflow data. Rather, a reasonable estimate of aggregate recession behaviour can be extracted with readily derived topographic metrics and an estimate of the time history of recharge and subsurface properties.

4.2 Methodology

In this work, we attempt to bridge the gap between the physics-based response of an array of hillslopes and the classical power-law recession curve at the basin scale by treating the subsurface as a physically-informed space and using simulation results from such a framework to inform the large-scale lumped reservoir concept. Here, the subsurface is conceptualized according to the hillslope-storage Boussinesq (hsB) equation (Troch et al., 2003) which assumes saturated, unconfined flow through hillslopes of variable width. The hsB characterizes the complex three-dimensional converging or diverging flow nets within a variable width hillslope (either pinching off downslope to a smaller width, or expanding downslope to a larger width) using a one-dimensional partial differential equation:

$$f\frac{\partial h(x,t)}{\partial t} = -\frac{1}{w(x)}\frac{\partial Q(x,t)}{\partial x} + N \tag{4.2}$$

$$Q(x,t) = -w(x)h(x,t)K\left(\cos\theta\frac{\partial h(x,t)}{\partial x} + sin\theta\right) \tag{4.3}$$

where x is the distance from the downslope (streamside) end of the hillslope [L], t is time [T], θ is the hillslope bedrock slope angle [°], h is groundwater head relative to bedrock [L], K is hydraulic conductivity [L/T], $w(x)$ is the hillslope width function characterizing the variation of width along the hillslope [L], Q is the flow through the hillslope [L³/T], N is the uniform recharge rate [L/T], and f is porosity [-]. This work uses the computationally efficient hsB Proxy (Ranjram and Craig, 2021) to produce solutions to the hsB for wedge-shapes hillslopes (i.e., those with linear width functions). The hsB applies to individual hillslope elements, and so we conceptualize the subsurface as a collection of independent, parallel reservoirs releasing water to the basin outlet. The hillslope elements are derived directly from topography, as detailed in Ranjram and Craig (2021). Extracting hillslopes from a basin results in distributions of four hillslope properties: the length (L) along the primary axis; the constant bed slope of the hillslope (θ); the width at the downslope end of the hillslope (W_b); and a non-dimensional parameter representing the upslope width of the hillslope as a fraction of the downslope

width (X), which can also be considered a measure of the degree of hillslope convergence (upslope width greater than downslope width, $X > 1$) or divergence (upslope width smaller than downslope width, $X < 1$). Solving the hsB at all hillslopes produces individual flow profiles that can be aggregated to produce a basin-scale subsurface flow response. In this way, we can upscale the smaller (hillslope) scale subsurface flow response into a larger (basin) scale recession response.

Thus, there are three general controls on the basin-scale subsurface-induced recession response in this conceptualization: (1) the topography of the basin, which informs the distribution of hillslope properties; (2) the recharge time series applied in the basin, which provides the boundary condition inducing subsurface flow; and (3) the subsurface properties, specifically the porosity and hydraulic conductivity of hillslopes. To generate an upscaling relationship, we require an understanding of the influence of each of these components on aggregate recession response.

4.2.1 The basin signature

The problem is first simplified by evaluating the topographic controls of the basin in isolation; this is achieved by an experimental set up in which a single recharge event (a daily rate applied over a single day) is applied to a basin with homogenous subsurface properties (hydraulic conductivity = 1 m/hr, porosity = 0.3) and the aggregate response derived. To produce a robust experiment, thirty basins from the CAMELS database are evaluated; benefiting from the detailed compilation of forcing data, streamflow data, and basin boundary shapefiles within this database. These thirty basins represent the smallest basins in the database by area (4.1-19.6 km^2), which provides two benefits: first, as a potentially useful statistical control in case basin area is a powerful distinguishing feature of aggregate response between basins; and second, to maximize the appropriateness of spatially homogenous precipitation data.

The behaviour of a basin is summarized by the aggregate recession response of the basin to a single recharge event, determined by solving the hsB at all hillslopes in the basin extracted from digital elevation data. This response is divided into three components: the initial flow value, a fast-phase power-law recession response, and a slower-phase power-law recession response, with the transition between the two phases chosen as when the fast-phase flow response reaches 50% of the initial value (Figure 4.1a). The slow-phase is assumed to end when the flow reaches 1% of the initial flow value. In this work, we herein refer to this response as the 'signature' response, as it is unique to each basin and so is a useful metric for distinguishing basin behaviour. We here chose to use two phases for the signature response as a single power law fit consistently led to discrepancy with the derived basin-scale signature (e.g., Figure 4.1b). The signature response is thus summarized by five coefficients: the initial flow value and the two coefficients each for the fast- and slow-phase recession power laws.

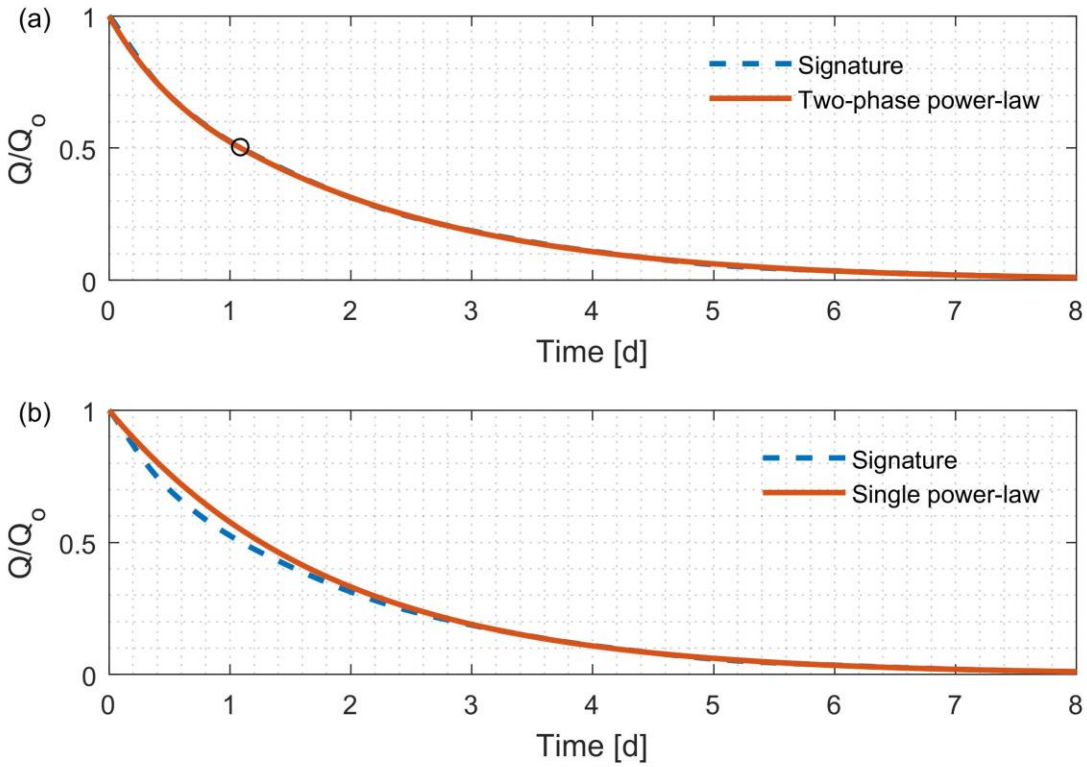

Figure 4.1. The signature response. Blue dashed lines indicate numerically-derived signature response. Red line indicates the signature response as represented by power-law functions. (a) An example of the signature response and its representation using a two-phase response, with the 50% flow value marked. (b) An example of the signature response and its representation using a single power-law.

Solving for the signature response in a basin across a large number of recharge events reveals a clear structure between the five signature coefficients and recharge magnitude, as follows: the initial flow value is a fixed fraction of the recharge magnitude; the exponential (c_2) coefficients of the two power-law responses are insensitive to recharge magnitude and hence are controlled solely by basin topography; and the multiplier coefficients (c_1) of the two power-law responses exhibit a log-linear variation with recharge magnitude. An example of this variation is presented in Figure 4.2.

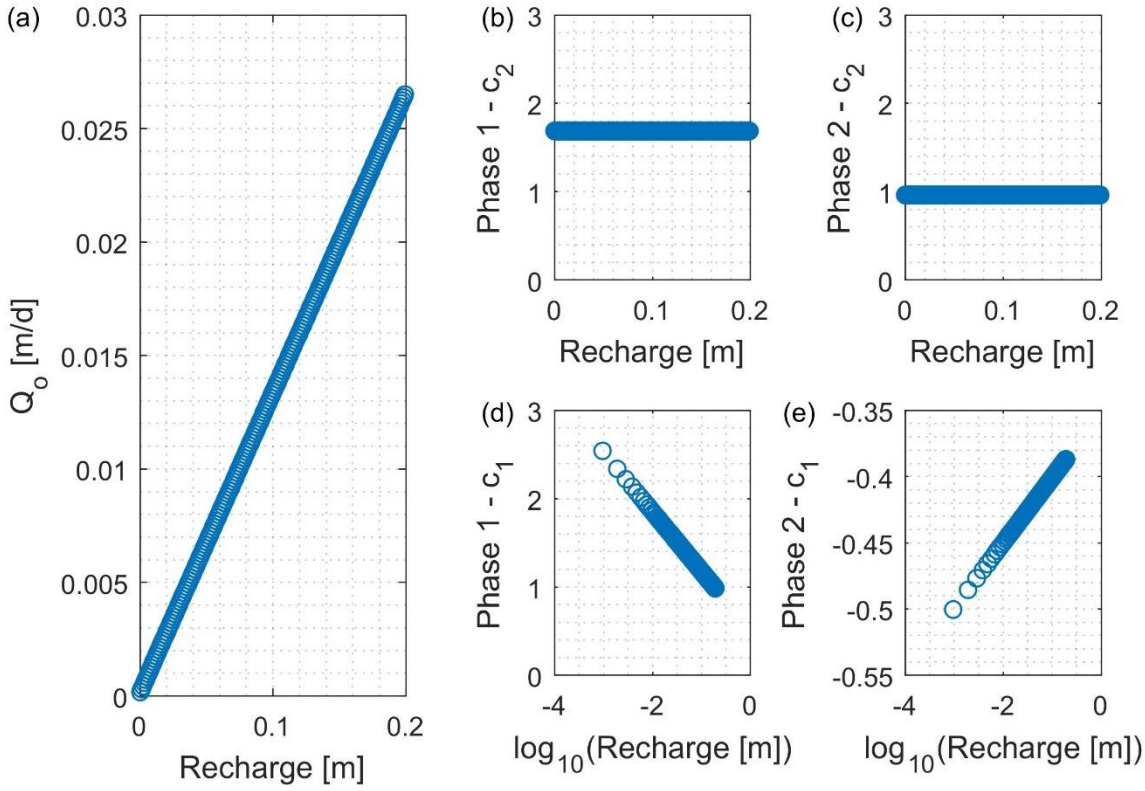

Figure 4.2. Variation of Qo and the coefficients of the two power-law responses with recharge for a single basin with homogenous subsurface properties. Each point represents the coefficient derived from the signature response produced by the associated daily recharge rate applied over a single day.

The log-linear c_1-recharge responses ($c_1 = c_{1,m} \log R + c_{1,b}$, Figure 4.2d and 4.2e) exhibit a useful secondary structure: the slope of these log-linear relationships is a fixed function of the associated fast- and slow-phase c_2 value ($c_{1,m} = 1 - c_2$). As such, the unique c_1 value at a given recharge magnitude can be derived using two recharge-independent parameters: $c_{1,b}$ and the associated c_2 value. Thus, the five coefficients of the signature response can be derived for any single-day recharge event using five recharge-independent parameters: (1) the slope of the Q_o-recharge relationship (Figure 4.2a), herein referred to as the Q_R parameter; (2, 3) the two fixed c_2 values unique to the fast-phase and slow-phase power-law responses (Figure 4.2b,c); and (4,5) the two fixed $c_{1,b}$ values unique to the fast-phase and slow-phase power-law responses. Because these five parameters are independent of recharge, and because this experimental set up uses identical, homogeneous subsurface parameters for all hillslopes in the basin, these five signature parameters are a valuable set of independent variates to test predictive relationships that can translate topographic properties of a basin to recession behaviour. That is, this converts the problem of mapping basin properties to recession properties into one of mapping basin properties to signature parameters. Herein, the use of 'signature parameters' refers to this set of five topography-dependent, recharge-independent variates.

The efficacy of the signature approach is tested as follows. First, the facsimile of the signature response generated by the derived signature parameters is compared to the numerical-modelling-

derived signature response (i.e, the comparison in Figure 4.1a, now quantified across thirty basins and hundreds of recharge events). This simple test of the single recession response to a single recharge event evaluated for two-hundred individual recharge events (1 – 200 mm/d) across the thirty test basins is used to demonstrate that the treatment of the signature response by the five signature parameters is appropriate. Second, the efficacy of the signature parameters in producing reasonable facsimiles of the transient subsurface stormflow response induced by a time series of recharge inputs to a basin is tested. To ensure a reasonable time series of recharge, thirty independent synthetic recharge time series are tested, each extracted from a single year of precipitation input from each of the thirty CAMELS basins, filtered to extract only warm-weather precipitation (temperatures greater than 0°C), with the assumption that the precipitation and recharge rates are identical. Although this one-to-one equivalency will lead to an exaggeration of recharge magnitudes, the timing of recharge events will be appropriately correlated to the timing of precipitation events. The expected transient response is derived by numerical modelling, and the signature-derived response is constructed by superposition of individual signature responses to the time series of recharge events.

High quality agreement between numerical modelling results and the responses generated by the five basin-specific signature parameters will indicate that the signature parameters can reasonably replicate the expected transient subsurface stormflow response in a basin. With this baseline established, upscaling relationships can be derived by mapping basin properties to signature parameters. The upscaling relationships are conceptualized as linear models of aggregate hillslope parameters, and the quality of these models is defined by their ability to produce high quality agreement (NSE>0.9) to the numerically-derived aggregate transient subsurface stormflow response in a basin. More robustly, the efficacy of these upscaling relationships is evaluated in a test of twenty validation basins which did not inform the derivation of the upscaling relationships.

With upscaling relationships in hand reasonably capable of replicating the expected transient subsurface stormflow response, a simple scaling behaviour is evaluated. This scaling behaviour enables the signature response associated with the reference conductivity (1 m/hr) used to derive the upscaling relationships to be scaled to any other homogeneous conductivity value.

4.3 Results

The efficacy of the signature approach is first demonstrated by simply comparing the signature response to a single recharge event derived by numerical modelling to the signature response produced by the extracted signature parameters, as illustrated by the x markers in Figure 4.3. In such a comparison, the derived signature parameters produce a signature that replicates the simulated signature responses at very high NSE values (> 0.97) across two-hundred simulated recharge events (1 – 200 mm/d). Note that the NSE value comparing the two signature values are insensitive to the recharge magnitude, and hence only a single NSE value is reported for each of the two-hundred simulated signature responses. Two outlier basins, Basin 7 and 22, exhibit lower NSE values than the rest of the basin (<0.99). These basins also exhibit the 'slowest' signature behaviour, with two of the

lowest Q_R values (<0.03) and with slow-phase c_2 coefficients less than one, indicating a much lower rate of recession at a given flow rate as compared to other basins in the data set.

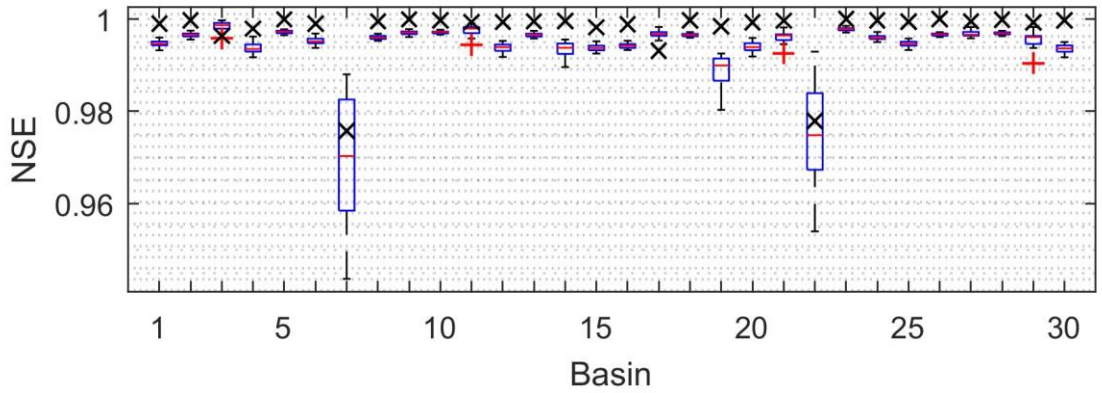

Figure 4.3. Signature performance as compared to numerical solutions. x symbols indicate the NSE value comparing the signature response generated by numerical simulation against the signature response generated by the five signature values. Box-and-whisker plots indicate the distribution of 30 NSE values comparing the numerical solution to the signature-derived hydrograph for 30 individual recharge time series

The second, more robust evaluation of the signature approach is to test the efficacy of the signature parameters in predicting the hydrograph generated in response to a time series of recharge inputs. Such a test evaluates the quality of a superposition of transient signature responses. Thirty time series of recharge are independently applied to each CAMELS basin, with recharge inputs derived from CAMELS precipitation data, as described previously. The metrics of this comparison are the distribution of thirty NSE values in each basin comparing the numerically-derived transient response and the signature-derived transient response generated in response to each of the thirty recharge time series. The distributions of these NSE values in each basin are presented in Figure 4.3. The high quality of NSE values (>0.98 in all basins except the previously highlighted basins 7 and 22, which are >0.94) across 30 unique time series of recharge verify that the signature response can reasonably approximate the numerically simulated response, even for a continuous time series of recharge. Thus, the ability to predict the five signature parameters using information about topography alone would allow for the generation of high-quality facsimiles of numerically-derived recession responses without solving the groundwater flow equation: such predictive relationships are therefore capable of translating topographic metrics into an upscaled, aggregate basin-scale subsurface stormflow response.

4.3.1 Upscaling relationships

To understand the relationship between the signature parameters and basin topography, the correlation between each of the five signature parameters and aggregate descriptions (e.g., mean, sum, and cumulative probability values) of the four hillslope parameters controlling hillslope-scale subsurface flow (e.g., length, bed slope, downslope width, and convergence-divergence parameter) was evaluated. This analysis was completed by deriving linear regression models capable of predicting signature parameters which produce high quality NSE values (>0.9) across the thirty time series of recharge inputs, as compared to the numerically-derived hydrograph. That is, the basin recession is

now compared against upscaling models driven solely by topographic characteristics of the basins, but using the same basins the regressions were trained upon, akin to calibration.

The derivation of upscaling relationships originates from the very strong log-log linear correlation (R^2 = 0.993, p << 0.05) between the Q_R parameter and mean bed slope in each basin:

$$\log_{10} Q_R = -2.01 + 0.845 \cdot \log_{10} E[\theta] \tag{4.4}$$

where $E[\theta]$ is the mean hillslope angle in degrees. Using this simple relationship to predict the Q_R parameter in each of the thirty basins results in a high quality of agreement between the thirty predicted and simulated recession hydrographs produced from the thirty time series of recharge inputs. The efficacy of this simple univariate regression model led to a structured evaluation of a large number of regression models, each relating the remaining four signature parameters to a single descriptor of the aggregate hillslope response (specifically, the mean parameter values and the cumulative distribution function (CDF) probabilities of thirteen to twenty-four threshold values for each of the four hillslope parameters, with the threshold values determined by considering the range of derived values in each percentile of the basin CDFs). However, none of the derived linear models could adequately predict signature parameters to the necessary quality. Rather than blindly regressing a large number of multivariate linear models to attempt to extract a high-quality response, the analysis instead transitioned to capturing the recession behaviour of the slow-phase response, which is inherently responsible for a longer portion of the recession time series and hence contributes more information to the agreement measured by the NSE.

To conceptualize the behaviour of basins in a more tangible manner, a new metric, 'hillslope speed' was derived. Hillslope speed is a second-order topographic metric (informed by topography but not directly extractable from digital elevation data), defined as the time it takes each hillslope in a basin to drain to 1% initial flow following a 10 mm/d recharge event over a single day at the reference basin conductivity of 1 m/hr (such that basins which skew to earlier time release water faster and thus have a higher "speed"). Characterizing hillslope speed requires a single preparatory numerical simulation, but the computational cost is comparable to the cost of deriving hillslopes from elevation data. An additional metric, ΣW_b was also introduced, being the sum of the downslope widths of all hillslopes, itself then equivalent to twice the distance of the overland flow network plus the width of any upstream headwater hillslopes. Careful analysis of the slow-phase coefficients and the associated characteristics of the hillslope distributions, including the new hillslope speed distribution and the ΣW_b parameter, led to the derivation of two useful linear models:

$$c_{1,b}^{PS} = -6.69 + 0.602 \cdot P[Sp \leq 5d] + 1.23 \times 10^{-6} \cdot \Sigma W_b + 6.05 \cdot P[X \leq 4] \tag{4.5}$$
$$c_2^{PS} = -1.50 - 0.246 \cdot P[Sp \leq 5d] + 4.35 \times 10^{-7} \cdot \Sigma W_b + 2.83 \cdot P[X \leq 4] \tag{4.6}$$

where the superposed label PS indicates slow-phase; $P[X \leq 4]$ is the probability of hillslope X values less than or equal to four; and $P[Sp \leq 5d]$ is the probability of hillslope speed values less than or

equal to five days. These regression models work in conjunction with the Q_R model to produce high quality comparisons in all basins except for Basins 3, 7 and 22. These three basins are outliers in terms of hillslope speed, with $P[Sp \leq 5d]$ of 3%, 0.9%, and 1.4%, respectively, compared to the fourth slowest basin which has $P[Sp \leq 5d] > 20\%$. Noticeably, these three basins also exhibit the smallest mean slopes in the experimental set of basins, both in terms of the derived hillslope properties (<2°) and the CAMELS metric for catchment mean slope (< 2.6 m/km). Thus, although a useful slow-phase predictive model could not be produced which accounts for these basins, these basins provide an obvious exclusion rule for the predictive model: any basin with mean hillslope angle less than two degrees cannot be simulated with these upscaling relationships. Further, because these basins are so shallow in terms of topographic gradient, there is useful conceptual context for the difficulty of the upscaling relationships to account for their behaviours: with a subtle topographic gradient, a driving head gradient may not occur, and the mechanism of subsurface stormflow may be unimportant in such basins.

The derivation of a fast-phase upscaling relationship proved most challenging. In evaluating the fast-phase coefficients, no regression model was able to predict these values at high quality: where linear model quality was sufficient, the prediction produced by these models in conjunction with the previous models (Equations 4.4-4.6) was not. However, the failure of linear regression in producing viable coefficients was related to a more over-arching difficulty: the challenge of assigning a unique set of power-law coefficients to a response than can be almost linear. Figure 4.4 illustrates the simulated and predicted behaviour of two basins with the worst agreement between solved fast-phased coefficients and the coefficients predicted by a set of simple linear regression models:

$$c_{1,b}^{PF} = -0.225 + 0.0103 \cdot E[\theta] \tag{4.7}$$

$$c_{2}^{PF} = 1.42 - 0.0103 \cdot E[\theta] \tag{4.8}$$

where the superposed label PF indicates fast-phase. These images demonstrate that even a major disagreement between fast-phase coefficients can have a minor impact on overall fit. Note that attempts to simulate the fast-phase as a linear response were untenable, as some basins require a concave response shape to ensure appropriate selection of the 50% flow value transition to the slow-phase. Indeed, these simple predictive models (Equations 4.7 and 4.8) were found to produce very high-quality predictions of simulated recession behaviour when combined with Equations 4.4-4.6, demonstrating that a simple power-law response centred at $c_{1,b}^{PF} = -0.225, c_{2}^{PF} = 1.42$ and varying according to a fixed fraction of mean slope can produce coefficients which suitably represent the fast-phase response. That is, these models represent a compromise between the regression approach which provides estimates of these coefficients as informed by the last-squares methodology of the regression process, and the practical desire to produce a predictive model capable of generating a good match to the numerically-derived hydrographs. Because derived power-law coefficients can deviate largely but still produce a similar fast-phase response, the fast-phase coefficients are predicted by simple linear models as a function of mean basin slope that exhibit unsatisfactory regression characteristics, but

produce a signature response that can be used to approximate the derived hydrograph at a high quality when combined with the previous regression models.

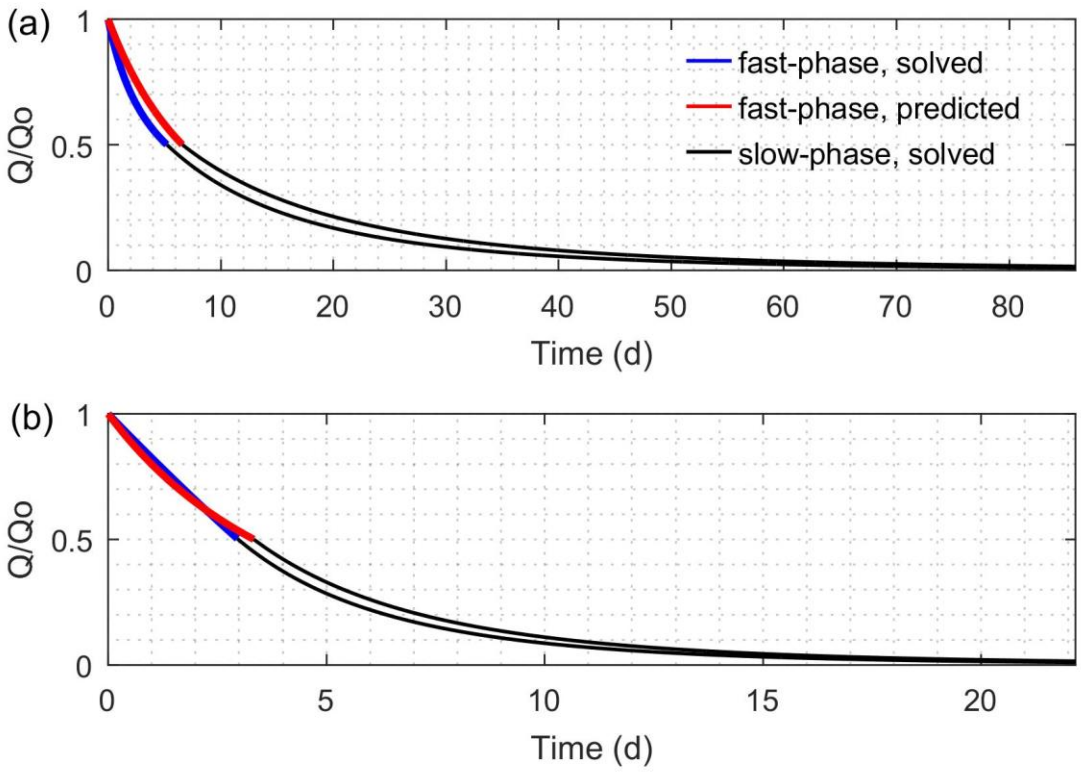

Figure 4.4. The challenge of characterizing fast-phase power laws. These curves illustrate the signature response of (a) Basin 11 and (b) Basin 12, using the fast-phase coefficients derived from the numerical solution (blue) and the fast-phase coefficients derived from the upscaling models (red). Basins 11 and 12 have the largest discrepancy between solved and predicted coefficients. Black lines illustrate the fast-phase response using the coefficients derived from numerical solutions.

The efficacy of the upscaling relationships (Equations 4.4-4.8) is first evaluated by their capacity to predict the numerically simulated transient hydrograph behaviour, illustrated in Figure 4.5 Here, twenty-seven basins exhibit NSE values > 0.96 across the thirty recharge time series, and the three outlier basins, based on their exceptionally shallow topographic gradient, are illustrated for context. However, the validity of these models is more robustly evaluated by their capacity to approximate the behaviour of a set of basins that did not inform these predictive models. Figure 4.6 demonstrates the predictive capacity of Equations 4.4-4.8 in twenty additional CAMELS basins (areas 19.7-28 km^2) not used for training the upscaling models. The high quality of agreement between the predicted and simulated hydrographs (NSE > 0.92; hydrograph of lowest NSE value illustrated in Figure 4.6b) demonstrates the utility of the topography-predicted signature coefficients and is a powerful indicator of the validity of these models in approximating derived behaviours.

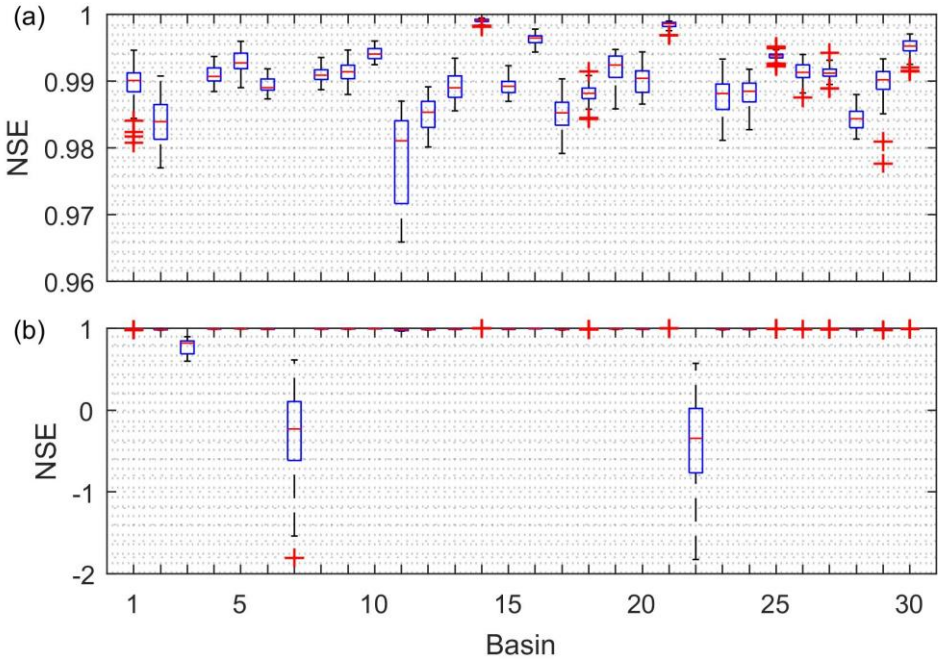

Figure 4.5. Predictive performance of the upscaling relationships in thirty basins used to generate the relationships. (a) the distribution of NSE values comparing the numerical solution to the signature-derived hydrograph with parameters obtained from the upscaling relationships in Equations 4.4-4.8, for thirty individual recharge time series. (b) The same plot extended to the poorer predictive extents of Basin 3, 7, and 22.

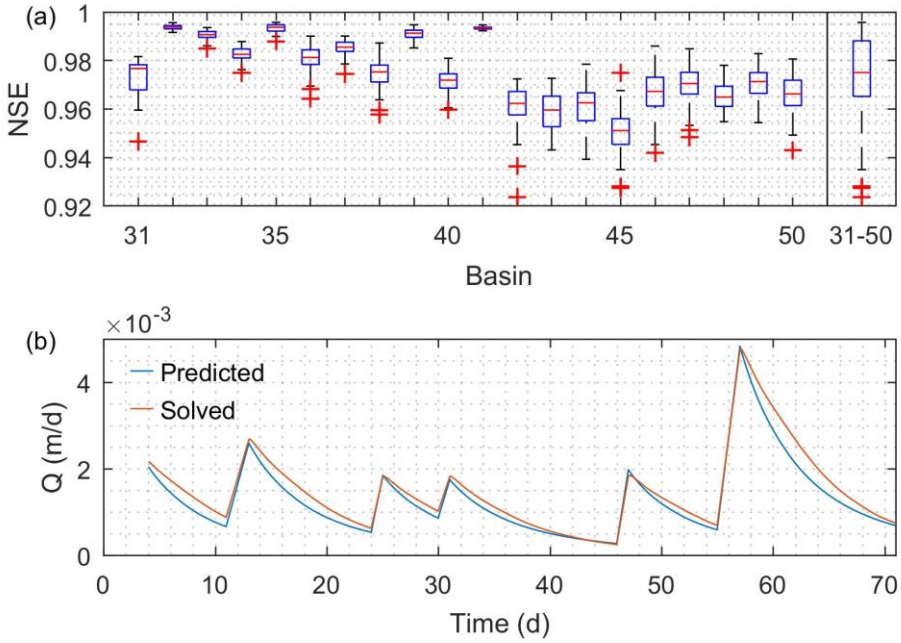

Figure 4.6. Predictive performance of the upscaling relationships in twenty validation basins. (a) Box-and-whisker plots indicate the distribution of NSE values measuring the quality of agreement between numerically simulated and predicted hydrographs resulting from thirty individual recharge time series inputs, in a set of twenty validation CAMELS basins (labelled basins 31-50, to distinguish from Basins 1-30 used to generate the upscaling relationships). The collective NSE distribution across basins 31-50 is also illustrated for reference. (b) The hydrograph associated with the lowest NSE value (Basin 42, NSE = 0.924).

4.3.2 Effective conductivity scaling

The recharge-independence of the signature coefficients and appropriateness of superposition of the hsB response (Ranjram and Craig, 2021) enables the application of the upscaling relationships to any recharge time series in a basin with mean slope greater than two degrees. The final control on the recession response, the subsurface conductivity, is thus left to be evaluated. To simplify the analysis, the control of hydraulic conductivity alone is considered, as opposed to a more complex accounting of the combined control of conductivity and porosity. In addition, we focus on basin-scale effective conductivity, implying a single homogeneous conductivity value at all hillslopes. Although the heterogeneity of conductivity is a powerful control on large-scale recession behaviour (e.g., Harman et al., 2009), the added complexity of accounting for such heterogeneity is considered beyond the scope of this work.

When the basin subsurface is homogeneous, the signature response exhibits a well-behaved scaling behaviour relative to the behaviour at the reference conductivity of 1 m/hr, as follows,

$$Q_K = Q_{sig}\left(K^{1-0.0053 \cdot E[\theta]}\right) \tag{4.9}$$

$$t_K = t_{sig}(K^{-1}) \tag{4.10}$$

where K is the new effective hydraulic conductivity (m/hr); Q_K and t_K are the new scaled flow and time values; Q_{sig} is the default signature response at K = 1 m/hr; and t_{sig} is the associated time series of the default signature response. These relationships demonstrate a simple vertical and horizontal scaling proportional to the new conductivity, where the magnitude of vertical scaling is also dependent on the mean slope of the basin. The efficacy of this scaling function is demonstrated by comparing the simulated signature response to a single recharge event (10 mm/d) at various homogenous conductivity values with the signature predicted by the scaling rules across the set of forty-seven basins with mean slope greater than two degrees (Figure 4.7, NSE > 0.965). Note that the tested homogenous conductivity values range from 0.05-0.75 m/hr. Although this is a limited subset of conductivity values, the resultant NSE distributions in Figure 4.7 do not exhibit a trend with conductivity, suggesting that the scaling rules hold across conductivity values. Additionally, at K = 0.05 m/hr the timelines of the signature response suggest a significant disconnect between applied recharge and the signature response as indicated by the time to 1% initial flow drainage ranging from 90-1800 days (median of 218 days) across the 47 basins. The linear scaling in Equation 4.10 makes it clear that lower homogeneous conductivity estimates would result in even longer recession timelines, implying that basin recession response is not controlled by the saturated, homogenous subsurface flow mechanics utilized in this work when conductivity values are low.

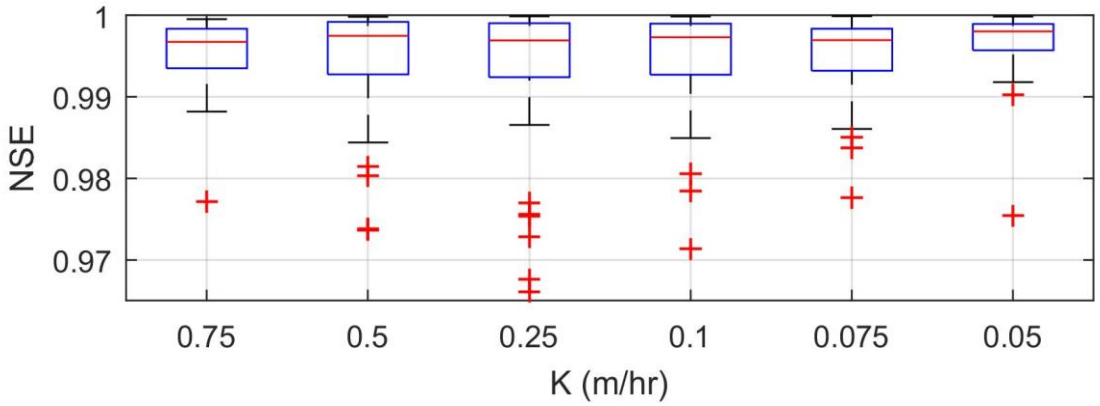

Figure 4.7. Efficacy of scaling rules (Equations 4.9 and 4.10). NSE values compare the simulated basin signature response and the predicted signature response using the scaling rules for a single recharge event (10 mm/d). The box-and-whisker plots summarize the distribution of NSE values across 47 basins at various effective conductivity values.

The validity of the upscaling relationships and the effective conductivity scaling rules have thus far been demonstrated in the context of numerical simulation. In what follows, these models are compared against observed recession behaviour.

4.3.3 Comparison to observed behaviour

To more robustly analyze the quality of the derived upscaling relationships, the observed recession behaviour of a single CAMELS basin is evaluated. The Sevenmile Run basin near Rasellas, Pennsylvania (labelled 'Basin 33' in the pool of experimental basins) is a forested catchment with a catchment mean slope of 15.8 m/km receiving an average of 849 mm of precipitation per year (Addor et al., 2017), and has the following derived signature topographic metrics: $P[X \leq 4]$ = 0.986, $P[Sp \leq 5]$= 0.395; ΣW_b= 318,450 m; and $E[\theta]$ = 6.53°. Thirty years of observed streamflow recessions in this basin are illustrated as black points on a log-log plot of flow (Q) and the first derivative of flow in time (-dQ/dt, the rate of recession) in Figure 4.8. Figure 4.8 also depicts the single recession function regressed through the entire data set (in green; Equation 1), representing the classical approach (Brutsaert and Nieber, 1977) to extracting subsurface-induced recession coefficients, as well as 1,788 transient recession events (in red), extracted according to the threshold used by Karlsen et al. (2019), itself a modification of the recession criterion in Shaw and Riha (2012). This threshold only considers recession events lasting greater than seven days, and excludes the first two days and last day of data when regressing power-law coefficients.

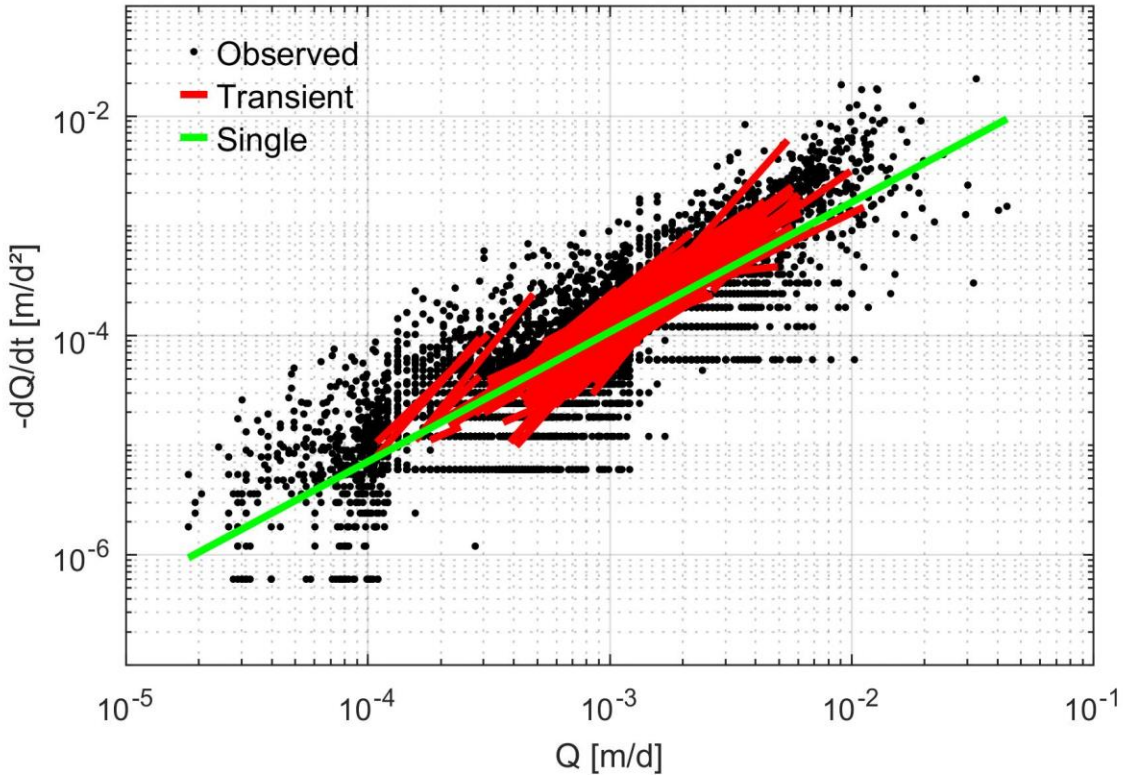

Figure 4.8. Thirty years of observed recession behaviour in the Sevenmile Run basin. Individual observed recession pairs illustrated in black. The single recession power-law (Equation 4.1) derived through the entire observed data set is illustrated in green. Individual transient recession power-laws illustrated in red

The efficacy of the upscaling relationships is here evaluated by their capacity to produce recession coefficients that respect the observed recession behaviour. To remove any bias towards our signature conceptualization, we extract single-phase recession responses, as opposed to extracting a fast- and slow-phase response. Thirty years of warm-weather precipitation data (i.e., precipitation data associated with atmospheric temperatures greater than 0°C) are used as an input recharge time series, and the upscaling relationships are used to produce recessions with the filtering rule of Karlsen et al. (2019) applied to ensure consistency. The efficacy of these predicted recessions is demonstrated by the quality of agreement between transient recession coefficients, presented in Figure 4.9. The signature-assumed conductivity (1 m/hr) produces minimal agreement between observed and predicted recession behaviour, but the agreement can be improved by a manual calibration of the effective conductivity to 0.5 m/hr, also illustrated in Figure 4.9.

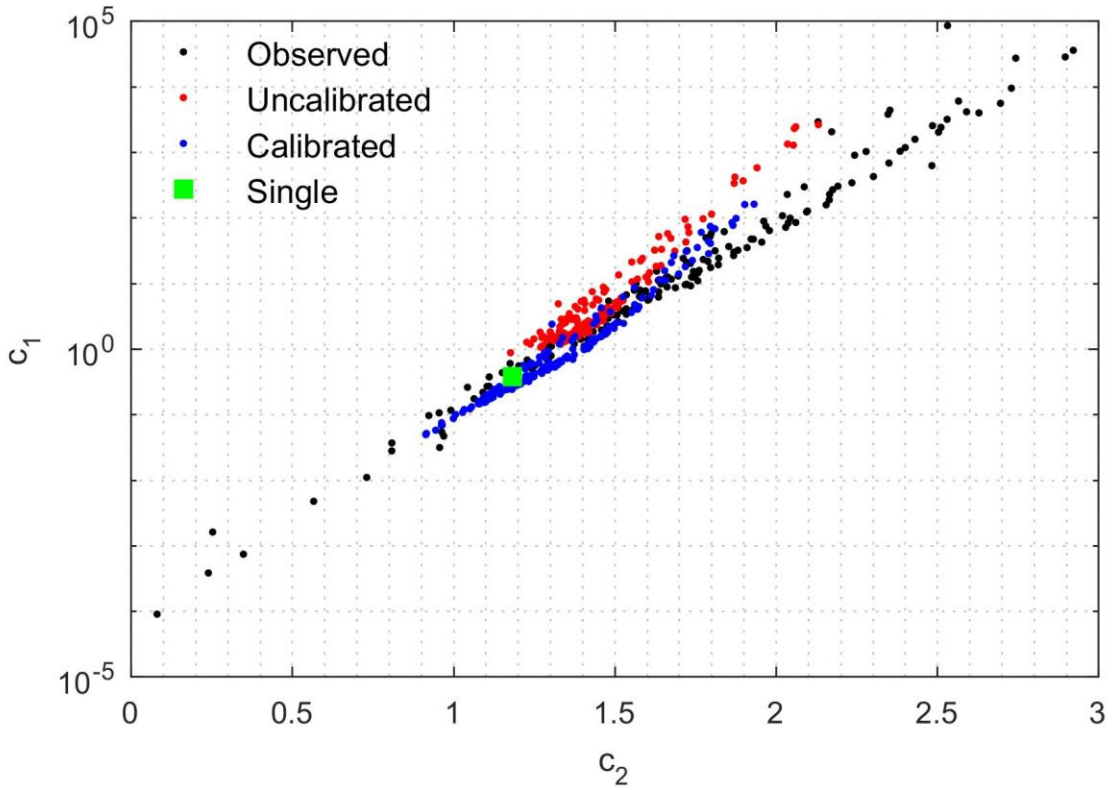

Figure 4.9. Transient recession coefficients in the Sevenmile Run basin. Black points indicate observed transient recession coefficients, derived according to the threshold established by Karlsen et al. (2019) and Shaw and Riha (2012). The single recession function derived through all thirty years of recession data is illustrated in green. The red and blue points illustrate the transient recession coefficients extracted by the upscaling relationships, using the signature default (1 m/hr) and manually calibrated (0.5 m/hr) hydraulic conductivity, respectively.

Figure 4.9 demonstrates that the manually calibrated upscaled recessions respect the observed recession behaviour, as evidenced by the agreement between upscaled and observed transient recession coefficients, although the upscaling relationships do not capture the full range of observed transient recession coefficients. This deficit is consistent across other tested basins, with calibrated recession coefficients consistently unable to simulate the wider range of observed recession characteristics, as summarized by the transient c_1 and c_2 coefficients (Figure 4.10; note that three basins with less than ten recessions lasting longer than seven days are excluded from this analysis). However, the utility of the upscaling relationships is clearly apparent: the upscaled transient recession behaviour is produced using only information about topography, recharge, and conductivity, and hence, have particular value in predicting recession behaviour in ungauged basins (Hrachowitz et al., 2013) where it would be impossible to characterize observed behaviour. The inability of the upscaling relationships to generate recession coefficients that capture the full extent of observed recession coefficients is in part due to the limitations of Boussinesq-based reservoirs, which have previously been shown to produce c_2 coefficients between 0 and 2 (Harman et al., 2009). Conceptually, this suggests that the larger variability in observed transient recession behaviour may require reservoirs that are not limited to the homogenous, saturated subsurface flow mechanics utilized in the derivation

of the upscaling relationships. That is, other controls, including unsaturated mechanics, saturation-excess and infiltration-excess limits on recharge, and subsurface heterogeneity may thus be fundamental to characterizing this full range of variability, and such mechanics should be integrated into future upscaling exercises.

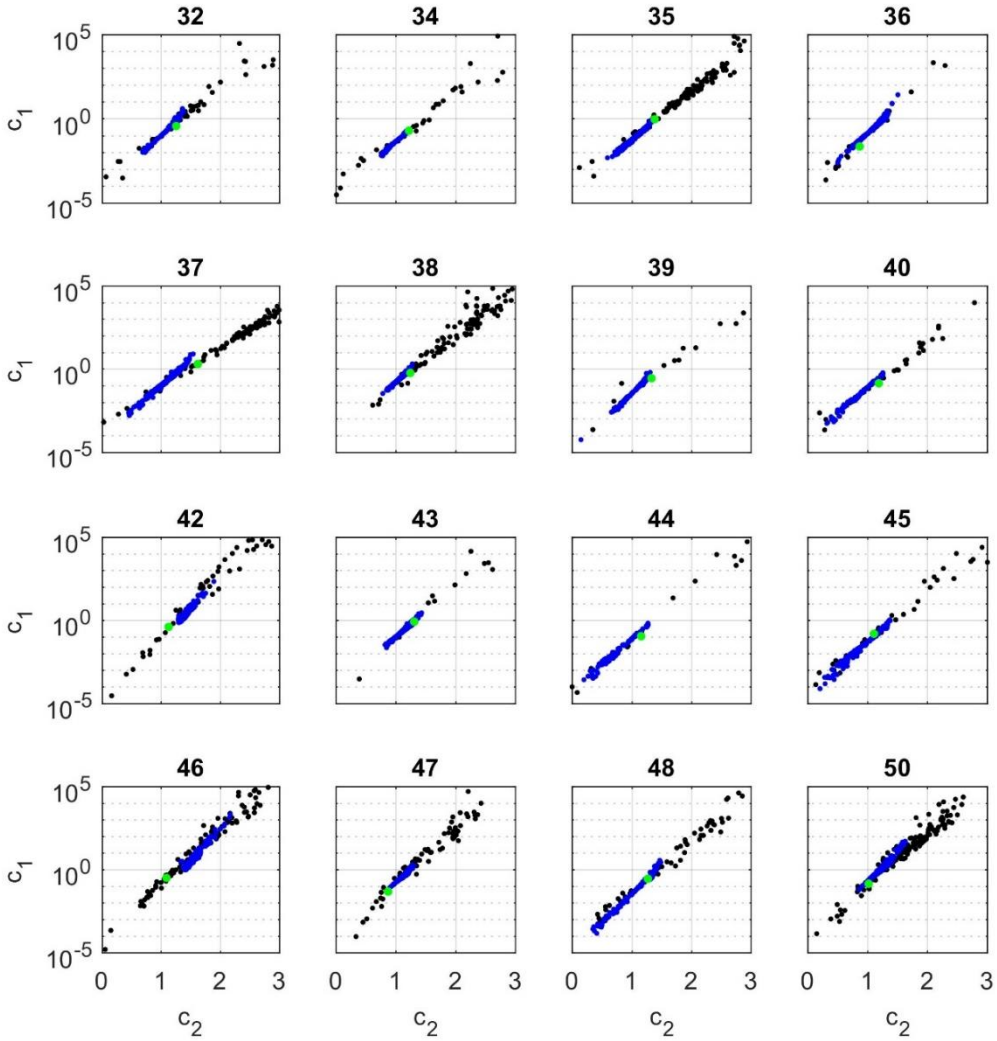

Figure 4.10. Observed and upscaled transient recession coefficients in sixteen verification basins. Black points indicate observed recession coefficients. Blue points indicate upscaled, calibrated recession coefficients. Green points indicate the single recession fit through entire data set. Basin number labels are indicated above each plot.

Although the upscaling relationships introduced here are unable to fully capture the range of transient recession coefficients observed in the tested basins, their capacity to produce coefficients that respect observed values represents a step forward in linking the small-scale heterogeneous response of subsurface reservoirs to the basin-scale aggregate recession response. Unlike earlier approaches, the upscaling relationships account for a diversity of reservoir shapes and sizes, and the resultant responses are based on subsurface physics, as opposed to abstracted lumped representations with no direct connection to landscape characteristics. These upscaling relationships enable the prediction of recession coefficients using only information about basin topography and an estimate

of recharge time history, as opposed to previous investigations into transient behaviour which must extract the non-unique recession responses from flow data. This makes the upscaling relationships particularly relevant in ungauged basins, where such flow information is unavailable. The relationships demonstrate a quantifiable correlation between mean basin slope, the distribution of converging-diverging hillslopes, the length of the surface drainage network, and the second-order hillslope speed metric, which are all straightforward to derive from topographic data.

4.4 Conclusion

A novel set of upscaling relationships have been developed that are capable of transforming simple metrics of basin topography, determined from the distribution of small-scale hillslope units, into a basin-scale description of subsurface-flow-induced recession behaviour. Critically, these upscaling relationships are independent of recharge magnitude, and so any time series of daily recharge can be converted into a set of transient recession coefficients. Additionally, simple scaling rules have been developed to handle changes in basin-scale effective conductivity. The following objectives have thus been achieved:

1. Upscaling relationships capable of producing subsurface-flow-induced recession behaviour in basins using only four topographic metrics (mean hillslope angle; the probability of hillslopes with divergence/convergence metric less than or equal to four; the probability of 'hillslope speed' less than or equal to five days; and the sum of all downslope hillslope widths) and a time series of recharge have been developed (Equations 4.4-4.8). These upscaling relationships successfully replicate numerical simulations when mean hillslope angle is greater than two degrees.
2. Simple scaling rules have been developed to handle the change in the upscaled behaviour as the homogeneous basin-scale conductivity is changed (Equations 4.9-4.10)
3. The efficacy of these predictive relationships has been demonstrated with respect to numerical simulations in twenty basins not used to train the upscaling relationships. The practical efficacy has been evaluated by comparing predicted transient recession behaviour to observed behaviour in seventeen basins within the CAMELS data set. Although predicted recession coefficients are within the range of observed coefficients, the full extent of observed recession coefficients are consistently not captured. This suggests that the mechanics of saturated subsurface flow through hillslopes with homogeneous subsurface properties used in this work are but a subset of the mechanics which generate transient variability in recession behaviour.

The upscaling relationships presented here may be potentially augmented to respect a variety of subsurface flow and lateral runoff mechanisms, and to attempt to account for heterogeneity in subsurface properties between hillslopes. These relationships will be ported into a hydrological modelling framework, to support a physical justification for the power-law coefficients used to simulate the subsurface-flow-induced recession behaviour at the basin scale.

4.5 References

Addor N, Newman AJ, Mizukami N, Clark MP. 2017. The CAMELS data set: catchment attributes and meteorology for large-sample studies. version 2.0. Boulder, CO: UCAR/NCAR. DOI: 10.5065/D6G73C3Q

Ali M, Fiori A, Bellotti G. 2013. Analysis of the nonlinear storage–discharge relation for hillslopes through 2D numerical modelling. *Hydrological Processes* 27: 2683–2690 DOI: 10.1002/HYP.9397

Bogaart P, Rupp D, Selker J, Velde YVD. 2013. Late-time drainage from a sloping Boussinesq aquifer. Water Resources Research 49: 7498–7507 DOI: 10.1002/2013WR013780

Brutsaert W, Lopez JP. 1998. Basin-scale geohydrologic drought flow features of riparian aquifers in the southern Great Plains. *Water Resources Research* 34 (2): 233–240 DOI: 10.1029/97WR03068

Brutsaert W, Nieber J. 1977. Regionalized drought flow hydrographs from a mature glaciated plateau *Water Resources Research* 13 (3): 637-643 DOI: 10.1029/WR013i003p00637

Chen B, Krajewski W. 2016. Analysing individual recession events: sensitivity of parameter determination to the calculation procedure. *Hydrological Sciences Journal* 61: 2887–2901 DOI: 10.1080/02626667.2016.1170940

Clark MP, Rupp DE, Woods RA, Tromp-van Meerveld HJ, Peters NE, Freer JE. 2009. Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale. *Hydrological Processes* 23 (2): 311–319 DOI: 10.1002/hyp.7154

Gao M, Chen X, Liu J, Zhang Z, Cheng Q. 2017. Using Two Parallel Linear Reservoirs to Express Multiple Relations of Power-Law Recession Curves. *Journal of Hydrologic Engineering* 22: 4017013 DOI: 10.1061/(ASCE)HE.1943-5584.0001518

Harman CJ, Sivapalan M, Kumar P. 2009. Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resources Research* 45 (9) DOI: 10.1029/2008WR007392

Hrachowitz M, Savenije HHG, Blöschl G, McDonnell JJ, Sivapalan M, Pomeroy JW, Arheimer B, Blume T, Clark MP, Ehret U, et al. 2013. A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological Sciences Journal* 58 (6): 1198–1255 DOI: 10.1080/02626667.2013.803183

Huyck AAO, Pauwels VRN, Verhoest NEC. 2005. A base flow separation algorithm based on the linearized Boussinesq equation for complex hillslopes. *Water Resources Research* 41 (8): 1–18 DOI: 10.1029/2004WR003789

Karlsen RH, Bishop K, Grabs T, Ottosson-Löfvenius M, Laudon H, Seibert J. 2019. The role of landscape properties, storage and evapotranspiration on variability in streamflow recessions in a boreal catchment. *Journal of Hydrology* 570: 315–328 DOI: 10.1016/J.JHYDROL.2018.12.065

Moore RD. 1997. Storage-outflow modelling of streamflow recessions, with application to a shallow-soil forested catchment. *Journal of Hydrology* 198 (1–4): 260–270 DOI: 10.1016/S0022-1694(96)03287-8

Ranjram M, Craig JR. 2021. Use of an Efficient Proxy Solution for the Hillslope-Storage Boussinesq Problem in Upscaling of Subsurface Stormflow. *Water Resources Research* 57 (4): e2020WR029105 DOI: 10.1029/2020WR029105

Rupp D, Rupp DE, Selker JS. 2006. On the use of the Boussinesq equation for interpreting recession hydrographs from sloping aquifers. *Water Resources Research* 42 DOI: 10.1029/2006WR005080

Shaw SB, Riha S. 2012. Examining individual recession events instead of a data cloud: Using a modified interpretation of dQ/dt-Q streamflow recession in glaciated watersheds to better inform models of low flow. *Journal of Hydrology* 434: 46–54 DOI: 10.1016/J.JHYDROL.2012.02.034

Tallaksen LM. 1995. A review of baseflow recession analysis. *Journal of Hydrology* 165 (1–4): 349–370 DOI: 10.1016/0022-1694(94)02540-R

Thomas BF, Vogel R, Famiglietti J. 2015. Objective hydrograph baseflow recession analysis. Journal of Hydrology 525: 102–112 DOI: 10.1016/J.JHYDROL.2015.03.028

Troch PA, Paniconi C, Emiel van Loon E. 2003. Hillslope-storage Boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. Formulation and characteristic response. *Water Resources Research* 39 (11) DOI: 10.1029/2002WR001728

Vogel RM, Kroll CN. 1992. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resources Research* 28 (9): 2451–2458 DOI: 10.1029/92WR01007

Zecharias YB, Brutsaert W. 1988. Recession characteristics of groundwater outflow and base flow from mountainous watersheds. *Water Resources Research* 24 (10): 1651–1658 DOI: 10.1029/WR024i010p01651

Chapter 5

Conclusions

Understanding how small-scale governing constitutive relationships and distributed small-scale heterogeneity aggregate to produce large-scale hydrological effects is foundational to the exercise of watershed modelling. Upscaling, the conversion of small-scale hydrological relationships to large-scale hydrological descriptions, is thus an essential component of modelling, even when done implicitly. In some distributed modelling applications, the governing constitutive relationships are resolved at a fine scale and the relevant landscape heterogeneity is explicitly represented, implying that the bulk basin response generated by the model directly accounts for the small-scale behaviour of a basin (an exact averaging approach). In other applications, the small-scale behaviour is assumed to apply at the basin-scale with no alterations (the naïve approach). In between these conceptual extremes lies a suite of upscaling approaches that are incorporated into many modern hydrological and land surface models, and are robustly classified and catalogued in Chapter 2 of this thesis. Upscaling relationships are plentiful where small-scale governing constitutive relationships exist and are parameterized by readily measurable parameters, where large-scale aggregate information (e.g., streamflow) can be collected at the point scale, and/or where the science has come to an agreement over the expected spatial distribution of the most relevant parameters. In other cases, the lack of readily observable hydrologic behaviour at convenient scales of human observation, and the difficulty of prescribing and understanding system heterogeneity, hamper the generation of upscaling insights and algorithms. Although advancements in computational processing power may suggest that all upscaling exercises be eventually resolved by finely-discretized distributed models, upscaling investigations provide an opportunity to directly assess the controls and closure relationships of hydrological process, enabling a deeper understanding of the fundamental scaling relationships of hydrologic systems. In this way, upscaling should continue to be an important component of hydrological research.

Extending on the ontological and historical classification of upscaling in watershed modelling presented in Chapter 2, the upscaling behaviour of recharge-induced subsurface flow (i.e., subsurface stormflow) through hillslopes, resulting in a simple large-scale characterization of the recession behaviour of a basin, was explored in Chapters 3 and 4. By explicitly resolving and aggregating the subsurface physics of hillslopes (a bottom-up exact averaging approach), experimentally-derived large-scale basin behaviour was extracted, and connections between this behaviour and aggregate metrics of basin topography (a top-down empirical approach) were derived. However, the computational expense of resolving hillslopes with physically-based descriptions of subsurface flow is not trivial. Chapter 3 presented a novel modelling tool, the hsB Proxy, which produces rapid solutions to the hillslope-storage Boussinesq (hsB) equation describing saturated unconfined subsurface flow through wedge-shaped hillslopes with linear width functions. The computational efficiency of this tool was demonstrated, as was its utility in upscaling problems, where it was used to produce large-scale

recession behaviour informed by hillslopes with physical dimensions derived from topography, as well as in calibrating a single effective hillslope that could reasonably reproduce numerically-simulated behaviour. In Chapter 4, the hsB Proxy was used to produce large-scale results that enabled the derivation of upscaling relationships connecting basin topography to the large-scale recession behaviour. The resultant upscaling relationships, simple linear functions relating signature basin recession characteristics to the mean hillslope angle, the probability of hillslope convergence/divergence, the probability of hillslope speed, and the sum of hillslope downslope widths, were able to replicate numerical simulations at high quality and were used to produce transient recession coefficients which respected observed behaviour.

The expansive synthesis and classification of computational upscaling approaches in hydrology in Chapter 2 and the two-step generation of a new set of upscaling relationships in Chapters 3 and 4 are significant contributions of this work. In particular,

1. The review of computational upscaling approaches in hydrology (Chapter 2) provides a formal classification scheme for upscaling methodologies in hydrology and demonstrates how these methodologies have been specifically applied to generate upscaling insights. This new classification scheme includes a formal set of definitions and a robust mathematical treatment, and serves as a streamlined ontological framework for upscaling studies in hydrology.
2. This review provides a novel summary of the upscaling approaches that have been applied to critical hydrological processes, and in evaluating the common challenges and successes amongst hydrological processes, provides a valuable starting point for researchers interested in assessing and closing the gaps in upscaling approaches.
3. The hsB Proxy introduced in this work uniquely approximates the computationally expensive results of numerical hillslope drainage models and enables the types of studies presented in Chapters 3 and 4, where hundreds or thousands of hillslope simulations are required. Critically, these hillslopes need not be defined abstractly by the modeller, but rather can be directly derived from digital elevation data. These derived hillslopes, and the physics of subsurface flow expressed by the hillslope-storage Boussinesq equation (hsB), were combined to produce upscaling experiments with minimal conceptual abstraction.
4. The computationally efficient hsB Proxy allows for empirically-derived (rather than theoretically-assigned) transient recharge inputs to be considered in the problem of upscaling hillslope subsurface flow to the aggregate recession behaviour of a basin. Thus, the hsB Proxy further minimizes conceptual abstraction in this upscaling problem by eliminating the often-used simplifying assumption of a steady depletion curve in hillslopes, therein enabling insight into the impact of transience in recession behaviour.
5. This work was thus the first to directly connect the bulk transient recession response of a basin to distributed hillslope properties. This connection is formalized by new upscaling relationships capable of predicting the aggregate recession response as a function of readily-derived terrain metrics.

However, the results presented in this thesis are not without limitations. The treatment of hillslopes as the small-scale dimension of the upscaling problem explored in Chapter 3 and 4 is well justified by the literature, and the characterization of these hillslopes as variable-width areas derived directly from topography, solved with a physically-based treatment of subsurface flow and able to handle transient recharge inputs, is a novel configuration, representing am appropriately advanced level of complexity in these studies. However, despite these complexities, the governing constitutive relationship is limited to a characterization of saturated subsurface flow, and so cannot reflect the more realistic (but much more computationally expensive) understanding of unsaturated mechanics, which may be a significant control on flow behaviour. Further, the upscaling relationships generated in Chapter 4 are confined to applications in basins considered effectively homogenous, such that critical controls of subsurface heterogeneity were not accounted for (although the timescales of hillslopes are heterogeneous, as a function of their shape and slope).

In the future, the upscaling relationships introduced in Chapter 4 may be improved by additional testing investigating the effect of hillslope-scale heterogeneity in conductivity and porosity on large-scale recession behaviour. This is not a computationally trivial exercise, but the known role of heterogeneity means that such investigations are likely to lead to useful conclusions. Further, the upscaling relationships may be improved by repeating the experiment in Chapter 4 with a variably saturated subsurface flow model, although this is also not computationally trivial (indeed, the application of a computationally simpler saturated flow model required the generation of a completely novel tool in Chapter 3 to produce the thousands of simulation results required to generalize the upscaling behaviour of basins). Finally, the upscaling relationships in Chapter 4 and the hsB Proxy in Chapter 3 should be ported into a hydrological modelling platform to ensure that the upscaling conclusions developed in this thesis find a practical realization in watershed models. Once these tools are readily available in hydrological models, a more robust accounting of the physics of subsurface flow can be readily included in hydrological modelling problems without the associated computational burden, and the clear utility of upscaling exercises can be further reinforced.

With time, patience, and increased computational faculties, it is the ambition of this thesis that the conceptual framework utilized here for upscaling hillslope subsurface flow (that is, using numerical models to produce bottom-up aggregated large-scale results which are then top-down reconsidered to produce new upscaling models) is applied to a number of hydrological processes, and that a new generation of researchers can build off the work conducted here to close the gaps between our small-scale understanding of hydrological processes and the large-scale application of relevant algorithms in our hydrological models. It is the second ambition of this thesis that these researchers need not dive into the fractured and disconnected upscaling literature to navigate their starting point, but rather can begin their study much more accelerated by digesting the catalogue and ontology of upscaling processes in the review paper resulting from this thesis work.

References

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

Brutsaert W, Nieber J. 1977. Regionalized drought flow hydrographs from a mature glaciated plateau *Water Resources Research* 13 (3): 637-643 DOI: 10.1029/WR013i003p00637

Clark MP, Rupp DE, Woods RA, Tromp-van Meerveld HJ, Peters NE, Freer JE. 2009. Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale. *Hydrological Processes* 23 (2): 311–319 DOI: 10.1002/hyp.7154

Harman CJ, Sivapalan M, Kumar P. 2009. Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resources Research* 45 (9) DOI: 10.1029/2008WR007392

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

Vereecken H, Kasteel R, Vanderborght J, Harter T. 2007. Upscaling Hydraulic Properties and Soil Water Flow Processes in Heterogeneous Soils: A Review. *Vadose Zone Journal* 6 (1): 1–28 DOI: 10.2136/vzj2006.0055

Vereecken H, Weihermüller L, Assouline S, Šimůnek J, Verhoef A, Herbst M, Archer N, Mohanty B, Montzka C, Vanderborght J, et al. 2019. Infiltration from the Pedon to Global Grid Scales: An Overview and Outlook for Land Surface Modeling. *Vadose Zone Journal* 18 (1): 1–53 DOI: 10.2136/vzj2018.10.0191