Pragmatic Groundwater-Surface Water Model Coupling with Unstructured Grids

by

Leland Scantlebury

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Master of Applied Science

in

Civil Engineering (Water)

Waterloo, Ontario, Canada, 2021

Author's Declaration

I hereby declare that I am the sole author of this 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.

Abstract

Faced with an array of water issues exacerbated by a rapidly changing climate, hydrologists and hydrogeologists have increasingly found themselves needing to simultaneously model the groundwater and surface water domains together. Historically, for convenience and due to computational limitations, they have been modeled separately, with tools evolving based upon the different needs and questions driving researchers and practitioners in each domain. The tools emerging to solve these new problems range from highly complex, fully coupled, parallelized software solutions requiring enormous computational resources, to comparatively simple combinations of existing models sharing fluxes between the domains. Both groups generally have utilized relatively inflexible representations of the surface-water domain, often with a fixed level of complexity that prevents explorations of model structural uncertainty and process algorithmic skill. In this thesis, a loosely coupled groundwater-surface water modelling framework is presented that allows for adjustable model complexity in both domains. This is accomplished through pairing MODFLOW-USG, a recent version of the industry-standard MODFLOW family of modular groundwater modelling codes that allows for unstructured model grids, with Raven, a state-of-the-art surface water modelling framework supports flexible representations of hydrologic processes, forcing interpolation, and spatial discretization schemes. The resulting software, compiled into a single executable, is aimed at modelling watersheds at the regional scale. Recharge estimated by Raven is directly entered into the MODFLOW-USG flow solution. River-groundwater interactions are handled through a novel sub-grid river package added to MODFLOW-USG, called the polyline boundary junction (PBJ) package. The PBJ method evaluates boundary conditions along individual segment locations within a grid's dual Delaunay triangulation and geometrically distributes the resultant fluxes to the appropriate Voronoi and/or rectangular cells. Groundwater heads are interpolated along the segment to handle head-dependent flux calculations. The resulting river fluxes are added or subtracted from the Raven river channel water balance, allowing for a closed simulation of the hydrologic cycle. The new coupled Raven framework is demonstrated on the Alder Creek watershed in Southern Ontario and shown to produce physically realistic flows between the surface and subsurface domains.

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. James R. Craig, for his support, inspiration, and mentorship during this project. I knew you were going to be a great adviser when I joined your research group, but your enthusiasm for your work and care for your students is above and beyond. I think we have created something pretty neat, and I look forward to future collaborations.

My Alder Creek model built upon past modelling efforts; Thanks to Andrew Wiebe, Reynold Chow, Martinus Brouwers, Paul Martin, and David Rudolph for tracking down old models, files, and publications, as well as sending me what data they could.

I would also like to thank my many friends and coworkers at S.S. Papadopulos & Associates. Somehow you all tricked me into being the type of millennial who acquires more degrees despite already having a career. In particular, I would like to thank Chris Muffels and Vivek Bedekar, who I talked to often during this project and who were always there to answer questions or just chat. I would be remiss without also thanking Matt Tonkin, Chris Neville, Marinko Karanovic, Erica DiFilippo, and Charles Andrews for their continued mentorship and support. I am grateful to our associate John Doherty, who introduced me to many of the philosophical underpinnings of my views on modelling and uncertainty in his publications, classes, and over a couple beers.

I was lucky enough to join an incredible hydrology research group at Waterloo. Special thanks to Élise, Mark, and Genevieve for making Becca and me feel welcome in a new country – I miss you all! Rob, I am glad we got to overlap and work on RavenR together. Thanks to Ming for helping me build my Alder Creek model. To Mahkameh, Shaghayegh, Hongren, Hannah, and everyone else in the group, I am bummed COVID took in-person hangout time from us. Bryan and Julie, I am excited for the project you have given me to finish out my time at Waterloo, I am glad to get to work with you both closely.

Thanks to my family for supporting me through many different majors during my undergrad years. Had I not been able to try so many things out I would not have found my dream career in hydrology. There really is not a better gift you can give your kids than the opportunity to find and realize their dreams. Thank you for your continued love and support throughout this and my future endeavors.

And finally, thanks to my fiancée Becca, for leaving your job and moving to Canada with me, for listening to long rants about my work, for putting up with being separated by a border during a pandemic, for making sure I got outside and saw sunlight, and for supplying me with just about every type of support that exists. I love you and I am excited for our next adventure.

Table of Contents

List of Figures

List of Tables

Chapter 1 - Introduction

Despite being comprised of processes that interconnect the two, the modelling of surface water and groundwater systems has historically been carried out separately, often by practitioners in separate fields (Staudinger et al., 2019). The origin of this division can be traced back to differences in the questions being asked, and to simplifications born to make those questions tractable. Differences in the rate of movement of surficial and subsurface waters, combined with differing spatial scales of interest (Furman, 2008) have also contributed to the divide. Thus, software used in surface water and groundwater modelling efforts often treat features from the other domain as boundary conditions, where water is added or removed from the system in a prescribed manner.

Advances in computing power have encouraged the development of increasingly complex modelling software. Concurrently, concerns regarding managing depleting water resources and a changing climate (possibly further depleting water resources) have pushed hydrologic modelling towards considering larger systems with less certainty of stationary processes. These forces in tandem have stimulated the production of powerful software capable of modelling physical groundwater and surface water systems together, with varying degrees of complexity in the coupling of the systems. At one end of the spectrum are fully coupled modelling schemes, which simultaneously solve a single system of equations for both the surface and groundwater water systems; at the other end are loosely coupled schemes, in which uni- or bi-directional fluxes are calculated and passed between the systems. Depending on the software, feedbacks from the fluxes may or may not be represented (Barthel & Banzhaf, 2016; Furman, 2008). Fully coupled modelling schemes are very computationally expensive, leading to much longer model run times, which in turn can impede model calibration and uncertainty quantification efforts.

This thesis introduces a loosely coupled groundwater-surface water modelling framework, expanding the existing Raven hydrological modelling framework (Raven) (Craig et al., 2020) by incorporating the unstructured groundwater modelling software MODFLOW-USG (Panday et al., 2013). Numerous surface water models have been coupled to the industry-standard MODFLOW family of codes, however the latest iterations of these codes supporting unstructured grids are yet to be featured in any published groundwater-surface water schemes. Additionally, existing groundwater-surface water modelling schemes have fixed representations of hydrological processes, while Raven supports a wide variety of interchangeable algorithms allowing for trillions of possible model configurations (Craig et al., 2020). This allows the practitioner to explore the sensitivity and structural uncertainty of a wide variety of model configurations. No groundwater modelling software has been previously coupled to Raven, although representations of surface water-groundwater interactions have been modeled in the past (Snowdon,

2015). The expanded Raven framework presented here continues the vision of flexible modelling with a vast array of process representations – the conceptualizations of groundwater-surface water (GW-SW) processes explained within this thesis serve as but a starting point for the connections between Raven and a groundwater flow code.

Additionally, this thesis introduces a new MODFLOW-USG package for representing sub-grid linear features (e.g., streams) on grids for which a Delaunay dual exists. The package can be used alone with MODFLOW-USG, but here will be primarily discussed in its incorporation into the Raven framework.

The ability of the coupled framework to model groundwater-surface water interactions is demonstrated through the construction of an integrated model of the Alder Creek watershed in Southern Ontario. This area has been well-studied (e.g., Frind et al., 2014) and was the subject of a previous model/methods comparison study (Chow et al., 2016), making it an appropriate watershed for testing new hydrological modelling software.

1.1 Goals and Objectives

The primary goal of this thesis is to develop and test a new groundwater-surface water modelling software comprised of Raven and MODFLOW-USG. A secondary goal is to develop and test a new method of representing polylinear features such as rivers on unstructured grids that also supports coupling with Raven or other vector-based surface water models. Both goals will be illustrated through case study examples.

From these goals, three main objectives follow:

1. To expand the Raven hydrological modelling framework to include MODFLOW-USG, creating a practical, loosely coupled groundwater-surface water modelling framework applicable to real-world water resource management problems.
2. To develop and test a novel way of representing polylinear hydrologic features (e.g., streams and rivers) as boundary conditions on unstructured grids that is grid independent.
3. To demonstrate the expanded Raven hydrological modelling framework's utility in modelling real-world systems via an application to the Alder creek watershed.

1.2 Thesis Organization

This thesis is broken in five chapters. This first chapter introduces the material presented within.

The second chapter contains background on hydrological modelling, groundwater-surface water interaction, and existing groundwater-surface water modelling software.

The third chapter covers the software being coupled, the methods being used to couple them, and the handling of various hydrological features. In particular, it presents the unstructured grid linear feature package used to model groundwater-river interactions in the expanded version of Raven.

The fourth chapter examines and compares the polyline boundary junction (PBJ) package, created for MODFLOW-USG and incorporated into the coupled GW-SW modelling framework, to similar existing MODFLOW packages.

The fifth chapter demonstrates the efficacy and utility of the expanded Raven framework through an example model of the Alder Creek watershed in Southern Ontario.

The sixth chapter provides a conclusion to the thesis and provides recommendations for further works.

Chapter 2 - Background

The following chapter serves to provide the reader with the necessary context to understand and evaluate the significance of the work being presented in this thesis. The first section explains why GW-SW modelling is increasingly important, and the following sections detail the current state of GW-SW modelling and some of the popular existing software applications.

2.1 The Need for Groundwater-Surface Water Modelling

The best case for why groundwater and surface water systems should be modeled together is simple: in the real world, they do not function independently. Precipitation percolates down into the subsurface, rivers lose water to and gain water from groundwater aquifers, and pumping from wells can impact water levels in nearby streams and lakes. In a review of isotope streamflow contribution studies, Jasechko (2019) found that, in a majority of reported cases, groundwater discharge was the most significant contributor to streamflow. The nature of GW-SW interactions varies across different regions and at different scales (physical and temporal), but the two systems are never acting entirely separately. So, the question becomes, why are they ever modeled apart?

2.1.1 A Tale of Two Disciplines

Staudinger et al. (2019) provides an interesting explanation of the separation, tracing the division back to the separate disciplines of hydrology (focusing on surface waters, connected primarily to engineering) and hydrogeology (focusing on groundwater aquifers, connected primarily to geology). Their discussion is influenced by Barthel (2014), which attempts to make sense of how hydrology can both contain *groundwater hydrology* (which is generally indistinguishable from *hydrogeology*) and yet also be seen by many who call themselves hydrologists as a separate field. Semantics aside, both papers identify differences in terminology, education, foci, timescales, relevant observations, and software, among others, that have historically prevented the fields from working together in an integrated manner.

To briefly expand upon some of the differences relevant to modelling, surface water hydrology is primarily concerned with catchments, while hydrogeology is focused on aquifers (Barthel, 2014b). Catchments (also called drainage basins, and in North America, watersheds) are an area of land where all the surface water runoff drains to a common point. Aquifers are subsurface bodies of water inhabiting the pores of sediment and/or fragmented rock layers. The differences in these systems are well highlighted in their most commonly collected observational data: surface water hydrologists principally measure streamflow at a point, essentially the total water volume leaving the catchment (excluding any losses to groundwater); Hydrogeologists measure water levels at observation wells, which in the aggregate, combined with knowledge of the subsurface, can provide estimates of aquifer extent and volume.

Historical records of these two data sources are generally used to evaluate the performance of numerical models ("history matching") (Doherty, 2015) that seek to represent the movement of water in these systems due to known stresses. Vitally, groundwater levels can be measured at any number of points within an aquifer to provide local information, while streamflow measurements are integrative: additional points downstream add incremental responses from the additional areas upstream. Streamflow represents an aggregate of hydrological processes and events within the catchment; a single groundwater level does not provide spatially aggregated data, but since many can be collected from different points in an aquifer simultaneously (and more easily than streamflow measurements), spatial patterns can be more easily identified. In either case, additional site data (bathymetry, stratigraphy, topography, precipitation, etc.) is vital for providing context to the primary observations.

Groundwater and surface water systems also can have spatial boundaries that do not align, leading to issues in GW-SW modelling. Catchments provide easy closed boundaries for surface water models, but aquifer boundaries can be difficult to define, depending on the size (particularly depth) of the system being modeled (Condon et al., 2020; Staudinger et al., 2019). While it is common practice to include fluxes in or out of the model domain as perimeter boundary conditions in groundwater modelling (Anderson et al., 2015, p. 134) difficulties in estimating such fluxes can add considerable uncertainty to the model and have impacts on calibration (Beven, 2005; Doherty & Welter, 2010). For instance, groundwater recharge in many regions is the primary source of water entering the aquifer system, but is difficult to estimate or measure (Anderson et al., 2015, p. 232) and estimates are often heavily dependent on the density and reliability of the rain gauge network (Wiebe & Rudolph, 2020). The uncertainty of such inputs, in value and spatial extent/distribution, can greatly increase the uncertainty of model outputs (Jyrkama & Sykes, 2006; Refsgaard et al., 2007).

Additionally, it is common to use major hydrological features (e.g., lakes or rivers) as perimeter boundaries of groundwater models (Anderson et al., 2015, p. 137), a strategy almost always incompatible with catchment boundaries. Revising groundwater model boundaries to match catchment boundaries may force modelers to select inferior boundaries subject to flow changes (e.g., groundwater divides as opposed to geologic features (Anderson et al., 2015, p. 136)) as their perimeter boundary conditions. As much of hydrological modelling relies on a water balance (conservation of mass), careful consideration of model boundaries is essential to capturing fluxes entering/leaving the system.

The temporal scale of relevance differs between groundwater and surface water systems, starting with the observation that groundwater moves much slower than streamflow (Sophocleous, 2002), and generally surface water hydrological processes occur within shorter time periods (Barthel, 2014b). This is reflected in the practice of modelling these systems – often groundwater models can be run with coarser

(longer) timesteps than surface water models. Surface water studies also often assume that the water balance "resets" each year, with a *water year* running roughly from September to October in the Northern Hemisphere. This is possible since the storage of surface water systems is generally very small, especially when compared to their fluxes (Barthel, 2014b; Staudinger et al., 2019). In contrast, groundwater systems are dominated by their water storage terms and often have very long residence times.

Historically, there have been a very large differences in the strategies employed to model these systems. Groundwater modelling has traditionally been done exclusively with physically-based models, since the groundwater flow equation is (a) known and (b) linear (Staudinger et al., 2019). Physically-based models became mainstream in surface water modelling much later, however, and more controversially (Beven, 1989). This is due to the much greater range of hydrological processes above ground, compounded by many differing plausible descriptions of these processes (Craig et al., 2020; Singh & Woolhiser, 2002). Worse, many of these processes are only understood at the point scale, and their usefulness and validity at the scales commonly needed for modelling is questionable (Beven, 1989; Kirchner, 2006). Lumped models historically have performed very well in surface water modelling, acting as "black boxes" opaquely simulating many hydraulic processes in a catchment using a few tuning parameters. While the lack of physically identifiable parameters can hinder experimentation with the model (e.g., what does it mean that I've doubled α_1?), parsimonious parameterization aids with addressing inevitable model non-uniqueness (Beven & Binley, 1992; Kirchner, 2006).

Issues of scale appear in groundwater modelling as well. However, the empirical Darcy's law lends itself much better to scaling (de Marsily et al., 1992). Effective (or block, or upscaled) values of hydraulic conductivity are commonly used to transform point measurements to model grid scale, incorporating heterogeneity, either through upscaling mathematical relationships or through estimation during the calibration process (de Marsily et al., 2005; Doherty et al., 2010; Wen & Gómez-Hernández, 1996).

As discussed by Staudinger et al. (2019), separate modelling of groundwater and surface water systems has, oddly, resulted in neither hydrology nor hydrogeology fully taking on modelling of the vadose zone. Instead, a wide variety of methods have been deployed to simplify its representation based on application region, scale, and purpose (Barthel, 2006; Harter & Hopmans, 2004). As the primary connection between the two domains, the vadose zone has generally served as an entry (groundwater) or exit (surface water) point for water in water balances of the systems. Thus, it serves as one of the main challenges in coupling GW-SW models, particularly at common scales of interest (Barthel, 2006). The primary challenge is the governing equation in the vadose zone (Richards' equation) is non-linear, not valid at larger scales (Harter & Hopmans, 2004), and can prove computationally demanding. Many promising simplifications exist (see Harter & Hopmans, 2004) but all require careful consideration of

their assumptions, making them potentially problematic to use in traditional, fixed-structure modelling software.

2.1.2 A Single Resource

A common reason for modelling GW-SW systems separately is that, historically, many of the questions posed to these models simply do not necessitate invoking the complications explored in the last section. While that assumption may or may not be valid, depending on the problem, increases in the understanding of GW-SW interactions (Sophocleous, 2002; Winter et al., 1998) and computing power (Zhou & Li, 2011), compounded by global climate change and other anthropogenic impacts on the environment, have further enabled and necessitated the modelling GW-SW together, especially at the regional scale (Barthel, 2014a).

The need for GW-SW integrated management is well explained in Winter et al. (1998) "Ground Water and Surface Water: A Single Resource". There, the authors explore the intertwined nature of surface water bodies such as wetlands, lakes, and streams with groundwater systems, and how human development and extraction of these resources further affects these systems. The report highlights that, due to the overwhelming complexity of climate systems, the seasonal and long-term changes to GW-SW systems are hard to predict. Scholarly papers since then have discussed the need for and utility of regional-scale hydrological modelling (Barthel, 2014a; Barthel & Banzhaf, 2016; Zhou & Li, 2011); the inability of models to accurately predict extreme events & represent long-term dynamics (Fowler et al., 2020; Seibert, 2003); impacts of climate change on groundwater (T. R. Green et al., 2011; Holman et al., 2012); and the dangers of modern pollution contaminating deep, ancient waters (Jasechko et al., 2017). To synthesize, many of contemporary pressing questions require a "single resource" approach to modelling, along with careful thought of the uncertainties introduced through our simplifications. In modelling, there will always be simplifications, but capturing the dynamic, non-linear interactions present in our world requires accurate representations of those processes (Furman, 2008; Kirchner, 2006).

2.2 Common Groundwater-Surface Water Connections in Numerical Models

Surface water and groundwater systems are modeled with different partial differential equations, necessitating explicitly defined boundary conditions at their interface (Furman, 2008). These boundary conditions, typically expressed for the groundwater model, generally fall into one of three types (following the definitions in Jazayeri & Werner (2019)):

Type 1. *Dirichlet conditions* or *specified head boundary* where a water level is set to a known value. If the value does not fluctuate over time, it is commonly called a *constant head boundary*.

These conditions are often used at the edge of groundwater models when no physical (e.g., geologic) boundaries exist but heads are known or can be approximated.

Type 2. *Neumann conditions* or *specified flux boundary* where the derivative of the head (rate of change) is set to a known value. If that value is zero, it is commonly referred to as a *no-flow boundary* which is often the default boundary at the edge of groundwater models to represent groundwater divides. Specified flux boundaries are often used in GW-SW connections to represent recharge. In the groundwater domain they are commonly used to represent pumping wells.

Type 3. *Robin conditions* or *head-dependent boundary* where the flux in or out of the boundary is determined based on the difference between a water level at or near the boundary and an external water level. Generally, Darcy's law is used to calculate a flow across the boundary, given the pressure gradient, in the form:

$$Q = -\frac{KA}{L}(\Delta h) \tag{1}$$

where Q is the flow, K is the hydraulic conductivity [L/T], A is a cross-sectional area, L is the distance over which the head change (Δh) is measured.

Additionally, combinations of these elements are often used, termed *mixed boundary conditions*. A conditional statement, usually involving the water level at the boundary, may be used to determine which type of boundary condition occurs – e.g., flow (Robin conditions) only occurs through the boundary when the water table reaches a certain height. Below that height, no flow occurs (Neumann conditions). However, these "threshold" conditions can complicate parameter estimation by degrading the smoothness of the objective function surface (Kavetski et al., 2006). The standard condition for coupling surface water and groundwater fully integrated (and some tightly coupled) codes is the Type 3 condition across an extremely thin resistance layer coupled with enforcement of head continuity (Ebel et al., 2009). Since the system is treated as a continuum in such simulation frameworks, such a condition is technically a *continuity* condition rather than a boundary condition.

In this section the common hydrologic interfaces and their mathematical representations are explored in the context of loosely coupled models. More complete reviews, with a focus on integrated flow equations, can be found in Furman (2008) or Morita & Yen (2000).

2.2.1 Rivers and Streams

Streams often have a large influence on the water levels and flow directions in the aquifers beneath them. Generally, in the context of groundwater modelling, the interactions are categorized into three different scenarios (Figure 2-1): groundwater can be flowing into the stream (gaining stream), the stream

water can be flowing into the groundwater (losing stream), and the stream can become hydrologically disconnected from the water table, but water still percolating down through the subsurface (losing stream that is disconnected from the water table).

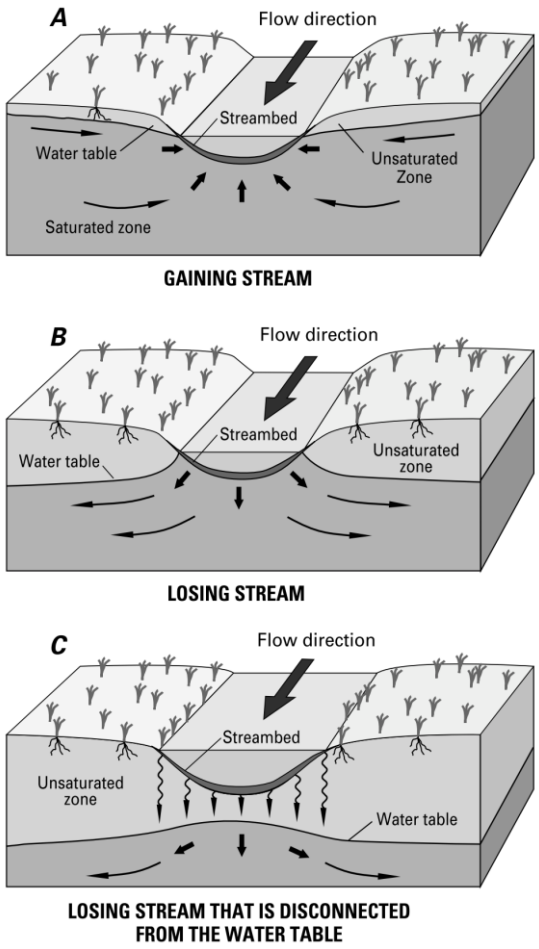

Figure 2-1 – The three main categorizations of stream-aquifer interactions: (A) stream gaining from groundwater, (B) stream losing to groundwater, and (C) stream losing to groundwater but disconnected from the water table. Figure from *Reilly (2001*).

GW-SW interaction flow direction also can vary greatly along a stream and change over time (Conant Jr., 2004; Winter et al., 1998). The left side of the stream can be gaining while the right side is losing (a "flow-through" stream), or groundwater flow can move parallel to the stream. Water can move in and out of the hyporheic zone (a porous region adjacent to the stream where groundwater and surface water mix). The depositional patterns of sediments, from both current and historic stream conditions, can form incredibly complex conductivity distributions (Woessner, 2000).

The intent and objectives of the modelling project, along with data availability, should dictate the complexity of the stream-aquifer interaction representation in a model. All four of the boundary condition types detailed in Section 2.2 (considering *mixed* conditions as the fourth) can be used to represent streams

in groundwater models (Reilly, 2001). In coupled systems, however, the existence of readily available stream levels from the surface water component often leads to the use of head-dependent or mixed boundaries to represent stream-aquifer fluxes. The conceptualization in several MODFLOW packages (Harbaugh, 2005), which serves as a good comparison point for other models, is a head-dependent boundary when the stream is connected to the aquifer:

$$Q = C_{RIV}(h_{str} - h_{GW}) \approx \frac{K_b A}{t_b}(h_{str} - h_{GW}) \tag{2}$$

where Q is the flow from the stream to the aquifer, C_{RIV} is the streambed conductivity [L/T], h_{GW} is the groundwater head at/near the river, h_{str} is the stream. MODFLOW conceptualizes C_{RIV} as the streambed conductivity, K_b, multiplied by the streambed area, A, over t_b, the thickness of the streambed sediments (Harbaugh, 2005). The common sign convention is that Q is positive when water is flowing into the aquifer. When the groundwater level goes below the bottom elevation of the stream, MODFLOW considers the aquifer to be hydrologically disconnected from the stream. This results in percolation from the stream to the aquifer. This is modeled using:

$$Q = C_{RIV}(h_{str} - z_{str}) \tag{3}$$

where z_{str} is the elevation of the bottom of the stream (i.e., Eq. [2] but with z_{str} with zero pressure head replacing h_{GW}). This disconnected losing stream conceptualization notably completely ignores the unsaturated zone, for example, Equation [3] does not consider the distance water must travel from the bottom of the stream to the aquifer below, the local pressure head in the unsaturated zone, or the time associated with percolation.

A further simplification of stream-aquifer interactions can be made by assuming that the stream is exclusively gaining from groundwater, as may happen in wet/humid environments. Flow occurs when the groundwater elevation rises above the topographic minimum elevation (e.g., a stream bed elevation), mediated by a conductance parameter. When the groundwater is below the minimum elevation, no flow goes to the stream:

$$\begin{aligned} h_{GW} &> z_{str} \qquad Q = C_{RIV}(h_{GW} - z_{str}) \\ h_{GW} &\leq z_{str} \qquad Q = 0 \end{aligned} \tag{4}$$

This is a mixed boundary condition, operating as either a head-dependent boundary or a specified flux boundary (with a flow rate of zero). In MODFLOW, this is called the “drain” condition because it can be used to represent, for example, agricultural pipe drains (Harbaugh, 2005) which generally become dry when they are above the water table. However, in certain settings, it can be appropriate to use it to represent streams (Reilly, 2001; e.g., Snowdon & Craig, 2016) as well as other surface water features

such as springs, wetlands, or canals (Harbaugh, 2005; Reilly, 2001) which go dry without the presence of groundwater, as may be the case in highly conductive formations.

The representations of streams in groundwater models using Eq. [2 & 3] are not without their strong critiques (Barthel & Banzhaf, 2016; Brunner et al., 2010, 2017; Mehl & Hill, 2010; Morel-Seytoux et al., 2014, 2017, 2018; Nemeth & Solo-Gabriele, 2003). The simplest criticism comes from the MODFLOW manual itself: the representation assumes most of the head loss occurs in a single, homogenous, confining ("clogging") layer at the bottom of the stream, which is uncommon. Various papers have explored issues of grid scale (Mehl & Hill, 2010; Morel-Seytoux et al., 2014; Nemeth & Solo-Gabriele, 2003), proposed improved river conductance formulations (Mehl & Hill, 2010; Morel-Seytoux et al., 2018; Nemeth & Solo-Gabriele, 2003), and derided the simplistic disconnected losing stream conceptualization (Brunner et al., 2017; Morel-Seytoux, 2020), but the MODFLOW stream flux conceptualization remains in many hydro[geo]logical modelling codes. This is for a variety of practical reasons, largely stemming from the complex flow processes in streambeds complicated by heterogeneity, scale problems, and complex hyporheic zone interactions (Sophocleous, 2002) which make measurement-based conductance values difficult, if not impossible to use (Mehl & Hill, 2010). Compensating for the complex physics and scale issues by treating the river conductance as an empirical term to be determined through calibration (Harbaugh, 2005) has proven a more popular route.

2.2.2 Infiltration, Recharge & the Vadose Zone

While surface water bodies such as rivers and lakes are the clearest location of GW-SW exchange, a hidden zone of unsaturated porous media – the vadose zone- connects the groundwater and surface water domains nearly everywhere else on land. The vadose zone is roughly 0-10 m thick and mediates the interrelationships between rainfall, snowmelt, infiltration, surface runoff, evaporation, root uptake, and groundwater recharge (Harter & Hopmans, 2004). Due the complexity of the governing equation of flow in the vadose zone (Richards' equation), flow to and from this domain is commonly used as a boundary condition in surface water and groundwater models, respectively, if not treated entirely as a "black box" (Harter & Hopmans, 2004; Staudinger et al., 2019). Here, infiltration refers to water that enters the topsoil into the vadose zone, and recharge refers to water that enters an aquifer from the vadose zone. Recharge is generally conceptualized as a specified flux boundary in groundwater models and a state-dependent loss term in surface water models. In coupled GW-SW modelling, representing the processes that facilitate the movement and distribution of water in the vadose zone is one of the primary challenges over single-domain modelling (Niswonger et al., 2006).

Vadose zone dynamics are dominated by different processes at different scales (Corwin et al., 2006). Here, the focus is limited to processes at the regional scale, where the vast differences between the

horizontal (large) and vertical (small) dimensions of the vadose zone begin to become important. Harter & Hopmans (2004) identified four main functions of the vadose zone at the regional scale:

> 1. To separate precipitation and applied irrigation water into infiltration, runoff, evapotranspiration, interflow and groundwater recharge;
> 2. To store and transfer water in the root zone between the atmosphere above and the deeper vadose zone or groundwater below, including interflow;
> 3. To store and transfer water in the 'deep vadose zone', that is, between the root zone above and groundwater below;
> 4. To store, transfer, filter, adsorb, retard and attenuate solutes and contaminants before these reach the ground water.
>
> (Harter & Hopmans, 2004)

Most surface water models treat F1 and F2 in at least a partial fashion, as the partitioning of these water balance components is critical for estimating streamflow. Stand-alone groundwater models tend to avoid representation of these processes altogether, though high moisture content in the vadose zone can significantly augment water table response to recharge (Gillham, 1984). However, Richards' equation is widely seen as the gold standard for representing these functions. It is generally understood to be valid at the representative elementary volume (REV) scale (e.g., Darcy-scale, or sometimes "local" scale) (Vereecken et al., 2019) but the nonlinear nature of the complex partial differential equation suggests at the regional scale the equation is likely not valid (Beven & Germann, 2013; Harter & Hopmans, 2004). While coupled models are generally aimed at a wide variety of scales, those applied at regional, or even continental scales still often employ some form of Richards' equation (e.g., Hwang et al., 2014) with grid cells much larger than the REV scale. Modern so-called "fully coupled" GW-SW modelling codes often implement a full three-dimensional (3D) formulation of the Richards' equation (Brunner & Simmons, 2012; Furman, 2008) which can significantly increase model run times, in part due to Richards' equation generally requiring smaller (shorter) time steps and finer model discretization (Niswonger et al., 2006).

Other loosely coupled GW-SW modelling codes often employ one- or two-dimensional analytical or numerical solutions to Richards' equation (Furman, 2008). These simplifications do not remedy the issues of scale associated with typical lateral model grid sizes (e.g., Craig et al., 2010), but do provide more computationally efficient estimates of runoff, infiltration, and recharge. Some of these solutions, such as the analytically-derived Green-Ampt equation (W. H. Green & Ampt, 1911) only provide an estimate of infiltration flux into soil. To obtain recharge from an infiltration estimate, estimates of evapotranspiration and other vadose zone processes can be used to form a soil water budget (Anderson et al., 2015, p. 232). Recharge is usually estimated as the remaining (residual) water after all soil water losses are calculated.

The primary downsides to this approach are difficulties incorporating when the water table rises and saturates the soil zone from below, causing saturated excess overland flow, and inaccurate timing of recharge due to water movement through the vadose zone not being modeled (Anderson et al., 2015, p. 232; Hunt et al., 2008). Both issues propagate to the soil water balance, changing the recharge estimates, and thus the entire water balance of the GW and SW domains.

To capture the movement of water in the vadose zone, the kinetic wave approximation (Lighthill & Whitham, 1955) can approximate Richards' equation in one dimension (1D) and be solved by the method of characteristics (e.g., Niswonger et al., 2006). The Brooks-Corey (Brooks & Corey, 1964) or Van Genuchten soil characteristic functions can be used to estimate the relation between unsaturated hydraulic conductivity, moisture content, and pressure head. This formulation allows for the simulation of wetting and drying fronts, and can incorporate simplified versions of ET (Figure 2-2; Niswonger et al., 2006).

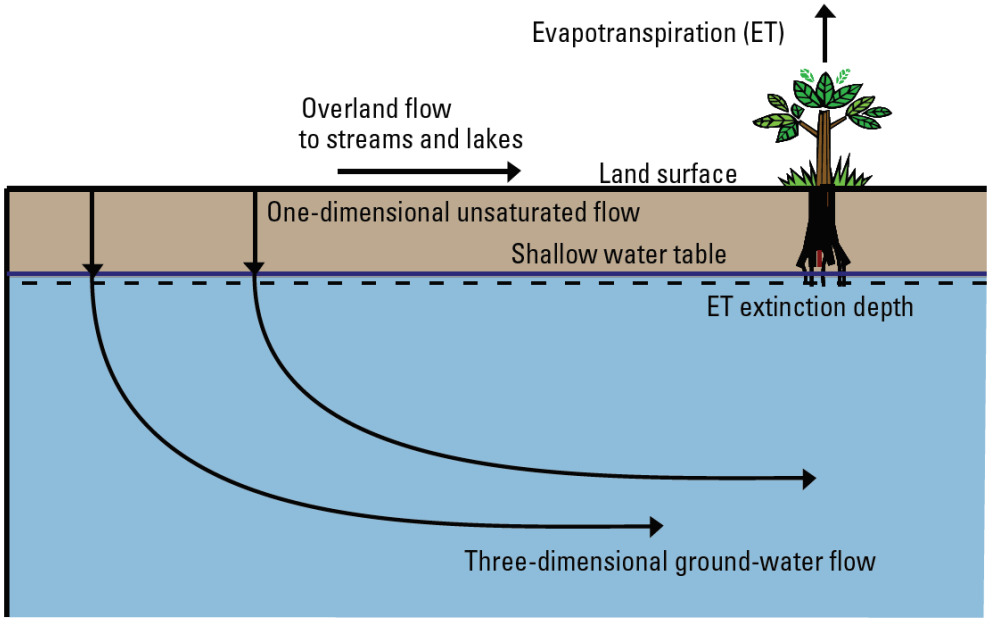

Figure 2-2 – Flow through the vadose zone approximated as one-dimensional (1D) flow from the land surface to the aquifer below. Evapotranspiration (ET) occurs above an extinction depth. Figure from *Niswonger et al. (2006)*.

Many other approaches to simulating the vadose zone and approximating/solving Richards' equation exist and are being developed (See, for example, those named in Farthing & Ogden, 2017; Furman, 2008; Panday & Huyakorn, 2008). Of particular note, the flexible surface-subsurface coupling options of Panday & Huyakorn (2004) allow for fully coupled flow, iteratively solved flow, or time lagged flow (Furman, 2008), allowing the user to compare different levels of complexity.

2.2.3 Evapotranspiration

Evapotranspiration (ET) is a combination of two intertwined processes removing water from the surface and shallow subsurface systems: evaporation and transpiration. Evaporation is the process of liquid water changing to the gas state and becoming atmospheric vapor, which can occur on surface water bodies or within the soil. Transpiration is the release (evaporation) of water vapor from plant leaves, but in the context of evapotranspiration is generally understood to be the removal of water from the subsurface by plant roots. Evaporation is also very important in lakes and wetlands, but here soil evaporation in the context of coupled models is focused on.

Various conceptualizations of evapotranspiration exist based on various land uses, forcing data requirements, regional climatic considerations, scale, and other factors (Xiang et al., 2020). Often, ET is calculated as a potential evapotranspiration (PET), a maximum amount of ET that can occur with unlimited water supply (i.e., not considering the actual water availability). In MODFLOW (e.g., Harbaugh, 2005) ET is represented as a distributed sink, with a maximum [P]ET rate and ET extinction depth (see Figure 2-2) set by the user. The maximum rate is used when the water table rises above the ground surface, and a loss rate of zero is used when the water table is below the extinction depth. Between the two extremes, the ET is calculated as a linear function based on the water table elevation (head-dependent boundary):

$$\mathrm{ET} = \mathrm{PET} \cdot \frac{h_{GW} - (z_{max} - d_{ext})}{d_{ext}} \tag{5}$$

where ET is the ET flux from the soil, PET is the user-input maximum ET rate, h_{GW} is the groundwater head (water table elevation), z_{max} is the ground surface elevation, and d_{ext} is the extinction depth. ET fluxes are usually given as volume per unit of water table surface area ([L/T], e.g., mm/day).

The MODFLOW formulation is overly simple, not accounting for water stored or moving through the vadose zone. The unsaturated-zone flow package (Niswonger et al., 2006), which uses a kinematic wave approximation to Richards' equation (see Section 2.2.2), more appropriately uses the ET rate to decrease the water content of a trailing wave. When the water table is above the extinction depth, and the specified ET rate is greater than the water content of the unsaturated zone, Eq. [5] is used to remove water from the water table. This conceptualization is likely more appropriate for wet environments, where negative pressure gradients caused by drying upper soils is less common (Niswonger et al., 2006).

As an example of surface water representation of ET, the Variable Infiltration Capacity (VIC) hydrologic model (Wood et al., 1992), which is frequently coupled to other models (e.g., Sridhar et al.,

2018) uses a power-law relationship to represent soil ET as a function of soil saturation in a finite topsoil strata:

$$\mathrm{ET} = \mathrm{PET} \cdot \left(1 - \left(1 - \frac{\phi_{soil}}{\phi_{max}}\right)\right)^{\gamma} \tag{6}$$

where ϕ_{soil} is the soil saturation, ϕ_{max} is the maximum soil saturation, and γ is a power law parameter. An assortment of other soil ET, as well PET, algorithms are collected in the Raven User's and Developer's Manual (Craig & the Raven Development Team, 2020), but most link actual ET to the degree of soil saturation, consistent with the simplistic MODFLOW approach noted above.

2.3 Existing GW-SW Modelling Software and Classifications

Software and strategies for GW-SW modelling come in a vast array of complexities, from simple spreadsheet models to advanced, parallelized models that treat the entire hydrologic system as a single set of equations (Tolley et al., 2019). Different problems require different solutions, depending on the problem scale, data availability, region, scope, budget, etc. More complex solutions do not guarantee more accurate results (Furman, 2008), but they do often guarantee longer model run times and/or demand greater computational resources.

The discussion here follows the GW-SW software comparisons of Furman (2008) and Barthel & Banzhaf (2016). These reviews are not exhaustive, and new software is always being developed, and existing software upgraded and expanded. Note that often both modelling software, and the collection of input files a modeler uses with this software, are both commonly referred to as "models". Here a choice has been made to distinguish the two, with *software* being the coded instructions commonly compiled into an executable, and *model* being a specific set of input files for a modelling software that represent a site or problem. Coupled models have been developed for specific sites (e.g., Tolley et al., 2019) but here the discussion is restricted to non-site-specific software.

Table 2-1 contains a selection of popular and/or influential GW-SW modelling software packages along with various classifications regarding licensing, coupling classification, and the various groundwater (GW), surface water (SW), and unsaturated zone (UZ) schemes implemented. Each of the classifications will be briefly discussed.

Table 2-1 – Selected GW-SW modelling software and various classifications, compiled from personal research as well as Furman (2008) and Barthel & Banzhaf (2016).

Software Name	Reference Paper(s)	Licensing	Coupling	GW scheme	SW scheme	UZ-Scheme
HydroGeoSphere	(Brunner & Simmons, 2012; Harter & Morel-Seytoux, 2013; Park et al., 2009)	Commercial	Full	3D CVFE	2D	3D Richards'
ParFlow	(Ashby & Falgout, 1996; Kollet & Maxwell, 2006)	Open Source	Full	3D FV	2D	3D Richards'
PIHM	(Qu & Duffy, 2007)	Open Source	Full	2D FV	1D	1D Richards'
FIHM	(Kumar et al., 2009)	Unknown	Full	3D CVFE	2D	3D Richards'
SHUD	(Shu et al., 2020)	Open Source	Full	2D FV	1D	1D Richards'
InHM	(VanderKwaak, 1999)	Unknown	Full	3D CVFE	2D	3D Richards'
MODHMS	(Panday & Huyakorn, 2004)	Commercial	Full/Iterative Loose*	MODFLOW[1]	1D	3D Richards'
MIKE SHE	(Hughes & Liu, 2008; Refsgaard & Storm, 1995)	Commercial	Loose	3D FD/ Lumped	2D	1D/2-layer water balance
GSFLOW	(Markstrom et al., 2008)	Open Source	Iterative Loose	MODFLOW[1]	PRMS (1D)	1D Kinetic Wave
IWFM	(Dogrul et al., 2018; Harter & Morel-Seytoux, 2013)	Open Source	Iterative Loose	3D FE/FD[4]	1D Kinetic Wave	1D
SWR-MODFLOW	(Hughes et al., 2015)	Open Source	Iterative Loose	MODFLOW[1]	SWR (1D, 2D)	1D Kinetic Wave[2]
SWAT-MODFLOW**	(Bailey et al., 2016)	Open Source	Loose	MODFLOW[1]	SWAT	SWAT/1D Kinetic Wave[3]
VIC-MODFLOW	(Sridhar et al., 2018)	Unknown	Loose	MODFLOW[1]	VIC	None
MODBRANCH	(Swain & Wexler, 1996)	Open Source	Iterative Loose	MODFLOW[1]	1D Saint-Venant	None

* MODHMS supports multiple coupling types.
** Several different authors have coupled SWAT to MODFLOW (Kim et al., 2008; Sophocleous & Perkins, 2000). Included here is a recent, open-source development.
[1] MODFLOW uses a 3D FV scheme.
[2] Via MODFLOW UZF package (Niswonger et al., 2006), compatibility implied.
[3] Recharge is calculated by SWAT, MODFLOW UZF package (Niswonger et al., 2006) compatibility is implied.
[4] IWFM uses FE in the horizontal and FD in the vertical.

Open-source software is defined by being free and having its source code freely available. This allows for it to be used, edited, examined, shared, and redistributed by anyone. This is common with software developed by United States government agencies (e.g., MODFLOW and other USGS variants) and open source software is a requirement for modelling projects in certain regions (e.g., California Department of Water Resources, 2016). Commercial software is proprietary, sold at a price, and generally no source code is available to the public. Nothing inherent to the licensing makes one category superior to the other: both categories contain extensively tested, peer-reviewed software. However, open-source software is often seen as beneficial as it allows models to be more easily shared and reviewed by different stakeholders.

Coupling strategies between the GW and SW domains have been categorized into three categories:

- **Fully** coupled modelling software solves the GW and SW domains together as a single set of equations for each timestep.
- **Iterative Loose** coupled modelling software solves each domain separately and iteratively (typically, the SW domain is solved first) until the fluxes between the two (e.g., recharge, discharge to rivers) are within a certain numerical tolerance.
- **Loose** coupled modelling software solves the GW and SW domain flows separately. Nearly every software handles this differently, but generally this group is categorized by a single iteration per time step. To accomplish this, feedback between the systems is lagged, creatively simplified, or small timesteps are used.

Notably, Furman (2008) and Morita & Yen (2000) would refer to the above "loose" coupling classification as *uncoupled*, citing that no feedback is used to "correct" the solution of the initial system. This seems like an unfair distinction: the connected models *are* passing fluxes back and forth each time step. Considering many of these models are run with (or require) relatively fine temporal discretization (e.g., daily or less), feedbacks are incorporated, just in a less precise manner. We reserve the term "uncoupled" for GW-SW modelling studies that run an entire simulation for one domain and then pass the solution fluxes to another domain, most common in recharge estimation for groundwater modelling.

The groundwater (GW), surface water (SW), and unsaturated zone (UZ) schemes are either the (1) common software package representing the domain or (2) dimensionally of the equation being used the model the domain. Several of the modelling packages use MODFLOW as their groundwater scheme and therefore, to represent the unsaturated zone, could potentially use the UZF package (Niswonger et al., 2006) or the variably-saturated flow process (VSF) depending on the compatibility of the implemented coupling and/or MODFLOW version.

Notably, none of the MODFLOW-coupled codes listed utilize either of the new, unstructured grid versions of MODFLOW: MODFLOW-USG (Panday et al., 2013) or MODFLOW 6 (Hughes et al., 2017). These control volume finite difference (CVFD) models allow for flexible grids able to better conform to real-world boundaries and features, and facilitate intelligent, gradual refinement in areas of interest. Combined with powerful sparse-matrix solvers, they are considered state of the art and represent the future of groundwater modeling.

Another noticeable hole in the literature (and software options) for joint GW-SW modelling is the coupling of flexible models of both domains. While MODFLOW generally represents a flexible GW model, none of the pairings identified in Table 2-1 represent state-of-the-art flexible SW models, such as SUPERFLEX (Fenicia et al., 2011) or SUMMA (Clark et al., 2015), which facilitate model structure-based experiments to form numerous model combinations. Raven (Craig et al., 2020) takes these

developments even further, integrating additional modular options for discretization, routing, and interpolation, with a robust process library (see Section 3.1.1). No existing GW model has been previously coupled to Raven.

2.3.1 Appropriate Complexity: A Critique of Fully Coupled Modelling Software

ParFlow and HydroGeoSphere (HGS), fully coupled GW-SW software, are undoubtably cutting edge tools that have been used to model complex, difficult problems. While some of the discussion in this section pertains to other fully coupled software detailed in Table 2-1, these two are singled out due to their prominence. These codes were created with the same general intent as the one in this thesis: to model the flow of water through GW-SW systems. However, those codes have taken a very different approach from the one taken here, aiming to solve much more difficult equations simultaneously, and extending the subsurface grid up to the surface domain. While this ostensibly results in more accurate representations of the natural world, in practice it often results in increased data requirements, longer run times, and issues with reliability (Doherty & Moore, 2019; Furman, 2008; Tonkin et al., 2020).

ParFlow and HGS grew out of academic research and are the result of trying to put our very best understandings of the hydrologic cycle into useable software. Both are written to take advantage of parallel processing (i.e., distributing the computational burden across many processor threads or computers), and use state of the art matrix solvers (Brunner et al., 2010; Kuffour et al., 2019). This allows model runtimes to scale (shorten) with the number of processors available, allowing them to take advantage of supercomputers with many available cores (ParFlow was originally run on the CRAY T3D, which could have between 32 and 4096 processors (Ashby & Falgout, 1996; Kessler & Schwarzmeier, 1993) – still many more than most computers commonly used for modelling today, although each processor was much slower).

The question is not whether these codes are exciting: they are undoubtedly impressive. The question is whether the complexities of ParFlow and HGS are necessary for the bulk of practical GW-SW modelling questions, or if the complexities conversely serve as an impediment to answering them. The appropriate level of complexity in models for any given problem is an ongoing, ever-evolving debate (Clark et al., 2015; Clement, 2011; Doherty & Moore, 2019; Ferré, 2017; Gómez-Hernández, 2006; Guthke, 2017; Hill, 2006; Hill et al., 2016; Hunt et al., 2007; Hunt & Zheng, 1999; Markstrom et al., 2016; Simmons & Hunt, 2012; Weijs & Ruddell, 2020). An answer to this question is beyond the scope of this thesis (and ever-changing, but well explored in the cited literature) but a discussion is warranted given that the software presented in this thesis both (1) presents a *less* complex GW-SW modelling framework than these prominent codes, and (2) advocates for the advantages of a flexible, modular modelling framework that can be used to build models of varying complexity.

The general consensus of the recent literature is that model complexity needs to be tempered by data availability and quantification of predictive uncertainty (Doherty & Moore, 2019; Guthke, 2017; Hill et al., 2016). Complex models, whether they are complex due to project scope (e.g., continental models), discretization (density of model input/solutions), or fidelity of process representation (e.g., solving 3D Richards' equation versus an approximation), come with increased run times, expanding the difficulty of calibrating the model (e.g., parameter estimation) and reducing and/or exploring the model uncertainty, both of which require a large number of model runs. Additionally, data requirements increase with greater model complexity, again either due to project scope, discretization (e.g., more temporal and spatial data), or due to processes representation (e.g., recharge versus full vadose-zone modelling). Data requirements can either come from the modelling side as model inputs, or from the calibration side, as observations used for model performance evaluation (history matching). Both have impacts upon the modelling calibration/parameter uncertainty process: rarely are model inputs known at every required location in the model, leading to additional parameters to be estimated; and model observations are necessary for mathematically constraining the parameter estimation process (Doherty & Moore, 2019). When calibration takes place with less historical data than the number of parameters being estimated, the model is said to be overparameterized, which leads to a non-unique model – i.e., multiple sets of parameter values can result in equally optimal model fit to historical data (Anderson et al., 2015, p. 378; Beven, 2006). The excess parameters, resulting possibly from excess model complexity, allow the modeler or calibration procedure to train the model to the *noise* of the data, rather than signal of the data (Guthke, 2017). This can be dangerous, as the model may do a fantastic job of matching the historical data it is given, but since it is overfit that period, it will perform poorly when making future predictions (Fowler et al., 2020; Guthke, 2017; Seibert, 2003).

Of course, the sole purpose of modelling is not only decision support or hindcasting – models are also created for purely epistemic research purposes, allowing scientists to utilize their best understanding of a complex system to perform experiments impossible in real-world settings and develop new theories (Clement, 2011). However, science is built upon making falsifiable claims, and model complexity can also interfere with a researcher's ability to rigorously, or transparently, test hypotheses (Hill et al., 2016). Software may calculate "correct" answers to synthetic problems with ideal inputs, but structural issues may emerge at different scales or with varying discretization, for example, that can only be discovered with investigations into sensitivity and uncertainty (Beven, 2005; Doherty & Welter, 2010). Sensitivity and uncertainty analyses require repeated application of similar simulations, rendering the computational burden of fully coupled approaches even less feasible.

Comparisons of GW-SW modelling software, including HGS and ParFlow, have focused upon model accuracy at solving synthetic problems (Kollet et al., 2017; Maxwell et al., 2014; Sulis et al., 2010) rather than run times and predictive uncertainty. Differences in modeler skill, level of effort, bias, and approach (e.g., "modeler uncertainty") do influence modelling outcomes (Ferré, 2017; Hämäläinen, 2015; Hunt et al., 2020; Linkov & Burmistrov, 2003; Matott et al., 2009). However, considering many practical problems in hydrology, and in GW-SW modelling, involve hypothesis testing and/or decision support modelling of complex natural systems, the ability to solve the problem within a specified tolerance limit is worthless information unless the uncertainty in that prediction (e.g., risk of a "bad thing" happening) can be adequately addressed. Chow et al. (2016), a baseflow contribution modelling software and methods comparison study between HGS, WATFLOW, FEFLOW 7.0, and MODFLOW 2000, exemplifies this tradeoff between model complexity and calibration. While HGS was the only software they used to model the unsaturated zone, the long model run times prevented them from performing automated calibration on the model. "HGS is computationally demanding and required approximately 18 h to reach a steady-state solution, while WATFLOW, MODFLOW, and FEFLOW could reach a steady-state solution within 1.6 min, 18 s, and 17 s, respectively.," the authors write. "The difference in run times played a key role in model calibration." The results from HGS were not more accurate than the other models, however, it should be noted that the two-way exchange fluxes calculated from the HGS model served as an input to the other models (Chow et al., 2016).

The GW-SW modelling software presented in this thesis has been designed in reaction to the practical issues associated with deploying complex, fully coupled GW-SW modelling software discussed in this section. The flexibility built into the framework enables stepwise increases in model complexity, process hypothesis testing, adaptable discretization, and balanced abstraction (Craig et al., 2020). Speed (e.g., model run times) has been prioritized in the software development process to facilitate parameter estimation, sensitivity analysis, and other uncertainty-related explorations. The complexity built into HGS and ParFlow will be necessary for certain problems, but for many others a less complex solution may be more optimal.

2.4 Flexible Process Representation in GW-SW Models

The Raven hydrological modelling framework is distinguished by its ability to represent hydrological processes using a myriad of algorithms in a flexible and robust manner (Craig et al., 2020). In the last section, the impact of software/model complexity on predictive uncertainty focused primarily on model *parameters*. Another source of uncertainty present in hydrological modelling is structural error (Beven, 2005, 2016; Doherty & Welter, 2010), which here is loosely defined as the error that occurs due to the modelling software not being a perfect proxy for the natural world. Particularly in process-based

surface water modelling, numerous algorithm options have emerged highlighting a wide diversity of conceptualizations of the hydrologic cycle, varying by complexity, application region, discretization scheme, process representation, scale, dimension, data requirements, numerical methods, among many other divisions (Clark et al., 2011, 2015; Craig et al., 2020). Taking these various approaches as hypotheses, or coupled hypotheses (Beven, 2002; Clark et al., 2011), it is clear there is no single dominant understanding of how to best mathematically represent all hydrological systems. This becomes troublesome when practitioners seek out software to model their particular problem: they are often either forced to inherit an existing hypothesis, the ramifications of which may be opaque to the modeler, or they may opt to write yet another piece of hydrological modelling software. Neither is ideal, which has more recently led to a rise in research-focused flexible modelling software allowing for the evaluation of various process hypotheses and evaluation/exploration of structural error (e.g., Clark et al., 2015). The Raven hydrological modelling framework is unique in additionally allowing flexibility in numerical schemes, discretization, and interpolation methods complete with a vast process library of interchangeable options (Craig et al., 2020) combined with being well documented and relatively easy to use.

The groundwater modelling world has its own history of flexible modelling software, largely centering around MODFLOW, the **mod**ular ground-water **flow** model, originally developed in 1984 to consolidate the U.S. Geological Survey (USGS) various groundwater modelling codes into a single software that was easy to use and expand (Harbaugh, 2005). The groundwater modelling field benefitted from the existence of a relatively simple linear equation (the groundwater flow equation) describing water movement through their domain. This enabled the community to focus more on the physics of interaction at the boundaries of the saturated domain. Processes and capabilities (called "packages") could easily be added to MODFLOW by including additional files, and developers could add new packages without modifying the core groundwater flow equation or solver. For instance, several different ways of conceptualizing stream-aquifer relations are available in MODFLOW-USG: the drain package (DRN), the general-head boundary package (GHB), the stream package (STR7), the river package (RIV), the connected linear network package (CLN), and the streamflow routing package (SFR2) (Panday et al., 2013). Now in its sixth version (Hughes et al., 2017), MODFLOW continues to add more capabilities and flexibility with each version, with independent developers adding their own packages as well.

Several of the existing GW-SW software packages in Table 2-1 allow for multiple levels of complexity. Some, using MODFLOW to represent the subsurface, inherit some of its packages as flexible representations of hydrologic processes. Some of the more advanced options come from the commercial products, MIKE SHE and MODHMS. MIKE SHE, in particular, contains modules for representing climate, stream, snow, irrigation, and vadose zone processes (Hughes & Liu, 2008), combined with a GUI

that facilitates adjustable spatial and temporal model discretization. MODHMS strongly resembles a greatly expanded version of MODFLOW, including multiple ways of connecting the SW and GW domains (Panday & Huyakorn, 2004). Users can select a time-lagged loose coupling, an iterative coupled solution, or a fully coupled solution. Panday & Huyakorn (2004) state that the fully coupled scheme is the most robust, efficient, and theoretically correct of the three, but that the loosely coupled options advantageously offer different time step sizes to be used for the different domains.

The work presented in this thesis can be thought of as adding a new option for representing groundwater flow into the Raven hydrological modelling framework: a control-volume finite-difference model of the subsurface with all the features of the popular groundwater modelling software MODFLOW. Previously, Raven did not have a conceptualization of the subsurface that allowed water to flow between subbasins, so this enables the modeler to test new additional hypotheses of intra-watershed flow. While only a simple set of connections between the surface and subsurface has been developed so far, the work presented here serves to build the initial infrastructure to test different hydrologic connections and even coupling methods between the two domains. This, importantly, allows the Raven tenets of flexibility and intelligent complexity to extend to this new subsurface element.

Chapter 3 - Methods

3.1 Modelling Software Being Coupled

This section provides some basic background on the surface water (SW) and groundwater (GW) software that has been loosely coupled in this thesis. In Section 2.4, the choice of the modelling software was briefly examined in context of their flexibility among the diversity of other options available. Here the focus is on the specifics of how these codes represent their domains and the processes within, with the intent to provide the reader with an insight into the software sufficient to understand the nature of their coupling.

3.1.1 The Raven Hydrological Modelling Framework

Raven is a generalized surface water hydrological modelling framework built to facilitate stepwise model complexity, conceptual/numerical hypothesis testing, examining model structural uncertainty, and compare algorithmic representations of the natural world (Craig et al., 2020). The extreme flexibility in model structures, discretization schemes, and interpolation choices, along with a robust library of hydrologic process algorithms makes it unique in the world of surface water modelling. It can emulate other existing modelling software including the UBC watershed model, GR4J (Perrin et al., 2003), MOHYSE (Fortin & Turcotte, 2006), HMETS (Martel et al., 2017), and HBV-EC (Hamilton et al., 2000; Lindström et al., 1997), and perhaps more interestingly can be used to expand these popular models without ever needing to write a line of code. In a sense, it turns the full power of hydrologic models over to the modeler, and away from just those with software development skills. It is open source and written in C++ with an object-oriented design that facilitates an easy addition of processes to the existing hydrologic process library.

The flexibility and power of Raven also make writing a succinct description of the formulation difficult. The reader is directed to Craig et al. (2020) as well as the comprehensive software manual (Craig & the Raven Development Team, 2020) for a more complete explanation with important context to the features.

In Raven, a watershed or catchment is composed of subbasins, which can then be sub-discretized into hydraulic response units (HRUs) (Figure 3-1). Land use, vegetation cover, terrain, and soil profile are defined at the HRU scale, allowing for Raven to be used as a semi-distributed model (Craig & the Raven Development Team, 2020) where spatial variability is represented by classes. HRUs are treated as homogenous units with a uniform response to climatic forcings (Craig et al., 2020). Raven can also represent watersheds with a lumped discretization (Figure 3-1 d), where effectively the entire model is a single subbasin/HRU. Lumped models often use averaged catchment parameters and generally represent

empirical and conceptual models, trading heterogeneity for fast run times (Sitterson et al., 2017). Raven can also emulate grid-based semi-distributed models (Figure 3-1 b, c) by defining subbasins as grid cells with sub-grid HRUs, or even defining the subbasins and HRUs at the same grid scale (effectively creating a fully distributed model). Since HRUs are defined by their membership in subbasins, and subbasins are connected solely through an upstream/downstream relationship, spatial discretization, and geometric relationships between HRUs is largely arbitrary to Raven (with the caveat that latitude/longitude coordinates are used to distribute weather station data to HRUs). The geometry of HRUs, with the exception of total area, is not an input to Raven.

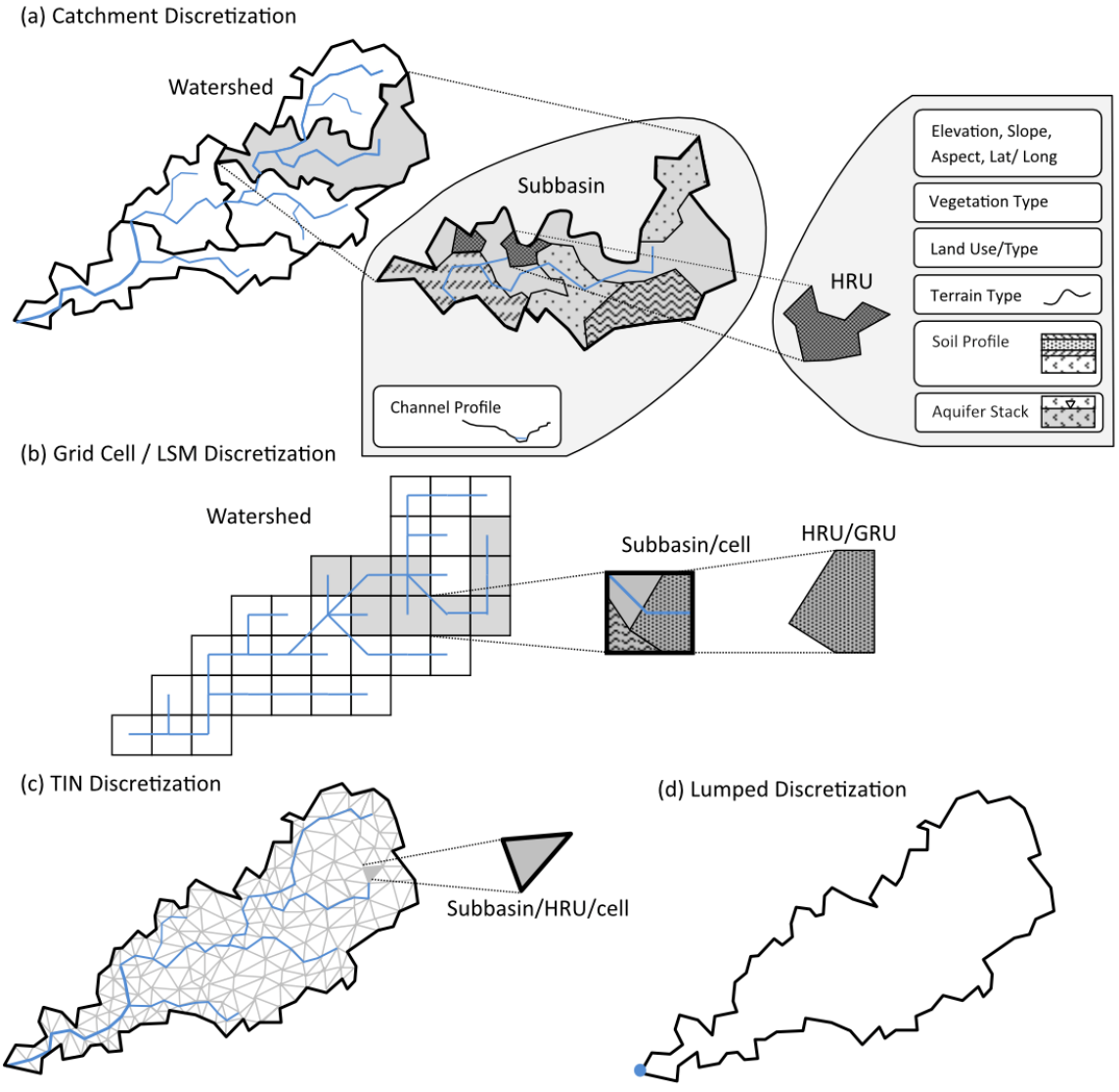

Figure 3-1 – Raven features flexible watershed discretization options, including (a) semi-distributed by subbasin, (b) semi-distributed by grid cells, (c) triangulated irregular network and (d) lumped (homogenized). Figure from Craig et al., (2020), reproduced with permission.

Water and energy balances are solved for each HRU, separating and distributing precipitation into HRU storage compartments (e.g., canopy, snow, runoff, soil, groundwater) based upon the hydrological processes enabled and the chosen solver. The distribution of precipitation into rain and snow can either be input directly as forcing data or calculated by various algorithms. In-catchment routing moves water from the various HRU stores into the basin stream channel via convolution-based methods, replacing the need for representing all the many complicated (and often unknown or unknowable) processes than move water internally within the basin. In-channel routing moves water within/between subbasin channels either using convolution-based, empirical (e.g., Muskingum-Cunge), complex (e.g., kinematic wave) or other routing options. Reservoir and lake routing can be included as well (Craig & the Raven Development Team, 2020).

Raven can be run with daily or finer (shorter) time step intervals. Forcing data does not need to be provided at the same interval as the model is run but must cover the entirety of the model duration. Raven forcings are interpreted as constant rates during the specified time interval, thus the temporal discretization does not have to match. Processes in Raven are able to extract the necessary values over the period they require (Craig et al., 2020).

Prior to this work, groundwater existed as an HRU-level storage compartment in the form of either a "deep groundwater" compartment or multi-layer aquifers with limited storage. Baseflow representation would typically be treated by a linear or power-law constitutive relation between flow rate and storage; aquifer heads were not explicitly represented. As such, there was no need to characterize distributed hydraulic conductivity or specific yield with the prior conceptualization. Subsurface water movement between HRUs had no representation, however, the HRU aquifers could "fill up", limiting/preventing further infiltration. Losses to pumping wells could not be simulated, as stream interactions were limited solely to baseflow; stream losses to groundwater and the head-dependence of this flux were not included in the prior representation.

3.1.2 MODFLOW-USG

Finite-element groundwater models (e.g., FEFLOW, IWFM) long have had the ability to use flexible meshes for spatial discretization, while finite-difference models like MODFLOW were limited to rectilinear grids consisting of rows and columns. MODFLOW-USG (Panday et al., 2013) is the first control volume finite-difference version of MODFLOW, allowing for flexible grids in the horizontal plane. This allows for the grid to better represent irregularly shaped boundaries (e.g., geological features, streams, lakes) as well as the ability to seamlessly vary cell (or in MODFLOW-USG parlance, "node") size through the model domain. MODFLOW-USG is also capable of supporting traditional rectilinear MODFLOW models as well as supporting transport simulations.

The following will provide a brief general background on what the various MODFLOW-family of codes (e.g., 2000 (Harbaugh et al., 2000), 2005 (Harbaugh, 2005), NWT (Niswonger et al., 2011), USG (Panday et al., 2013), Version 6 (Hughes et al., 2017)) do, but generally follows Panday et al. (2013) and thus is skewed towards the more recent versions. MODFLOW primarily solves the groundwater flow partial differential equation describing the flow of constant density water through a porous media:

$$\nabla \cdot (K \nabla h) = S_s \frac{\partial h}{\partial t} + W \tag{7}$$

where K is the hydraulic conductivity, h is hydraulic head, S_s is specific storage, t is time, and W is a volumetric sink/source per unit volume, which may be used to represent (e.g.,) pumping wells (Harbaugh, 2005; Panday et al., 2013). The model domain is discretized into volumetric nodes (cells) associated with aquifer/porous material properties (e.g., hydraulic conductivity, anisotropy, specific storage, specific yield). Various boundary conditions (as explained in Section 2.2) can be used represent sinks and sources such as wells, streams, and recharge as well as actual hydraulic boundaries at the edge of the model domain (e.g., no flow, constant flux). The volumetric flow between two nodes, *n* and *m*, is a function of their inter-cell conductance, C_{nm}, and the difference in their hydraulic heads:

$$Q_{nm} = C_{nm}(h_m - h_n) \tag{8}$$

For each time step, MODFLOW solves for the unknown heads ($\boldsymbol{h}$) in the model simultaneously using the matrix equation:

$$\boldsymbol{A}\boldsymbol{h} = \boldsymbol{b} \tag{9}$$

where $\boldsymbol{A}$ is the matrix of conductance terms between connected nodes, and $\boldsymbol{b}$ is the known right-hand-side (RHS) vector containing all the terms independent of the unknown heads at the end of the time step (generally, the sum of node storage, and non-head dependent changes in storage). On a per-node basis, the general form of the balance equation for node n is:

$$\sum_{m=1}^{p} C_{nm}(h_m - h_n) + HCOF_n(h_n) = RHS_n \tag{10}$$

where *m* is any of the *p* nodes connected to node n; *HCOF* is the sum of all the head-dependent coefficients related to node *n* (e.g., conductance terms from the river package). MODFLOW packages essentially operate by adding coefficients to *HCOF* and adding/subtracting volumes of water from the RHS vector to satisfy boundary conditions or otherwise represent internal fluxes.

If the aquifer system being solved is confined, then Eq. [9] is linear. However, MODFLOW models often include the unsaturated zone and thus the saturated thickness must be solved for, because C_{nm} may

be a function of h_n. This requires the use of more advanced solving techniques iterating over possible values of $\boldsymbol{h}$ until convergence is met. Various solver packages are available, depending on the version of MODFLOW and various options used. MODFLOW-USG contains a new sparse matrix solver (SMS) with an optional Newton-Raphson approach that advertises faster convergence for nonlinear problems (Panday et al., 2013).

MODFLOW discretizes time into two different divisions: stress periods and time steps. Stress periods are the interval at which time-variant parameter updates occur, e.g., changes in pumping rates at wells. Time steps are a subdivision of stress periods and correspond to when the matrix solver is invoked. For example, MODFLOW models can be run with yearly average pumping values (yearly stress periods) and time steps corresponding to each month of the year. MODFLOW allows the modeler to input the length of each individual stress period and the corresponding number of time steps separately, allowing the interval at which the model writes output can change over time. Unlike Raven, MODFLOW has no internal understanding of dates or calendars and thus the stress period and time step durations are unitless and must be interpreted carefully by the modeler.

The general MODFLOW-USG flowchart is shown in Figure 3-2. It corresponds very closely to the well-documented (commented) Fortran code. Memory is conserved by only storing one stress period of parameters at a time, read in at the start of every stress period. Solver settings include maximum number of iterations to obtain convergence (controlling the diamond “close” operation in Figure 3-2), as well as convergence criteria, preventing MODFLOW-USG from getting stuck in an endless loop.

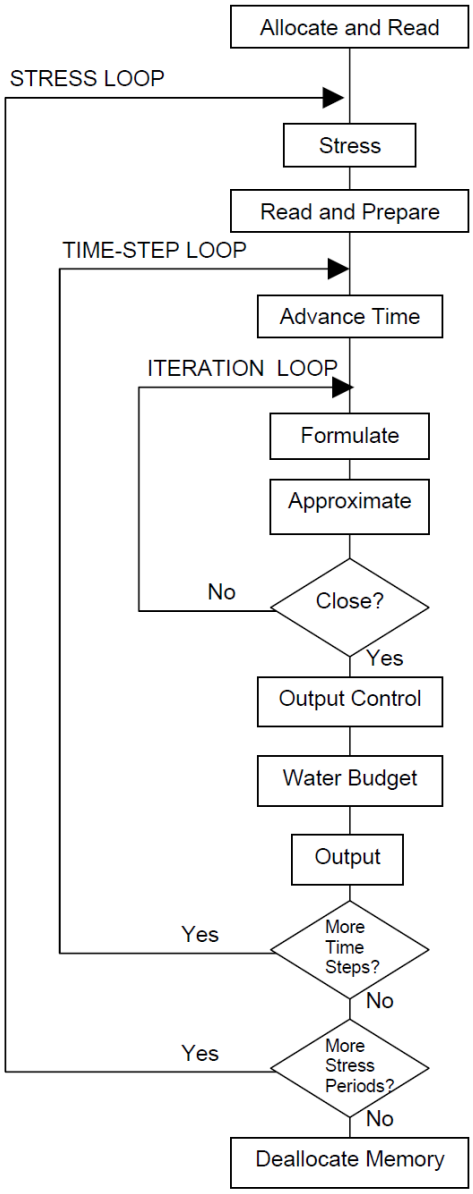

Figure 3-2 – General MODFLOW flowchart detailing the simulation of groundwater flow. From Harbaugh (2005)

3.2 River/Stream Representation – The PBJ Package

The second objective of this thesis is to develop and demonstrate a novel way of incorporating linear features (e.g., rivers, streams, horizontal wells, drains, barriers to flow) into unstructured grids. Here, "unstructured" is used in the sense of Panday et al., (2013): cells may connect to an arbitrary number of nodes in a 2D horizontal plane. As is explained in the formulation section, the method presented here specifically requires a grid for which a Delaunay dual triangular grid exists. Like MODFLOW-USG, for an accurate solution the cells should be convex and a line drawn between cell centers should cross the

shared face at a right angle (Panday et al., 2013). This section explains the method with which this thesis's second objective was accomplished and incorporated into both a specialized version of MODFLOW-USG and the coupled Raven framework that is the primary subject of this thesis. Chapter 4 contains verification and benchmarking tests of the method.

Groundwater modelling codes supporting unstructured grids (such as MODFLOW-USG and MODFLOW 6) allow for flexible discretization which can incorporate the locations of wells, streams, and other boundary features while varying cell sizes around these features and other areas of interest. However, in some modelling applications (for instance, at the regional scale), it is not always possible or advantageous to explicitly incorporate all features within the model grid discretization. Additionally, linear features like streams often still are simplified (i.e., straightened) to reduce additional grid refinement caused by small meanders. Here, a new method of incorporating complex polyline boundary conditions with arbitrary geometry in Voronoi or standard rectangular grids has been implemented into MODFLOW-USG. The module, called the Polyline Boundary Junction (PBJ) package, allows linear features to be represented appropriately at a subgrid-scale. It moves the fundamental unit of boundary condition discretization from the node (cell) to the linear segment.

Specifying river parameters at the segment scale provides several workflow advantages, specifically within the context of our coupled Raven framework. For one, information about river conductance, streambed elevation, and stage is often given at a finer resolution than regional-scale models can incorporate and is typically provided in terms of linear distance along a stream network. Secondly, vector-based shapefiles used in geographic information system (GIS) programs are almost always the basis of both river discretization in groundwater models and aid in watershed delineation in surface water models. An associated R package, *pbjR* (Scantlebury, 2020), has been created that facilitates easily going from shapefile to PBJ package, including writing the appropriate input file for MODFLOW-USG or Raven. Using one river shapefile for both SW and GW representations can streamline tasks and analyses.

3.2.1 PBJ Formulation

The formulation is reliant on a complementary triangular grid connecting the centroids of the model grid cells. Unstructured grids are often generated using a Delaunay triangulation between specified and/or generated points, which is the "dual graph" of the resulting Voronoi grid, a set of triangles with vertices corresponding to Voronoi circumcenters. It is also possible to form a Delaunay triangulation between structured rectangular grid cells, but the unstructured example is explored here for generality. Grid generation software, such as AlgoMesh (HydroAlgorithmics, 2016), are usually capable of outputting both the Voronoi grid and the Delaunay dual of any generated grid.

Segments of a linear feature such as a river, stream, or lake perimeter can be intersected with the triangular grid (Figure 3-3). For each segment within a triangular cell, the start and end points are described using barycentric coordinates. A barycentric coordinate system has three coordinates, one for each point of the triangle. The value of a coordinate is one at its corresponding corner point and zero at the triangle side opposite the point. Within the area of the triangle the coordinates always sum to unity and are all positive (i.e., above zero). In the triangular dual of the model finite-difference grid, the triangle points align with points at which the flow solution is solved. The barycentric coordinates of the segment therefore can be understood as describing the geometric relationship of the segment to the three model nodes connected by the triangle.

Figure 3-3 – Example of a segment (AB) within the triangular dual grid, with barycentric coordinates labeled for the segment start/end and the three connected model nodes. Figure originally appeared in Scantlebury & Craig (2020).

Designating the points of the triangle as 1, 2, and 3 and the start and end of the segments as A and B, as shown in Figure 3-3, for each point (model node) the corresponding head-dependent instantaneous total flow over the length of the segment can be calculated using:

$$Q_n = \int_A^B C \cdot t_n(L) \cdot \big(h_{GW}(L) - h_{Ref}(L)\big) dL \tag{11}$$

where n is the point 1, 2, or 3, Q_n is the flow (L^3/T) (adopting a positive value when water is leaving the aquifer) distributed to cell n, C is the conductance of the segment, t_n is the barycentric coordinate for point n, h_{GW} is the groundwater head over the segment, and h_{Ref} is a reference head, which differs depending on the form of boundary condition (e.g., drain elevation, river stage). This formulation ensures that the total flow (Q_T) from the linear feature is properly distributed to the three connected nodes (Eq [12]) through the barycentric coordinates, which effectively act as weights on the total flow.

$$Q_T = Q_{n1} + Q_2 + Q_3 = \int_A^B C \cdot \big(h_{GW}(L) - h_{Ref}(L)\big) dL \tag{12}$$

Assuming a planar head distribution over the triangle, the head at any point p in the triangle can be calculated as a linear function of the three connected node heads (h_1, h_2, h_3) and the barycentric coordinates of the point ($t_{p_1}, t_{p_2}, t_{p_3}$):

$$h_p = h_1 t_{p_1} + h_2 t_{p_2} + h_3 t_{p_3} \tag{13}$$

Across the segment, we can also expect the conductance and reference head to vary. Assuming that these values are specified at the segment ends and linearly interpolated in between, A and B, we can simplify the integral in Eq. [11] using the trapezoid rule:

$$Q_n = \frac{C_A t_{A_n}\big(h_{GW_A} - h_{Ref_A}\big) + C_B t_{B_n}\big(h_{GW_B} - h_{Ref_B}\big)}{2} L \tag{14}$$

where the various variables of Eq. [11] are now explicitly being defined at the start (point A) and end (point B) of the segment. Heads at the start and end (h_{GW_A}, h_{GW_B}) are calculated using Eq [13]. Depending on how conductance is specified, the length L of the segment may not appear in the equation, i.e., Eq. [14] assumes a unit-length conductance. This equation must be computed three times, for each of the points (associated with a model node) of the triangle to obtain the proper distribution of the fluxes. The barycentric coordinates, t_n, will always add to one so the total flux through the segment (Eq [12]) can be calculated by replacing the coordinates in Eq. [14] with unity.

The MODFLOW-USG PBJ package supports three boundary condition types: river, drain, and constant head. The river mode and drain mode work exactly as the RIV and DRN packages of MODFLOW, respectively, as detailed in Rivers and Streams (Section 2.2.1). In MODFLOW, constant/variable head boundaries are generally implemented by not including the node in the system of equations (i.e., IBOUND < 0) and simply substituting the input specified heads for the solution values

(Harbaugh, 2005, secs. 4–2). Using the same strategy for the PBJ package was both invasive to implement and possibly ill-advised considering that three nodes are connected to every segment – the result would likely remove extraneous cells from the solution. Instead, a very large conductance value is used on both sides of the matrix equation (Eq. [9]) to force the solver to a specified value. This the method suggested for specified heads in finite-element models in Wang & Anderson (1982, p. 128). In the form of Eq. [10], this means HCOF is assigned to a very large number, and RHS is that same large number multiplied by the intended head value. The constant head option was not included in the Raven version of the PBJ package, as the formulation does not make it easy to determine the fluxes to/from the specified head segments.

Two different conductance options exist for the PBJ package, both specified at the beginning and end of every segment. The first is a standard riverbed conductance, as used in the MODFLOW RIV package (Harbaugh, 2005, secs. 6–8):

$$C_{RIV} = \frac{K_{RIV} L_{seg} W_{seg}}{M_{RIV}} \tag{15}$$

where C_{RIV} is the conductance, K_{RIV} is the hydraulic conductivity of the riverbed materials, M_{RIV} is the thickness of the riverbed materials, and L_{seg} and W_{seg} are the length and width of the segment, respectively (Figure 3-4). Use of this conceptualization in the river package was previously shown in Eqs. [2,3], along with critiques of this conceptualization in Section 2.2.1. Note that in the PBJ package C_{RIV} is able to vary linearly over the segment length, from point A to point B (Eq [14]).

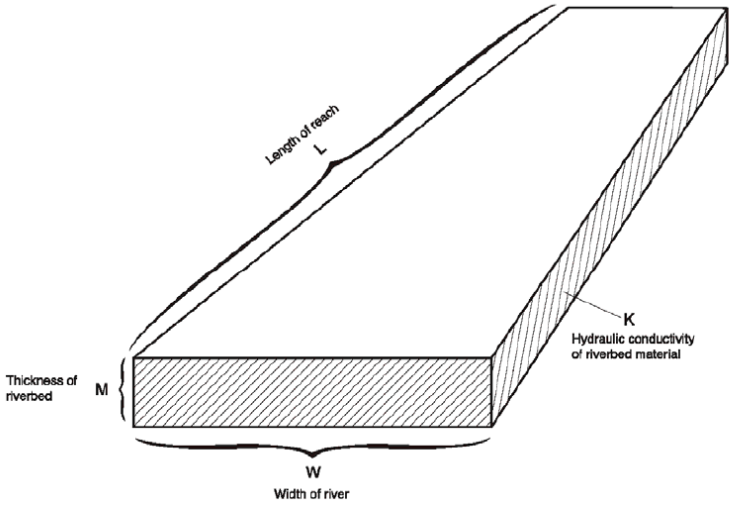

Figure 3-4 – MODFLOW idealization of the riverbed conductance formula, from Harbaugh (2005, secs. 6–8)

The second conductance conceptualization is simply a unit-length conductance, effectively identical to Eq. [15] but with the L_{seg} pulled out and moved to where it is seen in Eq. [14]. This feature was

created to facilitate model parameterization, as unit-length conductivities can be specified for entire reaches of rivers and streams (many connected segments). While our discussion here has focused on stream representation, these conductance formulations can be useful in many other linear feature contexts as well, such as tile drains. An improved formulation would likely vary conductance as a function of stream stage (e.g., Morel-Seytoux et al., 2018).

The intended workflow is for a modeler to be able to start with a polyline shapefile and easily produce a PBJ package input file. The following section briefly details this process and the *pbjR* R package developed to greatly simplify the process for practitioners.

3.2.2 From Shapefile to PBJ Package Input: *pbjR*

To avoid adding unnecessary complex subroutines to MODFLOW-USG (e.g., implementing shapefile readers), *pbjR* (Scantlebury, 2020) was developed in the R language (R Core Team, 2020) to perform the pre-processing using shapefiles, raster inputs, etc. and output the necessary inputs for both the *pbjR* package or the equivalent "GWRiverConnection" component for the coupled Raven framework. Additionally, the package contains analogous functions for creating/writing MODFLOW-USG drain (DRN) and river (RIV) packages, as utilized in the PBJ comparison tests in Chapter 4. The package functions are documented, in detail, in the R package itself. The workflow that follows will focus on producing a PBJ package input for MODFLOW-USG and/or Raven, although the functions for creating DRN and RIV packages will be briefly explained as well.

The *pbjR* package, at a minimum to get started, requires a polyline shapefile of a linear feature (a stream will be assumed here), a polygon shapefile of the model grid, and a polygon shapefile of the model grid dual Delaunay triangulation. Additionally, a raster digital elevation model (DEM) of the region can be used to obtain streambed elevations along the line, although alternatively a slope and known starting elevation can be used instead. The *pbjR* package heavily relies on the *sf* package (Pebesma, 2018) for reading and working with shapefiles. The R functions detailed below aid in producing a single final "DataFrame" (essentially, a table of values) comprised of segments and their corresponding nodes, barycentric weights, lengths, elevations, and conductances. Not all values are necessary for all boundary conditions (e.g., the constant head mode does not require lengths or conductances). For MODFLOW-USG, the reference elevations (i.e., river stage, specified head) and conductances can vary for each stress period. Raven assumes constant parameters and estimates its own river stages.

After the river, model grid, and Delaunay triangle shapefiles are read in, the bulk of the calculations are handled by the function `calc_stream_voronoi`, which takes those three shapefiles (read in through the sf package) as inputs. Stream segments are divided up based upon the triangles (e.g., Figure 3-3), and barycentric coordinates are calculated for the start and end of the segment. The points of the triangles are

mapped to the model grid nodes, and thus the relationship between each stream segment start and end, the Delaunay triangulation, and model grid is defined. The output of the function is a DataFrame where each row represents a stream segment, with barycentric coordinates for both segment ends (spatial geometry for each segment is stored in the DataFrame as well).

There exist a variety of common situations where the steps outlined above for `calc_stream_voronoi` break because the model grid-Delaunay triangle relationship does not match the ideal conditions assumed by the PBJ package and the function. Generally, these issues are encountered at the edges of the model but can occur any place where grid refinement creates unusually shaped cells. A particularly egregious example is shown in Figure 3-5.

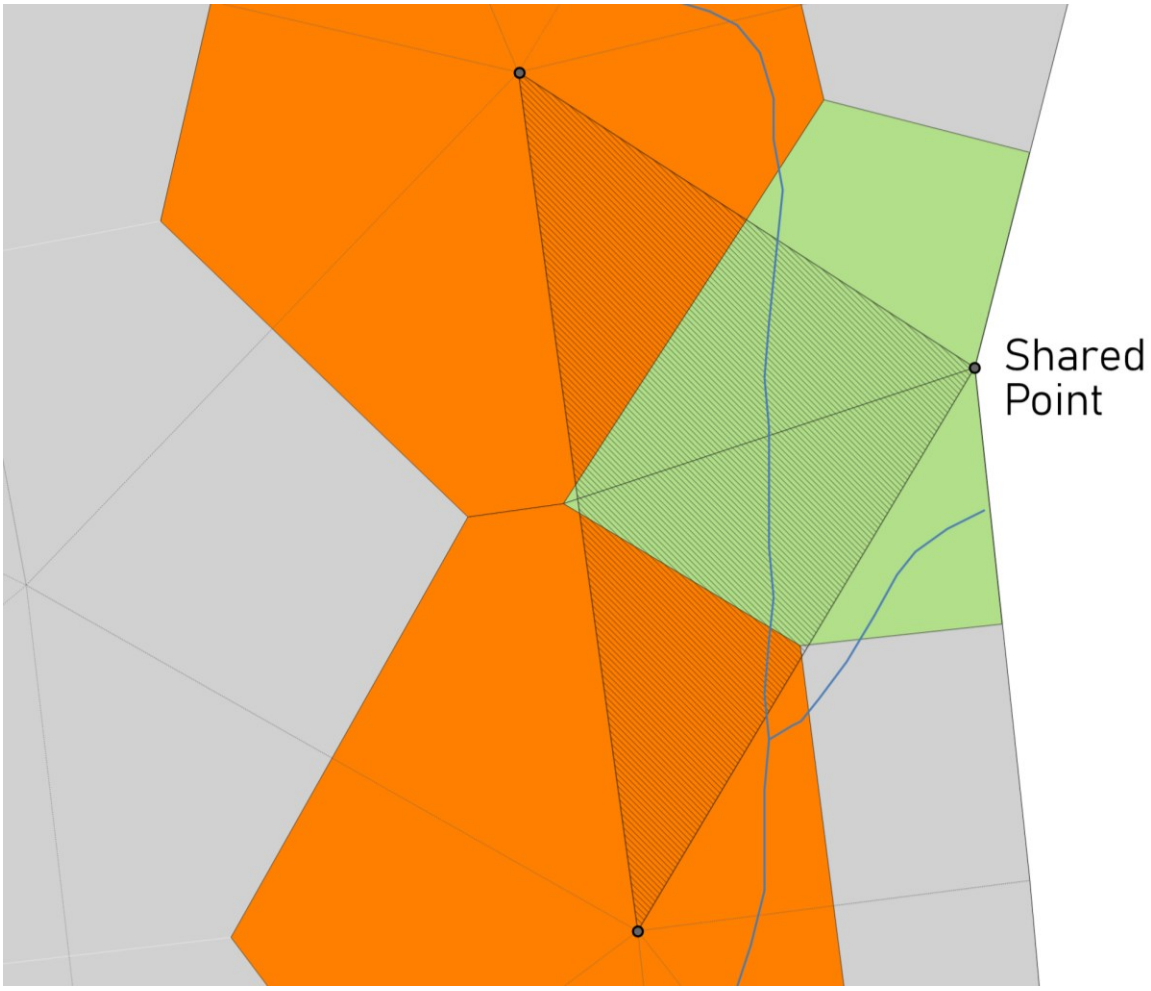

Figure 3-5 – Example Voronoi grid and corresponding Delaunay triangulation dual where a triangle (shaded) containing stream segments overlaps four nodes, creating issues for the PBJ formulation and *pbjR* package.

The shaded triangle in Figure 3-5 has two issues:

1. It overlaps four nodes, rather than the expected three, which causes one of its points to be at the intersection of two cells (rather than a cell center) making it hard for an algorithm to determine unique node ownership.
2. The segment must have no more than three nodes defined for the PBJ package, however, any combination of the three cells results in one of them not sharing a face with one of the others. This creates an issue within MODFLOW-USG when trying to add the off-diagonal conductances for the three nodes, since by default[1] only nodes with shared sides appear in the coefficient sparse matrix the PBJ package interacts with.

The `calc_stream_voronoi` function can address issue (1), but handling (2) has not been implemented. No fix from the PBJ package is going to rectify the issue properly; the grid is violating the assumptions of the package. The function attempts to return a suitable set of nodes by seeing if any of the nodes do not contain stream segments (often a decent fix for (1)), but if more than three nodes still remain, it will select the largest node. Either of these fixes can then result in (2).

To deal with (2) a manual fix is often necessary. This can usually be accomplished by reducing the connection list to two nodes through duplication (e.g., going from 1, 2, 3 to 1, 1, 2) and then entering a weight of zero for the duplicated node (e.g., 0.5, 0.0, 0.5). As long as the remaining two weights add to one, the flow solution should not be impacted noticeably. This error has only ever been observed at model edges. It likely can be entirely avoided by ensuring the groundwater model grid boundaries are several nodes away from any rivers or streams.

The *pbjR* package also contains functions for cleaning the segment list based upon a minimum length. The segments are removed, but their length (and line geometry) is added to another adjacent segment. This can help remove very tiny overlaps that occur in the segment/triangle overlap calculations.

Two functions exist for calculating segment start and end elevations for the segment list. The first, `stream_elev_from_slope`, is able to calculate stream elevations from a specified starting elevation and a streambed slope [L/L], assuming no stream branches. The second, `extract_stream_elev_from_raster`, utilizes a DEM of the region to determine the elevations of the start/end of each segment.

Conductances for the stream segment start and ends can be calculated using several different functions. The `calc_conductance_modflow` function uses the standard MODFLOW conceptualization (Eq. [15]), accepting values for streambed conductance, streambed material thickness, and stream width. A similar function, `calc_conductance_modflow_perLen`, performs the same calculation but without using the segment lengths, so that the returned conductances are per unit length. Two additional functions,

[1] This can be changed, but currently the PBJ package avoids altering the default sparse matrix setup.

unused in this work, implement experimental alternative conceptions from Mehl & Hill (2010) (Eqs. 3 and 4, as referenced in the publication).

Once all the segment-node connections and properties have been generated, the package can be used to write out input files for MODFLOW-USG or for Raven. The functions `write.PBJpackage` and `write.RavenGW` both essentially accept the same inputs: the segment DataFrame, output filename, boundary condition mode, and conductivity type. However, the Raven function also requires a subbasin ID for each segment in the segment DataFrame, and the MODFLOW-USG PBJ package function requires the number of stress periods in the model (it additionally, given additional DataFrame columns, can write different reference elevations and conductances for each stress period.

A new shapefile of the resulting segment properties can be easily exported using the `write_sf` function from the sf package. This can be very useful for error checking and visualization.

As mentioned before, the *pbjR* package also contains functions for developing RIV and DRN package input files for MODFLOW-USG. These functions perform operations analogous to the ones above but working with the nodes the polyline passes through rather than the individual segments. Cells where multiple lines/line sections pass through nodes are consolidated (as documented in Section 4.1) and cells with very small line overlaps can be removed. The resulting DataFrames can be written out to complete RIV and DRN files using `write.DRNpackage` and `write.RIVpackage`, respectively.

3.3 Model Coupling Strategy

The "strength" of coupling between GW and SW conceptualizations was discussed in Section 2.3, Existing GW-SW Modelling Software and Classifications, where *fully* coupled and *loosely* coupled solutions were explored. Similarly, the actual coupling of these *codes* can have various degrees of connection. For this purpose, software couplings will be defined as *united* and *separated*. Here a united coupling is software that internally shares data – e.g., the surface water system can directly access groundwater system state variables stored in the computer memory. This also can include couplings through a middle code such as an OpenMI (e.g., Fenske et al., 2009), which enables compatible codes to exchange memory. A separated coupling is when the two codes are distinct, unconnected executables and thus share data through either writing files (e.g., Tolley et al., 2019).To the author's knowledge, no work exists directly comparing types of software coupling within the context of environmental/hydrological modelling. However, the advantage of united coupled codes is they can easily access the data without either (1) the computational costs associated with writing files to disk or (2) codes not having full access to the state variables of the other code. The downside is it necessitates using a new version of a code (e.g., a special version of Raven that includes MODFLOW-USG) which can pose problems when, for instance, government contracts require pre-approved codes unmodified from a specific version.

The work presented here features a united version of the Raven and MODFLOW codes. The following sections explore the technical aspects as well as processes representations of the coupling. The last section briefly explores the routes for future expansion.

A primary component of the united connection is the spatial relationship between the Raven HRUs and the groundwater model nodes. This is handled through the Raven groundwater (rvg) file via an explicitly input node-HRU mapping that can be automatically generated by a R function. This function will eventually be included in the Raven support and analysis R package, RavenR (Chlumsky et al., 2020). The mapping is based upon area, and groundwater Voronoi nodes can belong to multiple HRUs, or in the case of non-overlapping sections, zero HRUs. The mapping consists of a list each present node-HRU combination and the percentage ("weight", from 0 to 1) of the node covered by the HRU. Therefore, volumetric flows from the SW system to the GW system ($Q_{SW\rightarrow GW}$) and from the GW system to the SW system ($Q_{GW\rightarrow SW}$) can be calculated using weighted averages:

$$Q_{SW\rightarrow GW} = \sum_{i=1}^{n} q_k \, w_{ik} A_i \tag{16}$$

$$Q_{GW\rightarrow SW} = \sum_{i=1}^{n} q_i \, w_{ik} A_i \tag{17}$$

where Q is the total volumetric flow [L^3/T], i is one of the n nodes connected to HRU k, q_k is the flux from HRU k, q_i is the flux from GW node i, w_{ik} is the percentage of area HRU k covers of node i, and A_i is the area of node i. For instance, recharge (Section 3.3.2) is calculated using a weighted average (Eq. [17]) for each node using the fluxes for each connected HRU. This formulation assumes that the HRU area, A_k, is equivalent sum of the overlapping nodal areas, i.e.,

$$A_k = \sum_{i=1}^{n} w_{ik} A_i \tag{18}$$

No checks currently exist in Raven to ensure this relationship is true – it is left to the modeler to ensure this is true. Raven does, however, check that node-HRU overlap percentages (w_{ik}) sum to one for each node and provides a warning where they do not.

It is worth noting that the node-HRU mapping means that HRUs, unlike in the uncoupled Raven, must have explicit spatial locations. However, like uncoupled Raven, they are not required to be spatially contiguous – i.e., an HRU may cover multiple locations within a basin.

MODFLOW packages generally only update stress rates (e.g., recharge, pumping) every stress period. However, to increase the flexibility of the Raven-MODFLOW connection this limitation has been

removed. Raven is capable of passing MODFLOW values at any interval or point that it requires updated rates. However, currently Raven is designed to pass MODFLOW updates at every time step. Various ideas and solutions are discussed in the Recommendations (Section 6.3), but the intent is that the groundwater model should be able to run at a multiple of the Raven time step size (e.g., months and days). Raven itself is locked at a maximum of single-day time steps.

3.3.1 The MODFLOW-USG Basic Model Interface

MODFLOW-USG is written in the Fortran programming language, while Raven is written in C++. To compile the two into a single executable the main MODFLOW-USG program was re-written into a series of *functions* that could be called by Raven. This is similar to a recent project undertaken by the USGS to create a MODFLOW 6 Basic Model Interface (BMI) (U.S. Geological Survey, 2020), so the same term will be adopted here, although the interface does not meet the official BMI specifications (i.e., Hutton et al., 2020). In Fortran, *subroutines* and *functions* are distinct units with different capabilities, however here *function* will be used to describe both since it is a more commonly understood term. The primary difference between the two is their method of returning values.

The MODFLOW-USG BMI (Figure 3-6) allows MODFLOW-USG to be run as a sequential set of functions. Variables normally scoped to the main program were moved into a new module, allowing the various loops within MODFLOW-USG to be controlled externally. Additional functions, on the right side of Figure 3-6, were created to allow access to USG variables and allow an external controller (i.e., Raven) to act as a MODFLOW package.

For example, `add_to_flow_eq()` allows the external controller to add/subtract values to/from the HCOF[2] and RHS arrays for a specific node (see Eq. [10]) and `add_to_flow_budget()` allows additions to the post-solution MODFLOW water budget arrays that are optionally printed in the listing file. Additional specific functions allow the external controller to query for specific node properties, such as node area (`get_node_area()`) and are well-documented in the code but omitted here.

This setup, notably, does not allow MODFLOW any control over Raven. Raven is able to call the Fortran functions it is given access to, but MODFLOW has no ability to query the internal state of Raven. This has important implications for development but does not pose any limits to the ability of the two united codes to communicate. Instead, it just dictates that all functions for sharing groundwater state variables must be written on the MODFLOW side but called from the Raven side. Examples already exist in the code for sharing numeric variables, arrays, and strings. However, multi-dimensional array sharing

[2] The conductance 2D-arrayhead coefficient matrix, named HCOF in normal MODFLOW, is referred to as AMAT in MODFLOW-USG

has been explicitly avoided due to the differences in how C++ and Fortran store arrays (column-major vs. row-major, respectively).

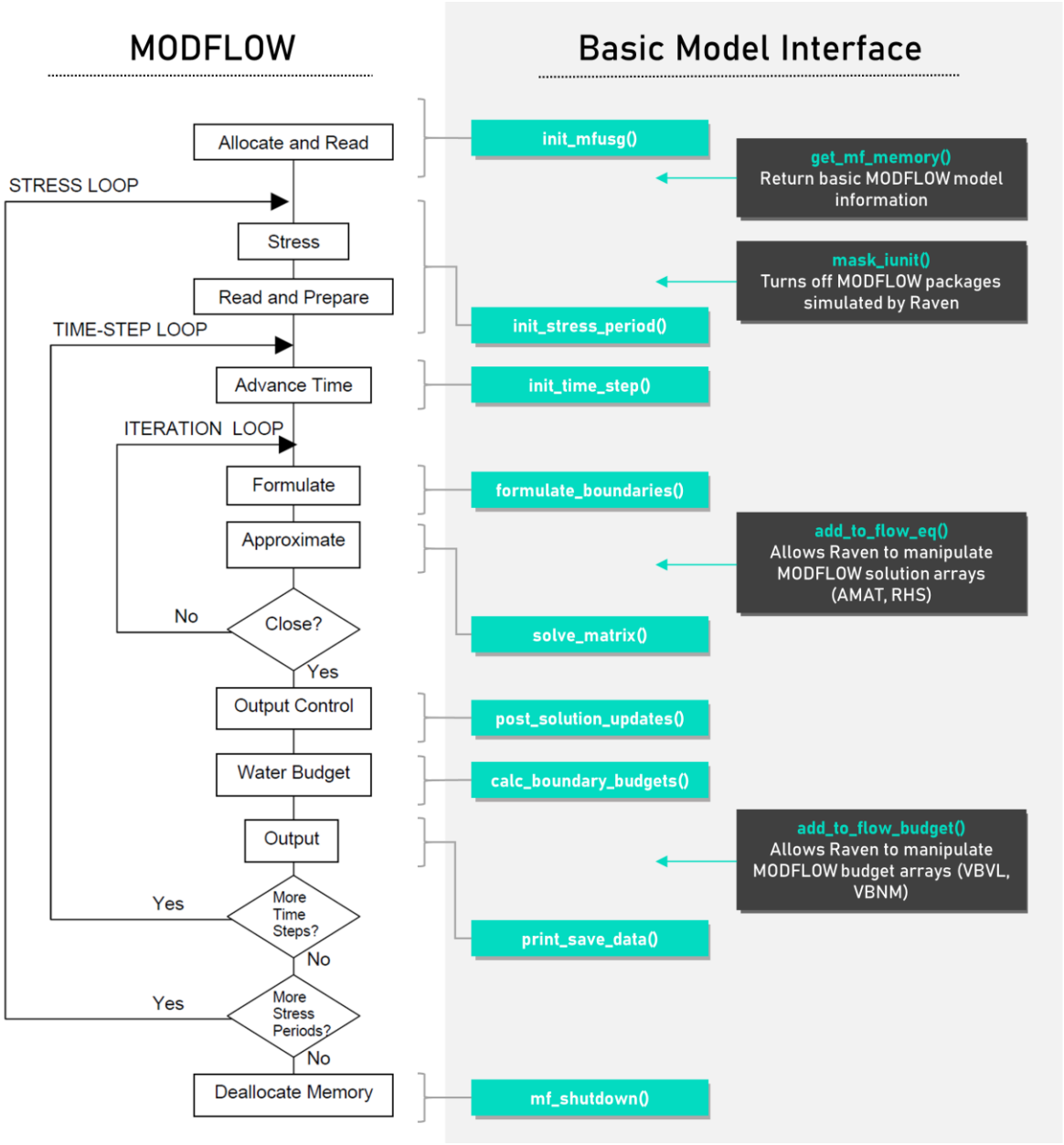

Figure 3-6 – Original MODFLOW simulation flowchart (left, Figure 3-2) and the Basic Model Interface (BMI) control functions (middle) that execute those sections of the MODFLOW code. Decisions (diamonds) that control the MODFLOW loop are handled externally. On the right are functions created for allowing the external program to act as a MODFLOW package. Minor functions for accessing and setting MODFLOW variables not shown.

To enable multi-language compilation for other developers (a priority since Raven is open-source software) regardless of platform, compiler version, and development environment, an instructional file for

the family of tools known as CMake was created (CMakeLists.txt). CMake is able to generate build files quickly and easily for a wide variety of integrated development environments (IDEs), making sharing the new coupled Raven framework between developers essentially no more complicated than before.

3.3.2 Groundwater Recharge Representations

Groundwater recharge – the process by which surface water enters the saturated zone of the subsurface – is one of the most important GW-SW processes in a coupled system. The ability to constrain groundwater recharge volumetric estimates based on precipitation, spatially varying by land use and soil properties, is a key upgrade from a stand-alone groundwater model. The coupled Raven framework has three ways recharge can be represented:

1. Any water moved to the "groundwater" HRU compartment in the arbitrary Raven conceptual model is deposited into the connected nodes prior to the GW system being solved. This is done using Eq. [17], where the depth of water in the compartment is converted to a volumetric flow rate using the node area.
2. A specific recharge hydrologic process was re-written in Raven, which can be used to move a specified flux to the GW system each time step. The rate can either be specified as a constant, or as time series data. Rates can be defined by HRU (as a forcing time series) or by node. The flux is distributed in the same manner as (1) above, using the nodal area to compute the total flow using Eq. [17].
3. Percolation from the PBJ package, representing a losing surface water feature. Just as in the RIV package in MODFLOW, water percolates from the stream channel down to the water table using Eq. [3]. The volume of water is removed from the channel storage after the GW system is solved.

Technically, all three representations can be present simultaneously, however (1) and (2) are redundant. Usually, (2) would only be used if recharge fluxes had been determined by another (external) model or the user was experimenting with constant recharge (to debug GW convergence issues, for example).

Inherently, (1) can be used to conceptualize recharge as the residual of the soil water balance (Figure 3-7). Raven uses a generalized soil representation which can represent the shallow subsurface using one or more layers (Craig & the Raven Development Team, 2020, p. 6). Layers have defined thicknesses as well as various properties which can be assigned through soil classes. While future versions of Raven may incorporate more complex integrations with the groundwater model water table elevation (e.g.), in its current form the vadose zone is simulated using Raven's existing formulation and processes.

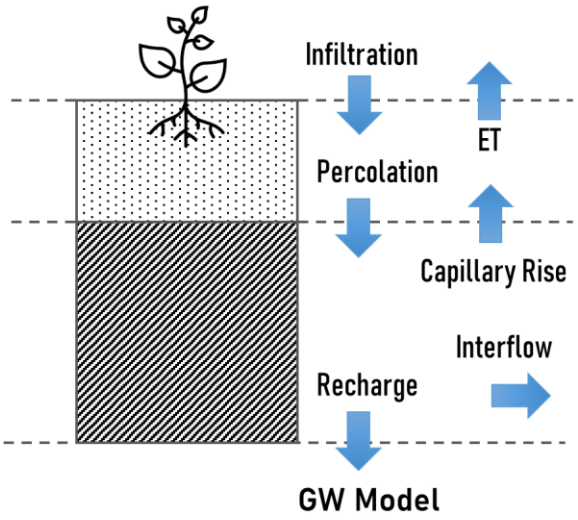

Figure 3-7 – Two-layer soil model with commonly represented Raven soil processes. Recharge can be conceptualized as the residual of the soil-water balance, or as a "demand" from the GW system.

3.3.3 Rivers & Streams

The PBJ package (Section 3.2) serves as the primary representation of GW-SW stream interactions in the coupled version of Raven. The primary difference in the Raven version of the PBJ package is the source of the river stage elevations used when in "River" mode (i.e., h_{Ref} in Eq. [14]). In the MODFLOW-USG PBJ package, the river stage elevations are required in the PBJ input file for each stress period, but when coupled to Raven, the PBJ package can access stream elevations at every timestep.

In Raven, channels are discretized at the subbasin level. The Raven version of the PBJ package (input into the .rvg file under the command "GWRiverConnection") therefore requires each segment to be assigned to a subbasin. As depicted in Figure 3-8, Raven calculates the river segment stages as the segment elevations added to the basin channel water depth. This is passed to MODFLOW-USG, where the elevations are used to calculate the groundwater flux to/from the river by the matrix solver. The flux solution is then returned to Raven, where it is totaled by basin and then added/removed to/from the basin channel storage.

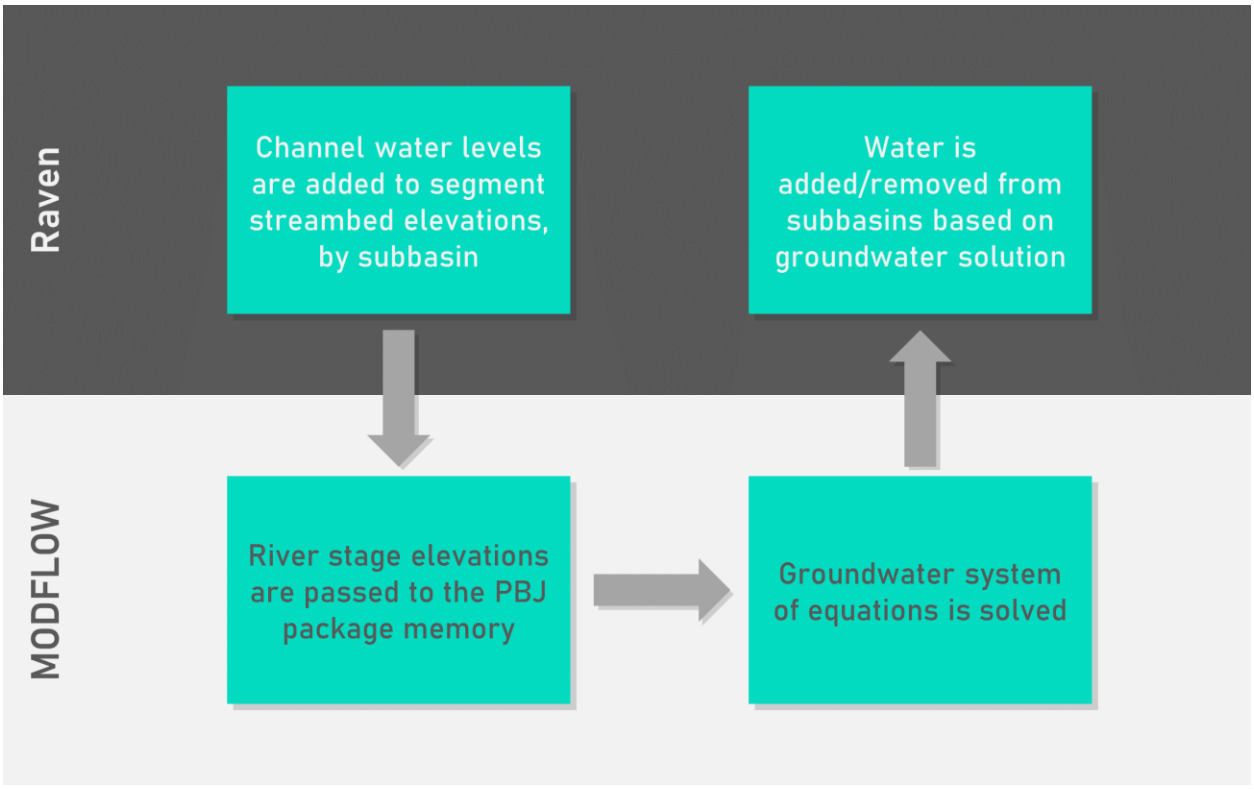

Figure 3-8 – Raven-MODFLOW-USG communication via the PBJ Package during a time step

Drains (as discussed in Section 2.2.1) can also be used to represent streams with the assumption of a purely gaining stream system. This simplifies modeled GW-SW interactions, potentially forming a more well-behaved coupled system (unidirectional stream flux) and ensuring MODFLOW cannot return a solution that fully empties the channel storage. In addition to drain being a boundary condition type available in the PBJ package, drains can be represented as a per-node process in Raven. They are implemented identically to how drains are conceptualized in MODFLOW (Harbaugh, 2005, secs. 6–12), but are controlled by Raven via the BMI, allowing water from the GW system to enter the SW water system at the HRU-level. The drain process may be ideal for representing intermittent (seasonal) streams or wetlands.

3.3.4 Accommodating Future Expansion

The primary contribution of this thesis is the development of the Raven to MODFLOW-USG coupling, which includes software infrastructure ready to be used to further expand the capabilities and processes in the coupled Raven framework. This section covers the GW-SW Process Abstract Base Class (ABC) which was added to the Raven code to facilitate hydrological processes that interface not just with HRU storage components, but with the groundwater state variables as well.

Classes, in object-oriented programming terminology, are effectively blueprints describing objects that can be created and manipulated within the program. Classes include state values (variables) and methods (functions). A class can "inherit" from another class (a "subclass"), acquiring most of the behaviors and properties of the parent "base" class. This setup allows programmers the ability to extend their software while reusing their existing code, among other advantages.

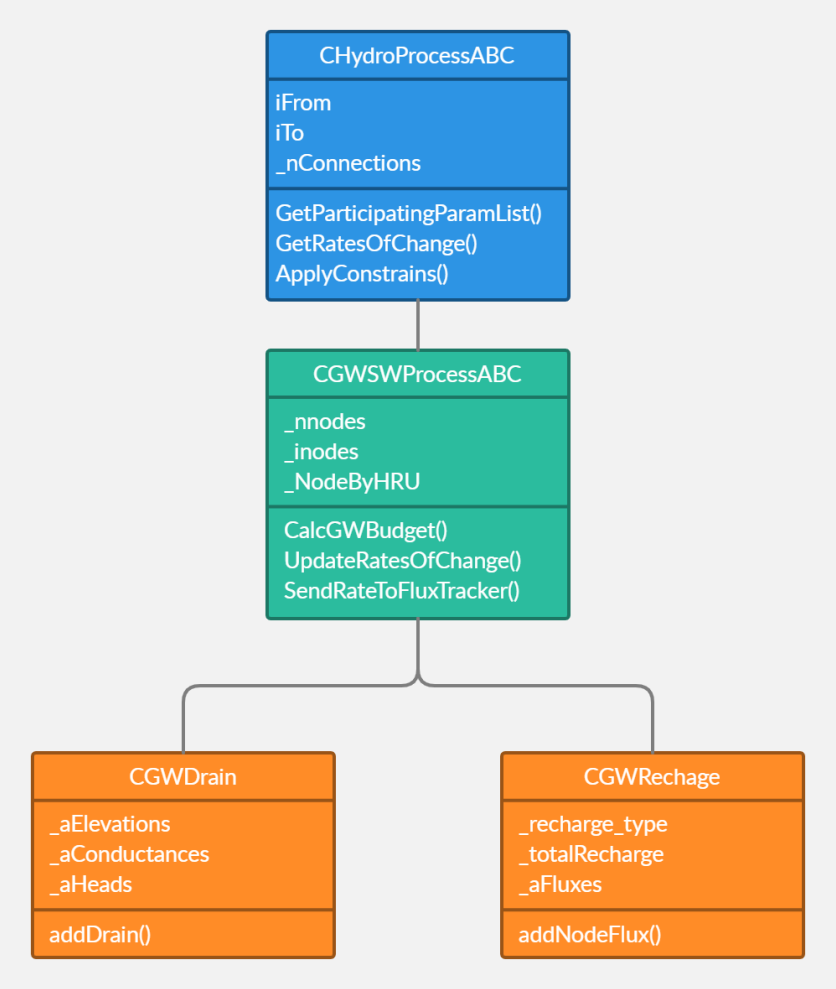

Figure 3-9 – Object-oriented class inheritance diagram for the new Raven GW-SW Process Abstract Base Class (ABC) and its current descendants. A subset of class member variables and methods are shown for each. The letter C in the class names denotes that they are classes.

The GW-SW Process abstract base class (GWSWProcessABC) is a subclass of the existing HydroProcessABC (Figure 3-9) in Raven, which is used as a base class for all hydrological processes. Here, the GW-SW Process ABC serves as the base class for all GW-SW processes in which Raven can act as a MODFLOW package. The GW-SW Process ABC contains additional variables for storing information about the groundwater model nodes associated with the process, as well as functions for working with fluxes to/from the groundwater model. GW-SW processes have no direct connection to the

groundwater model (e.g., they do not have access to the MODFLOW-USG BMI, Section 3.3.1), rather, they pass their per-node fluxes to a GW Model class that has been setup in Raven to track fluxes and interface with the BMI. By inheriting from the HydroProcessABC, these GW-SW processes can act as standard hydrological processes within Raven and move water between HRU compartments in the lateral process balance. However, since (at present) the GW solver is only run after all lateral exchange processes are calculated, any head-dependent GW-SW processes (e.g., drains) that add or subtract water from a surface water compartment are based upon groundwater levels from the previous time step MODFLOW-USG solution. An optional rudimentary correction system, only compatible with the Raven ordered series balance method, has been implemented to update compartment water balances after the GW model is solved.

Currently, only two GW-SW processes have been implemented: GWDrain and GWRecharge (Figure 3-9). The implementations of these have been discussed in previous sections (3.3.3 & 3.3.2, respectively). The classes are set up so that only one instance of them needs to be created: they are meant to represent their corresponding process for all nodes in the system with that process. However, there is no reason this must be the case; for instance, drain processes moving water to different HRU compartments (e.g., ponded rather than surface water) could be distinct.

The GW-SW Process ABC is envisioned as the primary expansion route of additional GW-SW processes in the coupled Raven framework. Ideas for additional processes and their implementation are discussed in Recommendations (Section 6.3)

Chapter 4 - PBJ Package Test Model

This chapter explores the results from several tests performed to (1) ensure the PBJ package works as intended and (2) compare its performance under variable grid resolution relative to common MODFLOW packages. The primary purpose of these tests is primarily to verify if the PBJ approach is a suitable replacement for the river and drain packages in MODFLOW-USG such that the package will be able to replicate their behavior in the coupled Raven framework. A secondary goal of the tests is to explore the effects of grid-scale changes on the calculated flows to and from the stream and compare those effects to the existing analogous MODFLOW packages. Given that the coupled Raven framework is aimed at regional scale modeling, in which ideally coarser grid scales (but finer time steps) can be used, the effects of grid scale upon the estimation of volumetric flows to/from the stream is important to explore.

4.1 Setup of Test Models

MODFLOW-USG groundwater flow models were created for testing the PBJ package against the standard MODFLOW River and Drain packages to compare performance, with the premise that the results from a finely resolved conventional model should be identical to that of a finely resolved PBJ model. The results herein only use the MODFLOW-USG PBJ package, and the model is not coupled to Raven. Reliance of the PBJ conductance parameter on grid cell size was also tested by generating these models at 14 different Voronoi grid resolutions (varying from ~13 m^2 to ~7000 m^2 in area). The model setup and river polyline was inspired by Mehl & Hill (2010) but recharge serves as the sole source of water entering the system rather than by using specified head boundaries. The model edges were all no-flow boundaries. To make the tests more challenging, wells were added at four locations such that they were centrally located in a MODFLOW cell at various grid resolutions (Figure 4-1).

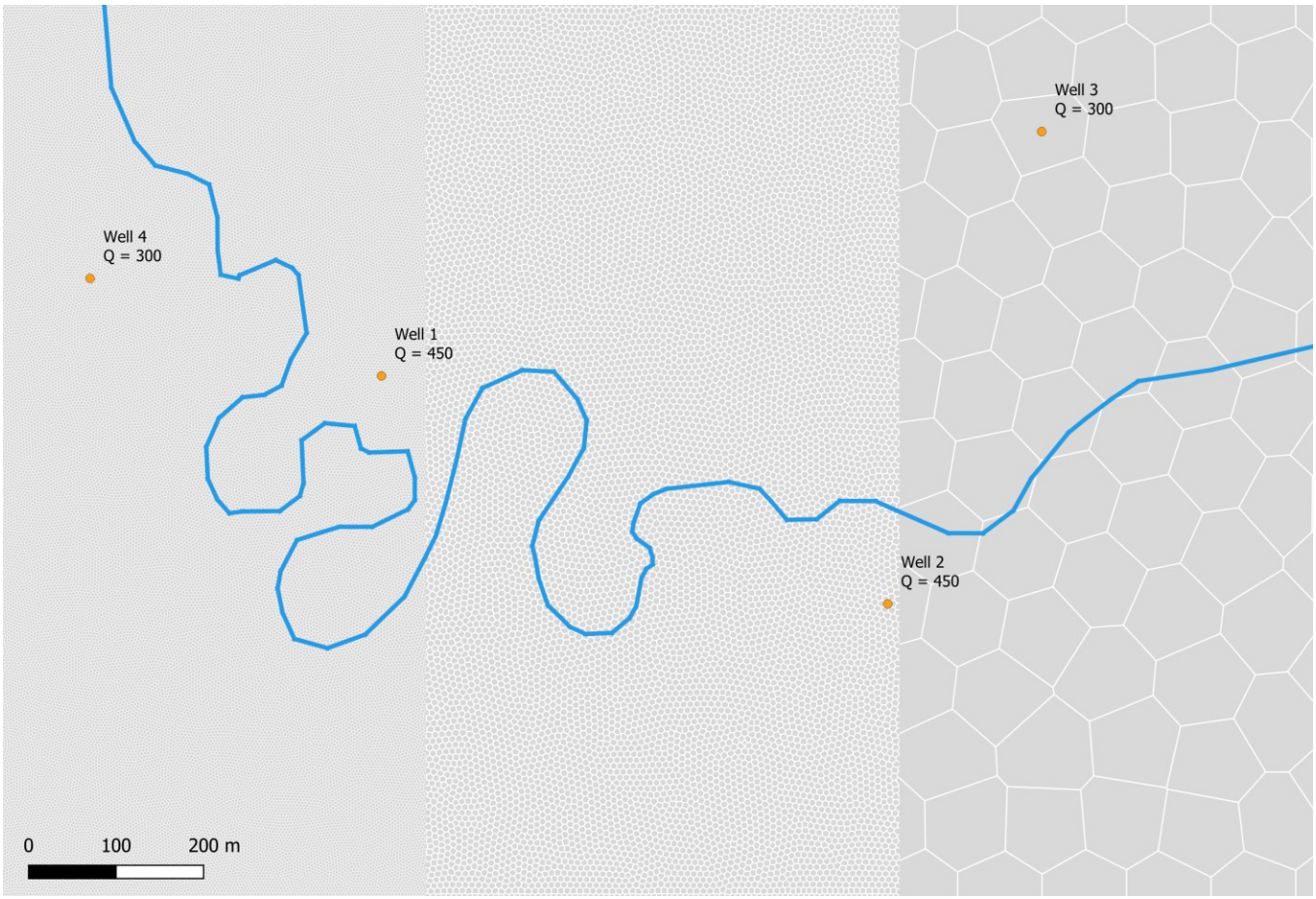

Figure 4-1 – Three of the 14 different model grids used to test the PBJ package, shown with the river polyline and wells. The number of nodes for the three grids are, left to right, 123,241 nodes, 32,011 nodes, and 232 nodes. These represent the finest grid, the second finest grid, and the coarsest grid. Pumping rates [m3/d] are shown for each well.

The 14 grids were generated using the AlgoMesh software program (HydroAlgorithmics, 2016) using only a model boundary and the locations of the four wells. Importantly, the river polyline was not used to discretize the model, since the PBJ package is specifically intended to correct for non-centered polylines. Settings were used to try to keep node areas consistent and well-optimized for the MODFLOW-USG solver. In lieu of being able to compare grid sizes by their rows and columns (like the rectilinear grids in Mehl & Hill (2010)), they are identified by their total number of nodes. The 14 different grids, their node count, and average nodal area are listed in Table 4-1. The model domain is 1029.4 by 1544.1 m with a uniform nodal top elevation of 55 m and a bottom at 0 m. Additional MODFLOW-USG files were generated using *flopy* (Bakker et al., 2016) as well as custom Python scripts. The Layer-property flow (LPF) package (Harbaugh, 2005) was used to specify aquifer properties for the model, controlling the flow between nodes. These properties were constant over the model domain and are displayed in Table 4-2. The initial head elevation for all nodes was 50 m.

Table 4-1 – PBJ test model properties for all 14 different grids

Grid	Nodes	Average Node Area (m^2)	Grid	Nodes	Average Node Area (m^2)
1	232	6851.3	8	8162	194.7
2	386	4117.9	9	14541	109.3
3	512	3104.5	10	19057	83.4
4	727	2186.4	11	22564	70.4
5	1393	1141.1	12	26609	59.7
6	2757	576.5	13	32011	49.7
7	5320	298.8	14	123241	12.9

Table 4-2 – Aquifer parameters for the PBJ test model

Parameter Name	Symbol	Value	Units
Horizontal Hydraulic Conductivity	K_x	1.000	m/day
Vertical Hydraulic Conductivity	K_z	0.100	m/day
Specific Storage	Ss	0.001	1/m
Specific Yield	Sy	0.300	-

PBJ package files for each grid were generated using the *pbjR* R package, described in Section 3.2.2. As in Mehl & Hill (2010), the river elevation started at 50 m on the left side of the model, dropping at a slope of 0.0015 m/m, down to just below 45 m at the east side of the model. Actual polyline shapefiles for the river were unavailable (the paper does not make clear if their synthetic model is based upon a real location) so the river was manually digitized in a GIS application from Figure 3 in Mehl & Hill (2010).

Comparison models were also built using the Drain (DRN) and River (RIV) packages instead of the PBJ package. The *pbjR* R package also contains functions for creating the necessary input files for these MODFLOW-USG packages.

The models had a single, unconfined layer and were run for two stress periods: an initial 30-day steady-state single step stress period, and a transient 365-day stress period with daily time steps. The recharge for the initial steady-state stress period was set to 2 mm/day and decreased to half (1 mm/day) for the transient stress period to allow for reactions to the change in the system to be observed. Pumping rates for the wells were constant over the simulation and are displayed in Figure 4-1 in m^3/d.

Although the model setup is loosely based on Mehl & Hill (2010), no attempt will be made to compare the results of these tests to theirs. The model setup was replicated simply due to its prominence in the MODFLOW river representation and scale issues literature.

4.2 PBJ Drain and River Test Results

This section evaluates the PBJ package results compared against equivalent setups using the MODFLOW-USG drain (DRN) and river (RIV) packages. As discussed in Section 2.2.1, the drain package can be used to model stream and rivers under the assumption that they are solely gaining from groundwater.

PBJ package input files were generated for all 14 models of varying grid size (Table 4-1). The stream conductivity was calculated using Eq. [15] with uniform values of K_{RIV} = 0.5 m/day, M_{RIV} = 0.5 m thick, and W_{seg}= 1 m wide. L_{seg} varied by the line segment length. Elevations at the segment start/ends were determined based on the stream slope of 0.0015 m/m. For the river boundary condition, a river stage is required, so a constant level of 1 m above the segment elevation was assumed.

DRN and RIV package input files were generated using the same parameters as the PBJ package. However, unlike the PBJ package, DRN and RIV parameters are specified by node, and therefore a single homogenous value must be used for each grid cell the package is active in. To calculate a node-scale equivalent streambed conductance, Eq. [15] was used, as above in the PBJ package inputs, but calculated with the entire length of the stream (L_{total}) passing through the node. The stream elevation was homogenized by calculating length-weighted average elevation for the node using the segment midpoint elevation (as calculated in the PBJ package):

$$h_{node} = \frac{1}{L_{total}} \sum_{s=1}^{nseg} L_s \frac{h_{A_s} + h_{B_s}}{2} \tag{19}$$

where h_{node} is the average streambed elevation for the node, s is one of the $nseg$ stream segments within the node, L_s is the length of segment s, and h_{A_s} and h_{B_s} are the elevations at the start and ends of segment s (see Figure 3-3). Note that the sinusoidal curves of the stream (Figure 4-1), particularly at coarse resolutions, cause the stream polyline to overlap nodes at non-consecutive locations.

Tests showed the PBJ package performs as intended, generally replicating the results of the DRN and RIV packages. Interestingly, however, as the grid sizes become coarser, the PBJ and MODFLOW solutions diverge, with the PBJ package at coarse resolutions better representing the fine-resolution MODFLOW head results. The top plot of Figure 4-2, for example, shows the along-stream node heads of PBJ drain mode and the MODFLOW drain package. These elevations reflect the actual GW level heads being used in the flow calculations. For the DRN package, this simply means the heads at each node where a DRN is active, but for the PBJ package this is a weighted average of the three nodes connected to the segment (Eq. [13]). The PBJ package, at the coarse scale, can better replicate the head distribution at the fine scale. Summary statistics describing the residual distribution of the along-stream GW levels

compared to the finest (grid #14) DRN package levels. The DRN package suffers from seemingly erratic fluctuations at the coarse scale, caused by the abrupt changes in node along the stream. The corresponding head contours for the finest (grid #14) and coarsest (grid #1) models can be seen in Figure 4-3.

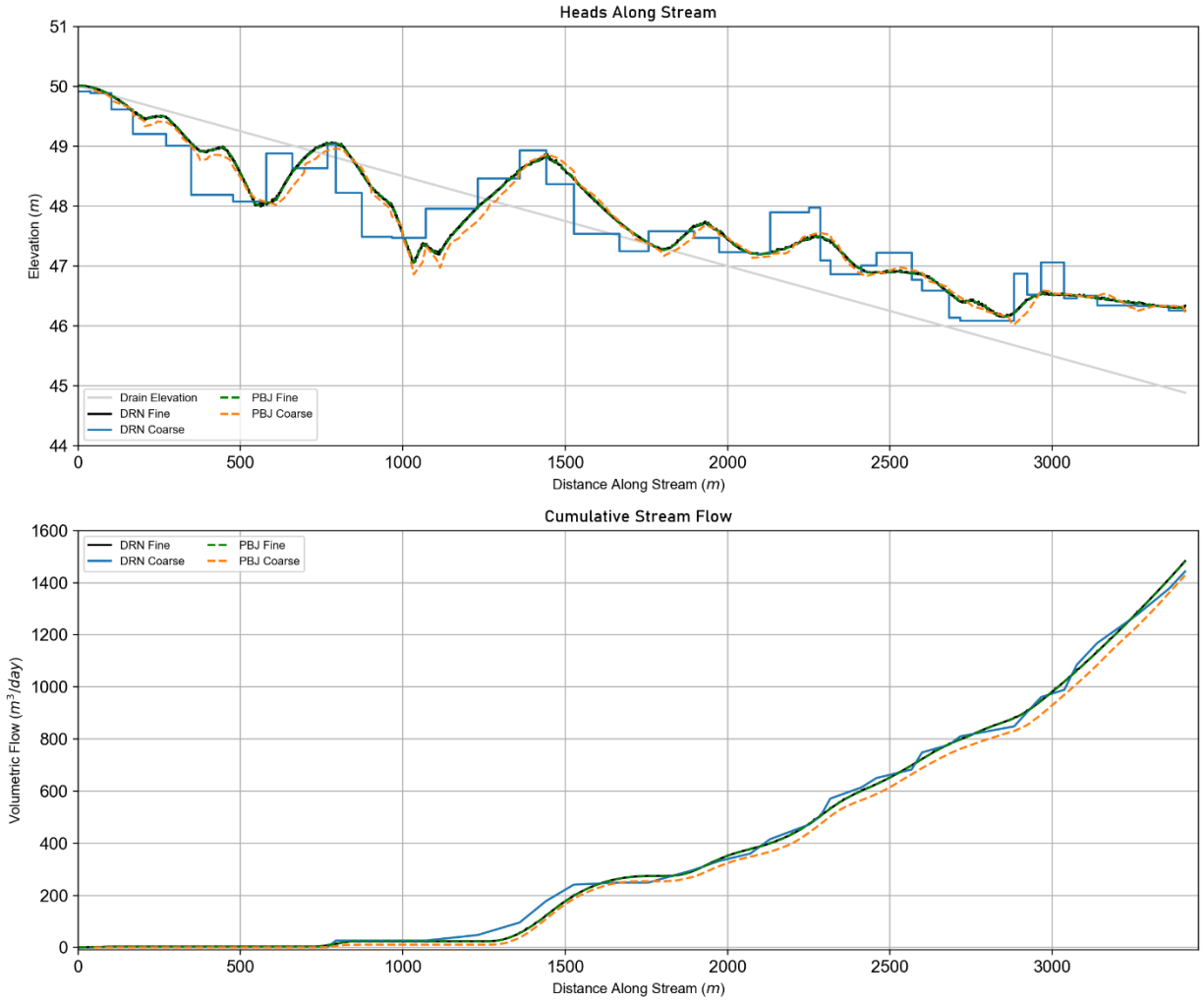

Figure 4-2 – Comparison of GW elevations and flows along the stream calculated by the PBJ and drain (DRN) packages, for the finest (grid #14) and coarsest (grid #1) grids. Heads and flows represent the final time step of the model.

Table 4-3 – Error metrics for the PBJ and DRN (drain) package water levels along the stream. Residuals refer to the absolute value of the difference from the DRN Grid 14 (Finest) groundwater levels.

Package/Grid	Number of Values	Average Residual [m]	Residual Standard Deviation [m]	Maximum Residual [m]
PBJ Grid 14 (Finest)	2043	0.012	0.011	0.068
DRN Grid 1 (Coarsest)	41	0.179	0.175	0.636
PBJ Grid 1 (Coarsest)	173	0.076	0.056	0.295

However, the bottom plot of Figure 4-2 shows the PBJ package delivers less flow to the stream than the fine solution over the total stream length. The effect as it varies with the grid size, for both drain and river modes, is explored in Figure 4-4. The DRN package, despite its seemingly erratic GW head fluctuations along the stream, shows better resilience of simulated cumulative flow to the changes in scale. Sensitivity tests, varying stream, well pumping rates, and model parameters are not detailed here but showed similar results. It is important to point out, however, that neither the PBJ nor DRN package is performing particularly terribly: the coarsest PBJ and DRN model flows deviate by about 3.7% and 2.7% from the finest scale solution during the final time step (Figure 4-4). For the River mode, that rises to 5.3% for the PBJ package and 3.1% for the RIV package, reflecting the greater variability in flux distribution along the stream.

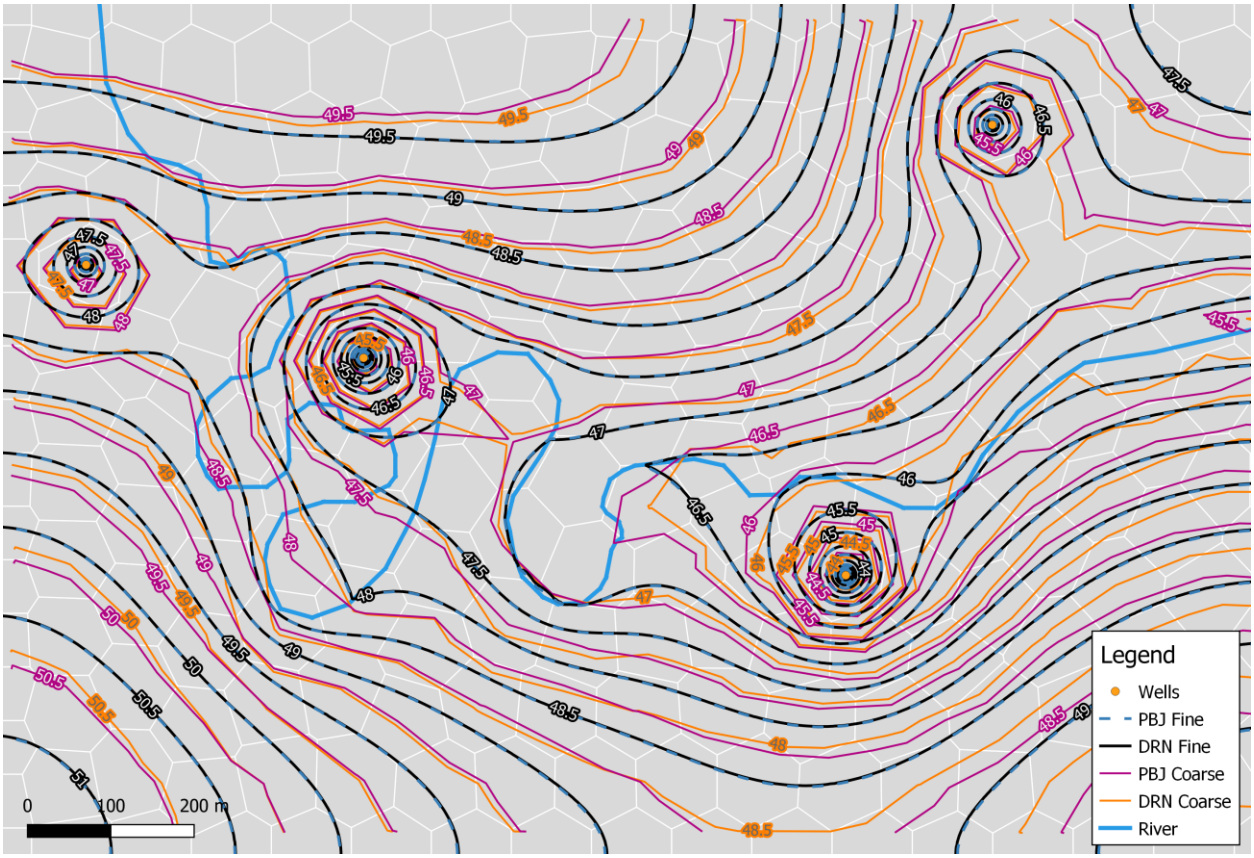

Figure 4-3 – Comparison of final time step head contours from the finest (grid #14) and coarsest (grid #1) grids run both with the PBJ package in Drain mode and the MODFLOW-USG drain (DRN) package. The two fine solutions almost perfectly align. River, coarsest (grid #1) Voronoi grid, and well locations shown for reference.

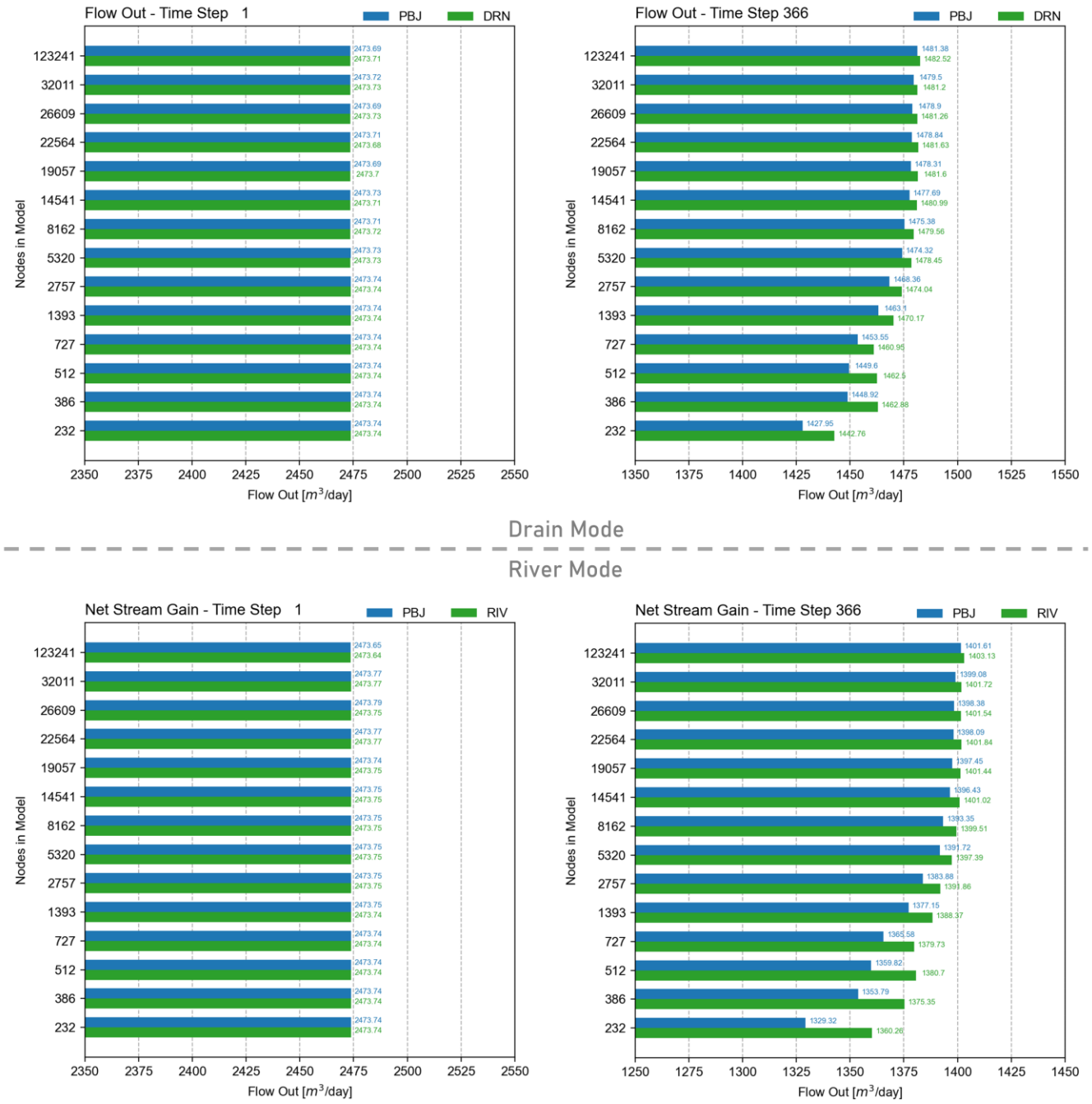

Figure 4-4 – PBJ and MODFLOW-USG drain (top) and river (bottom) flow comparisons after the first stress period (left) and final stress period (right). The drain comparison is simply flow to the stream, while the river comparison is the flow to the stream minus the flow *from* the stream (i.e., net stream gain). Flow *Out* on the x-axes refers to flow out of the aquifer. Note the right two plots have different horizontal scales.

Generally, the sensitivity to grid size effects shown here are attributed to the streambed conductance parameter's reliance on scale (Swain & Wexler, 1996) and its conceptualization (e.g., Morel-Seytoux et al., 2018; Nemeth & Solo-Gabriele, 2003). Obviously, a physically realistic streambed conductance

interpretation should be independent of node area. Here, however, as the node area increases, the node heads become averaged over a larger area and are increasingly less dominated by the node-stream relationship (Mehl & Hill, 2010; Swain & Wexler, 1996). For a head-independent conductance, the conceptualization would have to compensate for scale effects on GW head. However, given that river conductance is generally estimated as a calibration parameter (Mehl & Hill, 2010; Nemeth & Solo-Gabriele, 2003) having a scale-independent conductance value may not serve much practical use.

It is worth mentioning that most studies examining the grid scale effects on river conceptualizations use steady-state simulations and rectilinear grids (e.g., Di Ciacca et al., 2019; Mehl & Hill, 2010; Vermeulen et al., 2006). The results presented here warrant further experiments but do also show promise of using the MODFLOW-USG DRN and RIV packages with length-weighted average elevations and length-based streambed conductance values without the need to explicitly discretize the grid based on river features. More thoughts on the PBJ package and potential future improvements are discussed in Recommendations (Section 6.3).

4.3 Constant Head Test Results

The PBJ package in specified head mode (see Section 3.2.1 for specifics) was tested to ensure it operated as intended. The model setup was the same as in the previous section, but with the PBJ package using the streambed elevation as a specified head along the river. The PBJ package does an excellent job matching heads along the stream (Figure 4-5). The PBJ subgrid formulation forces the MODFLOW solution to follow the river heads very tightly, matching at a finer resolution than normally possible.

Comparison models were built using the MODFLOW CHD package to specify node heads instead of the PBJ package. The formulations of these are very different, and there are some differences in their head solutions at all model resolutions. This is not surprising considering the PBJ package works by forcing the solution to a head through large conductance values, while the CHD package merely removes the node from the solution and substitutes the specified head value.

The first notable difference was the run time. While for coarse models the run time between the packages was indistinguishable (run times under a second), the finest scale PBJ model took over nine minutes to run (compared to a CHD run time of just under one minute). The added time was spent entirely on the first time step and is likely related to the input initial heads. While different solver settings and initial conditions may have decreased the run time, the MODFLOW specified head formulation clearly set up the problem easier for the solver.

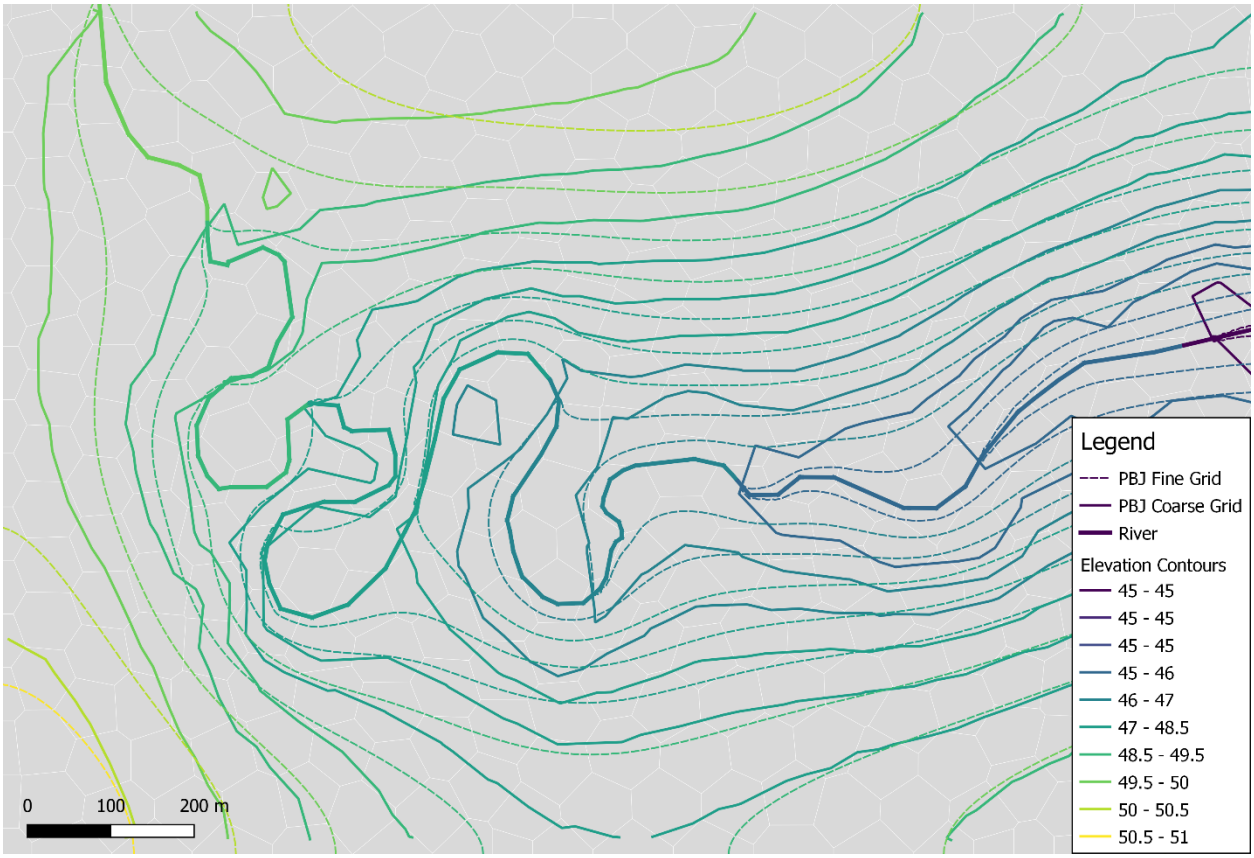

Figure 4-5 – Contours of finest and coarse (386-node grid) model head solutions, and the specified head river, all displayed using the same graduated color scheme. At the river, the GW heads consistently match the river heads, as shown by their matching colors. The 386-node grid is shown in the background.

The other notable differences can be understood through the two contour plots in Figure 4-6. The left plot depicts a location where the PBJ package at a coarse resolution does a great job matching the fine solution at the river, while the coarse CHD package is unable to match the heads. This shows one of the advantages of the PBJ package subgrid, segment-specified heads. On the right plot, however, we see a place where the two PBJ and one CHD package heads all (roughly) converge at a location along the river with a head of 46. However, away from the river, the coarse PBJ contours take a very different route than the other two. Likewise, in the left plot we see the coarse PBJ model has produced a mound in the upper-right corner not present in the other two solutions.

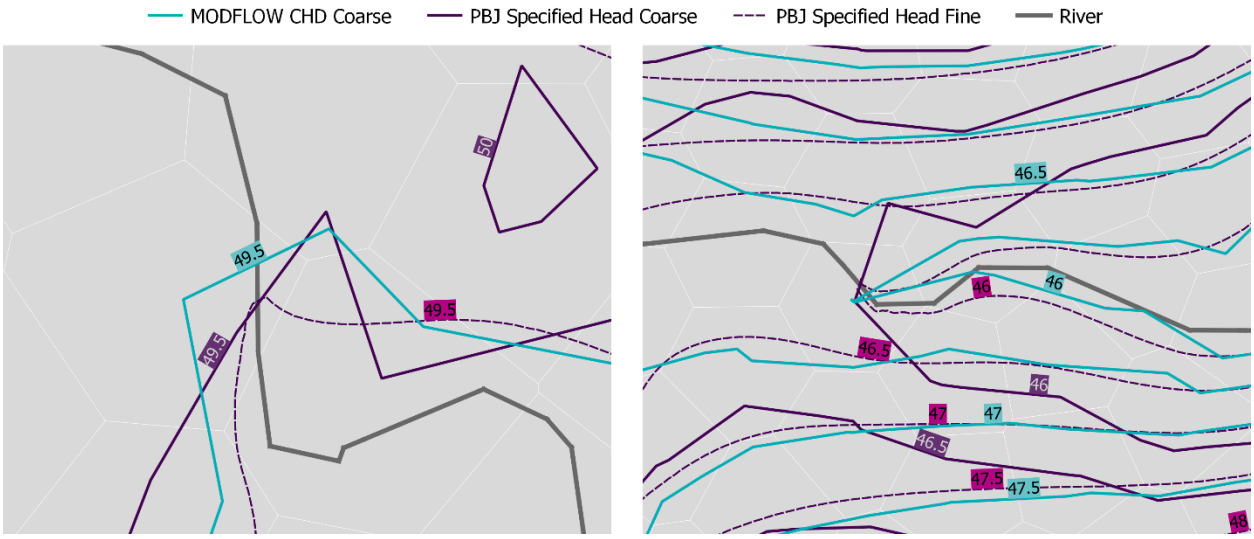

Figure 4-6 – Closeups of PBJ finest and coarse (386-node grid) model head contours, with coarse MODFLOW CHD (Specified Head) package model head contours, and the specified head river. The 386-node grid is shown in the background.

The reasons for these differences have not been extensively investigated, but one hypothesis is the PBJ package's use of three nodes per segment may have unexpected effects on the model head solution at coarse scales. In effect, it is rigidly constraining a large number of cells surrounding the river; when using the CHD package, many of those cells are only affected by the river through the flow solution.

Chapter 5 - Example GW-SW Model: Alder Creek Watershed

This chapter explores the application of the new, coupled Raven hydrological modelling framework to a real-world watershed. The Alder Creek watershed in Southern Ontario, just west of the city of Kitchener (Figure 5-1) was used for the experiments. Initially, separate MODFLOW-USG and Raven subsurface and surface water models, respectively, were built and calibrated. Then, the two were combined to form the coupled GW-SW model. The sections of this chapter detail the model development steps taken and the results from the final model.

As stated in Section 1.1, the third objective of this thesis is to demonstrate the utility of the coupled modelling framework's utility in modelling real-world GW-SW systems. This chapter serves to meet that objective through demonstration that the software works as intended and moves reasonable fluxes of water between the groundwater and surface water systems. Development of an application-ready model of the Alder Creek watershed is considered out of the scope of this thesis, in part because of the limited available groundwater level data in this basin. To the extent possible, attempts have been made to keep the model realistic for the sake of demonstration. The model presented here may have limited predictive power but serves its intended purpose: to demonstrate the successful algorithmic implementation of the Raven-MODFLOW-USG coupling. The coupling is evaluated in (1) its ability to preserve mass locally and globally, and (2) its ability to produce reasonable estimates of recharge and baseflow fluxes within this basin.

5.1 Site Description

The Alder Creek watershed (Figure 5-1) is an area of approximately 78 km^2 west of the tri-city area of Waterloo, Kitchener, and Cambridge. The area is of hydrologic importance because of the high-recharge sand and gravel zones utilized by the Regional Municipality of Waterloo for their municipal well fields (CH2M HILL & SSP&A, 2003; Wiebe & Rudolph, 2020). Additionally, domestic wells draw water from the underlying unconfined Mannheim aquifer (CH2M HILL & SSP&A, 2003).

Alder Creek passes through open fields, residential neighborhoods, and wooded areas before flowing into Alder Lake (Figure 5-1), a man-made lake formed by the construction of a dam. The dam discharge is regulated by the Grand River Conservation Authority (GRCA) and the water level drop caused by the dam is approximately four meters (CH2M HILL & SSP&A, 2003). The creek continues south until it discharges into the Nith river. Several other smaller surface water bodies exist in the area, including a number of small kettle ponds, a trailer park pond, a public trout-fishing pond, and a five wetland areas identified by the GRCA (CH2M HILL & SSP&A, 2003). The only streamflow gauge with publicly available data in the watershed is the Alder Creek near New Dundee station (station number 02GA030),

which has been in operation since 1965. The station is located above Alder Lake and notably above several of the tributaries to Alder River (Figure 5-1).

Land use in the area is primarily agricultural ("tilled"), with light urban areas and aggregate extraction sites (Figure 5-2). Many of the forest/wetland ("swamp") areas of the watershed are adjacent to the river. Permeable sands and gravels cover nearly half of the watershed area (Figure 5-3) (Wiebe & Rudolph, 2020), but other regions are overlayed with silt and clay tills (e.g. Maryhill Till, Lacustrine deposits). The surficial geology of the area is known to be highly complex due to the historical advance and retreat of glacial ice sheets (Jones et al., 2009).

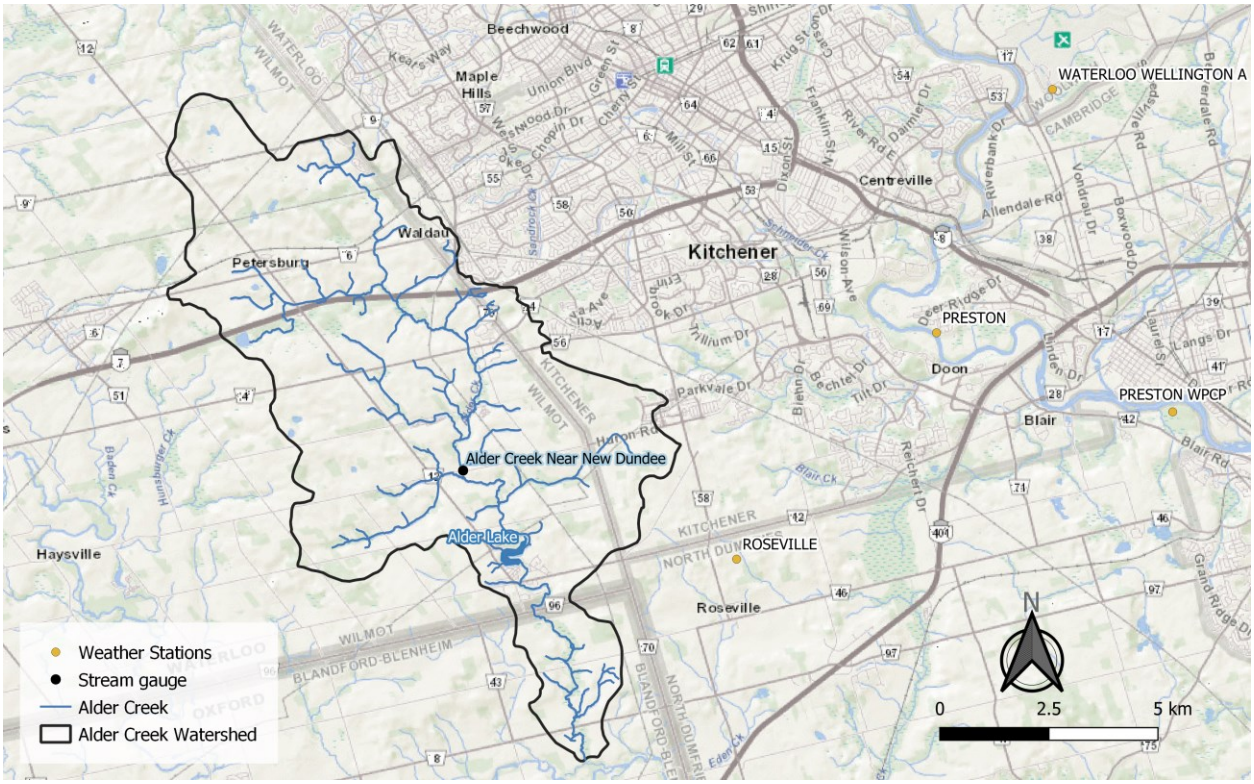

Figure 5-1 – Map of the Alder Creek watershed and surrounding region, with nearby weather stations.

A simple L'vovich water balance (Poncea & Shetty, 1995) was initially calculated for the watershed using stream flow and precipitation data from the 2001-2014 water years (October of the previous year to September). Volumetric streamflow data was sourced from the Alder Creek Near New Dundee gauge and precipitation data was downloaded for the Roseville station (ID: 6147188, Figure 5-1). Missing daily totals of precipitation were either filled using a nearby station's daily data (Waterloo Wellington, Waterloo Airport, or the Kitchener/Waterloo stations) or in lieu of any nearby data, the historical daily average value for the Roseville gauge. Due to the streamflow gauge being located well above the bottom of the watershed, a watershed area of 47.4 km^2 was used instead of the full 78 km^2 for the water balance

calculations. Baseflow was estimated from the daily streamflow data using the Lyne and Hollick filter approach from Ladson et al. (2013), as implemented in the hydrostats R package (Bond, 2019). Over the time period used here, the average streamflow (baseflow + surface runoff) was 138 mm, roughly the same as the long-term average of 140 mm reported in Wiebe & Rudolph (2020). The water balance is displayed in Figure 5-4.

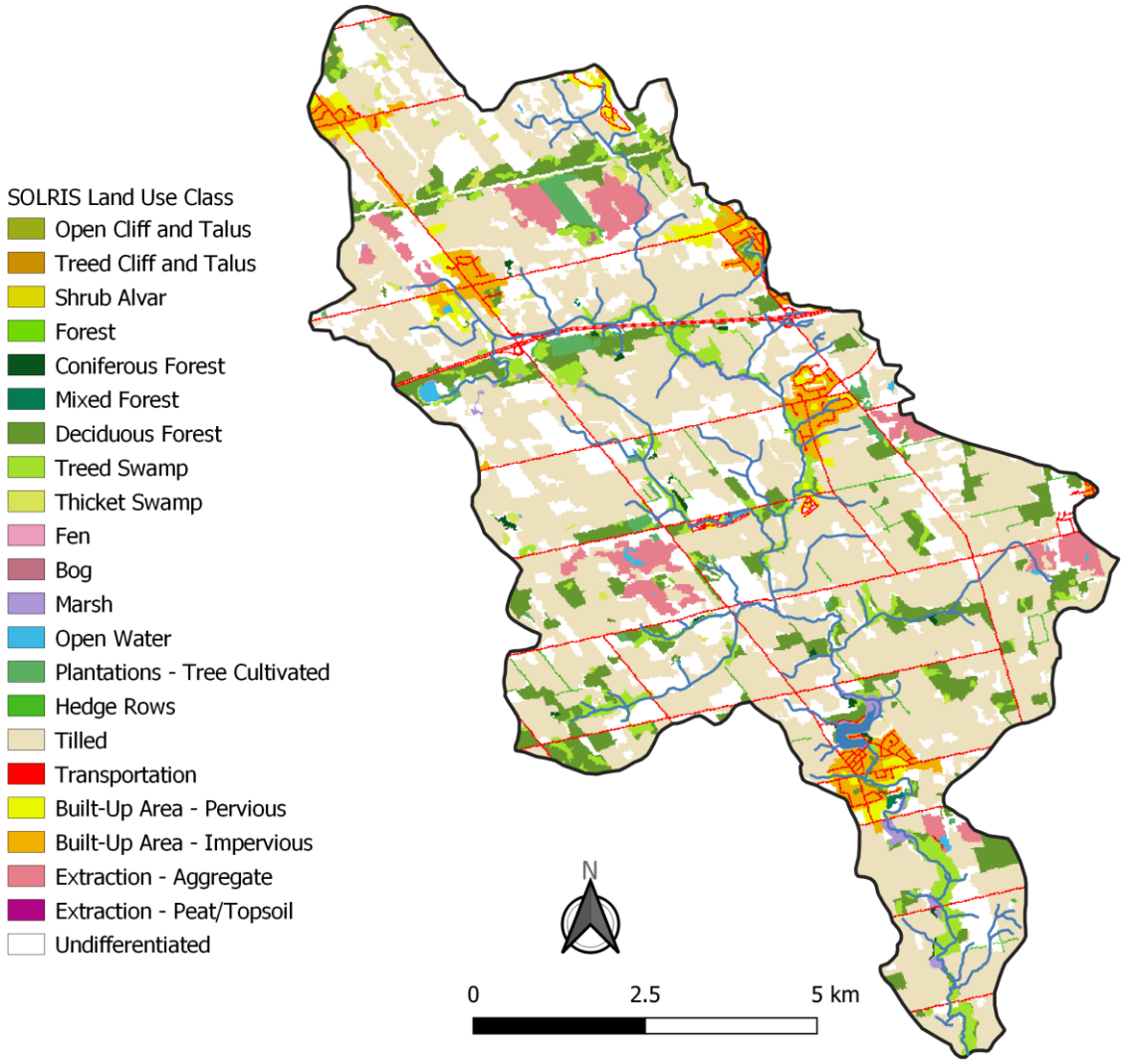

Figure 5-2 – Land use classifications from SOLRIS (MNRF, 2019) for the Alder Creek watershed.

According to the L'vovich water balance (Figure 5-4), the majority of the water entering the watershed as precipitation becomes ET. Estimates of actual evapotranspiration (AET) in the region are around 540 mm/year (see Wiebe & Rudolph, 2020), which is well under the ~810 mm estimated here. The difference between these numbers (270mm) is similar in magnitude to the long-term estimate of groundwater recharge reported in Wiebe & Rudolph (2020) of 321 mm/year. The simple L'vovich water

balance does not consider losses to (deep) groundwater, which may explain the discrepancy. The baseflow coefficient, or the amount of streamflow that is baseflow, estimated was 0.26 (26%), similar to Wiebe & Rudolph's (2020) estimates between 0.21 and 0.31. Their study considers 2014-2016, which is outside the years considered here (2000-2010) but the average precipitation is similar.

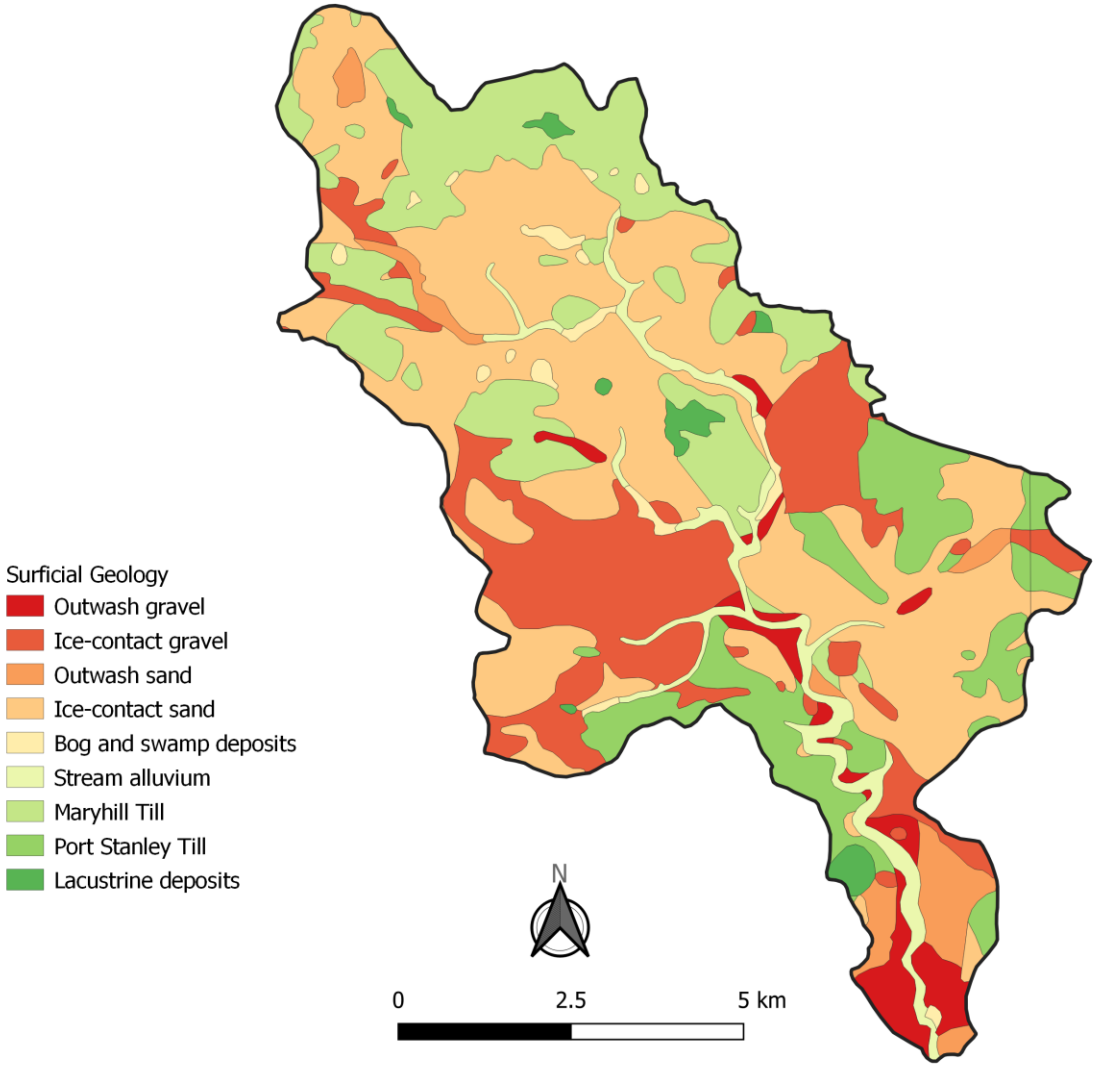

Figure 5-3 – Surficial geology map of the Alder Creek watershed, as classified in Ontario Geological Survey (2010)

The Alder Creek watershed has been the subject of, or included in, several groundwater modelling studies, on a variety of scales, using different software (see Frind et al., 2014; e.g., Chow et al., 2016). Alder Creek is within the Waterloo Moraine, which is stratigraphically complex and heterogenous, with a thickness ranging from 30 to over 100 m (Martin & Frind, 1998). As is further discussed in Section 5.2.1, the model horizontal grid was taken from a previous model, and the hydrostratigraphy was inferred from

a yet another. The groundwater model grid inherited used the watershed (catchment) boundary, which is not necessarily shared with the groundwater system (Staudinger et al., 2019). Previous studies of the area have used natural hydrogeologic boundaries sufficiently far from the watershed boundaries with the intent of limiting boundary conditions effects from interfering with model results (CH2M HILL & SSP&A, 2003). However, the model/methods comparison of Chow et al. (2016) used the watershed boundaries and the intent was this work could be (qualitatively) compared with their results.

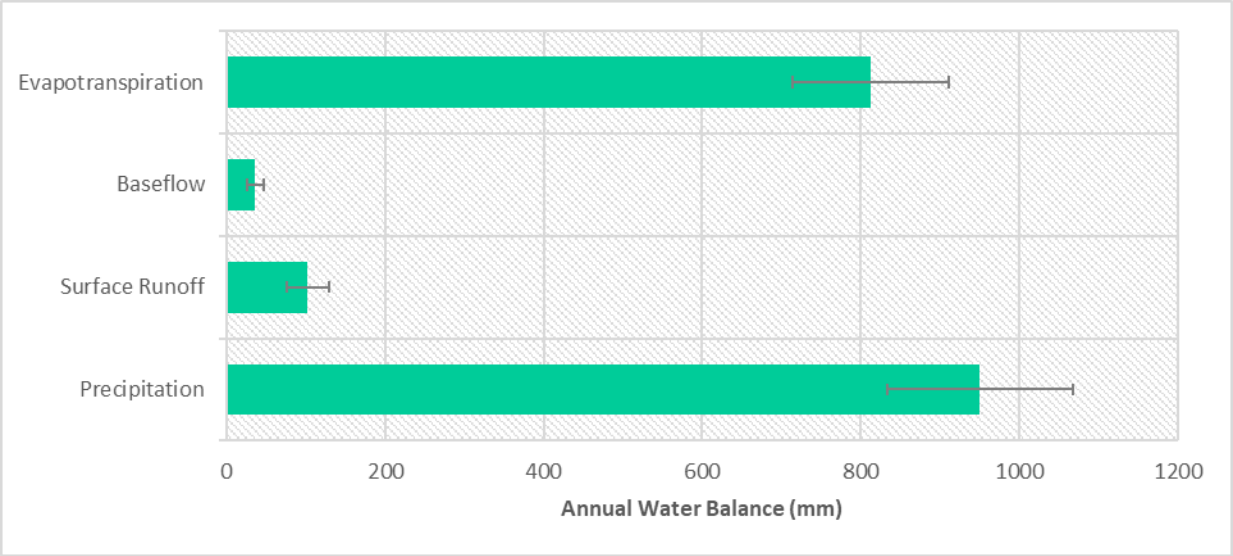

Figure 5-4 – L'vovich water balance, showing a 14-year (2001-2014) water-year estimate of how precipitation is partitioned in the watershed on average. Error bars show the interannual standard deviation.

5.2 Model Development & Calibration

This section details the development of the MODFLOW-USG and Raven models of the Alder Creek watershed, as well as how they were combined to form a coupled Raven model. The final subsection covers the calibration of the models.

5.2.1 Groundwater Model Development

The MODFLOW-USG groundwater model was based on the horizontal model grid of the HydroGeoSphere (HGS) model used in Jones et al. (2009) and Chow et al. (2016) to investigate GW-SW interactions in the Alder Creek watershed (Figure 5-5). Nodal coordinates from the HGS model were imported into AlgoMesh (HydroAlgorithmics, 2016) which was used to output basic discretization files for MODFLOW-USG. The original HGS mesh was triangular, but for the unstructured MODFLOW-USG model the Voronoi dual grid was used. This equates the both models solving at the same locations, although their formulations are very different (finite-difference vs. control volume finite-element).

Figure 5-5 – Voronoi unstructured model grid used for the GW model, based upon HGS triangular grid for a previous model of the area. Locations of pumping and observation wells, as well as Alder Creek, can readily be observed in the grid discretization.

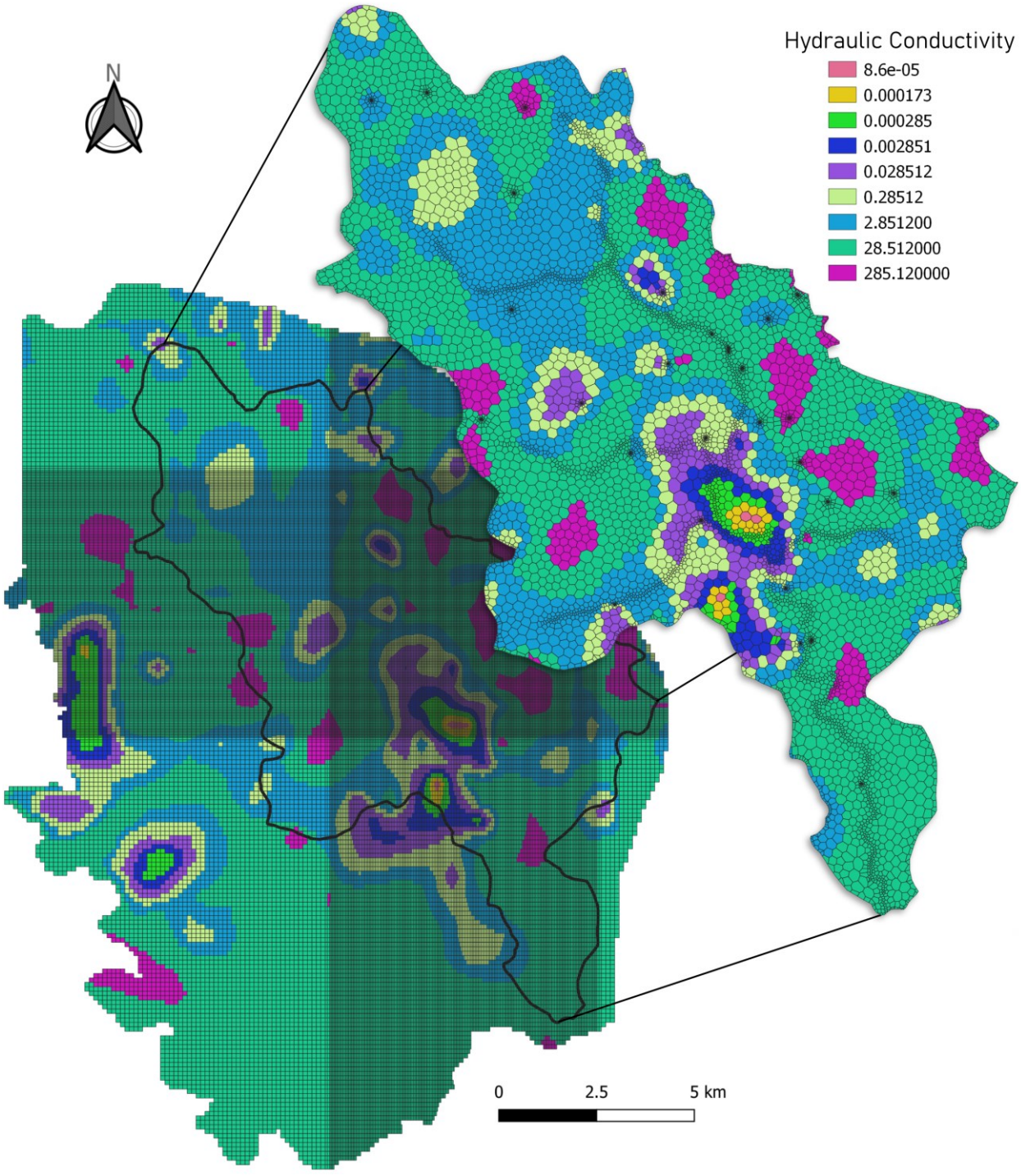

Figure 5-6 – Transfer of layer two hydraulic conductivities values (in discrete zones) from the rectilinear grid of CH2M HILL & SSP&A (2003) to the new, unstructured Alder Creek model via spatial majority of the overlapping cells. Scale bar is representative of rectilinear model.

Visual MODFLOW (i.e., MODFLOW 2000/2005) files were provided from a previous modelling effort extending the work of CH2M HILL & SSP&A (2003). These files were used to generate parameter maps (polygon shapefiles) for each of the four layers that could be used in a GIS program to produce spatial summary statistics. Hydraulic conductivity (Figure 5-6) in the existing model was classified into zones for the first two layers, but layers three and four were continuous surfaces (e.g., non-discrete values) of varying conductivity. Therefore, for the first two layers, the zones were transferred to the unstructured grid based on majority overlapping zone. For the bottom two layers, spatially averaged values were used to estimate conductivities for the unstructured grid. Vertical hydraulic conductivity values were consistently 10% of the horizontal conductivity values (vertical to horizontal anistrophy ratio of 0.1). These values were re-calculated for the unstructured grid based on the transferred horizontal conductivity values. The previous model was steady state so no storage parameters existed. Specific storage and specific yield were left as calibration parameters, with starting values based on literature values (Freeze & Cherry, 1979; Younger, 1993). The LPF package was used to simulate groundwater flow in MODFLOW-USG.

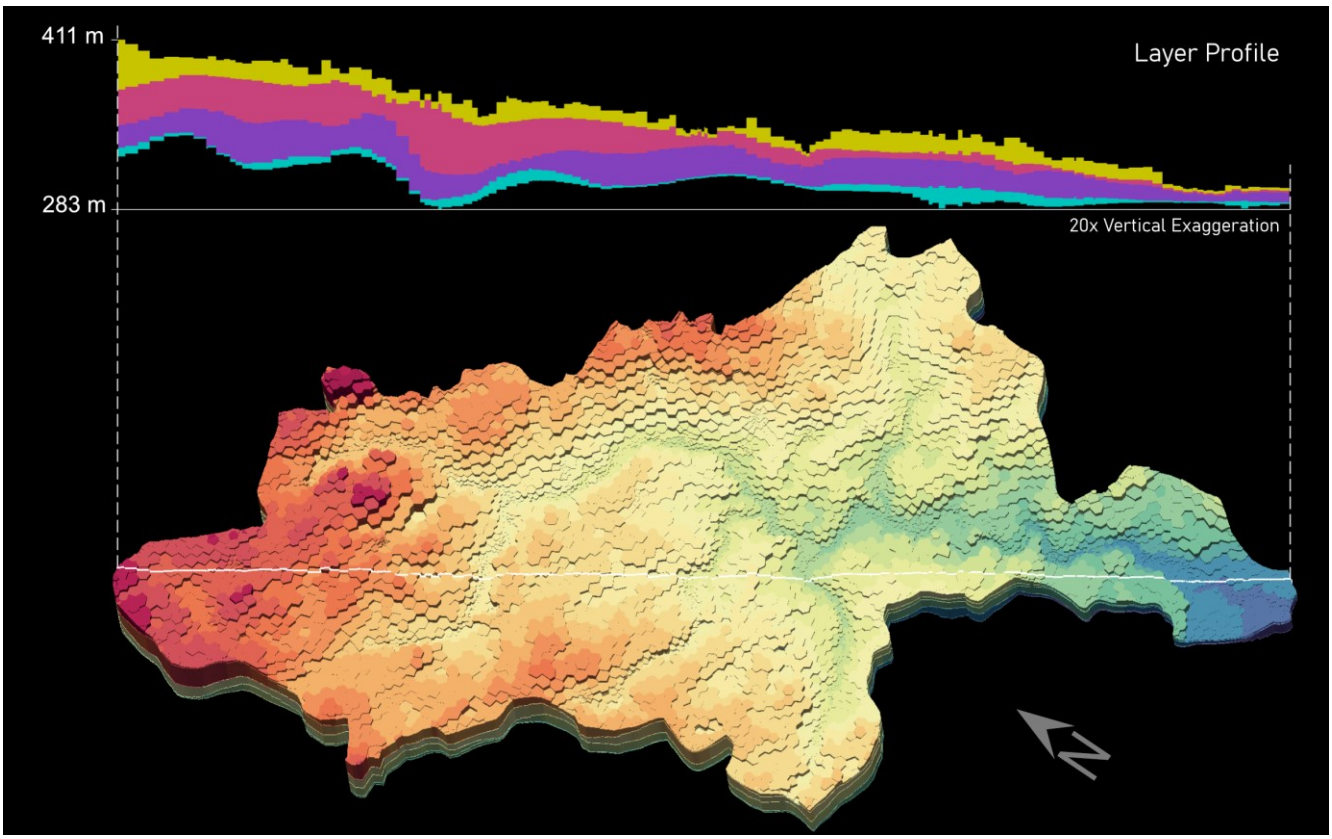

Figure 5-7 – Model layer thickness profile along a slice cutting through the longest portion of the model (top) and model layer top elevations (bottom). Figure created using GroundWater Desktop (www.groundwaterdesktop.com).

In the previous MODFLOW model, layers one and three were aquitards, while two and four were aquifers. The same setup was kept in the MODFLOW-USG model, with layer elevations being estimated as spatial averages from the previous model. Elevations for the model top (conceptualized as the land surface) were estimated using maximum elevations rather than average, additionally, some were adjusted based upon the DEM used to generate the streambed elevations (i.e., model top elevation was ensured to be above streambed elevations). All layers were set to be convertible between confined and unconfined conditions with upstream water-table weighting (LAYTYP =4) in the MODFLOW-USG LPF package (Bedekar et al., 2012; Panday et al., 2013). Model surface elevations and an example cross-section are shown in Figure 5-7.

The model simulation period was from October 1, 2000 to October 31, 2010. This was discretized into monthly stress periods, with daily time steps. This was to facilitate the eventual coupled Raven modelling, which is currently limited to daily (or smaller) time steps. The first month (October 2000) was run steady-state to serve as a warm-up period to reduce the effect of the initial heads (estimated based off of ground surface elevations).

As mentioned before, the watershed boundary used here is not aligned with hydrogeological boundaries. Therefore, constant head boundaries were added to the model (1) based upon the constant head boundaries used in CH2M HILL & SSP&A (2003) at the northern/southern ends of the model and (2) based on head contour maps in CH2M HILL & SSP&A (2003). Initially, only the northern and southern constant head boundaries were used, but early versions of the model were unable to reproduce the head distributions in CH2M HILL & SSP&A (2003), so more boundaries were added. Figure 5-8 displays the constant head nodes and elevations used in the model. The MODFLOW-USG Constant Head Boundary Package (CHD) was used to represent the 105 constant-head nodes into the model.

There are ten municipal pumping wells located in the Alder Creek Watershed (Chow et al., 2016) although an additional three lay either on the border or just outside the border of the watershed/model domain (Figure 5-8). Pumping rates for these wells are not publicly available, however, monthly averages were graphically compiled for the years 2000-2011 in the *Region of Waterloo Tier Three Water Budget and Local Area Risk Assessment* (Matrix Solutions Inc & SSP&A, 2014). Relevant scatterplot time series pumping rates were digitized manually using WebPlotDigitizer (Rohatgi, 2020). Far from an ideal way to obtain pumping data, the digitized data add additional uncertainty to the model water balance. Dry-node-related convergence issues led to wells W7 and W8 being simplifed to just the location of W7 and the pumping rate replaced with a simulation-period average. Private (residual) wells exist within the modeled region, but pumping data are neither available nor likely of a great enough magnitude to substantially affect the model results.

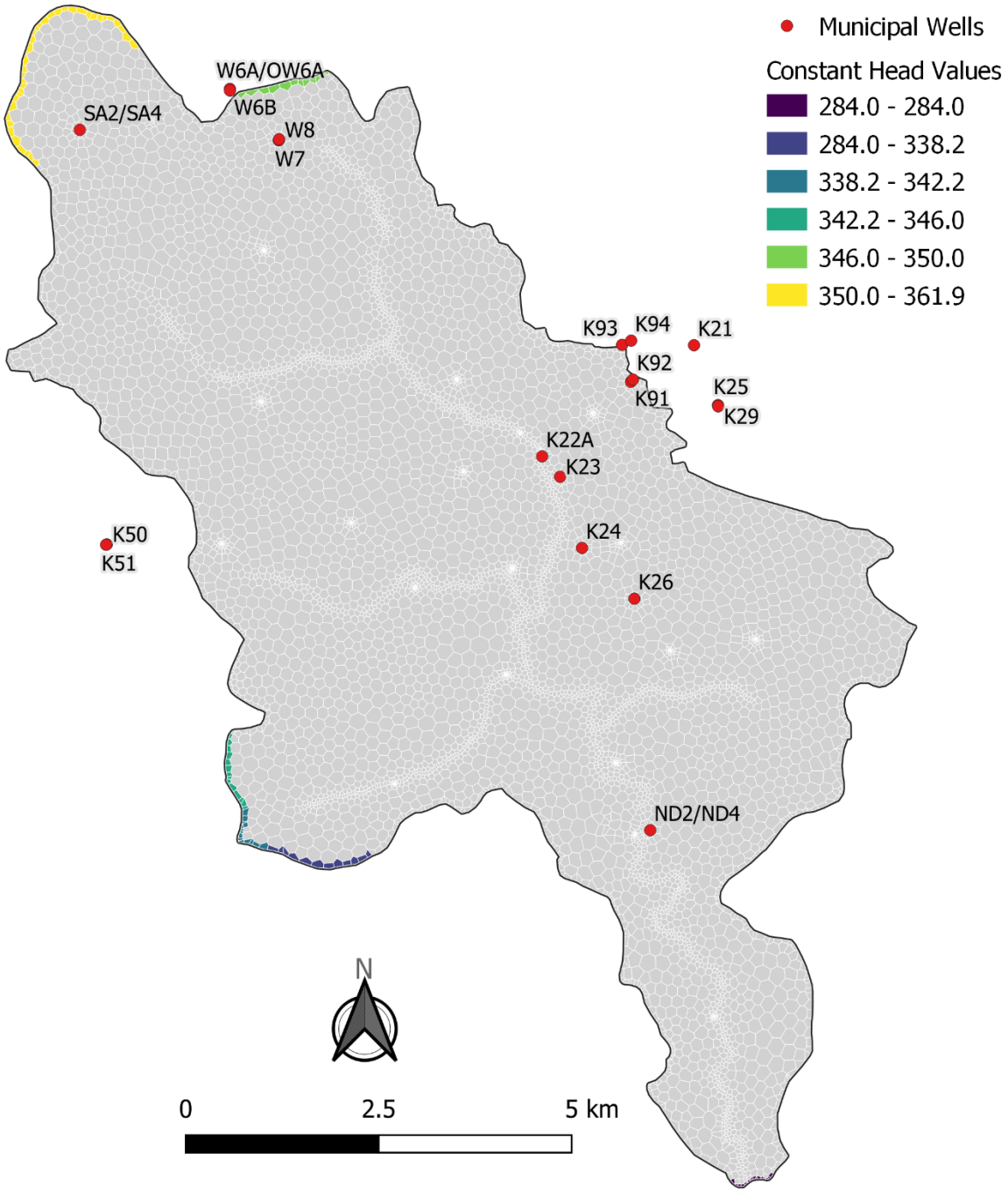

Figure 5-8 – Constant head node locations and values (elevations, in meters) with municipal pumping wells in and around the model domain. At the very bottom of the model are a strip of constant head nodes set to 284 m.

Alder Creek was represented in the model using the Polyline Boundary Junction package (Section 3.2) running in the unidirectional "drain" mode. The *pbjR* R package was used to generate the PBJ package input file from a simplified polyline shapefile of Alder Creek from the Enhanced Watercourse shapefile included in the Ontario Integrated Hydrology dataset (MNRF, 2019). A digital elevation model

(DEM) of Southern Ontario was used to determine streambed elevations along the stream. As is discussed further in Section 5.2.4, the stream was divided up into different zones based on stream order and location in the model and the streambed conductivities calibrated by zone. The spatial extent of Alder Creek present in the shapefile used here differs slightly from the one used by Jones et al., (2009) when they discretized their HGS grid, which has been inhereted in this effort (Figure 5-9).

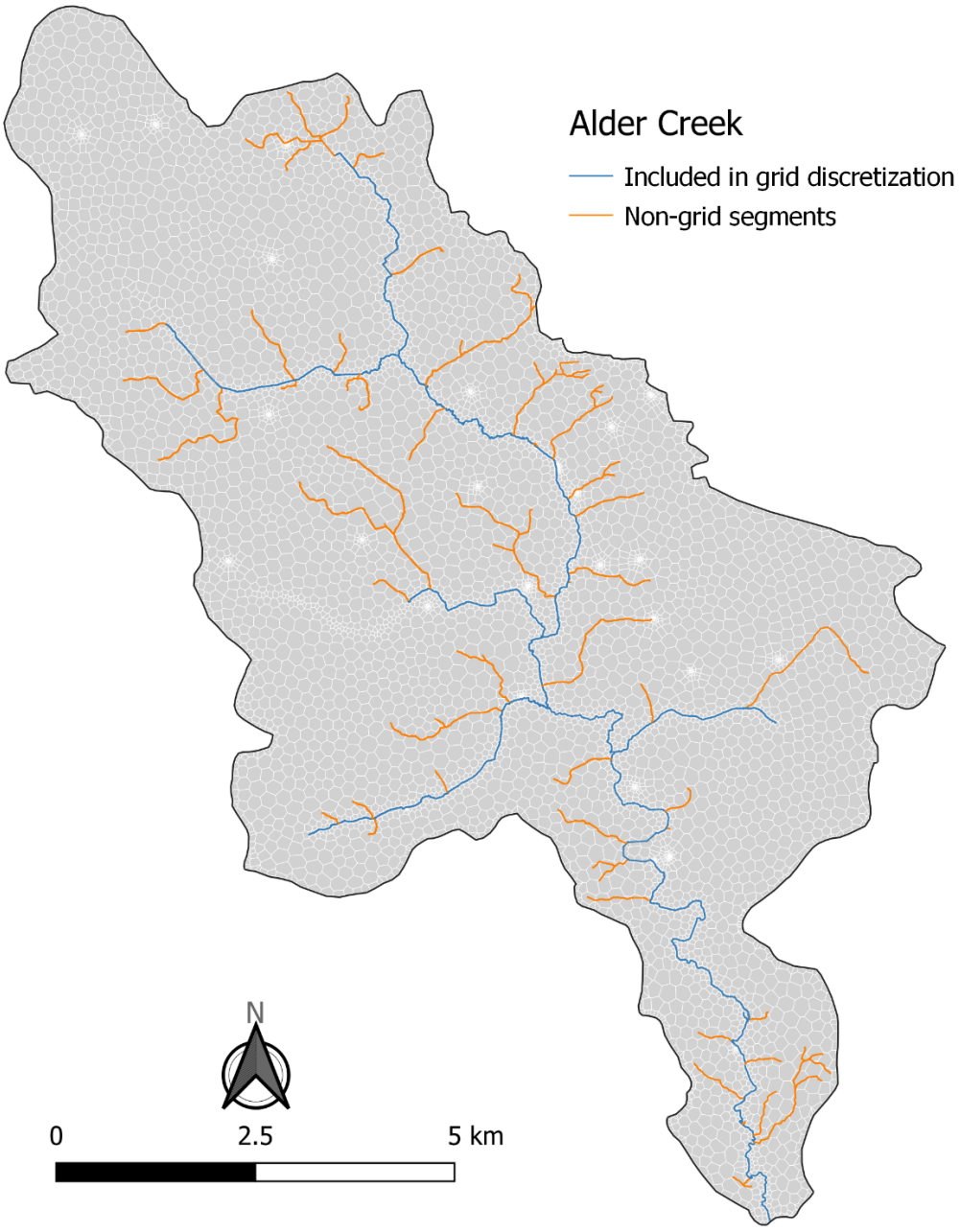

Figure 5-9 – Difference in spatial extent between the stream network used in the PBJ package and the one "baked in" to the unstructured grid. Note the section that exists in the grid, but not the PBJ package, on the mid-west section of the model domain.

Since the PBJ package includes a correction for off-center steams in nodes, this is not (numerically) an issue. However, it does raise some questions as to whether certain stream segments represent non-perennial streams (Shanafield et al., 2021), or if the shapefile is entirely accurate. This model, with the PBJ package in drain mode, can incorporate this uncertainty since drains only allow groundwater to flow to the river when the water table reaches the streambed elevation. In river mode, however, this could induce fictious flow to/from the aquifer. One western section of the grid (see Figure 5-9) contains an area of dense nodes, presumably a continuation of the stream, which is not reflected in the stream shapefile used here (this decision is question in Section 5.3.4). No lakes, including Alder Lake (see Figure 5-1) are explicitly represented in the groundwater model.

Historical time series data could only be obtained from two observation wells within the model domain (orange dots, Figure 5-10). The first well, W0000428-1, is part of the Provincial Groundwater Monitoring Network (Ontario Ministry of the Environment, Conservation and Parks, n.d.) and thus is publicly-available, but is screened in the lower aquifer (model layer 4) and is located on the edge of the model domain. The well OW8-61A was included in the *Tier Three Assessment* (Matrix Solutions Inc & SSP&A, 2014) in a water level change figure, calculated as deviations from an unknown initial level, showing the ability of their model to capture the observed trend. This figure was digitized in the same manner as the pumping data from the Tier Three Assessment. While important since the well is (1) in a central location (2) near the river and (3) close to pumping wells, digitized groundwater fluctuations (from an unknown initial level) are an inferior data source compared to the absolute heads for calibrating a GW-SW model. OW8-61A is screened within layer 2.

To better constrain the GW solution, the 25 long-term average water levels used in CH2M HILL & SSP&A (2003) and Chow et al. (2016) were included as calibration targets in the model (green dots, Figure 5-10). The CH2M HILL & SSP&A (2003) report implies all well water levels are “representative” averages from 1991-2000. Wells, average water levels, location in the model, and modeled layers are displayed in Table 5-1. Well OW8-61A, one of the transient wells, also appears in this dataset. The hope was having a long-term average for OW8-61A would help constrain the historical difference time series, which is calculated relative to an unknown initial water level.

Table 5-1 – Wells from previous modelling studies and their representative average water levels

Well Name	X-coordinate	Y-coordinate	Average Water Level	Model Node	Model Layer
AC1-01A	536156.0	4803609.0	330.10	931	3
AC1-01B	536156.0	4803609.0	332.05	931	2
AC2-01B	534626.0	4800802.0	336.98	4443	2
AC3-01A	537487.0	4801079.0	317.95	889	2
AC3-01B	537487.0	4801079.0	317.97	889	2
AC4B-01B	537741.0	4800160.0	317.52	2028	2
OW10-67A	532387.6	4803920.4	353.25	6300	3
OW2-61A	536253.5	4805339.0	332.58	776	3
OW2-77A	537920.2	4800200.1	313.49	2040	2
OW2-85A	537189.3	4805605.0	330.40	2172	3
OW3-61A	537095.0	4803851.9	325.27	359	2
OW8-61A	536545.3	4805108.0	328.37	1042	2
TW11-69A	537750.2	4803197.7	326.69	214	2
TW1-70A	538192.0	4802541.0	327.52	43	2
TW3-69A	537565.6	4803941.0	327.00	694	3
WM17-93A	532895.0	4805752.0	351.98	5073	2
WM17-93B	532895.0	4805752.0	352.08	5073	2
WM17-93C	532895.0	4805752.0	353.08	5073	1
WM18-93B	534070.0	4804188.0	349.22	4249	2
WM20-93A	535523.0	4804855.0	334.54	1923	2
WM22-93B	536072.0	4802225.0	326.30	1657	2
WM23-93A	539310.0	4802680.0	327.78	376	2
WM23-93B	539310.0	4802680.0	327.89	376	2
WM2-94C	535430.0	4806050.0	338.01	1230	2
WM-OW3AC-92B	534887.1	4803341.0	341.59	2489	2

The initial, non-coupled MODFLOW-USG model had constant recharge uniformly distributed over the model nodes. The recharge flux rate was 330 mm/year and came from the Raven Alder Creek surface water model. The flux rate was considered cromulent as it was similar to the estimate provided in the *Tier Three Assessment* (Matrix Solutions Inc & SSP&A, 2014) of 321 mm/year, as well as within the general ranges estimated by Wiebe & Rudolph (2020). The complex surficial geology of the region contradicts the assumption of uniform recharge used here, however, the performance of the model with this assumption was considered a useful comparison point for the final coupled model.

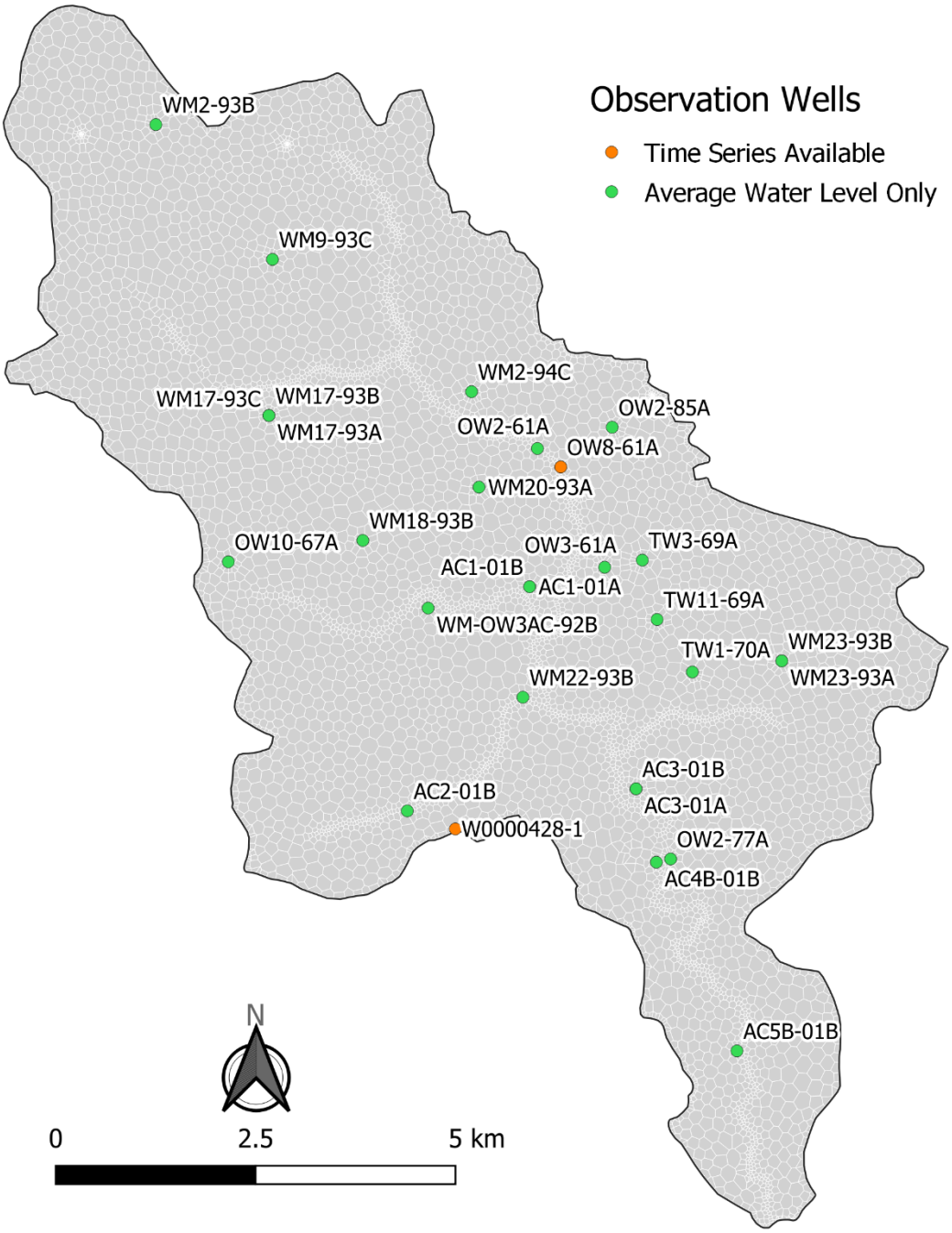

Figure 5-10 – Locations of observation wells used to evaluate model performance. Most well level data came from long-term averages included in past modelling efforts (green points). For OW8-61A, only the monthly change in hydraulic head is available. W0000428-1 is the only well with publicly available time series data in the modeled area.

5.2.2 Surface Water Model Development

Subbasins and river channel geometry for the Alder Creek watershed were generated using *BasinMaker* (Han et al. 2020). Subbasins were created at a "coarse" resolution (i.e., at every stream

branch) to facilitate the coupled GW-SW model, where stream levels for the PBJ package will only be available at the subbasin-level. Some additional basins were created based upon streamflow measurement points that ended up not having data within the final model time period. This led to a total of 140 subbasins, shown in Figure 5-11.

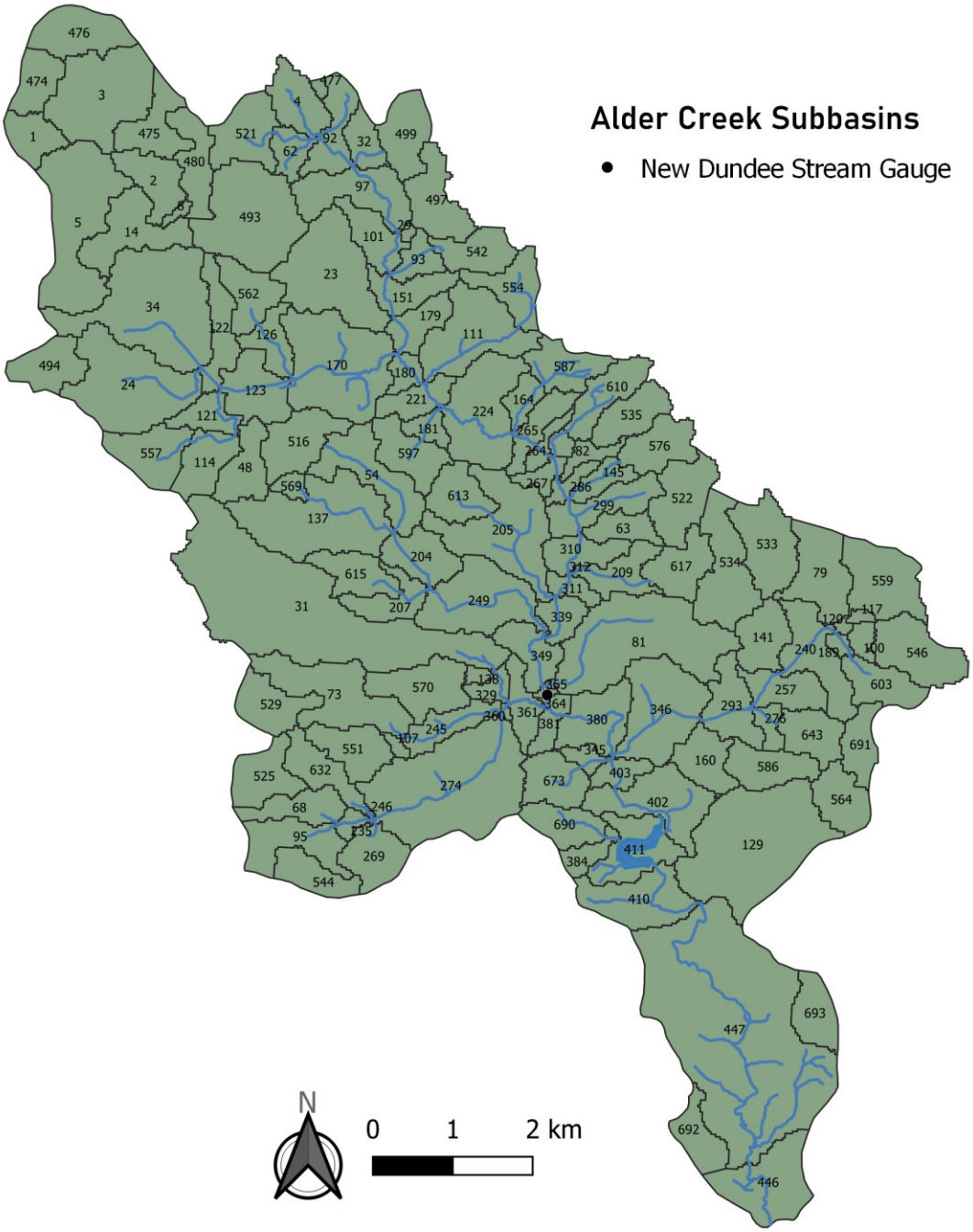

Figure 5-11 – Subbasins in the Alder Creek watershed Raven model. The one stream gauging station in the model is in subbasin 365, marked with a black dot. Subbasins are labeled with their non-sequential IDs.

Hydraulic Response Units (HRUs) for the watershed were created by combining similar land use classifications (Figure 5-2) together and dividing (subsetting) the resulting land use polygons by subbasin. This resulted in 425 HRUs (Figure 5-12), with six different classifications used for both land use and vegetation type. HRUs were spatially matched with the majority-overlap surficial geology (Figure 5-3) classification to determine their soil type. The Raven HRU/Basin Definition file (.rvh) was written from the resulting shapefile information using the *RavenR* package.

Compilation of precipitation data from the Roseville gauge, supplemented with nearby gauges and historical averages, was described in Section 5.1. Rainfall, snowfall, and daily min/max temperature were assembled in the same manner. Uncertainty in the meteorological forcing estimates (e.g., Wiebe & Rudolph, 2020) was considered outside the scope of this proof-of-concept modelling exercise.

The HMETS hydrological model (Martel et al., 2017) emulation within Raven (Craig & the Raven Development Team, 2020) was chosen as the starting point for representing the Alder Creek watershed. This simple template contains a two-layer soil model, linear baseflow and percolation, and handles in-catchment routing using gamma distribution convolution. However, the configuration does not utilize a GW compartment, so percolation from the second soil model to GW was added to represent GW losses from the SW system. This setup also better facilitated the eventual GW-SW coupling (see Section 5.2.3)

HMETS uses an uncorrected soil evaporation algorithm, which removes water from the soil at the rate demanded by the PET. For a more realistic soil evaporation representation, the GR4J algorithm was selected in Raven. The default HMETS PET algorithm is Oudin, but Hargreaves 1985 was substituted to move additional excess water. Likewise, to better represent the non-permeable till sections of the Alder Creek watershed, as well as the small kettle lakes in the region, an abstraction process was added to move a percentage of water to depression storage (e.g., puddles, ponds). Additionally, an open water evaporation algorithm was added to the Alder Creek model.

Given the available soil-related data in Wiebe & Rudolph (2020), the explicit Green-Ampt method (W. H. Green & Ampt, 1911) was chosen as the model infiltration algorithm. Additional literature values for wetting front capillary pressure and soil porosity were obtained from Rawls et al. (1983).

The only historical data used for evaluating the surface water model was the streamflow data from the Alder Creek at New Dundee station (Figure 5-1, Figure 5-11) obtained from the Water Survey of Canada (WSC) for the years 2000-2010 using the *RavenR* package.

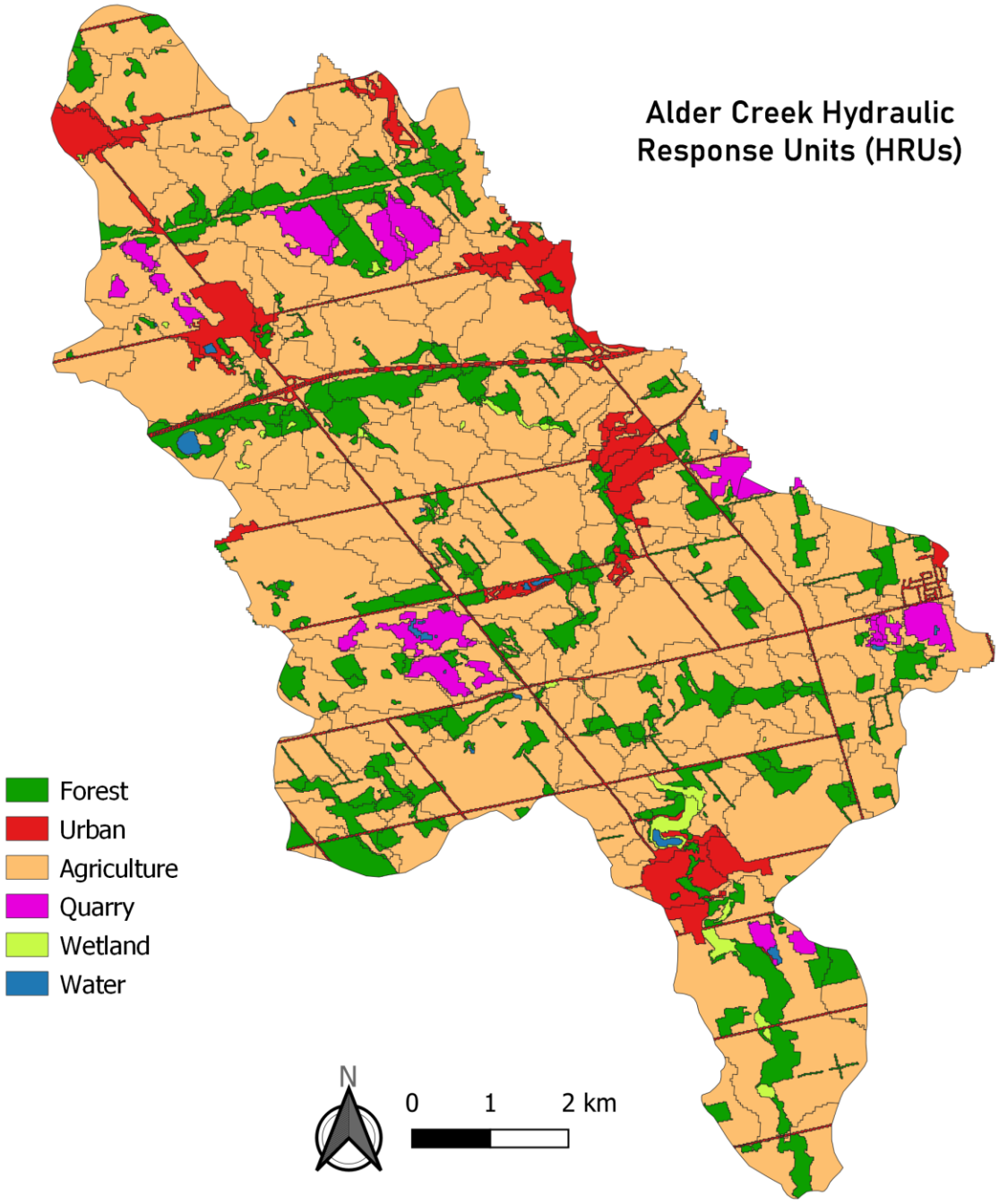

Figure 5-12 – Map of HRUs, and their simplified land use classes, in the model basin.

5.2.3 Coupled Model

To couple the models, Raven required the following additional information: (1) the location of the MODFLOW name file, (2) the HRU-node overlap relationship, and (3) the GW-River connection (PBJ

package) information. These were supplied in the Raven groundwater file (.rvg). For simplicity, the MODFLOW-USG model was copied into a subfolder of the Raven model.

The HRU-node overlap relationship (used in the Raven :OverlapWeights command) was generated by a custom R script using the shapefiles of the Voronoi groundwater model grid (Figure 5-5) and the surface water HRUs (Figure 5-12). As detailed in Section 3.3, the overlap relationship simply lists each HRU-node overlap combination that exists along with the percentage of the node covered by that HRU. However, the boundaries of the groundwater and surface models do not perfectly align. Despite the groundwater model being generated using a supposed watershed boundary, the DEM-based watershed delineation process creating the surface water model subbasins resulted in a different watershed boundary (Figure 5-13). This may be because the Alder Creek watershed boundary inherited by the groundwater model was developed prior to the release of modern high-resolution DEM products. When developing the subbasins (Figure 5-11 and Figure 5-13), some were clipped, extended, or – in three cases – added to ensure the model overlaps roughly similar. The goal was to ensure the surface water model boundary extended past the groundwater model boundary so that every groundwater node could be tied to a subbasin. Groundwater model nodes outside of a subbasin could have not received recharge from the Raven surface water model.

The GW-River connection (GWRiverConnection command) table was generated at the same time as the PBJ package for the MODFLOW-USG model, using a separate command in the *pbjR* R package. This produces a table with each line representing one segment of the stream. The only difference between this and the MODFLOW-USG PBJ package is that Raven additionally requires a column denoting the subbasin the segment belongs to.

In the uncoupled Raven model, baseflow was represented by the linear baseflow Raven process moving water from the second soil layer to the river. For the coupled model, this is handled by the GW-SW connection facilitated the PBJ package. Thus, the baseflow processes needed to be removed. The necessary change in model structure is diagrammed in Figure 5-14. The first soil layer baseflow process (meant to represent interflow) was unaltered. Note that the “Groundwater” box in the figure refers to the groundwater compartment in the uncoupled version, and the groundwater model in the coupled version. Notably, the HMETS Raven configuration our SW model was initially based on had no GW compartment (see Section 5.2.2).

Figure 5-13 – Overlap of the model spatial representations. Priority was placed on ensuring every node of the groundwater model was covered by a subbasin.

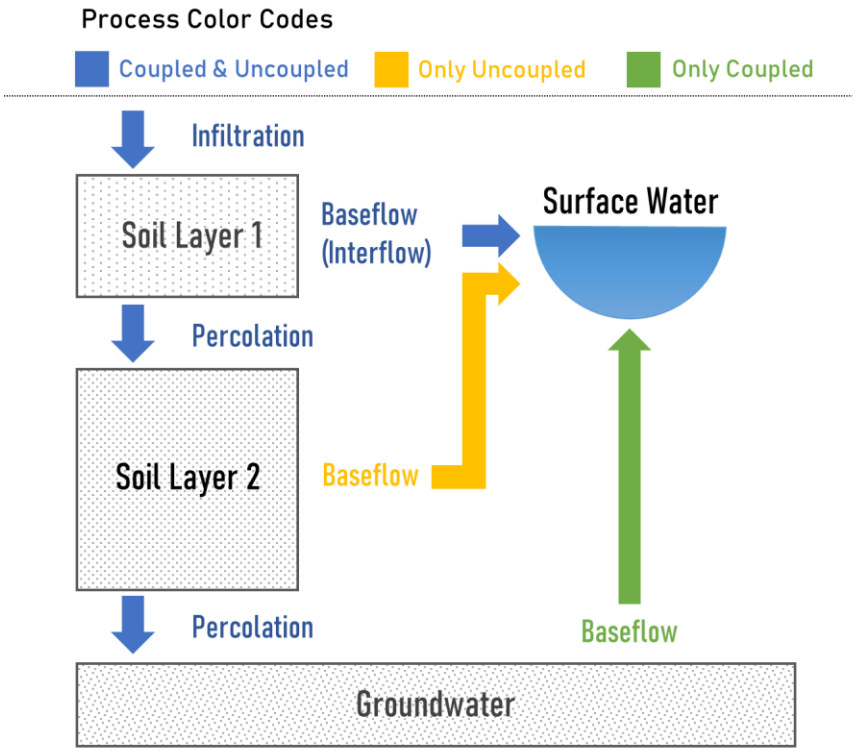

Figure 5-14 – Raven model structural changes in soil processes for the coupled & uncoupled versions.

5.2.4 Calibration of Models

All models (groundwater, surface water, and coupled) were calibrated using Ostrich (Matott, 2017) using the Parallel Dynamically Dimensioned Search (DDS) algorithm (Tolson et al., 2014) implementation. The parameterization and objective function for each of the models will be discussed in this section. As mentioned briefly before, the objective of calibration in this work was not to end with robust, predictive model realizations, but instead to estimate reasonable parameters that resulted in models performing well enough that the coupled Raven software functionality and representation of environmental processes could be evaluated. Given the poverty of data available for this project, the models are almost certainly overparameterized. As discussed later in Section 5.3.3, certain results highlighted issues with the model setup.

The strategy was to calibrate the surface and ground water models separately, with the hope that this would minimize calibration effort when the systems were connected. As discussed in Section 6.3, smart strategies for parameterizing, calibrating, and evaluating the uncertainty of coupled GW-SW models such as the one used here is an area of important future research.

Manual calibration was the starting place for the calibration for each of the models and was used to inform the parameterization for each of the models. Based on those experiences, tools were developed to allow Ostrich to intelligently manipulate the models for automated calibration.

Groundwater Model

Parameterization of the groundwater model was accomplished by utility programs created to adjust the streambed conductance and aquifer properties of the model. The utilities were written in Fortran and functioned by reading custom input files, used as templates in Ostrich, and writing out model input files.

The stream zones were broken up initially by stream order (i.e., level of branching) but during manual calibration additional zones were added to allow greater spatial variation in the model. A total of eight zones were used for the final calibration runs (Figure 5-15). The utility for adjusting the streams, SteamZoneAdjust, read in unit-length conductivity values for each zone and wrote them out to the PBJ package input file. Thus, Ostrich was used to estimate the unit-length conductivity for the eight zones. This lumped parameter included stream width, streambed conductance, and streambed confining layer thickness, as well as possibly compensating for some model structural issues.

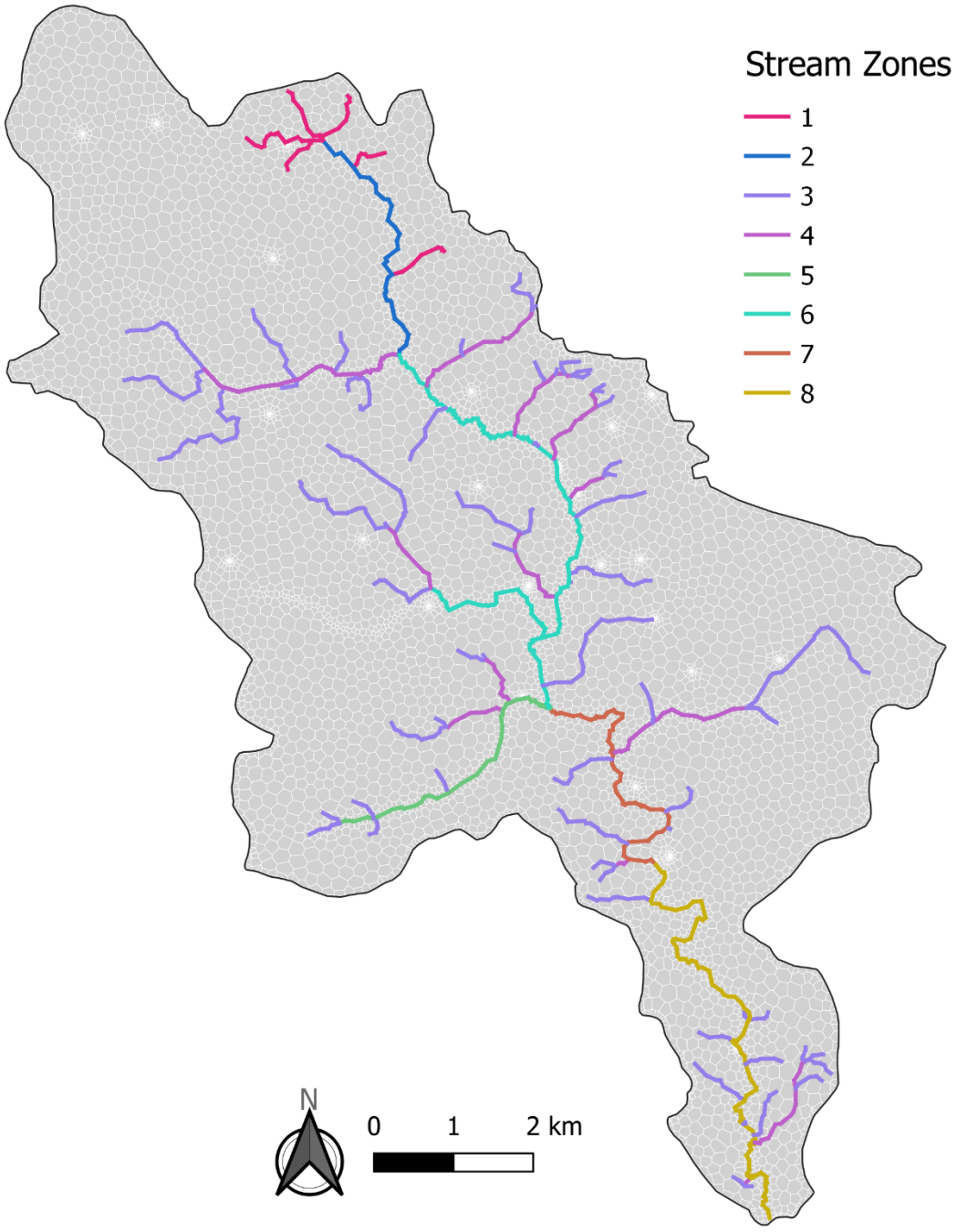

Figure 5-15 – Calibration zones for the PBJ package stream segments. Streambed conductivities were estimated by zone during the model calibration process.

The first two layers of the groundwater model were already divided into discrete soil zones (eight and nine zones, respectively, Figure 5-16). Those existing zones were leveraged to adjust the hydraulic conductivity and storage parameters for those layers. Layers three and four contained continuous (non-discrete) values and thus were adjusted using a single multiplier for parameter in each layer. Vertical

hydraulic conductivity was constant with a vertical to horizontal anistrophy ratio of 0.1. The utility for adjusting the parameters, ZoneAdjust, was programmed to be handle either multipliers or direct values for each parameter/layer combination. Experimentation during manual calibration, however, led to multipliers being the prefered method during automatic calibration. Conductivity and specific storage zones were allowed to vary an order of magnitude either direction; specific yield zones were adjusted between 0.1 and double their initial value.

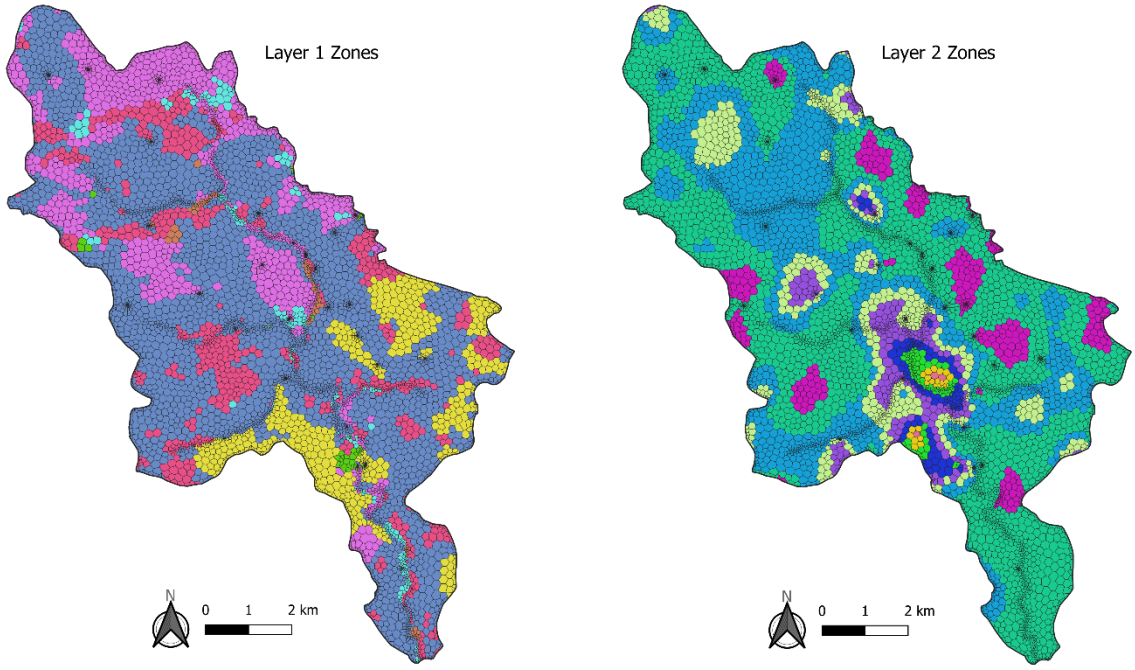

Figure 5-16 – Aquifer property calibration zones for the first two layers. Layer one has eight zones, while layer two has nine. Layers three and four were adjusted using a single multiplier for the entire layer.

The objective function for the model optimization was based on the sum of weighted squared residuals between observed and simulated values of (1) the head-difference times series at OW8-61 (2) the time series at W0000428-1 (3) the long-term average water levels in Table 5-1. This objective function is similar to the one used by PEST (Doherty, 2016) used to evaluate the solution. Flopy (Bakker et al., 2016) was used to import the model heads and a custom Python script calculated the residuals.

Surface Water Model

For the surface water model, parameters were adjusted through the Raven properties file (.rvp). Soil profile thicknesses, snow melt parameters, baseflow, runoff, percolation, and convolution parameters were all estimated by Ostrich after extensive manual calibration. As will be further discussed in Section 5.3, the model initially estimated much greater streamflow than observed at the stream gauge. This led to

additional processes being added, such as abstraction, and percolation to deep groundwater (initially, the conceptualization assumed losses to groundwater were balanced by baseflow gains).

In the final automated calibration runs, soil profile thicknesses were estimated in three groups based on permeability (high, medium, low). Baseflow and percolation coefficients for these soils were also estimated within these three categories. Snow melt parameters were treated as global and thus estimated all together. Runoff and gamma convolution (i.e., unit hydrograph) parameters varied by land use class (Figure 5-12).

In Ostrich, the Nash Sutcliffe Efficiency was used as the objective function to calibrate the stand-alone surface water model (e.g., maximize the NSE). Raven automatically calculates the NSE and writes it to a diagnostics file. The calibrated realization was then further manually adjusted to ensure a reasonable water balance compared to the estimates of Wiebe & Rudolph (2020).

Coupled Model

Only the groundwater model parameters were re-calibrated for the coupled model. Concurrent calibration of GW and SW domains was considered out of the scope of this project. The change from temporally and spatially constant recharge to spatially distributed daily recharge estimates made for a very different parameter estimation process. The methods used were identical to those discussed above for groundwater model calibration, and the objective function did not incorporate the surface water model output. However, one more stream zone was added (for a total of nine, Figure 5-17) to accommodate high residuals (model error) in the west side of the model. An alternate version of StreamZoneAdjust was written to output Raven groundwater input files (.rvg) instead of PBJ Package files.

While the additional baseflow contributions had an impact on the fit of the SW model, these impacts were not addressed through further calibration.

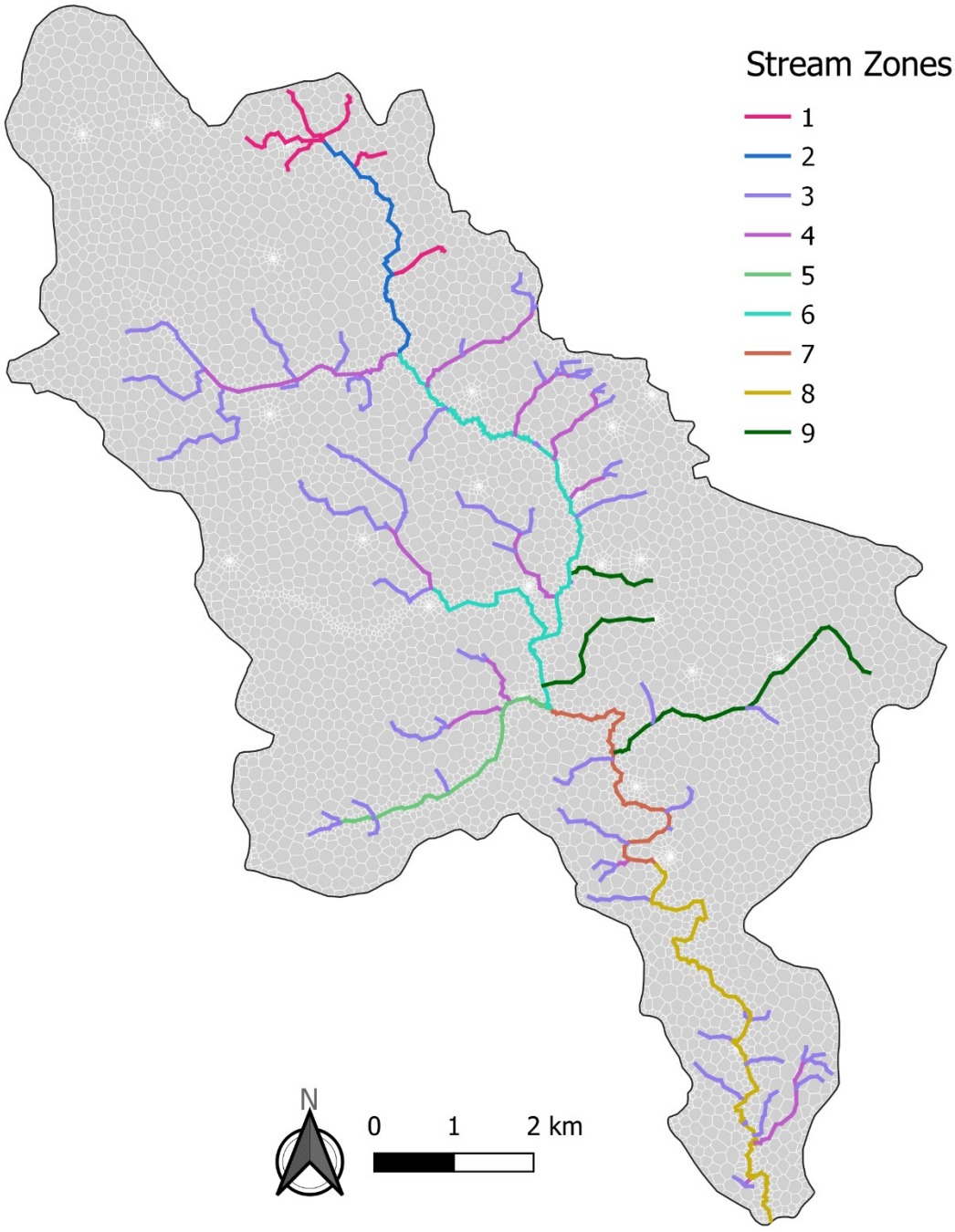

Figure 5-17 - Calibration zones for the PBJ package stream segments for the Raven coupled model. An additional zone (zone 9, dark green) was added in the west wing of the groundwater model to improve model performance in the region.

5.3 Model Results & Conclusions

The results of the model calibration efforts discussed in Section 5.2.4 are displayed and briefly discussed in this section. The model realizations resulting from calibration are poorly constrained by the

available data and likely have poor predictive power, but as mentioned in the introduction to this chapter, the intent is merely to show a practical application of the software to a real basin.

5.3.1 Uncoupled Groundwater Model Results

The groundwater model had poorly distributed data, particularly temporally, but the two wells with transient historical data were the primary focus of calibration. The model performance matching these wells, OW8-61 and W0000428-1 can be seen in Figure 5-18 and Figure 5-19, respectively. Error metrics for the wells are reported in Table 5-2, including Nash-Sutcliffe efficiency (NSE) (for the wells with time series), and root-mean-square error (RMSE).

Table 5-2 – Error metrics for the observation wells in the uncoupled groundwater model.

Well(s)	NSE	RMSE
OW8-61	0.84	0.413
W0000428-1	-32.6	1.25
Long-term Average	-	7.26

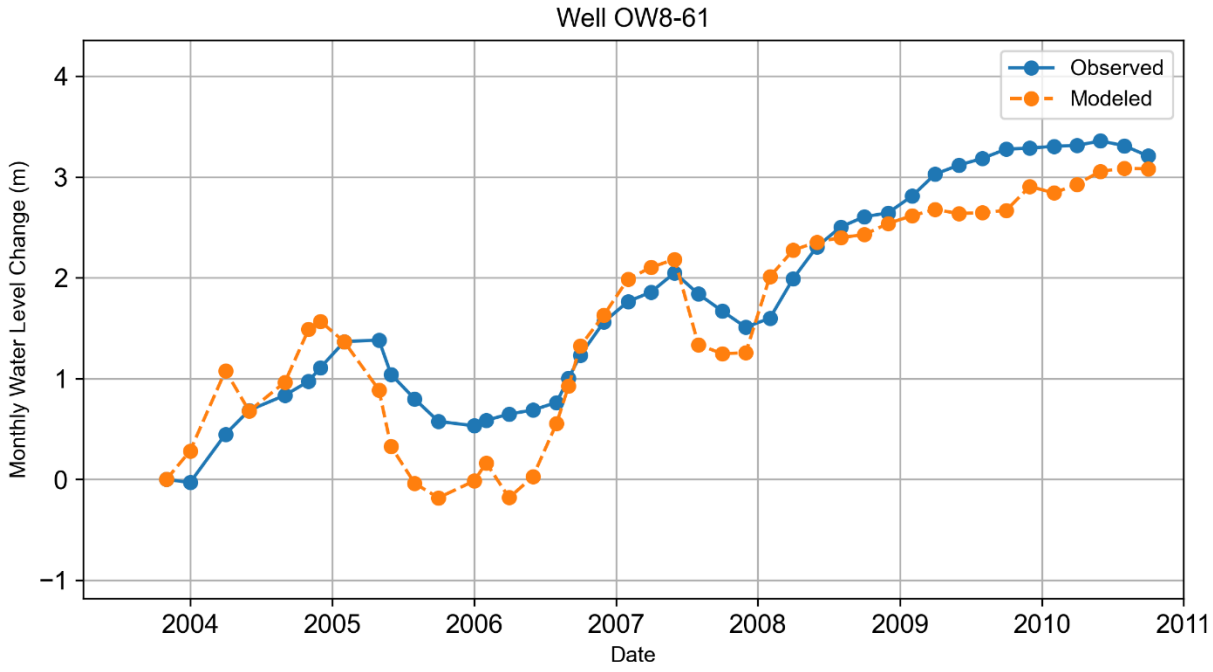

Figure 5-18 – Monthly water level fluctuation at well OW8-61, historical (observed) data and results from the uncoupled groundwater model.

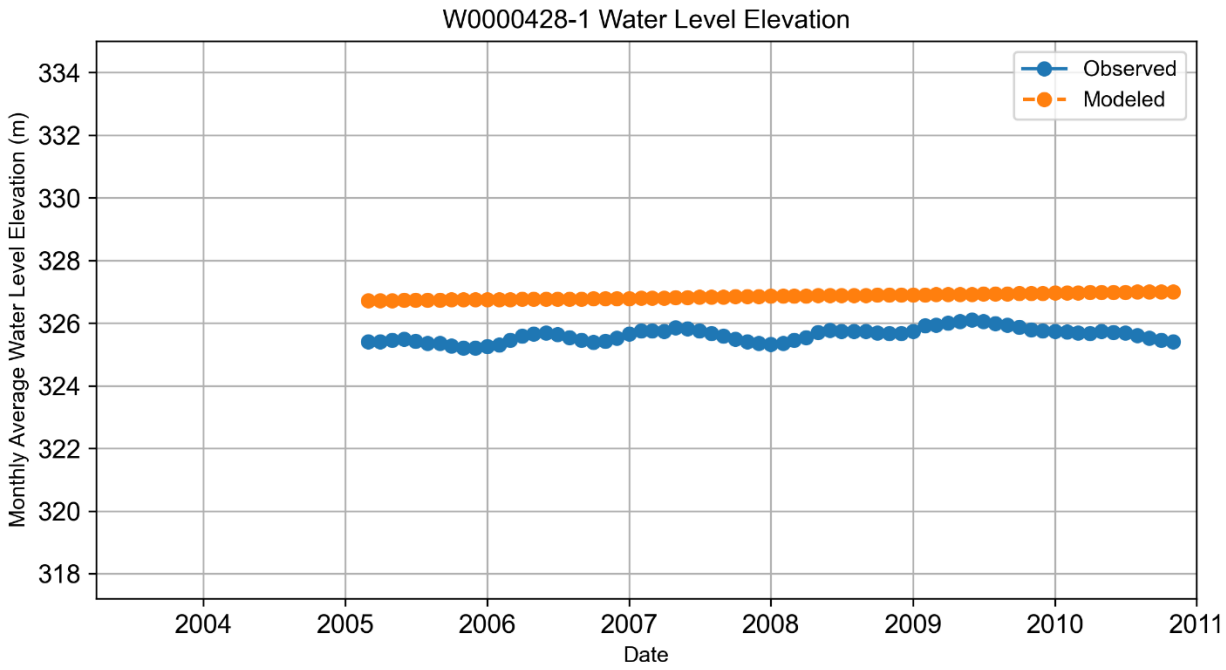

Figure 5-19 – Observed and modeled monthly water levels at well W0000428-1, located nearly at the edge of the groundwater model (see Figure 5-10).

The long-term average water level elevations from 25 wells (Table 5-1) were also used as targets. A scatterplot comparing the modeled versus observed averages of these wells is displayed in Figure 5-20. The scatterplot shows a clear high bias in the modeled water levels (as also seen in W0000428-1, Figure 5-19). There is both a computational and potentially physical reason for this bias.

A simple reason for the bias is simply that the long-term average wells were not weighted as highly in the calibration process as the wells with transient data. Had the wells been weighted more evenly, or higher, we could expect the bias to be closer to zero and the wells to be evenly scattered around the 1:1 line.

The reason the long-term average data was not weighted higher is that it is based upon data prior to the simulation period. As seen in the water level changes of OW8-61 (Figure 5-18) water levels increased during the years of simulation. However, it is not clear how representative the trend at OW8-61 was. For instance, well W0000428-1 (Figure 5-19) does not show much variation in water levels at all. The trends at OW8-61 are much more influenced by the river and local pumping conditions.

Given that the recharge rate was constant over the uncoupled groundwater model simulation, it seems very likely that some of the bias seen in Figure 5-20 is due the model/calibration trying to artificially meet transient pumping and water level targets with a fixed flux of water into the system. This

highlights the utility of a coupled SW-GW model where a more realistic, time-variant recharge rate can be used.

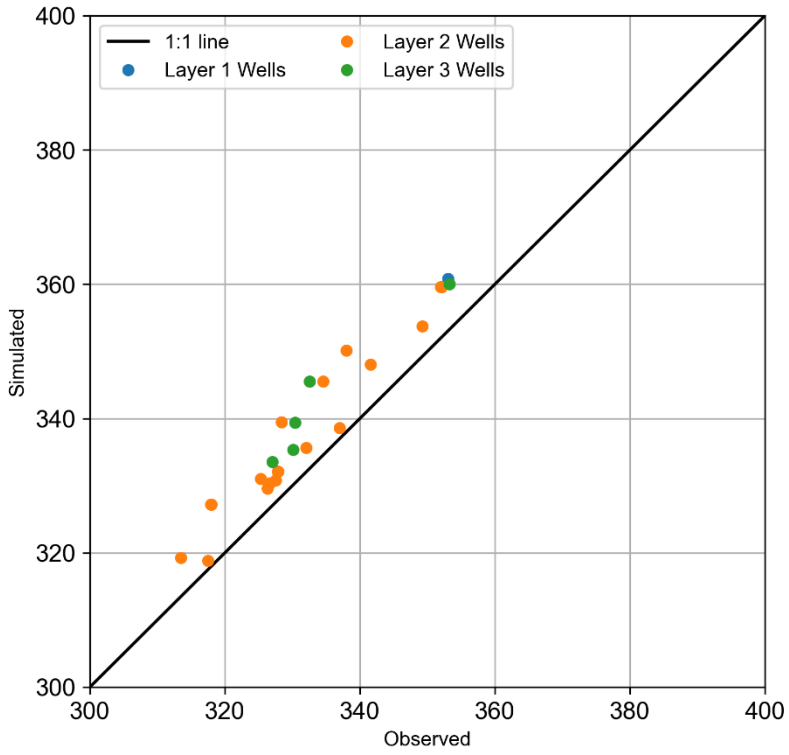

Figure 5-20 – Scatterplot of long-term average water level wells (see Figure 5-10) comparing observed and simulated values. The simulated values are biased high, likely as a result of being weighted lower in the calibration objective function compared to the wells with transient data.

5.3.2 Uncoupled Surface Water Model Results

The full hydrograph of the single streamflow gauge used for calibration is shown in Figure 5-21. The NSE for the model is around 0.45, which indicates the model performs better than the long-term average of the data, but the fit is generally poor (Ritter & Muñoz-Carpena, 2013). Difficulty was encountered in the calibration process in maintaining realistic fluxes out of the model to the groundwater system and atmosphere (based upon estimates and bounds from Wiebe & Rudolph (2020)) without compromising the NSE. Given the importance of the fluxes to the groundwater system in this exercise, maintaining a reasonable recharge to the deep groundwater store was prioritized. Generally, there was too much water in the system, and the model had difficulty matching the low flows (and low baseflows) seen at the Alder

Creek gauge. Figure 5-22 shows a subset of the simulation years (water years 2004 and 2005). The snow and snow melt storage compartments from Raven are shown as well, to demonstrate the yearly melt dynamics. The NSE also was intentionally impacted by biasing baseflow slightly low in the final uncoupled surface water model to facilitate anticipated baseflow from the groundwater model in the final coupled simulations.

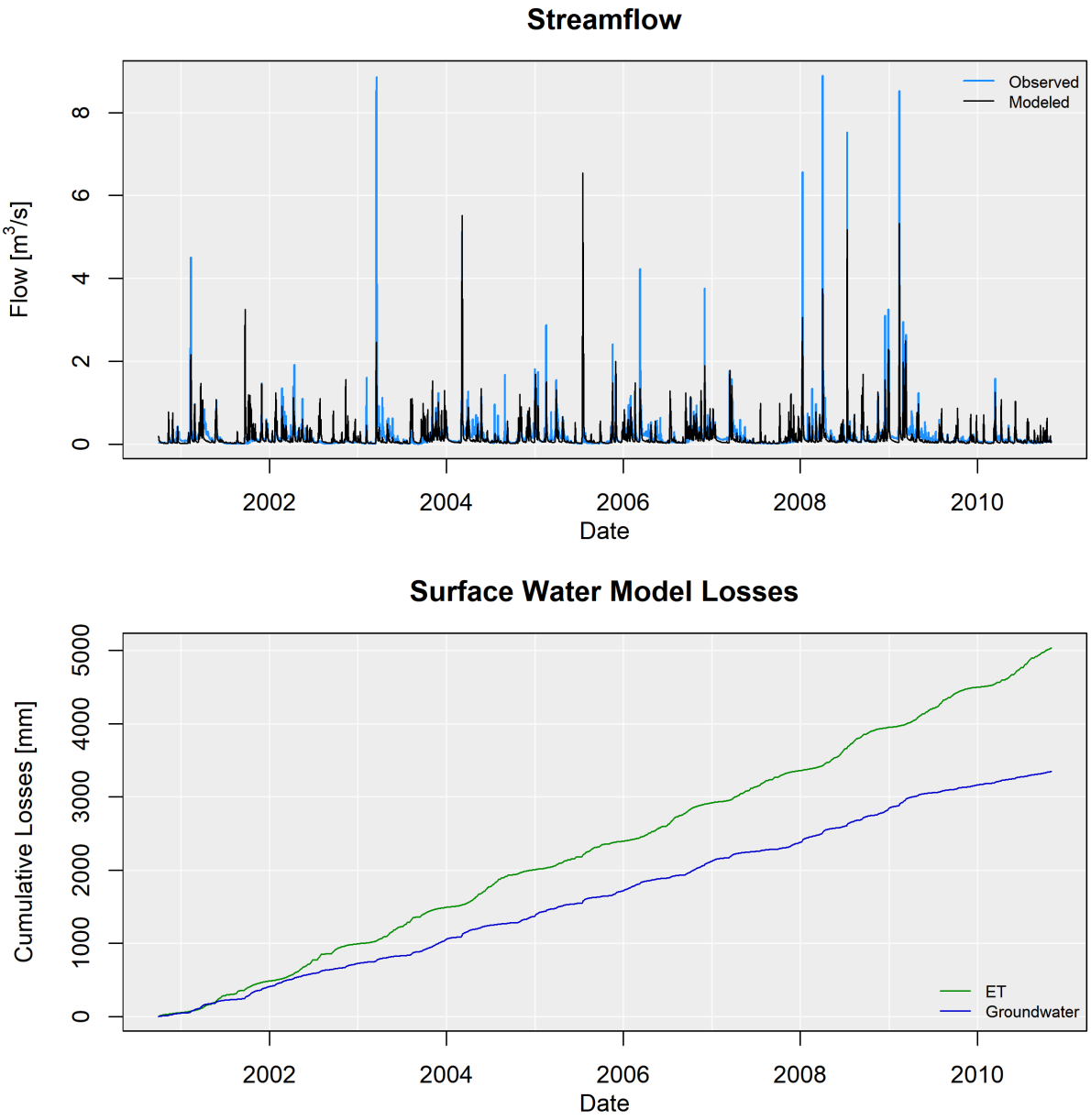

Figure 5-21 – Observed and modeled streamflow (top) for the entire simulation period. Cumulative losses to groundwater and atmosphere (ET) from the model are shown on the bottom plot.

The total average annual fluxes out of the surface water model were estimated at 332 mm/year to groundwater and nearly 500 mm/year to atmosphere (ET). This groundwater flux is within the range estimated by Wiebe & Rudolph (2020), but the losses to ET are higher. However, as noted before, their estimates are for the years 2014-2016.

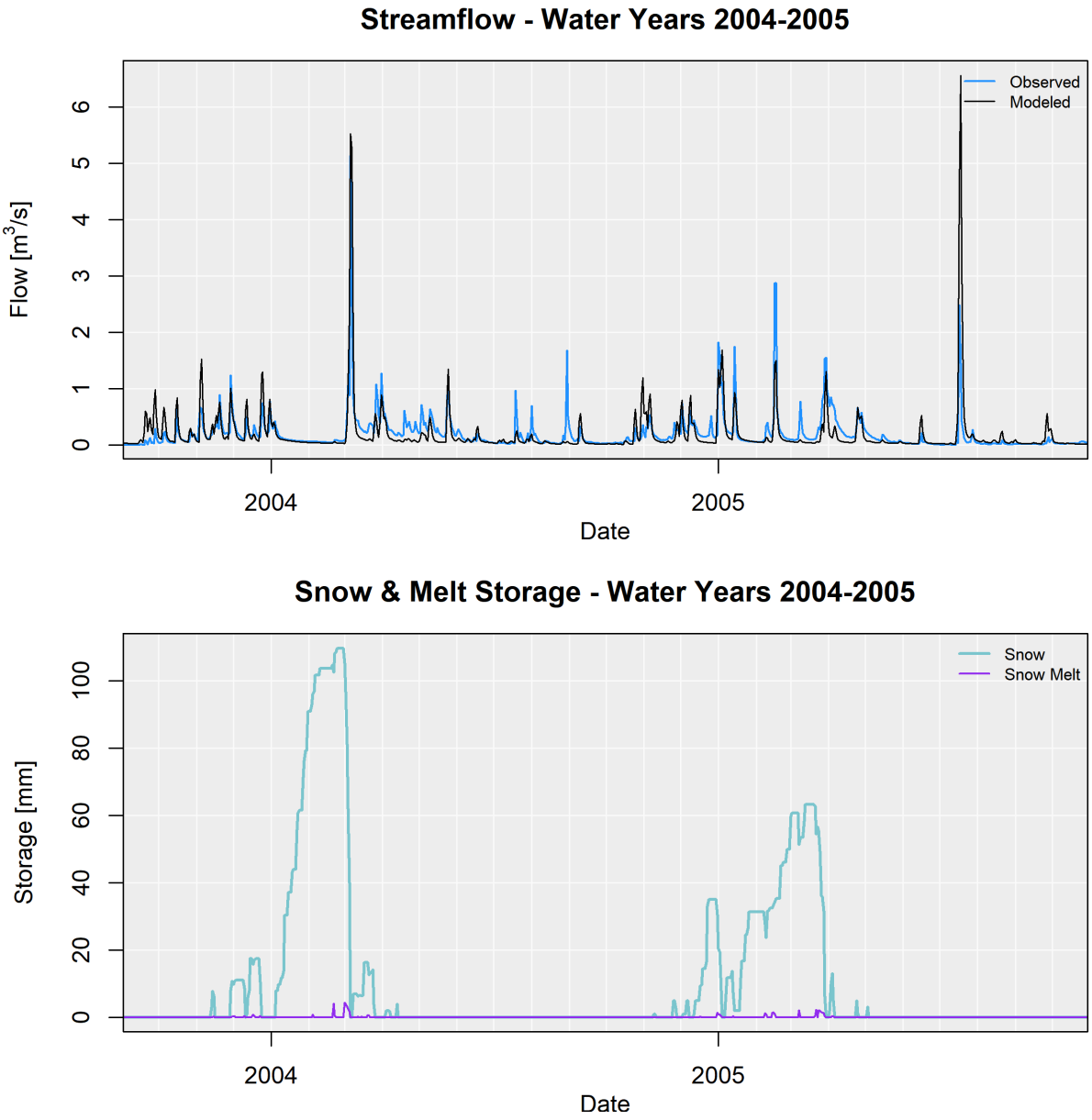

Figure 5-22 – Closeup of observed/modeled hydrograph for the water years 2004 and 2005 (top), and modeled snow dynamics (bottom) over the same period.

The years water years 2001-2011 had higher precipitation at the Roseville gauge, on average, than any of the scenarios used in their study. The spatial distribution of average water flux to the groundwater system (groundwater recharge) are shown in Figure 5-23. The recharge map resembles the HRU map (Figure 5-12), because recharge is determined by HRU. It is worth noting that while the average total flux to the

groundwater system was monitored during manual calibration, this does not necessarily ensure the spatial distribution was reflective of reality.

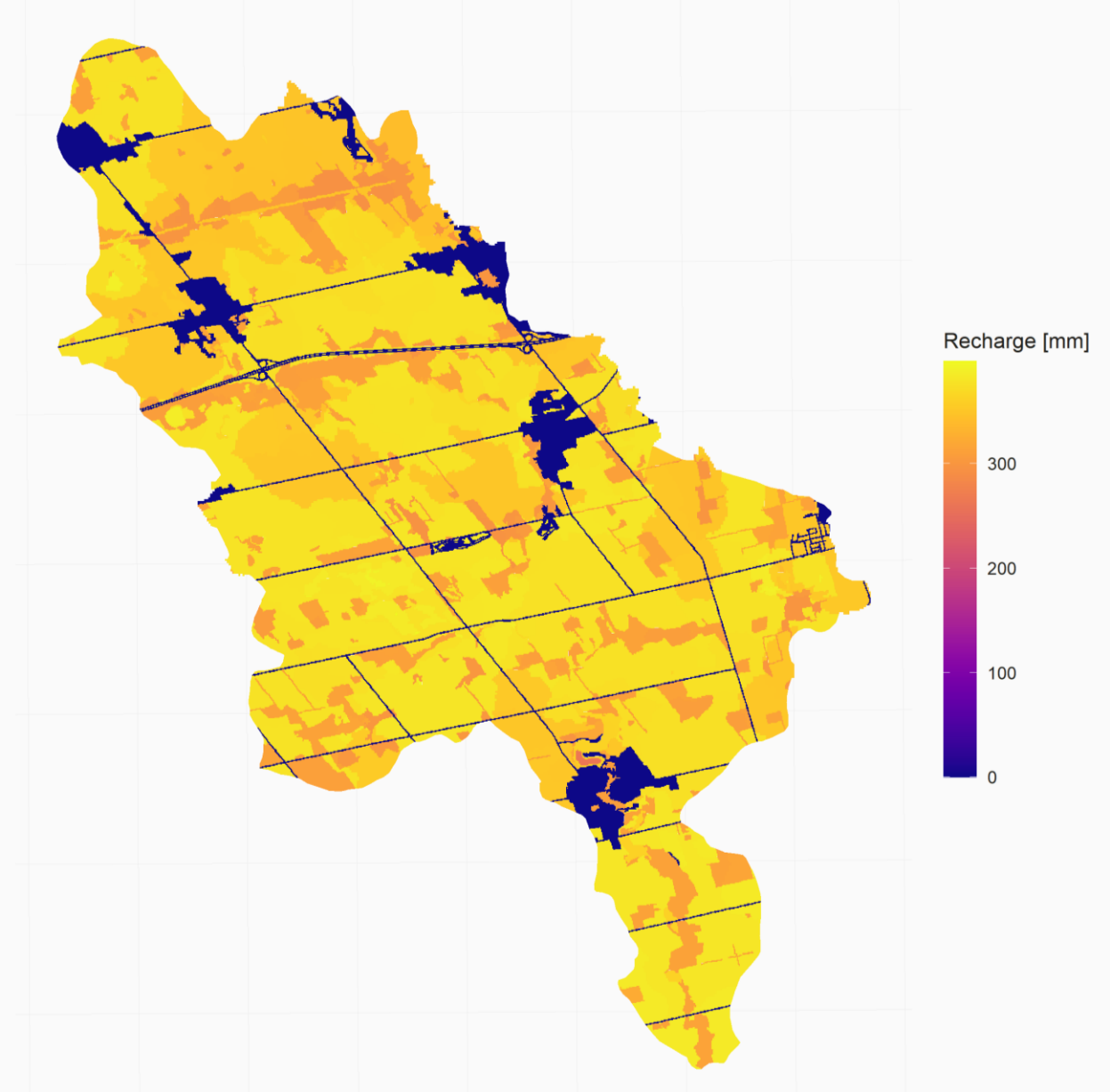

Figure 5-23 – Modeled recharge rates from the surface water model, for each HRU, passed down to the groundwater compartment.

5.3.3 Coupled Model Results

The coupled model, as briefly explained in Section 5.2.4, was only calibrated to the groundwater model targets. The Alder Creek model realization produced from this process is likely not a very realistic

model, but instead serves as a potential starting point for further calibration and investigation for future efforts. However, it is a great demonstration of the effects and potential advantages of the coupled GW-SW modelling framework, and thus we will discuss the results in detail.

The spatially distributed average net annual flux, calculated as recharge minus stream losses, is displayed in Figure 5-24. The figure color gradations have been purposely selected to showcase the (positive) spatial variation, however, a few errant PBJ segments have very high flows from the GW model. These outliers are obscured through the bin (range) selection but will be discussed in Section 5.3.4. For reference, the average PBJ node has a flux of about -3600 mm/year.

The model was able to generally capture trends at the two transient wells at the two observation wells with transient data (Figure 5-25 and Figure 5-26). Error metrics (NSE, RMSE) are reported in Table 5-3. For well OW8-61, for which only digitized water level fluctuations were available, the simulated values are biased high and display a peak during 2009 that are not reflected in the data. The NSE dropped from 0.84 with the uncoupled groundwater model to -0.14 with the coupled model. This well is an a very interesting location (Figure 5-10) which is both affected the stream, pumping wells, and in the coupled model, an area with very little recharge (the overlying HRU is urban, which is treated as fully impervious). Notably, the starting point for the observed data is unknown, but the residual computed from the long-term average (seen in Figure 5-28) is -4.2 m, suggesting even the starting point may be too high. As discussed more in the next section, both this study and previous studies have identified this area of the model as a net positive recharge region where pumping may pull water from the river. Model limitations arising from the assumption of a purely gaining stream (the selection of the drain package) and the use a completely impervious urban region above likely negatively impact the model fit at this well.

Table 5-3 – Error metrics for the observation wells in the coupled groundwater model

Well(s)	NSE	RMSE
OW8-61	-0.14	1.12
W0000428-1	-5.88	0.565
Long-term Average	-	4.99

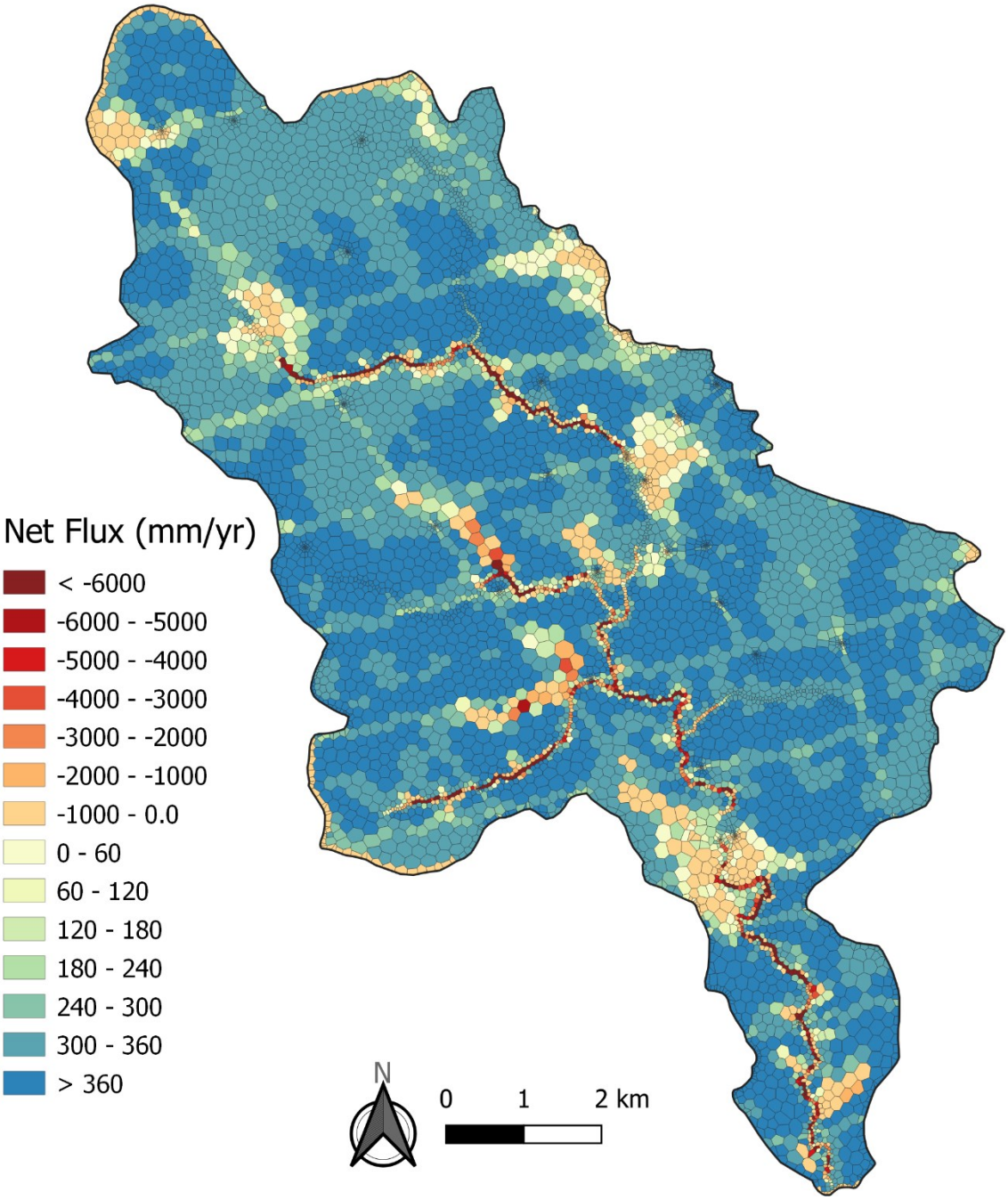

Figure 5-24 – Average net flux (recharge – losses to river) for the coupled Alder Creek model. Note the very different bin sizes for positive (cool colors) and negative (warm colors) values.

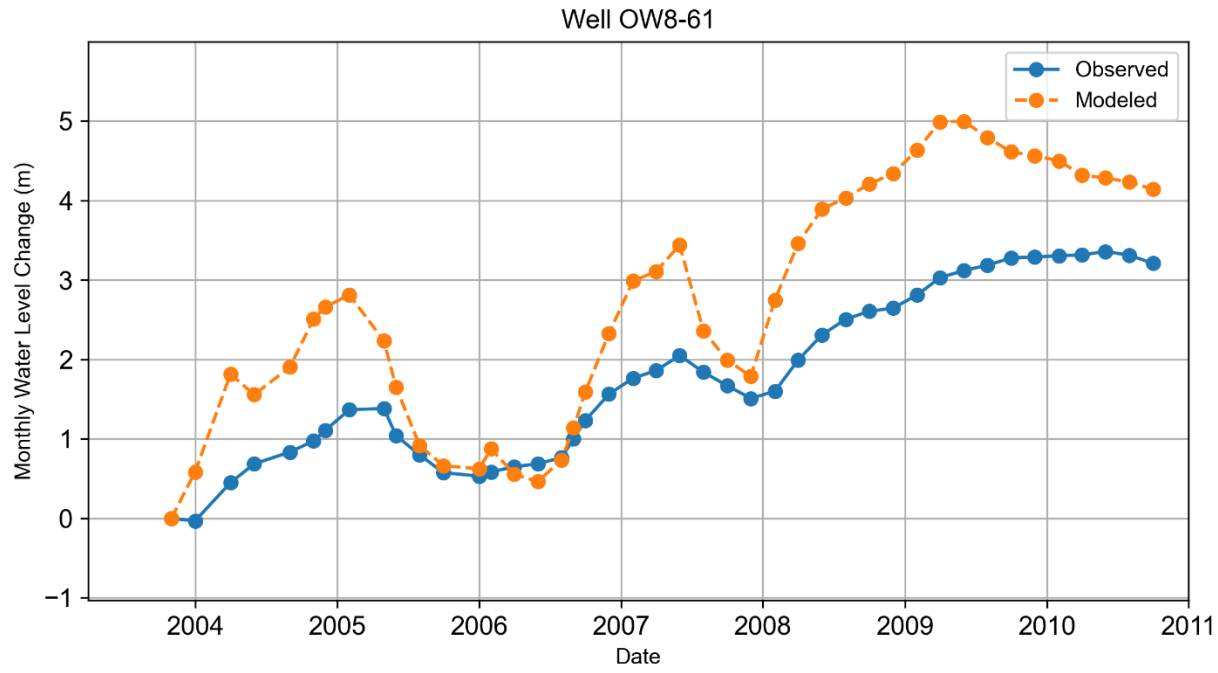

Figure 5-25 - Monthly water level fluctuation at well OW8-61, historical (observed) data and results from the coupled model. The fluctuations are much greater and more biased than the uncoupled results (Figure 5-18)

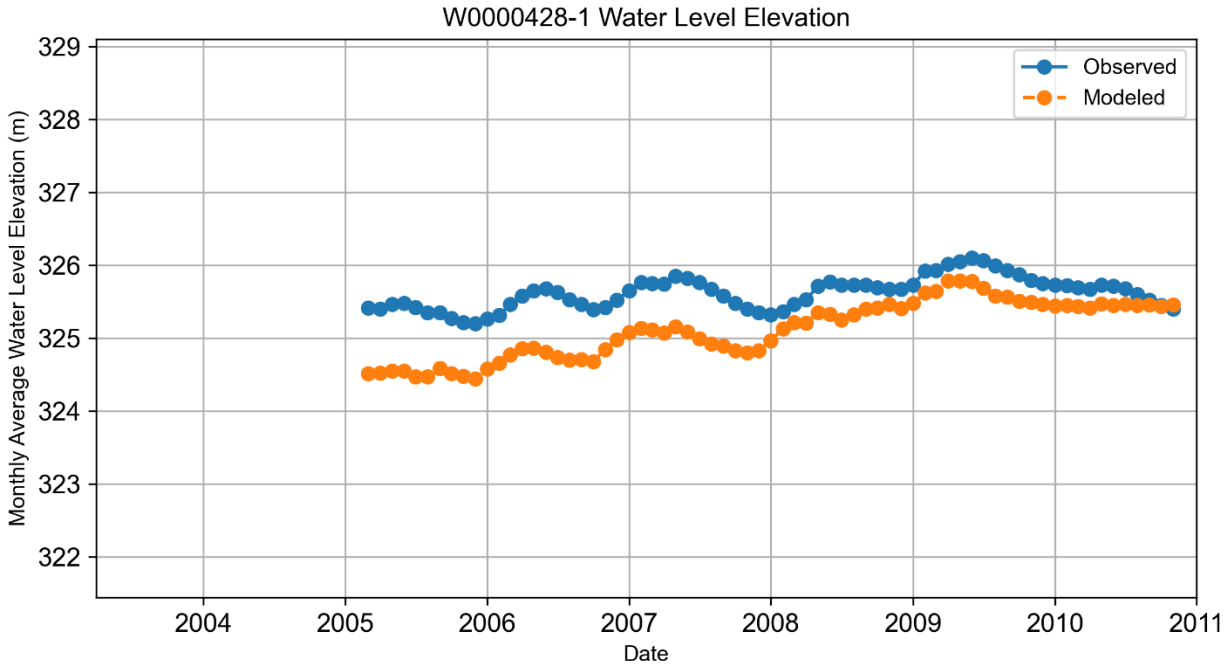

Figure 5-26 - Observed and modeled monthly water levels at well W0000428-1 in the coupled model. The modeled results capture the trends of the data much better than the uncoupled model (Figure 5-19). Note the different vertical scale used in this figure (smaller range compared to 5-18).

Well W0000428-1 (Figure 5-26) displays a comparatively good match between modeled and observed values. Of particular note – to the extent that it is fair to compare the performance of the uncoupled and coupled models – the coupled version captures the water levels trends in the data much better than the uncoupled model (Figure 5-19), which is essentially produced a straight line. While this is generally just reflective of the differences in recharge (daily estimations versus a constant value), it does highlight the advantage of the model-estimated recharge. The trend in the observed time series seems slightly lagged compared to the modeled values, possibly reflecting the assumption of recharge instantaneously moving from the surface water soil storage to the groundwater model.

The modelled water levels at the long-term average wells (Table 5-1) compared to the observed values are shown by layer in a scatterplot (Figure 5-27), a residual map (Figure 5-28), and the RMSE is reported in Table 5-3. Unlike the uncoupled GW model, the long-term well residuals are not all biased high. This possibly reflects the change to spatially varying recharge, although the lower water levels in the eastern wing of the model suggests the recharge distribution could use improvement. For the coupled model, a separate stream zone was added in this region to attempt to improve the model fit. The calibration estimated a very low conductance, effectively shutting off much of stream (Figure 5-24). The mismatch could also be caused by poor boundary assumptions (i.e., no-flow along the east side of the model).

The groundwater model produced much more baseflow than anticipated, which resulted in a reduced fit for the SW model results at the one streamflow gauge (shown for a subset of years in Figure 5-29). However, it does show the groundwater model is capable of delivering fluxes from the PBJ package to Raven. Both coupled and uncoupled model results are shown in the Figure 5-29. The final NSE of the modeled streamflow was 0.33, a drop from the uncoupled NSE of 0.45.

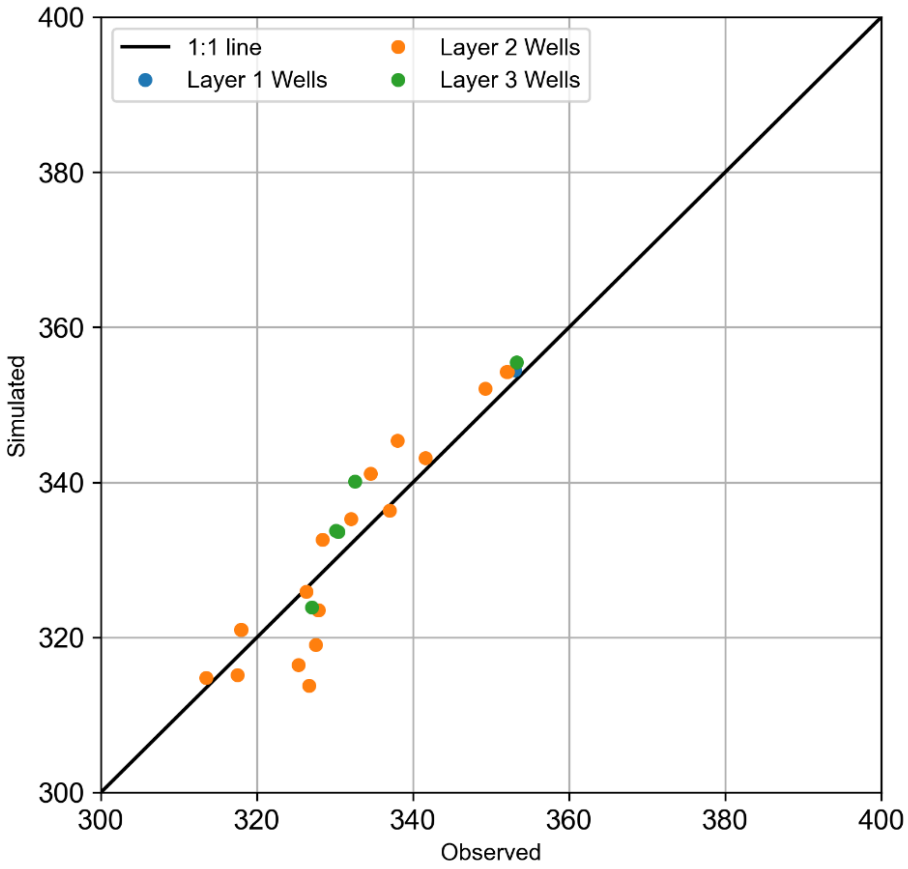

Figure 5-27 - Scatterplot of long-term average water level wells comparing observed and simulated values from the coupled model.

The bottom plot of Figure 5-29 displays the daily total recharge from the SW model to the GW model. Interestingly, it strongly mirrors the streamflow – reflecting both processes are essentially controlled by the soil model dividing water between the stream and the groundwater model, driven by the daily precipitation.

A mass balance for the entire simulation is presented in Table 5-4. This table is primarily calculated from the Raven and MODFLOW-USG local input/output mass balances. Recharge to portions of the SW model that extend past the GW model (see Figure 5-13) are included as “Raven Losses to GW”. The small (0.024%) mass balance error is similar to the uncoupled Raven mass balance error (0.023%) showing the coupled model does not greatly increase the mass balance error. The GW mass balance error reported by MODFLOW is 0.000018%.

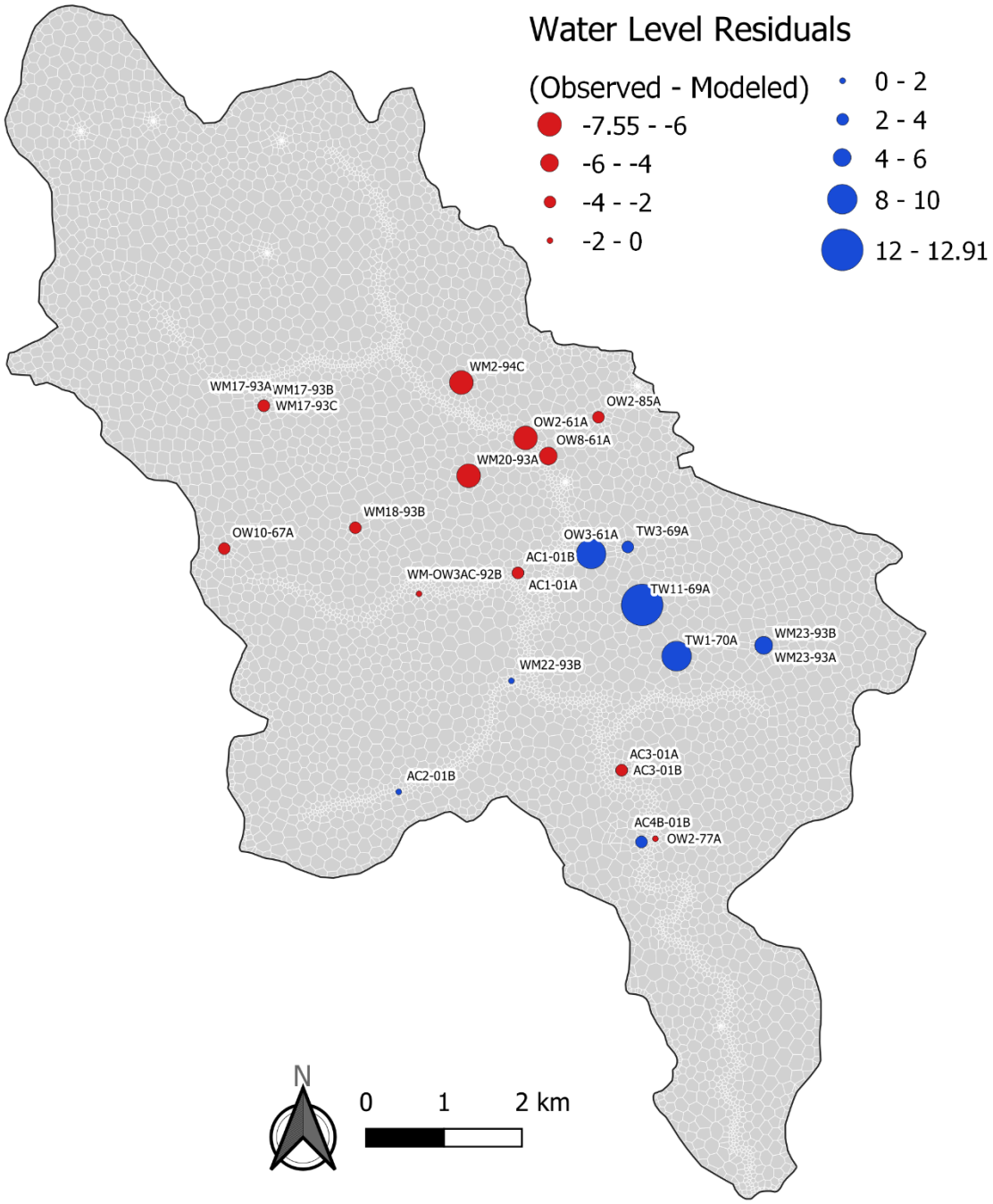

Figure 5-28 – Map of water level residuals for the long-term average observation wells for the coupled model.

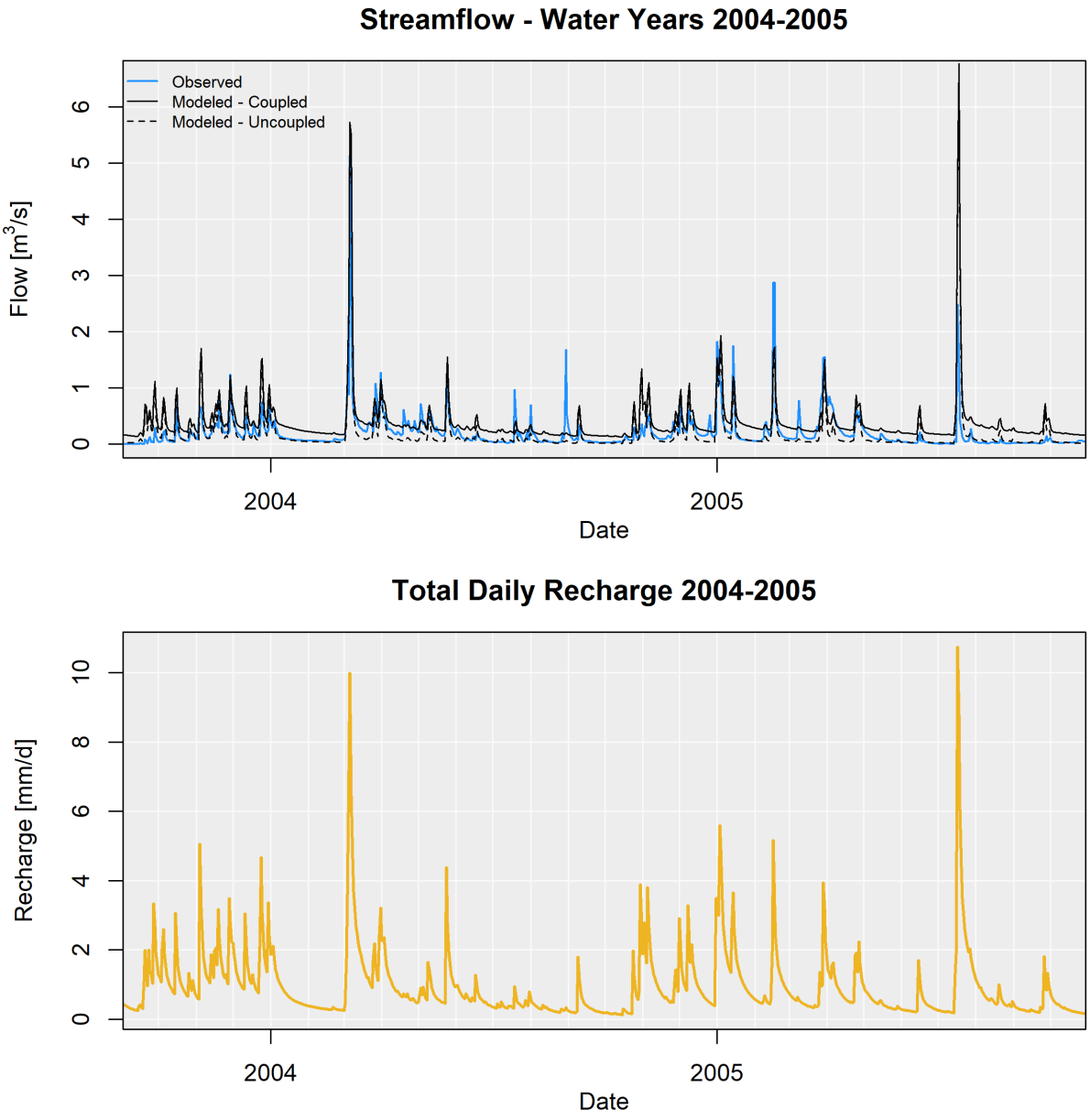

Figure 5-29 – Streamflow (top) for the coupled model (for the same time period of Figure 5-21), with the uncoupled streamflow (dashed line) shown for comparison. The groundwater model increases the baseflow to stream, reducing the fit of the surface water model. The lower plot displays the total SW flux to the groundwater model over time.

An inter-model mass balance (e.g., water volumes transferred between models) is shown in Table 5-5. An unusual sign convention is used: positive values indicate volumes from the model *sending* the water, negative values correspond to a *received* volume. Thus, the sum of the values in the error. Balances were separately calculated for recharge and groundwater loss to streamflow. The percent error was calculated as the absolute sum of both errors divided by the sum of the sent volumes. The inter-model mass balance error is very small and is likely only non-zero because of calculation differences between the Fortran and C++ portions of Raven (i.e., roundoff errors).

Table 5-4 – Global mass balance for the coupled Alder Creek model. All values are in cubic meters and rounded to the whole number.

WATER IN	
Precipitation	760007358
Constant Head	7162963
GW Storage	77319961
IN Total	**844490282**

WATER OUT	
Constant Head	13596980
Wells	80753726
Losses to Atmosphere	397951573
Streamflow Out	219173261
GW Storage	127937992
Losses to GW	4876090
OUT Total	**844289621**
IN - OUT	**200661**
Percent Error	**0.024%**

Table 5-5 – Inter-model mass balance for the coupled Alder Creek model. All values are in cubic meters and rounded to the nearest tenth.

Recharge from Raven	264635096.0
Raven Losses to GW	-4876090.3
Recharge to MODFLOW	-259759022.6
Recharge Error	**16.3**

Raven Stream Gain	-121953187.9
MODFLOW Stream Loss	121953188.2
Streamflow Error	**0.35**

Mass Balance Error	**16.6**
Percent Error	**0.000004%**

5.3.4 GW-SW Modelling Conclusions

The Alder Creek example shows that the coupled Raven GW-SW modelling framework is able to connect the two domains and pass fluxes between them in a physically reasonable manner. The Raven model generates spatially-distributed recharge fluxes that get passed down to the MODFLOW-USG model (Figure 5-23, Figure 5-24) and the groundwater model returns water that seeps into the river streambed directly to the river channel in Raven. The mass balance results (Table 5-4, Table 5-5) show that mass is conserved globally as well as between the two domains. Preservation of the mass balance and estimation of reasonable fluxes were the primary objectives of this modelling experiment. Moving from uncoupled models to a coupled model is as easy as generating an input file using shapefiles and a few R commands and telling Raven the location of the MODFLOW name file.

The calibrated Alder Creek model presented here was compared to results from prior studies in the region. Generally, the model fit is worse than prior modelling studies (CH2M HILL & SSP&A, 2003;

Chow et al., 2016; Jones et al., 2009; Matrix Solutions Inc & SSP&A, 2014), although, of those studies, only Matrix & SSP&A (2014) was a transient model. Matrix & SSP&A (2014) also had the benefit of much more data, particularly transient data, to inform model development and calibration.

Both Jones et al. (2009) and Chow et al. (2016) are studies that modeled this region using a HydroGeoSphere model, the grid of which is the geometric dual to the Voronoi groundwater model grid used here (see Section 5.2.1). While this, to some extent, gives us the ability to directly compare their results, those studies themselves do not directly agree, and both treated HGS primarily as a tool for generating steady-state GW-SW exchange estimates as inputs to other models. In both studies, a 200 mm/yr rainfall rate was assumed and run until steady flow conditions emerged (Jones et al., 2009). The models were calibrated manually to match long-term average water level values. It is unclear why 200 mm/year was chosen for the rainfall, as both the total annual precipitation (~900 mm/year) and estimated recharge rate (321 mm/year) (Wiebe & Rudolph, 2020) exceed that value. However, the complex HGS model, which included 87 layers of interpolated aquifer parameters, still serves as an important qualitative comparison for the coupled Raven Alder Creek model, given the much more complicated representation of GW-SW processes in HGS.

The net flux map is repeated in Figure 5-30, but with specific areas circled. The areas circled in green roughly highlight regions where the Raven model and the HGS model (Chow et al., 2016; Jones et al., 2009) showed net groundwater recharge along the stream (e.g. losing stream sections) (although both previous studies allegedly used the same HGS realization, their recharge maps seemingly disagree). The areas circled in orange highlight regions where Raven estimates no groundwater losses to streamflow but the HGS model predicted large losses to streamflow. The larger orange circle contains an area of the model where the stream shapefile did not identify that a stream existed, i.e., it was not included in the PBJ package. Recent aerial photos of the area do not show there to be a stream in this region, although perhaps one existed in the past. It is not clear what informed the previous studies decisions to include a stream in this area.

Interestingly, some of the regions in the Raven model where we assumed steam channels to exist (based on the stream shapefile) but that were not included in the HGS model grid (which clearly was discretized around known stream channels) are shown to have negative recharge in both the Raven model and the HGS model. These regions are not highlighted but can be identified by the larger warm-colored cells in Figure 5-30 in the mid-western side of the model. These areas somewhat align with known wetlands (Figure 5-2), where our assumption of drains to represent Alder Creek may be quite apt.

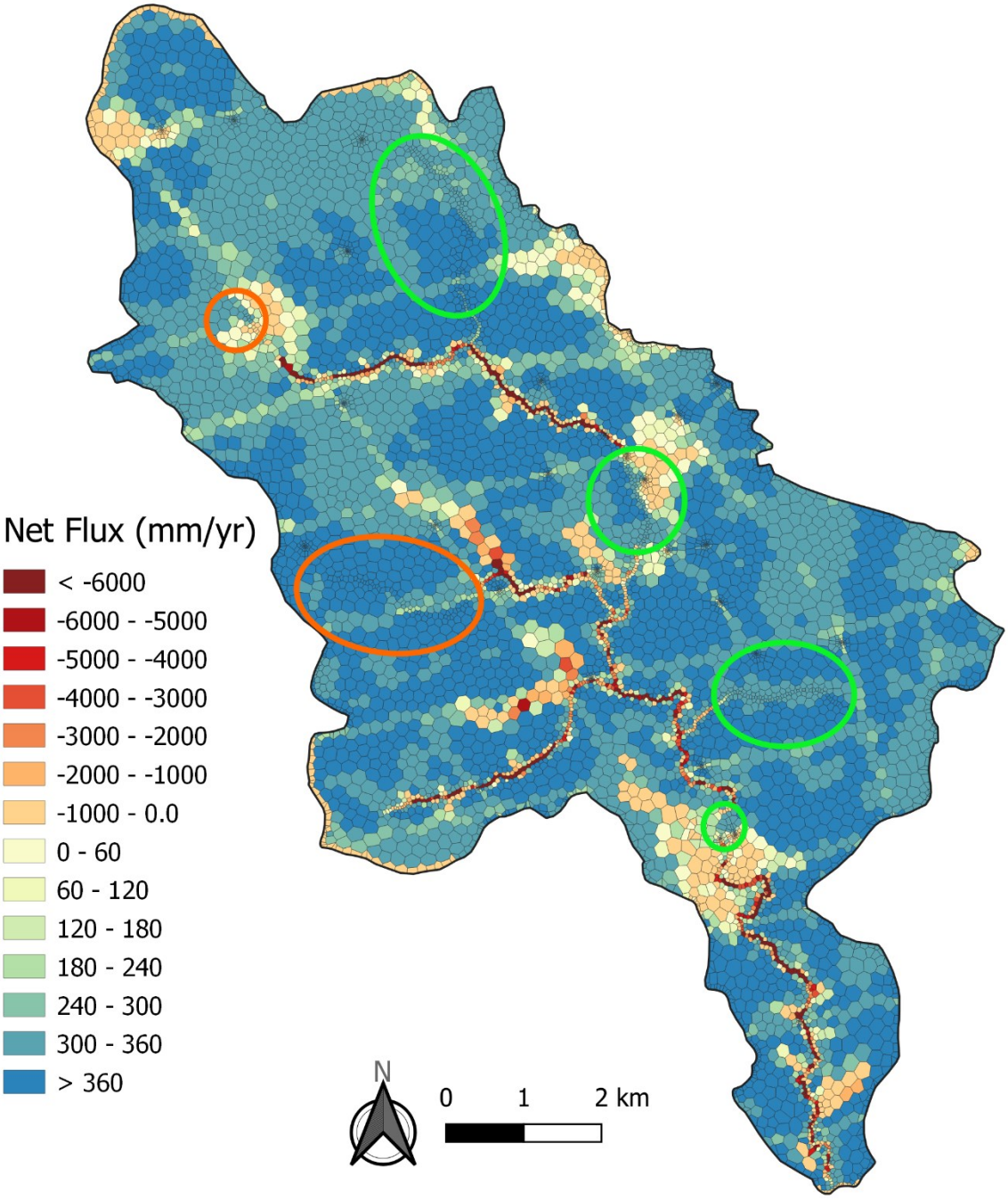

Figure 5-30 – Net flux (recharge – stream losses) with areas of interest discussed in the text circled. The four green circles correspond to results corroborated with prior studies modelling Alder Creek with HydroGeoSphere, while the orange circles are areas of conflicting results.

The Jones et al. (2009) and Chow et al. (2016) studies show fairly different ranges of exchange flux – 26280 mm/yr to -26280 mm/yr and 1000 mm/yr to -1000 mm/yr, respectively – to which the estimates here are closer to those in Chow et al. (2016). However, as mentioned before, there are some clearly erroneous regions in the coupled Raven model where extreme rates are present. For instance, one node in

the lower-west wing of the model has an average net exchange flux of nearly -67,000 mm/year. Nearby nodes have somewhat more reasonable fluxes, but much of that stream branch has fluxes in excess of -10,000 mm/year. This branch had a separate stream zone (Figure 5-17) for estimating conductivities, is near constant head cells moving water into the model, and is just south of the region noted in Figure 5-30 for lacking a gaining stream section. The high stream loss node is a particularly great outlier because the DEM-sourced elevation for the stream segments in the region dips slightly, creating a fictious low point in the stream. Further efforts with this model would require careful vetting of the relative elevations of the stream segments and perhaps better, data-informed constraints on the stream segment conductivities. This stream branch connects to the main Alder Creek branch below the stream gauge, so quantities of water delivered to it are poorly constrained by the existing surface water data.

Minor oddities aside, the Raven model was able to reproduce similar exchange fluxes (in both magnitude and spatial distribution) to those found in the HGS model. The coupled Raven model runs in about an hour, the HGS model was reported to take 18 hours (in 2016, but without temporally varying rainfall). Run times of the Raven model could be greatly improved by adding the capability for coarser groundwater model time steps (reducing the flux changes between time steps, see suggestions in Section 6.3). While the HGS model possesses many capabilities Raven does not, for the purposes of estimating exchange flux between the surface and groundwater domains in Alder Creek it is not clear what advantage HGS holds, particularly when run times (and feasibility for automated calibration and uncertainly analysis) are considered.

While the fits in our model are not very satisfactory, they easily could be improved with (1) more transient water levels and streamflow measurements, (2) better selection of model boundary/boundary conditions, (3) portions of the stream being modeled in river mode, rather than as drains, to better represent the GW/SW interactions, (4) hydraulic conductivities estimated using pilot points (Doherty, 2003; Doherty et al., 2010) to better represent sub-grid (and sub-zone) heterogeneity, instead of homogenous zones (Kitanidis, 2015) and (5) multi-objective calibration optimizing both surface and groundwater domain fits. In particular, one could explore the tradeoff between optimizing the fit of the groundwater and surface water models. None of this would specifically require any modifications to the software, although many of the software modifications suggested in Section 6.3 would likely improve the Alder Creek model as well.

Chapter 6 - Conclusions & Recommendations

This chapter serves to summarize this thesis, as well as explore the novelty of this work, and provide a roadmap for future Raven developers wishing to expand upon the coupled GW-SW software.

6.1 The PBJ Package

The Polyline Boundary Junction (PBJ) package is a new, novel way of incorporating sub-grid detail for linear features into a groundwater model or coupled GW-SW model. In this work, its use in representing gaining streams has been demonstrated both through the PBJ examples (Chapter 4) and the coupled Raven framework examples (Chapter 5).

Generally, rivers and streams are represented and input into the groundwater model at the grid scale, requiring the groundwater model spatial discretization to account for surface water features. The PBJ package instead allows for linear features to be included at the segment scale, allowing for explicit sub-grid heterogeneity and ease moving from GIS applications (shapefiles) to a package for MODFLOW-USG or the coupled Raven framework. Additionally, the PBJ package features a correction for off-center segments to appropriately add/remove water from nearby grid nodes. The formulation can work with any grid that has a complementary triangular grid connecting the centroids of the model grid cells (e.g., Voronoi, rectilinear) but the concept could also be expanded with alternate interpolation/distribution schemes as well.

The PBJ package was designed with coupled GW-SW models in mind and does not perform any flow routing between segments. The intent is for flow routing to be handled in a dedicated surface water model like Raven, which then can pass the spatially distributed water levels to the PBJ framework. The segment-based nature of the PBJ package allows for particularly interesting workflow opportunities in GW-SW modelling, enabling the modeler to use one shapefile for model discretization and development for both domains.

6.2 Coupled Raven Framework

The primary objective of this thesis was to develop a new, novel coupled groundwater-surface water modelling software. This was accomplished by coupling the Raven hydrological modelling framework with MODFLOW-USG to form a new, coupled version of Raven that compiles into a single executable. This allows for models built separately using Raven and MODFLOW-USG to quickly and easily used together with only minor changes to the Raven files and one additional groundwater Raven file. This additional file needs only to tell Raven about the overlap between the surface water and groundwater discretization overlap and the location of river segments.

This is a novel development for both Raven and MODFLOW-USG, neither of which have been previously coupled to a groundwater or surface water model, respectively. While many coupled GW-SW modelling solutions exist, and a large number utilize MODFLOW, no software has yet been developed providing a GW-SW modelling connection using either of the "unstructured" (i.e., non-rectilinear in the horizontal plane) control volume finite-difference versions of MODFLOW: MODFLOW-USG and MODFLOW 6. A relatively new advance in groundwater modelling, these software tools allow for flexible grid discretization around surface water features, geologic features, and areas of interest (Panday et al., 2013). Long considered the de-facto standard for groundwater modelling, many model creation and visualization tools have been built specifically for MODFLOW-based models.

Additionally, Raven is an extremely unique surface water model, giving the modeler access to flexible process representation, spatial discretization, catchment and in-catchment routing, interpolation and other features that enable structural uncertainty assessment and model comparison (Craig et al., 2020). Raven is unrivaled in its flexibility, which makes it a very powerful and unique choice for a coupled GW-SW model. Many existing coupled GW-SE solutions have fixed representations of the surface domain, if not entirely region-specific.

Both Raven and MODFLOW-USG are useful modelling software for practicing hydrologists, hydrogeologists, and engineers, with calibration-friendly run times and robust features. While many GW-SW modelling software options have focused on solving complex non-linear equations that lead to long run times, reducing available time for calibration without necessarily increasing model accuracy, the coupled Raven framework aims to provide modelers with a highly flexible and fast GW-SW modelling solution built on familiar tools.

6.3 Recommendations

The software developed in this research thesis is open source and only a starting point for the coupled GW-SW Raven hydrological modelling framework. A foundational structure has been created here to be built upon by future students and practitioners who can add further capabilities, flexibilities, and processes to the software. This section explores some of the most important areas of improvement needed in the coupled Raven framework informed by both the literature review and experiments done in this work. The new GW-SW Process (see Section 3.3.4), which is an expansion upon the existing Raven hydrological process, has added the capability for Raven process to interact with the groundwater model. This is expected to be the primary basis for future GW-SW process additions. Keeping with the spirit of Raven, an important addition will be adding multiple representations of the existing (and suggested) processes connecting the models.

The PBJ package is a unique way of representing rivers in groundwater model grids, and for connection GW and SW domains. However, many practitioners will likely be interested in seeing both more traditional representations (e.g., the river and stream packages of MODFLOW) and more complex representations (e.g. Morel-Seytoux et al., 2018). Given one of the strengths of Voronoi and other non-rectilinear grids is their ability to include surface water features without excessive impacts to the grid, the PBJ package may not always be the computationally correct choice for GW-SW studies (e.g., a model where it is desirable to discretize finely around a river). Additionally, MODFLOW-USG has novel capabilities for including flow from linear features using the CLN package (Panday et al., 2013) which could use flow/water level information from Raven. A diversity of river formulations would greatly add to the utility of Raven and allow a variety of experiments that have never been done before.

Additionally, the PBJ package itself could use a variety of improvements to better realize the goals of the package. Experiments in this work showed an increased reliance on grid resolution for the PBJ package compared to the river and drain packages for MODFLOW-USG (Section 4.2). While, as discussed, this is not a large issue given that the streambed conductance is nearly always a calibration parameter, a conductance formulation (1) more resistant to scale changes, or (2) that changed intelligently with scale or was even (3) grid-independent would allow for state-of-the-art experiments with grid resolution included in the calibration (e.g. Sbai, 2020) without having to re-estimate river conductance. Possible routes for such improvements could be: incorporating Raven's existing channel representation to calculate wetted perimeter and using a leakance coefficient (e.g., Morel-Seytoux et al., 2018); adding an option to use reach transmissivity to calculate the river flux (e.g., Nemeth & Solo-Gabriele, 2003), particularly when using coarse grids; experimenting with estimating sub-grid heads in the manner used in the ghost node correction (GNC) package (Panday et al., 2013; Panday & Langevin, 2012) and/or a clever scaling algorithm (e.g. Di Ciacca et al., 2019; Kollet et al., 2007). Few, if any, studies outside of this one have considered the unique challenges, advantages, and opportunities for simulating rivers and streams within unstructured GW-SW models.

An obvious weak point in many GW-SW models is their representation of the vadose zone (see Section 2.2.2) and the formulation here leans very heavily on the soil water balance capabilities of Raven and unsaturated flow representation in MODFLOW-USG. An obvious missing process representation in the coupled Raven framework is saturation-excess overland flow from flooded groundwater nodes. Currently, this can naively be approximated using drains to move water levels above the land surface to compartments in the surface water model. However, a more realistic (or ideally, a variety of) representations should be developed that consider current state variables (e.g., water levels, soil moisture). In particular, estimating the time delay of recharge arriving at the water table may be very important in

some possible applications (Hunt et al., 2008). MODFLOW-USG does contain a Richards' equation-based representation of the unsaturated zone (Panday et al., 2013; Panday & Huyakorn, 2008) that may be compatible with Raven with minimal modifications. Overall, the existing coupled framework has Raven and MODFLOW-USG are both simulating a version of the shallow subsurface without communicating with each other. This is perhaps the part of the coupled framework most urgently in need of development. Ideally, Raven's options would range in complexity and include options to allow either the groundwater model or the surface water model to fully control the unsaturated zone. Percolation from rivers through the subsurface could also be handled using the same, if not very similar, processes.

Lakes have almost entirely been ignored in the development of the coupled modelling framework. To some extent, many of the options and tools built in this work for rivers will also be applicable to lakes. Raven can simulate lakes and percolation to the groundwater system, but again, more intelligent processes aware of the states of both modeled domains would be far superior.

Another critical upgrade to the modelling framework would be the ability to run the surface water and groundwater systems at different temporal discretizations. Currently, like GSFLOW (Markstrom et al., 2008), the coupled Raven framework is locked at a maximum of one-day time steps. As mentioned in the real-world modelling experiment, this hampered the modelling experiments in Section 5.3. While the Raven SW component itself is limited to maximum one day time steps (appropriate for SW modeling) there is no need to also constrain the GW model: the steps need not match. This is a common problem in GW-SW modeling, and several solutions are summarized below.

The following assumes that the groundwater model would be the model run at coarser (e.g., monthly) time steps, given the relative speed of water movement in the systems. One minimally invasive option is to simply move the quantity water in the groundwater store at the end of the GW time step into the GW model (time-delayed recharge and river flux). This would work well in systems where recharge and baseflow do not vary heavily month to month. The variably-coupled GW-SW software MODHMS (Panday & Huyakorn, 2008) includes this option, among others. The compatibility of this assumption with certain GW-SW processes would have to be carefully considered.

However, another intelligent option for handling temporal discretization would be a simple iteration method that takes advantage of the speed of Raven. Since Raven time steps generally take a fraction of a second to run, they are far less expensive than a MODFLOW-USG time step solution. Raven could be run for the duration of the (longer) groundwater time step, with a long-term average (or alternative estimate) of baseflow from the groundwater model, calculating the daily fluxes to the groundwater model. Then, the groundwater system would be solved and estimated fluxes to the surface water system would be passed back to Raven. Raven would then re-simulate the previous (GW) time step, assuming the same fluxes

passed to the groundwater model, but correcting its own water balance by distributing the returning GW-SW fluxes evenly over the month. This computationally frugal method would not necessarily add much accuracy to the simulation, since feedbacks to the surface water balance correction would be ignored, but it would maintain the water balance and flux timing between the models. Admittedly, this would work very poorly in regions where the rivers are net losing (e.g., arid regions). Like Raven, this method is intended for relatively wet regions and use within regional models. Running the groundwater model at coarser time steps would more than make up for any time lost by having to run Raven twice per groundwater time step. This method would require a complicated modification to the Raven solver.

One important feature that would be relatively simple to implement is support for existing rectilinear "structured" MODFLOW models (e.g., pre-MODFLOW-USG versions). MODFLOW-USG already supports the reading of structured MODFLOW input files and internally converts row, column, layer values to a sequential node number. Although it has not been tested, structured models that do not have streams may already be compatible. Those with streams may require some additional updates for the PBJ package to be compatible (rectilinear grids have triangular dual grids, so the package conceptually is compatible, but the R package may encounter issues in creating the files). There is some evidence that bilinear interpolation is more robust than the barycentric coordinates for head interpolation with rectilinear grids (Muffels et al., 2019). Given the relatively slow adoption of unstructured grids, support for structured models may increase the adoption of the Raven framework for coupled GW-SW modelling.

Raven has a unique and very useful capability of calculating simple error metrics (e.g., NSE, RMSE) for gauges. This is particularly useful when calibrating the model, either manually or using automated software. Expanding this feature to the new groundwater representation would be highly useful as well. The user could input times series of water levels and Raven could use the MODFLOW-USG BMI to query the corresponding water levels to obtain the simulated values.

Lastly, as mentioned previously, the USGS has released MODFLOW 6 (Hughes et al., 2017) with unstructured grid support and more recently has created a MODFLOW 6 basic model interface (BMI) (U.S. Geological Survey, 2020) that replicates some of the capabilities developed here for MODFLOW-USG (see Section 3.3.1). The USGS and the developers of MODFLOW-USG have indicated, going forward, MODFLOW-USG will become deprecated in favor of MODFLOW 6. Thus, at some point, it may become advantageous to port the PBJ package to MODFLOW 6 and re-couple the Raven modelling framework using the MODFLOW 6 BMI.

References

Anderson, M. P., Woessner, W. W., & Hunt, R. J. (2015). *Applied Groundwater Modeling* (Second). Academic Press / Elsevier. https://doi.org/10.1016/B978-0-08-091638-5.00001-8

Ashby, S. F., & Falgout, R. D. (1996). A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. *Nuclear Science and Engineering*, *124*(1), 145–159. https://doi.org/10.13182/NSE96-A24230

Bailey, R. T., Wible, T. C., Arabi, M., Records, R. M., & Ditty, J. (2016). Assessing regional-scale spatio-temporal patterns of groundwater–surface water interactions using a coupled SWAT-MODFLOW model. *Hydrological Processes*, *30*(23), 4420–4433. https://doi.org/10.1002/hyp.10933

Bakker, M., Post, V., Langevin, C. D., Hughes, J. D., White, J. T., Starn, J. J., & Fienen, M. N. (2016). Scripting MODFLOW Model Development Using Python and FloPy. *Groundwater*, *54*(5), 733–739. https://doi.org/10.1111/gwat.12413

Barthel, R. (2006). Common problematic aspects of coupling hydrological models with groundwater flow models on the river catchment scale. *Advances in Geosciences*, *9*, 63–71. https://doi.org/10.5194/adgeo-9-63-2006

Barthel, R. (2014a). A call for more fundamental science in regional hydrogeology. *Hydrogeology Journal*, *22*(3), 507–510. https://doi.org/10.1007/s10040-014-1101-9

Barthel, R. (2014b). HESS Opinions “Integration of groundwater and surface water research: an interdisciplinary problem ?” *Hydrology and Earth System Sciences*, *18*, 2615–2628. https://doi.org/10.5194/hess-18-2615-2014

Barthel, R., & Banzhaf, S. (2016). Groundwater and Surface Water Interaction at the Regional-scale – A Review with Focus on Regional Integrated Models. *Water Resources Management*, *30*(1), 1–32. https://doi.org/10.1007/s11269-015-1163-z

Bedekar, V., Niswonger, R. G., Kipp, K., Panday, S., & Tonkin, M. (2012). Approaches to the Simulation of Unconfined Flow and Perched Groundwater Flow in MODFLOW. *Ground Water*, *50*(2), 187–198. https://doi.org/10.1111/j.1745-6584.2011.00829.x

Beven, K. (1989). CHANGING IDEAS IN HYDROLOGY- THE CASE OF PHYSICALLY-BASED MODELS. *Journal of Hydrologic Engineering*, *105*, 157–172.

Beven, K. (2002). Towards a Coherent Philosophy for Modelling the Environment. *Proceedings: Mathematical, Physical and Engineering Sciences*, *458*(2026), 2465–2484.

Beven, K. (2005). On the concept of model structural error. *Water Science & Technology*, *52*(6), 167–175. https://iwaponline.com/wst/article-pdf/52/6/167/434056/167.pdf

Beven, K. (2016). Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, *61*(9), 1652–1665. https://doi.org/10.1080/02626667.2015.1031761

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1–2), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, *6*(3), 279–298. https://doi.org/10.1002/hyp.3360060305

Beven, K., & Germann, P. (2013). Macropores and water flow in soils revisited. *Water Resources Research*, *49*(6), 3071–3092. https://doi.org/10.1002/wrcr.20156

Bond, N. (2019). *hydrostats: Hydrologic Indices for Daily Time Series Data* (R package version 0.2.7). https://cran.r-project.org/package=hydrostats

Brooks, R. H., & Corey, A. T. (1964). Hydraulic Properties of Porous Media. *Hydrology Papers, Colorado State University*.

Brunner, P., & Simmons, C. T. (2012). HydroGeoSphere: A Fully Integrated, Physically Based Hydrological Model. *Ground Water*, *50*(2), 170–176. https://doi.org/10.1111/j.1745-6584.2011.00882.x

Brunner, P., Simmons, C. T., Cook, P. G., & Therrien, R. (2010). Modeling surface water-groundwater interaction with MODFLOW: Some considerations. *Ground Water*, *48*(2), 174–180. https://doi.org/10.1111/j.1745-6584.2009.00644.x

Brunner, P., Therrien, R., Renard, P., Simmons, C. T., & Franssen, H.-J. H. (2017). Advances in understanding river-groundwater interactions. *Reviews of Geophysics*, *55*(3), 818–854. https://doi.org/10.1002/2017RG000556

California Department of Water Resources. (2016). Modeling Best Management Practice. In *California Department of Water Resources*.

CH2M HILL, & S.S. Papadopulos & Associates Inc. (SSP&A). (2003). *Alder Creek Groundwater Study: Final Report*.

Chlumsky, R., Craig, J. R., Brown, G., Scantlebury, L., Grass, S., & Lin, S. (2020). *RavenR: Raven Hydrological Modelling Framework R Support and Analysis* (R package version 2.0.2).

https://github.com/rchlumsk/RavenR

Chow, R., Frind, M. E., Frind, E. O., Jones, J. P., Sousa, M. R., Rudolph, D. L., Molson, J. W., & Nowak, W. (2016). Delineating baseflow contribution areas for streams – A model and methods comparison. *Journal of Contaminant Hydrology*, *195*, 11–22. https://doi.org/10.1016/j.jconhyd.2016.11.001

Clark, M. P., Kavetski, D., & Fenicia, F. (2011). Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, *47*(9), 1–16. https://doi.org/10.1029/2010WR009827

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., & Rasmussen, R. M. (2015). A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, *51*(4), 2498–2514. https://doi.org/10.1002/2015WR017200.A

Clement, T. P. (2011). Complexities in Hindcasting Models-When Should We Say Enough Is Enough? *Ground Water*, *49*(5), 620–629. https://doi.org/10.1111/j.1745-6584.2010.00765.x

Conant Jr., B. (2004). Delineating and Quantifying Ground Water Discharge Zones Using Streambed Temperatures. *Ground Water*, *42*(2), 243–257. https://doi.org/10.1111/j.1745-6584.2004.tb02671.x

Condon, L. E., Markovich, K. H., Kelleher, C. A., McDonnell, J. J., Ferguson, G., & McIntosh, J. C. (2020). Where Is the Bottom of a Watershed? *Water Resources Research*, *56*(3). https://doi.org/10.1029/2019WR026010

Corwin, D. L., Hopmans, J., & de Rooij, G. H. (2006). From Field- to Landscape-Scale Vadose Zone Processes: Scale Issues, Modeling, and Monitoring. *Vadose Zone Journal*, *5*(1), 129–139. https://doi.org/10.2136/vzj2006.0004

Craig, J. R., Brown, G., Chlumsky, R., Jenkinson, R. W., Jost, G., Lee, K., Mai, J., Serrer, M., Sgro, N., Shafii, M., Snowdon, A. P., & Tolson, B. A. (2020). Flexible watershed simulation with the Raven hydrological modelling framework. *Environmental Modelling and Software*, *129*(December 2019), 104728. https://doi.org/10.1016/j.envsoft.2020.104728

Craig, J. R., Liu, G., & Soulis, E. D. (2010). Runoff-infiltration partitioning using an upscaled Green-Ampt solution. *Hydrological Processes*, *24*(16), 2328–2334. https://doi.org/10.1002/hyp.7601

Craig, J. R., & the Raven Development Team. (2020). *Raven user's and developer's manual* (3.0.0). http://raven.uwaterloo.ca

de Marsily, G., Combes, P., & Goblet, P. (1992). Comment on ' Ground-water models cannot be validated', by L.F. Konikow & J.D. Bredehoeft. *Advances in Water Resources*, *15*, 367–369.

de Marsily, G., Delay, F., Gonçalvès, J., Renard, P., Teles, V., & Violette, S. (2005). Dealing with spatial heterogeneity. In *Hydrogeology Journal* (Vol. 13, Issue 1, pp. 161–183). Springer. https://doi.org/10.1007/s10040-004-0432-3

Di Ciacca, A., Leterme, B., Laloy, E., Jacques, D., & Vanderborght, J. (2019). Scale-dependent parameterization of groundwater–surface water interactions in a regional hydrogeological model. *Journal of Hydrology*, *576*(June), 494–507. https://doi.org/10.1016/j.jhydrol.2019.06.072

Dogrul, E. C., Kadir, T. N., & Brush, C. F. (2018). *Integrated Water Flow Model Theoretical Documentation*. https://water.ca.gov/-/media/DWR-Website/Web-Pages/Library/Modeling-And-Analysis/IWFM/Files/IWFM_2015_0_706/IWFM20150706TheoreticalDocumentation.pdf?la=en&hash=5BA427C5A46F80586317EEFCBA179E88DF9DE9A3

Doherty, J. E. (2003). Ground water model calibration using pilot points and regularization. *Ground Water*, *41*(2), 170–177. https://doi.org/10.1111/j.1745-6584.2003.tb02580.x

Doherty, J. E. (2015). *Calibration and Uncertainty Analysis for Complex Environmental Models*. Watermark Numerical Computing. https://doi.org/10.1111/gwat.12360

Doherty, J. E. (2016). *PEST User Manual* (6th ed.). Watermark Numerical Computing. www.pesthomepage.org

Doherty, J. E., Fienen, M. N., & Hunt, R. J. (2010). *Approaches to Highly Parameterized Inversion: Pilot-Point Theory, Guidelines, and Research Directions Scientific Investigations Report 2010-5168*.

Doherty, J. E., & Moore, C. (2019). Decision Support Modeling: Data Assimilation, Uncertainty Quantification, and Strategic Abstraction. *Groundwater*. https://doi.org/10.1111/gwat.12969

Doherty, J. E., & Welter, D. (2010). A short exploration of structural noise. *Water Resources Research*, *46*(5). https://doi.org/10.1029/2009WR008377

Ebel, B. A., Mirus, B. B., Heppner, C. S., VanderKwaak, J. E., & Loague, K. (2009). First-order exchange coefficient coupling for simulating surface water-groundwater interactions: parameter sensitivity and consistency with a physics-based approach. *Hydrological Processes*, *23*(13), 1949–1959. https://doi.org/10.1002/hyp.7279

Farthing, M. W., & Ogden, F. L. (2017). Numerical Solution of Richards' Equation: A Review of Advances and Challenges. *Soil Science Society of America Journal*, *81*(6), 1257–1269. https://doi.org/10.2136/sssaj2017.02.0058

Fenicia, F., Kavetski, D., & Savenije, H. H. G. (2011). Elements of a flexible approach for conceptual

hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, *47*(11), 1–13. https://doi.org/10.1029/2010WR010174

Fenske, J., Banta, E. R., Piper, S., & Donchyts Gennadii. (2009). Coupling HEC-RAS and MODFLOW using OpenMI. *Advances in Hydrologic Engineering*, *May 2014*, 8. http://www.hec.usace.army.mil/newsletters/HEC_Newsletter_Spring2013.pdf

Ferré, T. P. A. (2017). Revisiting the Relationship Between Data, Models, and Decision-Making. *Groundwater*, *55*(5), 604–614. https://doi.org/10.1111/gwat.12574

Fortin, V., & Turcotte, R. (2006). Le modèle hydrologique MOHYSE. *Note de Cours Pour SCA7420, Département Des Sciences de La Terre et de l'atmosphere, Université Du Québec à Montréal*.

Fowler, K., Knoben, W., Peel, M., Peterson, T., Ryu, D., Saft, M., Seo, K. W., & Western, A. (2020). Many Commonly Used Rainfall-Runoff Models Lack Long, Slow Dynamics: Implications for Runoff Projections. *Water Resources Research*, *56*(5), 0–3. https://doi.org/10.1029/2019WR025286

Freeze, R. A., & Cherry, J. A. (1979). *Groundwater*. Prentice-Hall.

Frind, E. O., Molson, J. W., Sousa, M. R., & Martin, P. J. (2014). Insights from four decades of model development on the Waterloo Moraine: A review. *Canadian Water Resources Journal*, *39*(2), 149–166. https://doi.org/10.1080/07011784.2014.914799

Furman, A. (2008). Modeling Coupled Surface-Subsurface Flow Processes: A Review. *Vadose Zone Journal*, *7*(2), 741–756. https://doi.org/10.2136/vzj2007.0065

Gillham, R. W. (1984). The capillary fringe and its effect on water-table response. *Journal of Hydrology*, *67*(1–4), 307–324. https://doi.org/10.1016/0022-1694(84)90248-8

Gómez-Hernández, J. J. (2006). Complexity. *Ground Water*, *44*(6), 782–785. https://doi.org/10.1111/j.1745-6584.2006.00222.x

Green, T. R., Taniguchi, M., Kooi, H., Gurdak, J. J., Allen, D. M., Hiscock, K. M., Treidel, H., & Aureli, A. (2011). Beneath the surface of global change: Impacts of climate change on groundwater. *Journal of Hydrology*, *405*(3–4), 532–560. https://doi.org/10.1016/j.jhydrol.2011.05.002

Green, W. H., & Ampt, G. (1911). Studies of soil physics, part I – the flow of air and water through soils. *Journal of Agricultual Science*.

Guthke, A. (2017). Defensible Model Complexity: A Call for Data-Based and Goal-Oriented Model Choice. *Groundwater*, *55*(5), 646–650. https://doi.org/10.1111/gwat.12554

Hämäläinen, R. P. (2015). Behavioural issues in environmental modelling – The missing perspective.

Environmental Modelling & Software, *73*, 244–253. https://doi.org/10.1016/j.envsoft.2015.08.019

Hamilton, A. S., Hutchinson, D. G., & Moore, R. D. (2000). Estimating Winter Streamflow Using Conceptual Streamflow Model. *Journal of Cold Regions Engineering*, *14*(4), 158–175. https://doi.org/10.1061/(ASCE)0887-381X(2000)14:4(158)

Han, M., Mai, J., Tolson, B. A., Craig, J. R., Gaborit, É., Liu, H., & Lee, K. (2020). Subwatershed-based lake and river routing products for hydrologic and land surface models applied over Canada. *Canadian Water Resources Journal*, *45*(3), 237–251. https://doi.org/10.1080/07011784.2020.1772116

Harbaugh, A. W. (2005). *MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model—the Ground-Water Flow Process*. U.S. Geological Survey Techniques and Methods 6-A16.

Harbaugh, A. W., Banta, E. R., Hill, M. C., & Mcdonald, M. G. (2000). MODFLOW-2000 , The U .S . Geological Survey modular graound-water model — User guide to modularization concepts and the ground-water flow process. *U.S. Geological Survey*.

Harter, T., & Hopmans, J. W. (2004). Role of vadose-zone flow processes in regional-scale hydrology: review, opportunities and challenges. *Unsaturated-Zone Modeling: Progress, Challenges and Applications*, *6*, 179–208.

Harter, T., & Morel-Seytoux, H. (2013). Peer Review of the IWFM, MODFLOW and HGS Model Codes: Potential for Water Management Applications in California's Central Valley and Other Irrigated. *Excellence and Consensus in Water and …*, 121.

Hill, M. C. (2006). The practical use of simplicity in developing ground water models. *Ground Water*, *44*(6), 775–781. https://doi.org/10.1111/j.1745-6584.2006.00227.x

Hill, M. C., Kavetski, D., Clark, M., Ye, M., Arabi, M., Lu, D., Foglia, L., & Mehl, S. (2016). Practical Use of Computationally Frugal Model Analysis Methods. *Groundwater*, *54*(2), 159–170. https://doi.org/10.1111/gwat.12330

Holman, I. P., Allen, D. M., Cuthbert, M. O., & Goderniaux, P. (2012). Towards best practice for assessing the impacts of climate change on groundwater. *Hydrogeology Journal*, *20*(1), 1–4. https://doi.org/10.1007/s10040-011-0805-3

Hughes, J. D., Langevin, C. D., & Banta, E. R. (2017). Documentation for the MODFLOW 6 Framework. *U.S. Geological Survey: Techniques and Methods 6-A57*.

Hughes, J. D., Langevin, C. D., & White, J. T. (2015). MODFLOW-Based Coupled Surface Water Routing and Groundwater-Flow Simulation. *Groundwater*, *53*(3), 452–463.

https://doi.org/10.1111/gwat.12216

Hughes, J. D., & Liu, J. (2008). MIKE SHE: Software for integrated surface water/ground water modeling. *Ground Water*, *46*(6), 797–802. https://doi.org/10.1111/j.1745-6584.2008.00500.x

Hunt, R. J., Doherty, J., & Tonkin, M. J. (2007). Are models too simple? Arguments for increased parameterization. *Ground Water*, *45*(3), 254–262. https://doi.org/10.1111/j.1745-6584.2007.00316.x

Hunt, R. J., Fienen, M. N., & White, J. T. (2020). Revisiting "An Exercise in Groundwater Model Calibration and Prediction" After 30 Years: Insights and New Directions. *Groundwater*, *58*(2), 168–182. https://doi.org/10.1111/gwat.12907

Hunt, R. J., Prudic, D. E., Walker, J. F., & Anderson, M. P. (2008). Importance of unsaturated zone flow for simulating recharge in a humid climate. *Ground Water*, *46*(4), 551–560. https://doi.org/10.1111/j.1745-6584.2007.00427.x

Hunt, R. J., & Zheng, C. (1999). Debating complexity in modeling. *Eos*, *80*(3), 29. https://doi.org/10.1029/99EO00025

Hutton, E., Piper, M., & Tucker, G. (2020). The Basic Model Interface 2.0: A standard interface for coupling numerical models in the geosciences. *Journal of Open Source Software*, *5*(51), 2317. https://doi.org/10.21105/joss.02317

Hwang, H., Park, Y., Sudicky, E. A., & Forsyth, P. A. (2014). A parallel computational framework to solve flow and transport in integrated surface-subsurface hydrologic systems. *Environmental Modelling and Software*, *61*, 39–58. https://doi.org/10.1016/j.envsoft.2014.06.024

HydroAlgorithmics. (2016). *AlgoMesh User Guide*. http://www.hydroalgorithmics.com/software/algomesh

Jasechko, S. (2019). Global Isotope Hydrogeology—Review. *Reviews of Geophysics*, *57*(3), 835–965. https://doi.org/10.1029/2018RG000627

Jasechko, S., Perrone, D., Befus, K. M., Bayani Cardenas, M., Ferguson, G., Gleeson, T., Luijendijk, E., McDonnell, J. J., Taylor, R. G., Wada, Y., & Kirchner, J. W. (2017). Global aquifers dominated by fossil groundwaters but wells vulnerable to modern contamination. *Nature Geoscience*, *10*(6), 425–429. https://doi.org/10.1038/ngeo2943

Jazayeri, A., & Werner, A. D. (2019). Boundary Condition Nomenclature Confusion in Groundwater Flow Modeling. *Groundwater*, *57*(5), 664–668. https://doi.org/10.1111/gwat.12893

Jones, J. P., Sousa, M. R., Frind, E. O., & Rudolph, D. L. (2009). Determining the influence of surface,

unsaturated and saturated processes on source water protection strategies: a multi-model study. *GeoHalifax Conference*.

Jyrkama, M. I., & Sykes, J. F. (2006). Sensitivity and uncertainty analysis of the recharge boundary condition. *Water Resources Research*, *42*(1), 1404. https://doi.org/10.1029/2005WR004408

Kavetski, D., Kuczera, G., & Franks, S. W. (2006). Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts. *Journal of Hydrology*, *320*(1–2), 173–186. https://doi.org/10.1016/j.jhydrol.2005.07.012

Kessler, R. E., & Schwarzmeier, J. L. (1993). CRAY T3D: a new dimension for Cray research. *1993 IEEE Compcon Spring*, 176–182. https://doi.org/10.1109/cmpcon.1993.289660

Kim, N. W., Chung, I. M., Won, Y. S., & Arnold, J. G. (2008). Development and application of the integrated SWAT–MODFLOW model. *Journal of Hydrology*, *356*(1–2), 1–16. https://doi.org/10.1016/J.JHYDROL.2008.02.024

Kirchner, J. W. (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), 1–5. https://doi.org/10.1029/2005WR004362

Kitanidis, P. K. (2015). Persistent questions of heterogeneity, uncertainty, and scale in subsurface flow and transport. *Water Resources Research*, *51*(8), 5888–5904. https://doi.org/10.1002/2015WR017639

Kollet, S. J., & Maxwell, R. M. (2006). Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, *29*(7), 945–958. https://doi.org/10.1016/j.advwatres.2005.08.006

Kollet, S. J., Sulis, M., Maxwell, R. M., Paniconi, C., Putti, M., Bertoldi, G., Coon, E. T., Cordano, E., Endrizzi, S., Kikinzon, E., Mouche, E., Mügler, C., Park, Y.-J., Refsgaard, J. C., Stisen, S., & Sudicky, E. (2017). The integrated hydrologic model intercomparison project, IH-MIP2: A second set of benchmark results to diagnose integrated hydrology and feedbacks. *Water Resources Research*, *53*(1), 867–890. https://doi.org/10.1002/2016WR019191

Kollet, S. J., Zlotnik, V. A., & Kollet, S. J. (2007). Evaluation of the streambed leakage concept in analytical models using data from three pumping tests. *Papers in the Earth and Atmospheric Sciences Earth and Atmospheric Sciences*. https://doi.org/10.1007/s10040-006-0156-7

Kuffour, B. N., Engdahl, N., Woodward, C., Condon, L., Kollet, S. J., & Maxwell, R. (2019). Simulating Coupled Surface-Subsurface Flows with ParFlow v3.5.0: Capabilities, applications, and ongoing

developmentof an open-source, massively parallel, integrated hydrologic model. *Geoscientific Model Development Discussions*, 1–66. https://doi.org/10.5194/gmd-2019-190

Kumar, M., Duffy, C. J., & Salvage, K. M. (2009). A Second-Order Accurate, Finite Volume-Based, Integrated Hydrologic Modeling (FIHM) Framework for Simulation of Surface and Subsurface Flow. *Vadose Zone Journal*, *8*(4), 873–890. https://doi.org/10.2136/vzj2009.0014

Ladson, A. R., Brown, R., Neal, B., & Nathan, R. (2013). A Standard Approach to Baseflow Separation Using The Lyne and Hollick Filter. *Australasian Journal of Water Resources*, *17*(1), 25–34. https://doi.org/10.7158/13241583.2013.11465417

Lighthill, M. J., & Whitham, G. B. (1955). On kinematic waves I. Flood movement in long rivers. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *229*(1178), 281–316. https://doi.org/10.1098/rspa.1955.0088

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, *201*(1–4), 272–288. https://doi.org/10.1016/S0022-1694(97)00041-3

Linkov, I., & Burmistrov, D. (2003). Model Uncertainty and Choices Made by Modelers: Lessons Learned from the International Atomic Energy Agency Model Intercomparisons. *Risk Analysis*, *23*(6), 1297–1308. https://doi.org/10.1111/j.0272-4332.2003.00402.x

Markstrom, S. L., Hay, L. E., & Clark, M. P. (2016). Towards simplification of hydrologic modeling: Identification of dominant processes. *Hydrology and Earth System Sciences*, *20*(11), 4655–4671. https://doi.org/10.5194/hess-20-4655-2016

Markstrom, S. L., Niswonger, R. G., Regan, R. S., Prudic, D. E., & Barlow, P. M. (2008). GSFLOW—Coupled Ground-Water and Surface-Water Flow Model Based on the Integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005). In *U.S. Geological Survey*.

Martel, J.-L., Demeester, K., Brissette, F. O., Poulin, A., & Arsenault, R. (2017). HMETS-A Simple and Efficient Hydrology Model for Teaching Hydrological Modelling, Flow Forecasting and Climate Change Impacts. *International Journal of Engineering Education*, *33*(4), 1307–1316. https://www.ijee.ie/latestissues/Vol33-4/16_ijee3473ns.pdf

Martin, P. J., & Frind, E. O. (1998). Modeling a Complex Multi-Aquifer System: The Waterloo Moraine. *Ground Water*, *36*(4), 679–690. https://doi.org/10.1111/j.1745-6584.1998.tb02843.x

Matott, L. S. (2017). *OSTRICH: an Optimization Software Tool, Documentation and User's Guide*

(Version 17). University at Buffalo Center for Computational Research. www.eng.buffalo.edu/~lsmatott/Ostrich/OstrichMain.html

Matott, L. S., Babendreier, J. E., & Purucker, S. T. (2009). Evaluating uncertainty in integrated environmental models: A review of concepts and tools. *Water Resources Research*, *45*(6). https://doi.org/10.1029/2008WR007301

Matrix Solutions Inc, & S.S. Papadopulos & Associates Inc. (SSP&A). (2014). *Region of Waterloo Tier Three Water Budget and Local Area Risk Assessment*. https://www.sourcewater.ca/en/source-protection-areas/Grand-River-Background-reports.aspx

Maxwell, R. M., Putti, M., Meyerhoff, S., Delfs, J.-O., Ferguson, I. M., Ivanov, V., Kim, J., Kolditz, O., Kollet, S. J., Kumar, M., Lopez, S., Niu, J., Paniconi, C., Park, Y.-J., Phanikumar, M. S., Shen, C., Sudicky, E., & Sulis, M. (2014). Surface-subsurface model intercomparison: A first set of benchmark results to diagnose integrated hydrology and feedbacks. *Water Resources Research*, *50*, 1531–1549. https://doi.org/10.1002/2013WR013725.Received

Mehl, S., & Hill, M. C. (2010). Grid-size dependence of Cauchy boundary conditions used to simulate stream–aquifer interactions. *Advances in Water Resources*, *33*(4), 430–442. https://doi.org/10.1016/J.ADVWATRES.2010.01.008

Ministry of Natural Resources and Forestry (MNRF). (2019). Southern Ontario Land Resource Information System (SOLRIS) Version 3.0. In *Forest Resources Inventory Program*. Ontario GeoHub (Accessed Oct 1, 2020). https://geohub.lio.gov.on.ca/

Morel-Seytoux, H. J. (2020). Seepage and Recharge under a Stream-aquifer Unsaturated Connection. *Physical Science International Journal*, *24*(3), 20–42. https://doi.org/10.9734/PSIJ/2020/v24i330181

Morel-Seytoux, H. J., Mehl, S., & Morgado, K. (2014). Factors influencing the stream-aquifer flow exchange coefficient. *Groundwater*, *52*(5), 775–781. https://doi.org/10.1111/gwat.12112

Morel-Seytoux, H. J., Miller, C. D., Mehl, S., & Miracapillo, C. (2018). Achilles' heel of integrated hydrologic models: The stream-aquifer flow exchange, and proposed alternative. *Journal of Hydrology*, *564*, 900–908. https://doi.org/10.1016/J.JHYDROL.2018.07.010

Morel-Seytoux, H. J., Miller, C. D., Miracapillo, C., & Mehl, S. (2017). River Seepage Conductance in Large-Scale Regional Studies. *Groundwater*, *55*(3), 399–407. https://doi.org/10.1111/gwat.12491

Morita, M., & Yen, B. C. (2000). Numerical methods for conjunctive two-dimensional surface and three-dimensional sub-surface flows. *International Journal for Numerical Methods in Fluids*, *32*(8), 921–

957. https://doi.org/10.1002/(SICI)1097-0363(20000430)32:8<921::AID-FLD993>3.0.CO;2-3

Muffels, C., Bedekar, V. S., & Kulbersh, M. (2019). Designing the Ghost Node Correction Package of MODFLOW-USG to Mitigate Local Flow Oscillations [Conference Poster]. *MODFLOW & More 2019*.

Nemeth, M. S., & Solo-Gabriele, H. M. (2003). Evaluation of the use of reach transmissivity to quantify exchange between groundwater and surface water. *Journal of Hydrology*, *274*(1–4), 145–159. https://doi.org/10.1016/S0022-1694(02)00419-5

Niswonger, R. G., Panday, S., & Ibaraki, M. (2011). MODFLOW-NWT, A Newton Formulation for MODFLOW-2005. In *Techniques and Methods 6-A37* (p. 44). U.S. Geological Survey.

Niswonger, R. G., Prudic, D. E., & Regan, R. S. (2006). Documentation of the Unsaturated-Zone Flow (UZF1) Package for Modeling Unsaturated Flow Between the Land Surface and the Water Table with MODFLOW-2005. *U.S. Geological Survey Techniques and Methods*, *6-A19*.

Ontario Geological Survey. (2010). *Surficial geology of Southern Ontario*.

Ontario Ministry of Natural Resources and Forestry (MNRF). (2019). *Ontario Integrated Hydrology Data*. Ontario GeoHub (Accessed Oct 1, 2020). https://geohub.lio.gov.on.ca/

Ontario Ministry of the Environment Conservation and Parks. (n.d.). *Provincial Groundwater Monitoring Network (PGMN)*. https://www.ontario.ca/environment-and-energy/map-provincial-groundwater-monitoring-network

Panday, S., & Huyakorn, P. S. (2004). A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Advances in Water Resources*, *27*(4), 361–382. https://doi.org/10.1016/j.advwatres.2004.02.016

Panday, S., & Huyakorn, P. S. (2008). MODFLOW SURFACT: A State-of-the-Art Use of Vadose Zone Flow and Transport Equations and Numerical Techniques for Environmental Evaluations. *Vadose Zone Journal*, *7*(2), 610–631. https://doi.org/10.2136/vzj2007.0052

Panday, S., & Langevin, C. D. (2012). Improving sub-grid scale accuracy of boundary features in regional finite-difference models. *Advances in Water Resources*, *41*, 65–75. https://doi.org/10.1016/j.advwatres.2012.02.011

Panday, S., Langevin, C. D., Niswonger, R. G., Ibaraki, M., & Hughes, J. D. (2013). MODFLOW – USG Version 1: An Unstructured Grid Version of MODFLOW for Simulating Groundwater Flow and Tightly Coupled Processes Using a Control Volume Finite-Difference Formulation. *U.S. Geological Survey*.

Park, Y.-J., Sudicky, E. A., Panday, S., & Matanga, G. (2009). Implicit Subtime Stepping for Solving Nonlinear Flow Equations in an Integrated Surface-Subsurface System. *Vadose Zone Journal*, *8*(4), 825–836. https://doi.org/10.2136/vzj2009.0013

Pebesma, E. J. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446. https://doi.org/https://doi.org/10.32614/RJ-2018-009

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

Poncea, V. M., & Shetty, A. V. (1995). A conceptual model of catchment water balance: 1. Formulation and calibration. *Journal of Hydrology*, *173*(1–4), 27–40. https://doi.org/10.1016/0022-1694(95)02739-C

Qu, Y., & Duffy, C. J. (2007). A semidiscrete finite volume formulation for multiprocess watershed simulation. *Water Resources Research*, *43*(8), 1–18. https://doi.org/10.1029/2006WR005752

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rawls, W. J., Brakensiek, D. L., & Miller, N. (1983). Green-ampt Infiltration Parameters from Soils Data. *Journal of Hydraulic Engineering*, *109*(1), 62–70. https://doi.org/10.1061/(asce)0733-9429(1983)109:1(62)

Refsgaard, J. C., & Storm, B. (1995). MIKE SHE. In *Computer Models of Watershed Hydrology*.

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process - A framework and guidance. *Environmental Modelling and Software*, *22*(11), 1543–1556. https://doi.org/10.1016/j.envsoft.2007.02.004

Reilly, T. E. (2001). System and Boundary Conceptualization in Ground-Water Flow Simulation. In *Techniques of Water-Resources Investigations of the United States Geological Survey* (Vol. 3).

Ritter, A., & Muñoz-Carpena, R. (2013). Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology*, *480*, 33–45. https://doi.org/10.1016/j.jhydrol.2012.12.004

Rohatgi, A. (2020). *WebPlotDigitizer* (4.4). https://automeris.io/WebPlotDigitizer

Sbai, M. A. (2020). Unstructured Gridding for MODFLOW from Prior Groundwater Flow Models: A New Paradigm. *Groundwater*. https://doi.org/10.1111/gwat.13025

Scantlebury, L. (2020). *pbjR: Modflow-USG Polyline Boundary Junction (PBJ) R Package* (0.1.0). https://github.com/scantle/*pbjR*

Scantlebury, L., & Craig, J. R. (2020). The Polyline Boundary Junction Package: A Novel Method for Incorporating Linear Features in Voronoi Grids. *AGU 2020 Fall Meeting Poster*.

Seibert, J. (2003). Reliability of Model Predictions Outside Calibration Conditions. In *Nordic Hydrology* (Vol. 34, Issue 5).

Shanafield, M., Bourke, S. A., Zimmer, M. A., & Costigan, K. H. (2021). An overview of the hydrology of non-perennial rivers and streams. *Wiley Interdisciplinary Reviews: Water*, *8*(2), 1–25. https://doi.org/10.1002/wat2.1504

Shu, L., Ullrich, P. A., & Duffy, C. J. (2020). Simulator for Hydrologic Unstructured Domains (SHUD v1.0): numerical modeling of watershed hydrology with the finite volume method. *Geoscientific Model Development*, *13*(6), 2743–2762. https://doi.org/10.5194/gmd-13-2743-2020

Simmons, C. T., & Hunt, R. J. (2012). Updating the debate on model complexity. *GSA Today*, *22*(8), 28–29. https://doi.org/10.1130/GSATG150GW.1

Singh, V. P., & Woolhiser, D. A. (2002). Mathematical Modeling of Watershed Hydrology. *Journal of Hydrologic Engineering*, *7*(4), 270–292. https://doi.org/10.1061/(ASCE)1084-0699(2002)7:4(270)

Sitterson, J., Knightes, C., Parmar, R., Wolfe, K., Muche, M., & Avant, B. (2017). *An Overview of Rainfall-Runoff Model Types An Overview of Rainfall-Runoff Model Types*. *September*, 0–29.

Snowdon, A. P. (2015). *Upscaling of Coupled Models with Topography-Driven Surface-Water/Groundwater Interactions* [University of Waterloo]. https://uwspace.uwaterloo.ca/bitstream/handle/10012/10489/Snowdon_Andrew.pdf?sequence=1&isAllowed=y

Snowdon, A. P., & Craig, J. R. (2016). Effective groundwater-surface water exchange at watershed scales. *Hydrological Processes*, *30*(12), 1849–1861. https://doi.org/10.1002/hyp.10759

Sophocleous, M. (2002). Interactions between groundwater and surface water: the state of the science. *Hydrogeology Journal*, *10*(1), 52–67. https://doi.org/10.1007/s10040-001-0170-8

Sophocleous, M., & Perkins, S. P. (2000). Methodology and application of combined watershed and ground-water models in Kansas. *Journal of Hydrology*, *236*, 185–201. https://doi.org/10.1016/S0022-1694(00)00293-6

Sridhar, V., Billah, M. M., & Hildreth, J. W. (2018). Coupled Surface and Groundwater Hydrological

Modeling in a Changing Climate. *Groundwater*, *56*(4), 618–635. https://doi.org/10.1111/gwat.12610

Staudinger, M., Stoelzle, M., Cochand, F., Seibert, J., Weiler, M., & Hunkeler, D. (2019). Your work is my boundary condition!: Challenges and approaches for a closer collaboration between hydrologists and hydrogeologists. *Journal of Hydrology*, *571*(February), 235–243. https://doi.org/10.1016/j.jhydrol.2019.01.058

Sulis, M., Meyerhoff, S. B., Paniconi, C., Maxwell, R. M., Putti, M., & Kollet, S. J. (2010). A comparison of two physics-based numerical models for simulating surface water-groundwater interactions. *Advances in Water Resources*, *33*(4), 456–467. https://doi.org/10.1016/j.advwatres.2010.01.010

Swain, E. D., & Wexler, E. J. (1996). A coupled surface-water and ground-water flow model (MODBRANCH) for simulation of stream-aquifer interaction. In *U.S. Geological Survey Techniques of Water-Resources Investigations, Book 6 Modeling Techniques*. https://pubs.usgs.gov/twri/twri6a6/

Tolley, D., Foglia, L., & Harter, T. (2019). Sensitivity Analysis and Calibration of an Integrated Hydrologic Model in an Irrigated Agricultural Basin With a Groundwater-Dependent Ecosystem. *Water Resources Research*. https://doi.org/10.1029/2018wr024209

Tolson, B. A., Sharma, V., & Swayne, D. A. (2014). Parallel Implementations of the Dynamically Dimensioned Search (DDS) Algorithm. *Environmental Software Systems*, *7*(Second Edition), 589–598. https://doi.org/10.13140/2.1.3773.6001

Tonkin, M., Hill, M., Maxwell, R. M., & Zheng, C. (2020). Groundwater Modeling and Beyond: MODFLOW-and-More-2019 Special Issue. *Groundwater*, *58*(3), 325–326. https://doi.org/10.1111/gwat.12999

U.S. Geological Survey. (2020). *MODFLOW 6 Release Notes: Version mf6.1.1.* https://water.usgs.gov/water-resources/software/MODFLOW-6/release_6.1.1.pdf

VanderKwaak, J. E. (1999). Numerical simulation of flow and chemical transport in integrated surface-subsurface hydrologic systems. *Department of Earth Sciences*, 218. https://uwspace.uwaterloo.ca/handle/10012/412

Vereecken, H., Weihermüller, L., Assouline, S., Šimůnek, J., Verhoef, A., Herbst, M., Archer, N., Mohanty, B., Montzka, C., Vanderborght, J., Balsamo, G., Bechtold, M., Boone, A., Chadburn, S., Cuntz, M., Decharme, B., Ducharne, A., Ek, M., Garrigues, S., … Xue, Y. (2019). Infiltration from the Pedon to Global Grid Scales: An Overview and Outlook for Land Surface Modeling. *Vadose Zone Journal*, *18*(1), 1–53. https://doi.org/10.2136/vzj2018.10.0191

Vermeulen, P. T. M., te Stroet, C. B. M., & Heemink, A. W. (2006). Limitations to upscaling of groundwater flow models dominated by surface water interaction. *Water Resources Research*, *42*(10), 1–12. https://doi.org/10.1029/2005WR004620

Wang, H. F., & Anderson, M. P. (1982). *Introduction to Groundwater Modeling*. Academic Press Inc.

Weijs, S. V., & Ruddell, B. L. (2020). Debates: Does Information Theory Provide a New Paradigm for Earth Science? Sharper Predictions Using Occam's Digital Razor. In *Water Resources Research* (Vol. 56, Issue 2). Blackwell Publishing Ltd. https://doi.org/10.1029/2019WR026471

Wen, X.-H., & Gómez-Hernández, J. J. (1996). Upscaling hydraulic conductivities in heterogeneous media: An overview. *Journal of Hydrology*, *183*(1–2), ix–xxxii. https://doi.org/10.1016/S0022-1694(96)80030-8

Wiebe, A. J., & Rudolph, D. L. (2020). On the sensitivity of modelled groundwater recharge estimates to rain gauge network scale. *Journal of Hydrology*, *585*(February), 124741. https://doi.org/10.1016/j.jhydrol.2020.124741

Winter, T. C., Harvey, J. W., Franke, O. L., & Alley, W. M. (1998). Ground Water and Surface Water: A Single Resource. In *USGS Publications*.

Woessner, W. W. (2000). Stream and fluvial plain ground water interactions: Rescaling hydrogeologic thought. *Ground Water*, *38*(3), 423–429. https://doi.org/10.1111/j.1745-6584.2000.tb00228.x

Wood, E. F., Lettenmaier, D. P., & Zartarian, V. G. (1992). A land-surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research*. https://doi.org/10.1029/91JD01786

Xiang, K., Li, Y., Horton, R., & Feng, H. (2020). Similarity and difference of potential evapotranspiration and reference crop evapotranspiration – a review. In *Agricultural Water Management* (Vol. 232, p. 106043). Elsevier B.V. https://doi.org/10.1016/j.agwat.2020.106043

Younger, P. L. (1993). Simple generalized methods for estimating aquifer storage parameters. In *Quarterly Journal of Engineering Geology* (Vol. 26).

Zhou, Y., & Li, W. (2011). A review of regional groundwater flow modeling. *Geoscience Frontiers*, *2*(2), 205–214. https://doi.org/10.1016/j.gsf.2011.03.003