DISSERTATION

EVALUATING SPATIAL AND TEMPORAL CONTROLS ON RECHARGE FLUXES IN A STREAM-ALLUVIAL-BEDROCK AQUIFER SYSTEM

Submitted by

Kristen Cognac

Department of Geosciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2023

Doctoral Committee

Advisor: Michael Ronayne

Ryan Bailey
Sara Rathburn
Lisa Stright

ABSTRACT

EVALUATING SPATIAL AND TEMPORAL CONTROLS ON RECHARGE FLUXES IN A STREAM-ALLUVIAL-BEDROCK AQUIFER SYSTEM

The dynamics and timescales associated with natural and induced recharge to aquifers dictate whether and for how long groundwater resources are sustainable. This dissertation contains three studies which apply groundwater flow and geostatistical modeling to evaluate spatial and temporal controls on recharge fluxes in a stream-alluvial-bedrock system. Each study is based on a recharge mechanism that occurs within the Denver Basin aquifer system, a regionally significant water supply for which long-term pumping and active aquifer depletion call for improved characterizations of recharge. While recharge is the theme of this dissertation, I don't attempt to directly estimate recharge for the Denver Basin, but rather to investigate and expose dynamics of recharge that are essential for accurate conceptualizations and estimates of recharge.

The first study investigates controls and timescales associated with streambed fluxes which are an important component of seepage recharge along mountain-front streams. Streambed fluxes are highly variable through time and space, having a range of implications for stream-aquifer processes. While spatial variations in streambed flux have been heavily characterized, temporal variability has been limited to short-term or low-frequency measurements. This study calculates high-frequency time series of Darcy-based streambed fluxes over a three-year period using water level and temperature inputs from shallow (<1.5m) nested streambed piezometers installed in two mountain-front streams in Colorado, USA.

Results reveal important conclusions about controls and patterns of temporal variability. Three predominant temporal scales of variability, sub-daily (<1day), daily (>1d; <1y), and interannual (>1y), are quantified through statistical measures. Sub-daily variability was related to ET, temperature-induced changes in hydraulic conductivity, and variable stream stage while daily variability was highly seasonal and related to specific events on the channel (e.g., beaver dams). The magnitude of sub-daily variability was significant compared to daily variability (ratio 0.03 to 0.7). Annual median fluxes at each site varied across years, but typically remained consistent in order of magnitude and direction. A strong linear correlation characterizes the relationship between the daily variability and the median annual flux at individual sites, highlighting how sites with greater fluxes also exhibit greater temporal variability. The temporal flux variations documented in this study have important implications for calculations and interpretations of hyporheic exchange and groundwater recharge. Results provide a basis for quantifying temporal variations in streambed fluxes and highlight the extent to which fluxes vary over multiple timescales.

Chapters 3 and 4 are organized to progress vertically downward within the system to investigate controls for inter-aquifer exchange between the alluvial and bedrock aquifer, an important component of recharge to the underlying bedrock aquifer system. In Chapter 3, the potential for and controls of hydraulic disconnection between the alluvial and bedrock aquifer are investigated. Hydraulic disconnection occurs when unsaturated conditions develop between a stream and water table causing seepage rates to stabilize with additional water table drawdown. In this study, I demonstrate that hydraulic disconnection can occur between an alluvial and bedrock aquifer when unsaturated conditions develop between the two water tables and inter-aquifer flow rates stabilize with subsequent drawdown. Variably saturated flow modeling is performed to simulate the effects of drawdown on alluvium to bedrock flow rates (A-B flow). Bedrock aquifer heterogeneity is represented through object-based geostatistical models that are conditioned to wellbore data from the Denver Basin aquifer system. The Monte Carlo framework includes 200 heterogeneity realizations across a range of sandstone fractions.

Results document the formation of unsaturated regions beneath the alluvium in all models, particularly where sandstone channels underlie thinner low-permeability mudstones. Three-dimensional heterogeneity creates complex saturation patterns that result in localized flow paths, spatially varying disconnection, and a gradual transition to hydraulic disconnection as the regional water table is lowered. Successive changes in A-B flow decrease over the course of simulations by 80% to 99% and final rates approach stability as indicated by changes of <1% between successive stress periods. Of the 200 models, 190 reach full hydraulic disconnection and 10 conclude with a transitional flow regime. Dynamic connectivity metrics developed within the study strongly explain flow results. I also evaluate the aspects of heterogeneity that are most likely to produce disconnection, highlighting several factors that influence disconnection potential.

Chapter 4 evaluates the potential for a beaver dam to drive flow across the alluvial-bedrock contact. Beavers construct dams which promote a range of surface and near-surface hydrologic processes, however, the potential for beavers to influence deeper aquifer dynamics is less often, if ever, considered. In this study I consider the potential for a beaver dam, specifically increased stream stage and width upstream of a dam, to drive deeper flow from an alluvial to bedrock aquifer. I utilize a numerical groundwater flow model to simulate the effects of the beaver dam on inter-aquifer exchange rates. The base case model is parameterized based on observations from a beaver dam constructed on Cherry Creek in 2020 and the stream-alluvial-

bedrock aquifer sequence in the Denver Basin in previous chapters. I also test whether the influence of the beaver dam is sensitive to the alluvial-bedrock contact depth, beaver pond depth, and hydraulic properties by simulating flow across a range of sensitivity scenarios.

Model results document an increase in alluvial to bedrock flow on the order of 0.5% to 4%, depending on the contact depth, beaver pond depth, and hydraulic properties. Changes in hydraulic head due to the dam propagate deep into the aquifer (>30m), highlighting the potential for deeper aquifer impacts. The effect of the beaver dam is greatest for shallow alluvial-bedrock contact depths, deeper pond depths, and lower hydraulic conductivity contrasts between the alluvial and bedrock aquifer. Overall, results document the potential for beavers to influence deeper aquifer fluxes where regional hydraulic gradients are downward, highlighting broader potential for beaver dams to enhance aquifer recharge in deeper aquifer settings.

ACKNOWLEGEMENTS

I am sincerely thankful to everyone who supported me on this journey through higher education. First and foremost, I would like to thank my advisor, Mike Ronayne, for his guidance, mentorship, and endless patience. His door was always open, and it has been a pleasure to have him as my mentor, collaborator, and resident expert for the past seven years. Thanks to my committee members Lisa Stright, Sara Rathburn, and Ryan Bailey for their insights and support regarding my research and graduate experiences.

This study was partially funded by a USGS National Institutes for Water Resources program grant, administered through the Colorado Water Center. I wish to thank Sorab Panday for helpful suggestions related to the application of MODFLOW-USG and Peter Barkmann at the Colorado Geological Survey for providing access to geophysical logs.

I have also benefited from the guidance and support of many collaborators over the years. Sincere thanks to Tom Sale, Sally Sutton, Bill Sanford, Bob Stoller, and Greg Butters at CSU for providing opportunities, financial support, and mentorship. Thanks to Heather Justus and Lauren Moore at Castle Rock Water for funding, motivation, and field support and thanks to the many hands that provided support in the field: Matt Tyrell, Cody Delgado, Isabella Ulate, Julia Grabowski, and Kelsy Oberkoetter.

Though life as a groundwater modeler can lack the camaraderie of more field- and lab-heavy disciplines, I have genuinely enjoyed friendships and collaborations with many geoscience and engineering graduate students during my time at CSU. A special thanks goes out to my peers in the GeoGSO, GAAP, DEI committee, URGE, and department coffee hours.

Thanks to Sean Bryan, Dan McGrath, Rickey Frierson, the geosciences DEI committee, and URGE participants for inspiring and collaborating with me and countless other students to make our department a more inclusive, equitable, and just space for future generations of geoscientists.

Lastly, thanks to my friends and family behind the scenes who have supported me in countless ways: my running buddies, Julia and Salt, my KCSU therapy sofa, Hans, my hours-long phone calls, Kel and Kelsy. Thanks to my family, Steven, Mom, Arrie, Ronelle, Silo, and Poppy for the endless support and patience. And finally, thanks to Klara for pushing, supporting, and motivating me every single day.

DEDICATION

I dedicate this dissertation to my father, Albert, who pushes me to live as I please.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Groundwater represents the largest available source of freshwater in the world, making up 99% of all non-frozen, non-saline water (Siebert et al., 2010; Famiglietti, 2014). Since the mid-1900's, global groundwater development has rapidly intensified to meet rising municipal, industrial, and agricultural water demands (Wada et al., 2010). With expanding populations and climate change, the utilization of and reliance on groundwater is expected to rise as surface water resources are progressively more stressed (Aeschbach-Hertig and Gleeson, 2012). As global dependence on groundwater increases, sustainable groundwater management is increasingly necessary to ensure that supplies can meet future needs.

Quantifying rates of aquifer replenishment is a critical step in sustainable management of groundwater supplies. Unlike surface water, which is replenished annually from precipitation and snowmelt, aquifers are recharged through diverse mechanisms that often occur over long time periods. These include natural processes like infiltration, inter-aquifer flow (e.g., mountain-block recharge), and diffuse percolation of precipitation, as well as human-facilitated recharge through infiltration ponds, well injection, or as induced recharge from pumping (e.g., capture). The dynamics and timescales associated with natural and induced recharge dictate whether and for how long groundwater resources are sustainable. As global aquifers undergo long-term pumping, induced recharge is increasingly being drawn into aquifers as a critical offset to storage losses and pumpage (Thaw et al., 2022).

In addition to water supply, recharge plays a key role in a host of processes involving the transport of nutrients, contaminants, and heat (Healy, 2002). With climate change and global development increasingly stressing ecosystems and resources, the ability to model and predict transport processes which influence water quality and habitat will be essential for preserving and remediating natural systems. For these reasons, it is increasingly important to quantify and understand mechanisms and rates of groundwater recharge.

This dissertation applies groundwater flow and geostatistical modeling to evaluate spatial and temporal controls for recharge fluxes throughout a stream-alluvial-bedrock aquifer system based on the Denver Basin of Colorado, USA. Chapter 2 begins at the stream-aquifer interface where I investigate controls and timescales associated with streambed fluxes, and important component of seepage recharge to Denver Basin aquifers. I expand characterizations of streambed flux variability to the temporal domain, an area which has been largely overlooked compared to its spatial counterpart. Findings highlight new insights about the drivers and timescales associated with streambed fluxes and associated implications for recharge and hyporheic exchange. This chapter was published with minor divergences in the journal Hydrological Processes in 2023 (Cognac and Ronayne, 2023).

Chapters 3 and 4 are organized to progress deeper within the system to evaluate recharge fluxes across the alluvial-bedrock contact. Flow from the alluvial to bedrock aquifer constitutes an important component of recharge to the bedrock aquifers of the Denver Basin, particularly as vertical hydraulic gradients increase from long-term pumping and associated head declines in the bedrock aquifers (Cognac and Ronayne, 2020). In Chapter 3, I employ geostatistical and variably saturated flow modeling to whether and how the development of unsaturated zones could impact flow rates from the alluvial to bedrock aquifer. Geostatistical models are developed to represent the geologic heterogeneity associated with the fluvial depositional environment, allowing for multiple realizations that are used to explore the sensitivity of results to aquifer heterogeneity. This chapter reveals insights about the multi-aquifer systems response to pumping when heterogeneity and saturation effects are considered. It also extends conceptualizations about hydraulic disconnection and provides what I believe is the first example of hydraulic disconnection between two aquifers. Chapter 3 was submitted for publication to journal Water Resources Research in July of 2023. It is currently in review.

Throughout the fieldwork for Chapter 2, I observed dozens of beaver dams being constructed and breached along East Plum and Cherry Creeks in Douglas County, Colorado. A few dams were constructed within 10 m of the existing piezometer network, providing an opportunity to study and capture hydrologic impacts from the dam. Because many of the shallow hydrologic effects of beaver dams have already been well characterized, in Chapter 4, I consider the potential for a beaver dam to drive deeper flow across the alluvial-bedrock contact. The Denver Basin provides a unique setting to consider deeper aquifer influence of beaver dams because vertical fluxes are predominantly downward. I evaluate the dam's influence using a groundwater flow model and explore the sensitivity of the influence to the alluvial-bedrock contact depth, beaver pond depth, and hydraulic parameters.

While recharge is the theme of this dissertation, I don't attempt to directly quantify recharge rates for a region, but rather investigate and expose dynamics of recharge that are critical for accurate conceptualizations and estimates of recharge. In this way, I contribute to the understanding of recharge mechanisms in ways that can be applied to estimating recharge in aquifers across the globe.

Each study is conceptualized to represent the streams, alluvial, and bedrock aquifers and of the Denver Basin in Colorado, USA. The Denver Basin aquifer system is one of many in the United States that is undergoing long-term pumping and associated depletion (Konikow, 2015). Accurate estimates of recharge are essential for developing sustainable management strategies and for understanding the useable lifespan of the aquifer system. Further, the Denver Basin aquifers underlie a surface water system that constitutes an equally critical water supply and is governed by a complex legal framework. It is essential to accurately quantify interactions between groundwater and surface water (e.g., recharge) to effectively allocate water resources. While existing regional groundwater models exist for this purpose, some aspects of recharge may be mis-represented due to coarse discretization and effective parameterization used in models (e.g., Paschke et al., 2011).

1.1 References

Aeschbach-Hertig, W., & Gleeson, T. (2012). Regional strategies for the accelerating global problem of groundwater depletion. Nature Geoscience, 5(12), 853-861.

Bredehoeft, J. D. (2002). The water budget myth revisited: why hydrogeologists model. Groundwater, 40(4), 340-345.

Cognac, K. E., & Ronayne, M. J. (2020). Changes to inter-aquifer exchange resulting from long-term pumping: implications for bedrock groundwater recharge. Hydrogeology Journal, 28(4), 1359-1370.

Cognac, K. E., & Ronayne, M. J. (2023). Multiple timescales of streambed flux variability in two perennial mountain-front streams. Hydrological Processes, 37(3), e14840.

Famiglietti, J. S. (2014). The global groundwater crisis. Nature Climate Change, 4(11), 945-948.

Konikow, L. F. (2015). Long-term groundwater depletion in the United States. Groundwater, 53(1), 2-9.

Paschke, S.S. ed., (2011), Groundwater availability of the Denver Basin aquifer system, Colorado: U.S. Geological Survey Professional Paper 1770, 274 p. doi: 10.3133/pp1770

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation–a global inventory. Hydrology and earth system sciences, 14(10), 1863-1880.

Thaw, M., GebreEgziabher, M., Villafañe-Pagán, J. Y., & Jasechko, S. (2022). Modern groundwater reaches deeper depths in heavily pumped aquifer systems. Nature communications, 13(1), 5263.

Wada, Y., Van Beek, L. P., Van Kempen, C. M., Reckman, J. W., Vasak, S., & Bierkens, M. F. (2010). Global depletion of groundwater resources. Geophysical research letters, 37(20).

CHAPTER 2: MULTIPLE TIMESCALES OF STREAMBED FLUX VARIABILITY IN TWO PERENNIAL MOUNTAIN-FRONT STREAMS

2.1 Introduction

Groundwater-surface water exchange is an important driver for many hydrological (Winter et al., 1998), biogeochemical (Boulton et al., 1998; Hayashi and Rosenberry, 2002; Boano et al., 2010), and ecological (Jones and Mulholland, 2000; Hancock et al., 2005) processes in streams and surrounding environments. Field estimates of streambed fluxes are regularly used to quantify aquifer recharge and discharge (Constantz et al., 1994; Conant, 2004; Constantz, 2008), hyporheic exchange (Cardenas et al., 2004), contaminant and water flow-paths and residence times (Essaid et al.,2008; Lewandowski et al., 2011; Briggs et al., 2019), and nutrient and temperature dynamics for aquatic habitats (Caissie, 1991; Alexander and Caissie, 2003; Tonina and Buffington, 2009; Van Grinsven et al., 2012; Gilmore et al., 2016). Unfortunately, reliable exchange rates are notoriously difficult to obtain, largely due to the significant spatiotemporal variations that characterize fluxes in field settings (Scanlon et al., 2002; Kalbus et al., 2006; Rosenberry and LaBaugh, 2008; Healy, 2010; Lautz, 2010).

Recently, authors have employed a range of techniques to characterize spatiotemporal flux variations, with an emphasis on linking variations to stream-aquifer processes. Briggs et al. (2012) used high-resolution distributed temperature sensors to monitor fluxes around beaver dams and identified spatially distinct, temporally varying hyporheic exchange patterns resulting from dams. Kennedy et al. (2009) combined tracer, age-dating, and Darcy methods to investigate nitrogen and water fluxes in an agricultural creek and found that variable streambed fluxes were a primary control for nitrogen concentrations. Essaid et al. (2008) employed nested piezometers and heat-based numerical models to investigate how spatiotemporal flux changes influenced flow paths and residence times of degrading agricultural contaminants. Malcolm et al. (2010) determined that variable nutrient concentrations in aquatic habitats were strongly related to variable flux rates, deeming flux variability an essential consideration for aquatic habitat thresholds.

Within these investigations, however, spatial heterogeneity has dominated results and discussions, revealing important insights about hyporheic flow paths, reach scale nutrient and carbon cycling, and other stream-aquifer processes. Spatial heterogeneity in fluxes has been attributed to variations in hydraulic conductivity and stream curvature (Cardenas et al., 2004), streambed topography, morphology, and bedforms (Harvey and Bencala, 1993; Gooseff et al., 2006; Hester et al., 2013; Sebok et al. 2015; Song et al., 2017; Cheng et al., 2019), and spatially variable groundwater gradients (Storey et al., 2003). Ranges of fluxes across study sites have typically spanned 1 to 2 orders of magnitude with complex spatial patterns of positive and negative values indicating zones of groundwater discharge and recharge (e.g., Conant, 2004; Hatch et al., 2010; Briggs et al., 2012).

Temporal variations in fluxes occur as either fluctuations and repeating patterns across different temporal scales, or short- and long-term trends controlled by natural or anthropogenic processes. Temporal dynamics can similarly reveal useful insights about streambed processes, though they are less often directly assessed. Previously, authors have documented temporal variations in streambed fluxes at sub-daily through seasonal timescales and have attributed them to: (i) variable stream stage from natural and human causes (Palkovics et al., 1975; Arntzen, 2007; Loheide and Lundquist, 2009; Argerich et al., 2011; Fritz and Gribovszki et al., 2010;

Griglio et al., 2013), (ii) variations in aquifer hydraulic head related to evapotranspiration, seasonality of recharge, and pumping (Kashara and Hill, 2006; Rosenberry and Pitlick, 2009; Gribovski et al., 2010), (iii) changes in hydraulic properties related to temperature, biological activity, or gas content (Constantz et al., 1994; Gribovski et al., 2010; Genereux et al., 2008; Song et al., 2010), and (iv) sporadic, seasonal, and long-term changes in streambed sediments including erosion and deposition (Kasahara and Hill, 2006; Keery et al., 2007; Genereux et al., 2008; Rosenberry and Pitlick, 2009; Hatch et al., 2010; Argerich et al., 2011; Simpson and Meixner, 2012; Briggs et al., 2012).

Despite numerous reports, few studies have rigorously described the range of temporal variations encountered at individual sites over time periods greater than a few months. Hyporheic exchange, contaminant transport, and nutrient cycling operate over a broad range of timescales, spanning from hours to years (e.g., Cardenas et al., 2008; Buffington and Tonina, 2009; Kennedy et al. 2009). Most field studies have employed either high-measurement-frequency, short-duration sampling schemes spanning days to a couple months (e.g., Arntzen et al., 2006; Käser et al., 2009; Briggs et al.,2012) or long duration, low-measurement-frequency sampling schemes such as bimonthly for 1 year (e.g., Kennedy et al., 2009). A few studies have spanned greater time periods ranging from 5 months to 1 year and have documented greater temporal flux variations (Essaid et al., 2008; Hatch et al., 2010; Anibas et al., 2016). Anibas et al. (2016) reported seasonal flux variations over a 5-month study period with bi-monthly monitoring that spanned 2 orders of magnitude at individual sites and cases where fluxes changed between positive and negative values (gaining and losing conditions) seasonally. Essaid et al. (2008) reported modeled flux results over a 9-month study period and documented seasonal variations in hydraulic gradients and fluxes, with flow reversals occurring during high-flow events. However, examination of fluxes and temporal variability over multi-year timescales is not common because continuous monitoring of streambed temperatures and heads for long durations is challenging. Long-term estimates of temporal variability are important for understanding seasonal dynamics and associated uncertainty in streambed flux estimates.

This study employs a high-frequency (10-min logging frequency), long-term (3 year) monitoring period to characterize temporal variability in streambed fluxes in two second-order mountain-front streams in Colorado, USA (***Figure 1***). Study sites are located along the western margin of the Denver Basin aquifer system where focused seepage along stream channels is potentially a major component of groundwater recharge (Scanlon et al., 2002; Wilson and Guan, 2004). The primary objectives of this study were to quantify temporal variability in streambed fluxes over short- and long-timescales using methods that have been applied to spatial variability. I identify the magnitude of temporal variability along with patterns and timescales that are relevant for streambed processes. I also investigate controls for different timescales and discuss the implications of temporal changes for recharge estimates, field data collection, and measurement uncertainty.

2.2 Data and methods

2.2.1 Description of Study Sites

Field investigations focused on two second-order, perennial streams in central Colorado, East Plum Creek (PC) and Cherry Creek (CC), that were previously identified as having predominantly losing conditions with potential for seepage recharge (Cognac and Ronayne, 2020). PC has headwaters in the Rampart Range of the Rocky Mountains, and along a topographic high extending east of the Rampart Range known as the Palmer Divide (***Figure 1a***).

The 17 km PC study reach begins approximately 0.8 km east of the mountain-front where PC emerges from Stone Canyon. At the upstream monitoring station, the drainage area of PC is approximately 32 km^2. CC lies within the adjacent drainage to the east, and the 3 km study reach begins approximately 40 km north of headwaters along the Palmer Divide. At the upstream monitoring station, the drainage area of CC is 540 km^2. From headwaters, both streams flow north and east toward the South Platte River along 0.5 to 1.5 km wide Quaternary alluvial valleys which dissect Tertiary sandstones and siltstones of the Denver Basin (Barkmann et al., 2015).

Both streams exhibit meandering morphologies and are characterized by pool-riffle sequences, sand-and-gravel streambeds, occasional fluvial bars, and floodplains that are confined by Quaternary alluvial terraces and incised banks. Both channels are incised within the modern floodplain and more broadly confined by alluvial valleys shaped by historic flooding. Recent incision was observed to be greatest at PC-DS and least at PC-MU. Floodplain vegetation is dominated by grasslands with interspersed phreatophyte communities of willow (*Salix spp.*) and cottonwood (*Populus deltoids*) (***Figure 1a***). Variable discharge and intense flooding are common for streams in this region due to orographic precipitation and convective storms along the Rocky Mountain front (Livers and Wohl, 2015). Both streams have recorded historic, catastrophic floods which caused significant damage and intensive geomorphic change (Matthai, 1969). Nearby USGS stream gages indicate daily mean discharges that range from 0.01 to 17 $m^3 s^{-1}$ and 0.02 to 10 $m^3 s^{-1}$ on PC and CC respectively during the study period, with annual minima occurring between August and February and annual maxima between March and July (U.S. Geological Survey, 2022). Study reaches were situated upstream of major urban developments along the mountain front zone but are downstream of publicly reported diversions. Data collected in this study are part of a broader effort to monitor water resources as underlying Denver Basin aquifers undergo significant, long-term pumping and water level declines.

Meteorological stations near the study sites report annual precipitation ranging from 12.5 to 77 cm (1962 to 2021), with study years (2018-2021) ranging from 25 to 42cm (Colorado Climate Center, Castle Rock, Station #051401). Average daily atmospheric temperature ranges are 15°C with maximum ranges up to 35°C. Large atmospheric temperature fluctuations are transferred to stream and groundwater temperatures, particularly along losing stream reaches with minimal shade cover and low stream discharge (Constantz, 2008). Stream and groundwater temperatures at study sites exhibit daily fluctuations of up to 17°C, and are notably greater than similar studies (i.e., 3-8 °C by Lautz, 2010; Caissie et al., 2014), thus providing an opportune setting to evaluate temperature effects on fluxes.

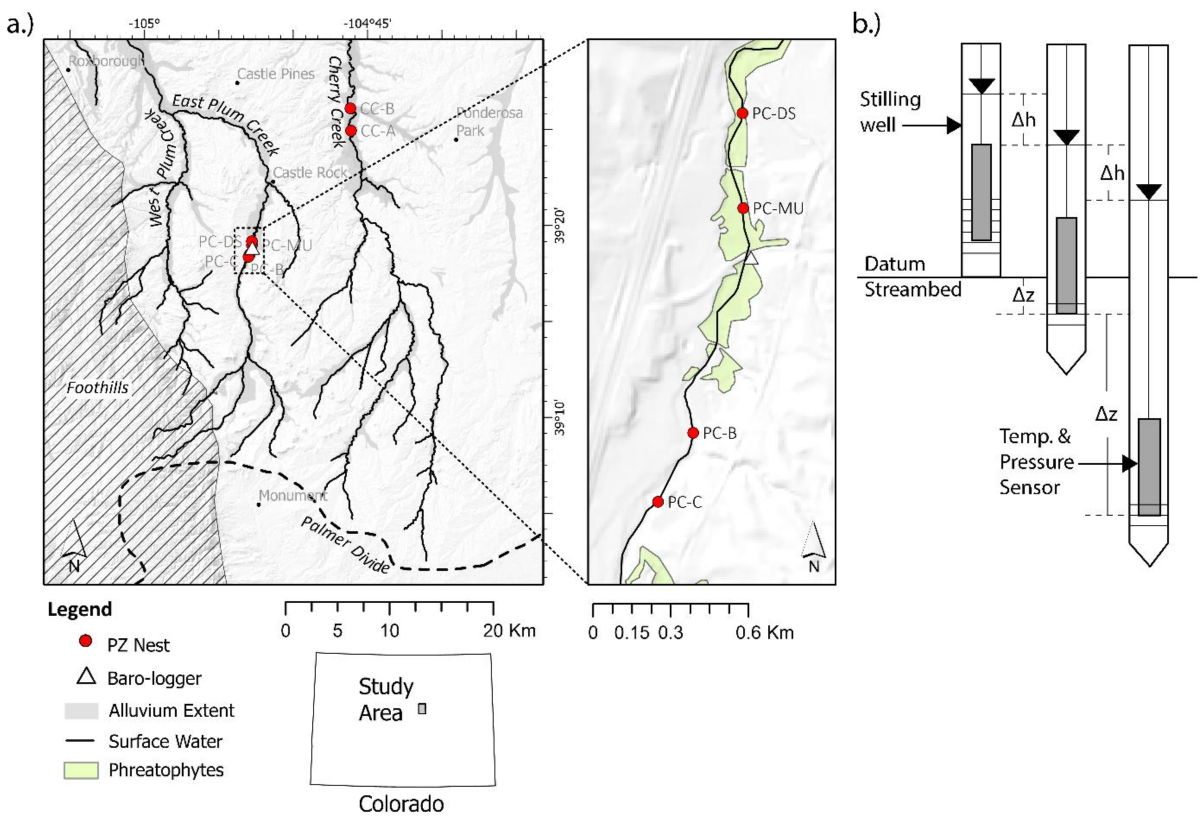

Figure 1. a.) Location of monitoring sites and b.) schematic of piezometer nest with stilling well.

2.2.2 Field data collection

Temperature and pressure data were obtained from a network of nested piezometers (PZs) installed within shallow streambed sediments along PC and CC study reaches (***Figure 1a***). Each PZ nest consisted of either two or three 1-inch-diameter (2.54 cm) slotted-steel pipes that were manually driven into the streambed to depths ranging from 0.05 m to 1.65 m (***Figure 1b***). 1-inch diameter steel pipes are preferred over PVC for measuring streambed temperatures as they minimize thermal buffering in the water column (Cardenas, 2010). At four nest locations, a co-located stilling well was installed to measure stream stage. The stilling well was constructed from slotted PVC pipe and anchored to the deepest PZ. At each nest location, tops of PZs and stilling wells were set at equal elevation using a level and regularly monitored to confirm consistency in the configuration. Nests were installed in the center of the channel preferably in straighter-planform areas which tended to contain riffles rather than pools. Piezometers were periodically developed to avoid clogging of the slotted intervals by purging with an inertial pump until water ran clear.

Pressure transducers (accuracy ± 0.25 cm) and temperature sensors (accuracy ± 0.05 °C) were deployed within the slotted intervals of stilling wells and PZs and set to log at 10-minute intervals. Barometric pressure was logged from a central location every 10 minutes (***Figure 1a***). Water levels were determined relative to the streambed datum by subtracting atmospheric pressure from the total pressure measured in PZ and stilling well transducers. Regular data downloads and site visits (4-6 times per year) recorded stream stage, depth to streambed, changes in stream conditions, and nearby animal activity. Depth to streambed was measured from the top of the deep piezometer riser, which stayed fixed through time, and was used to monitor erosion and deposition. Sieve analysis was performed on sediment samples collected in the vicinity of each PZ-nest. The data-collection period spanned from September 28, 2018, to September 28, 2021, which is approximately 3 water years, defined by the U.S. Geological Survey (USGS) as the period spanning October 1 to September 30 the following year. Data gaps are present for several locations due to staggered installation, flooding from beaver dams, and a stilling well going dry due to low flow. Discharge measurements were also collected with the intent of calculating seepage losses via streamflow differencing; however, large daily variations in stream discharge prohibited reliable differenced flux estimates.

2.2.3 Vertical hydraulic gradients and streambed fluxes

Streambed vertical hydraulic gradients (VHGs) were calculated as VHG = $\Delta h/\Delta z$, where Δh is the elevation difference between water levels in two piezometers, and Δz is the distance between midpoints of the piezometer slotted intervals. Positive VHGs reflect the potential for downwelling conditions, and negative for upwelling.   Because PZ nests were constructed with a welded bar between the shallow and deep piezometer, Δz remains fixed through time even if the streambed elevation changes (i.e., erosion or deposition). VHGs were calculated for the upper two streambed piezometers with average measurement depths of 0.05 m and 1 m below the streambed.

Vertical streambed fluxes were calculated for the same depth interval using a 1D form of Darcy's law: $q = -K(T)VHG$, where q is the water flux through the streambed (m s^{-1}), $K(T)$ is the temperature dependent hydraulic conductivity (m s^{-1}), T is water temperature (°*C*) and *VHG* is the vertical hydraulic gradient as described above. Streambed flux is the volume of water moving across an area of streambed over a given time and is sometimes reported as volume per area per time. The temperature dependent hydraulic conductivity (*K(T)*) was incorporated into

Darcy flux estimates to account for changes in *K* that occur due to variable water temperatures. It is well documented that temperature changes cause density and viscosity to vary, with viscosity having a significant effect on *K* (e.g., Constantz, 1982). In this study, *K(T)* was determined for each measurement time (10-min frequency) by calculating the temperature dependent viscosity $\eta(T)$ using the average temperature of the shallow and deep piezometer, and inserting $\eta(T)$ into Equation 2.1 for hydraulic conductivity, where *k* is intrinsic permeability (m^2), ρ is the water density ($kg\ m^{-3}$), *g* is acceleration due to gravity ($m^2\ s^{-1}$), and η is the viscosity of the water ($kg\ m^{-1}\ s^{-1}$). Water viscosity was estimated as a function of temperature (°C) using the Vogel equation (Equation 4) (Vogel, 1921). Intrinsic permeability was assumed constant in time for each location and was calculated using site-specific K and viscosity at a given temperature. The effects of temperature-related density changes on K are negligible (e.g., Constantz, 1982). However, the effects of temperature-related viscosity changes on K can be significant. Using equations 3 and 4, a 15 °C temperature fluctuation would produce a 40% change in K.

$$K(T) = \frac{k\rho g}{\eta(T)} \quad \text{(Eq. 2.1)} \qquad\qquad \eta = e^{-3.7188+\frac{578.919}{135.6+T}} \quad \text{(Eq. 2.2)}$$

The reliability of Darcy-based flux estimates has been questioned due to uncertainty in field-based K estimates related to spatiotemporal variability (e.g., Surridge et al., 2004; Genereux et al., 2008). Here, a two-step approach for estimating K is used to reduce uncertainty and improve flux results. First, K was estimated for the upper 10-15 cm of the streambed at each monitoring location through falling head tests (FHTs) conducted in the shallow streambed sediments surrounding the PZ nest using the Hvorslev (1951) equation (with m=1) as described by Landon et al. (2001). This method is appropriate for sand-and-gravel streambeds (Landon et al., 2001). FHTs involved driving a permeameter 10-15 cm into the streambed, imposing a hydraulic head above the stream water level, and measuring the rate at which displacement occurred. I found that 15-cm was lower limit of FHTs due to gravel that limited the depth to which permeameters could be driven into the streambed. FHT-based K estimates are reported as temperature-corrected values for T= 20°C using Equations 2.1 and 2.2.

K estimates were refined through numerical heat and flow modeling with subsurface water temperature as a calibration target. One-dimensional models were developed in VS2DH, a partially coupled heat and fluid flow model capable of simulating transient pressure and temperature conditions in porous media (Healy, 1990; Healy and Ronan, 1996). Numerical modeling involved solutions to the governing convection-conduction equation for heat transport, presented in one dimensional (1D) form as (Stallman, 1965; Goto et al., 2005; Hatch et al., 2006; Keery et al., 2007):

$$\frac{\partial T}{\partial t} = \kappa_e \frac{\partial^2 T}{\partial z^2} - q_{gw} \frac{C_w}{C} \frac{\partial T}{\partial z} \quad \text{(Eq. 2.3)}$$

where T is temperature (°C), t is time (s), κe is effective thermal diffusivity ($m^2 s^{-1}$), Cw is heat capacity of the fluid ($J\ m^{-3}\ °C^{-1}$), C is volumetric heat capacity of sediment and fluid ($J\ m^{-3}\ °C^{-1}$), z is depth into streambed (m), q_{gw} is the vertical groundwater (Darcy) flux ($m\ s^{-1}$). The effective thermal diffusivity, κe, is defined as:

$$\kappa_e = \frac{\lambda_e}{C} = \frac{\lambda_0}{C} + \beta \left| \frac{q_{gw}}{n} \right| \quad \text{(Eq. 2.4)}$$

where λ_e is effective thermal conductivity ($J\ s^{-1}\ m^{-1}\ °C^{-1}$), λ_0 is the bulk effective thermal conductivity of the fluid and sediment matrix ($J\ s^{-1}\ m^{-1}\ °C^{-1}$), n is the porosity (-), and β is thermal dispersivity (m).

In VS2DH, heat transport influences fluid transport through the temperature dependence of viscosity and saturated hydraulic conductivity. Details of the fluid and heat equations and iterative solver methods can be found in Lappala et al. (1987) and Healy and Ronan (1996). One-dimensional models of vertical water and heat flow were developed in VS2DH to simulate SW-GW fluxes for the four PZ nests with time series records of stage (stilling wells). Model domains spanned from the streambed to 1-1.6 m below the stream representing the midpoint of the slotted interval of the deepest PZ. Model grid spacing was set to 0.02 m in the vertical direction and 1 m in the horizontal direction (***Figure 2***), following the approach of Essaid et al. (2008). Sensitivity analysis confirmed that this grid spacing produced stable model results. Temperatures, heads, and water fluxes were generated through simulations in which time-variable boundary conditions were assigned from field data. The upper boundary condition was set to the measured stream stage and temperature, and the lower boundary condition was set to the deepest PZ hydraulic head and temperature.

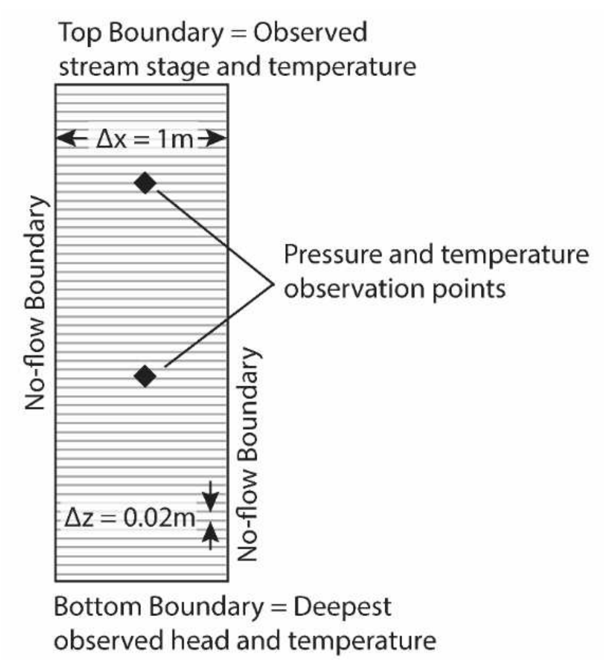

Figure 2. Schematic showing model setup.

For model calibration, simulation periods had a duration of 15 days. Multiple time periods were evaluated for each location to estimate site-specific values of hydraulic conductivity (K) and bulk effective thermal conductivity (λ_0). Calibration involved varying the λ_0 and K values until the best fit between modeled and observed temperature data was achieved. Prior to calibration, a sensitivity analysis was performed in which it was determined that modeled temperatures were highly sensitive to changes in hydraulic conductivity and less sensitive to thermal conductivity, thermal dispersivity, and sediment heat capacity in decreasing order. For this reason, hydraulic conductivity and thermal conductivity were prioritized as calibration targets, and other thermal parameters were estimated using published values from studies in similar hydrogeologic settings described subsequently. Thermal conductivity and hydraulic conductivity were calibrated simultaneously by minimizing a mean absolute error (MAE) surface between modeled and observed data. Calibrated hydraulic conductivity values were compared to field-based K estimates. The calibrated λ_0 values, ranging from 2.4 to 3.1, were within plausible ranges estimated using streambed sediment mineralogy and porosity from field samples, as described herein.

Independent estimates of λ_0 were obtained using the equation: $\lambda_0 = \lambda_f^n \, \lambda_g^{(1-n)}$, where λ_f is the thermal conductivity of the fluid (water), λ_g is the thermal conductivity of the grains, and n is the porosity (e.g. Hatch et al., 2006; Rau et al., 2014). For this study, λ_f was assigned a value of 0.588 J s^{-1} m^{-1} C^{-1} (Sengers and Watson, 1986) and $\lambda_g$was estimated using constituent mineral mass fractions to assign weights to published thermal conductivities for quartz and feldspar. PC and CC streambed sediments are sourced from weathered Pikes Peak Granite, which consists primarily of quartz and feldspar (Barker et al., 1974). Mass fractions of quartz (f_q) and alkali feldspar (f_{fe}) were calculated through manual sorting and weighing of streambed sediment samples. Resulting f_q ranged from 73 to 84 percent. λ_g was then estimated as: $\lambda_g = f_q\lambda_q + f_{fe}\lambda_{fe}$. Reported thermal conductivity values for quartz (λ_q) and feldspar (λ_{fe}) ranged from 6.15-11.3 and 2.33-2.49 J s^{-1} m^{-1} C^{-1} respectively, with average values of 7.69 and 2.35 J s^{-1} m^{-1} C^{-1} (Cermak and Rybach, 1982; Clauser and Huenges, 2013). The resulting range of thermal conductivity values for calibration was 1.5 to 3 J s^{-1} m^{-1} C^{-1}, which is similar to other studies (e.g., Lautz, 2012; Rau et al., 2015). Porosity was calculated for three sediment samples collected near each piezometer nest by differencing the mass of saturated and oven-dried samples and dividing by total sample volume and water density. Resulting porosity values ranged from 0.27 to 0.35 with a mean of 0.31 J s^{-1} m^{-1} C^{-1}. A low-temperature (45°C) dehydrator was used to dry samples for 3 weeks between mass measurements. Calibration ranges for λ_0 were limited to ±50% of the estimated range of values from mineral and porosity-based estimates.

The volumetric heat capacity of the sediment-fluid matrix was estimated as (Silliman et al., 1995): $\rho c = n\rho_w c_w + (1-n)\, \rho_s c_s$, where ρ_w and c_w represent the density and specific heat of the water (1000 kg m^{-3} and 4187 J kg^{-1} $°C^{-1}$) , and ρ_s and c_s represent the density and specific heat of the solids within the streambed. Using published values of ρ_s and c_s for quartz and feldspar (Cermak and Rybach, 1982), along with the site-specific mineral fractions, the mean volumetric heat capacity of the saturated sediment was calculated to be 2460 J m^{-3} $°C^{-1}$. which is within the range of previously reported values for sand and gravel sediments (Hamdhan and Clarke, 2010; Lautz, 2012; Caissie et al., 2014). Dispersivity (β) was tested for a range of values between 0.1-1% of the model domain. A value of 0.01 m was ultimately assigned for models to be consistent with exchange studies across similar streambed intervals (e.g., Rau et al., 2015).

Some authors have found the effects of β to be negligible and have disregarded this term in heat-based flux estimates (e.g., Keery et al., 2007; Rau et al., 2015).

A minimum of three monitoring depths were required for calibration, excluding locations PC-MU and PC-DS from calibration. Depth-specific water temperature and hydraulic head were calculated during simulations with time-varying boundary conditions assigned from field data. The upper boundary condition was set to the measured stream stage and temperature, and the lower boundary condition was set to the deepest PZ hydraulic head and temperature. Lateral boundary conditions were set to no-flow conditions. Notably, the model-estimated K values correspond to the same depth interval as VHGs. For model calibration, simulation periods were set at 15-day intervals during times when stage data were most reliable (e.g., not during high-stage events). Calibration involved varying thermal and hydraulic conductivity until a best fit between modeled and observed temperature in the streambed was obtained.

During calibration, the K value was permitted to vary over a range spanning multiple orders of magnitude; for each location, the lower bound of this range was the minimum field-estimated value for all sites divided by 10, and the upper bound was the maximum field-estimated value for all sites multiplied by 10. Calibration dates include 15-day periods where high-quality stage data were available. Unlike piezometers, which were securely anchored from driving steel pipes into the streambed, PVC stilling wells were anchored using zip ties and occasionally changed position during high discharge events. During and after disruptions, data were deemed unfit for calibration because it was typically impossible to determine if the change in position (of both the stilling well and transducer) was instant or gradual. Importantly, disruptions in stilling well data do not impact piezometer data or flux estimates. Examples of observed and modeled temperature values during calibration periods for locations PC-B and CC-B are shown in ***Figure 3***.

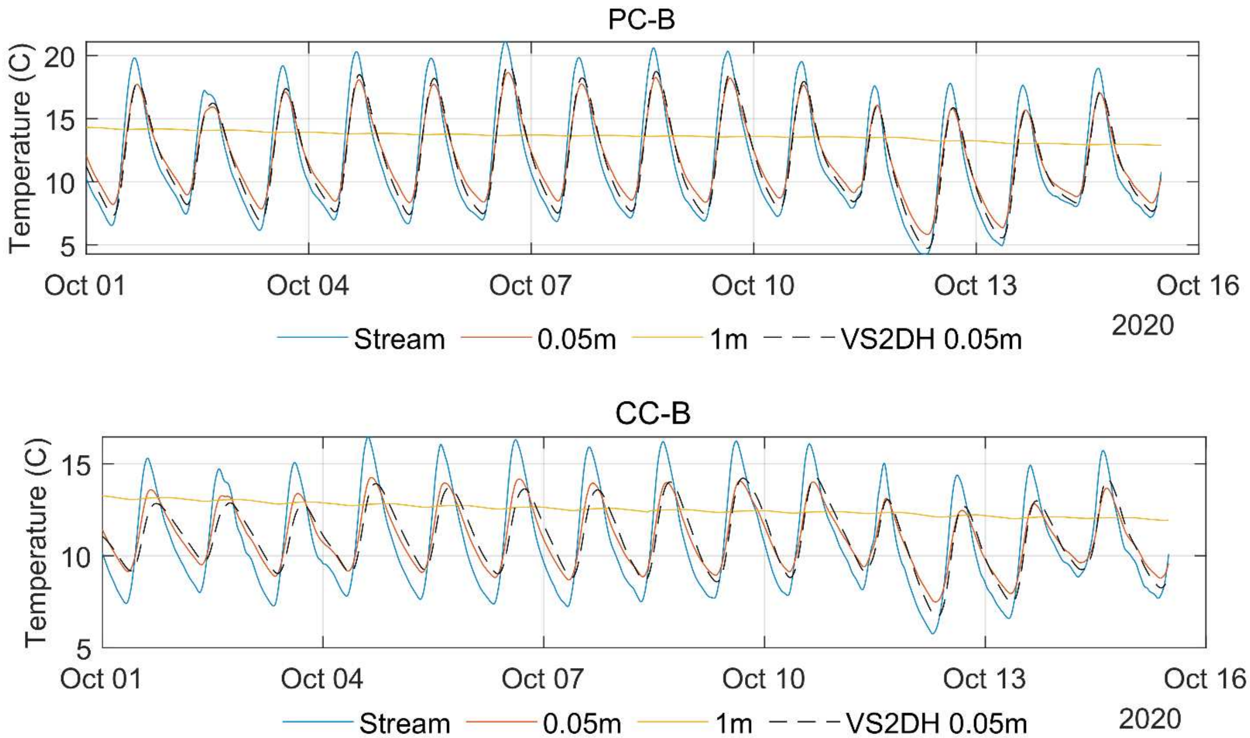

Figure 3. Example observed and modeled temperature during October 1 to October 15, 2020 calibration periods for locations CC-B and PC-B.

Model-calibrated K values were sensitive to the error function used to compare modeled and observed temperatures. This was determined to result from minor offsets in either phase or amplitude between the two quasi-sinusoidal signals causing variable errors depending on the function used. I prioritized matching amplitude between modeled and observed temperature rather than phase because amplitude of the temperature signal is more likely to be preserved in the case of thermal buffering, while phase can be lagged (e.g., Cardenas et al., 2010). Several objective functions were tested (e.g., RMSE, MAE, amplitude residual), and mean absolute error (MAE) was determined to produce the best results.

2.2.4 Statistical analysis of temporal variability

Flux distributions were affected by short-duration events that resulted in outliers, skewness, and multiple modes. Therefore, resistant statistical metrics for central tendency (i.e., median) and variability (i.e., median absolute deviation) were employed (Helsel et al., 2020). For each nest location and water year, the mean ($\bar{x}$), median (Med.), minimum, maximum, median absolute deviation (MAD), and quartile-based coefficient of variation (QCV) were calculated from 10-minute time series of fluxes and VHGs. The MAD and QCV were also calculated for sub-sets of flux data that were either resampled (i.e., daily mean) or detrended to evaluate variations at different temporal scales as described herein. While QCV is typically avoided for variables with positive and negative values, I report absolute value of the QCV to compare relative variations between sites, similar to other related flux studies that have utilized CV (e.g., Kennedy et al., 2008; 2009).

Three general timescales, sub-daily, daily, and interannual, (***Figure 4***), were selected for analysis based on timescales associated with previously identified physical controls on streambed fluxes (i.e., ET, temperature, etc.), and observed patterns in data. Sub-daily variability was defined as variations that occur throughout a given day relative to the daily median. Statistics for evaluating sub-daily variability were calculated with subsets of detrended fluxes that were produced by subtracting a daily moving average from 10-minute time series. The median absolute deviation ($MAD_{S\text{-}D}$), quartile-based coefficient of variation ($QCV_{S\text{-}D}$), and daily range were calculated for yearly sets of detrended fluxes.

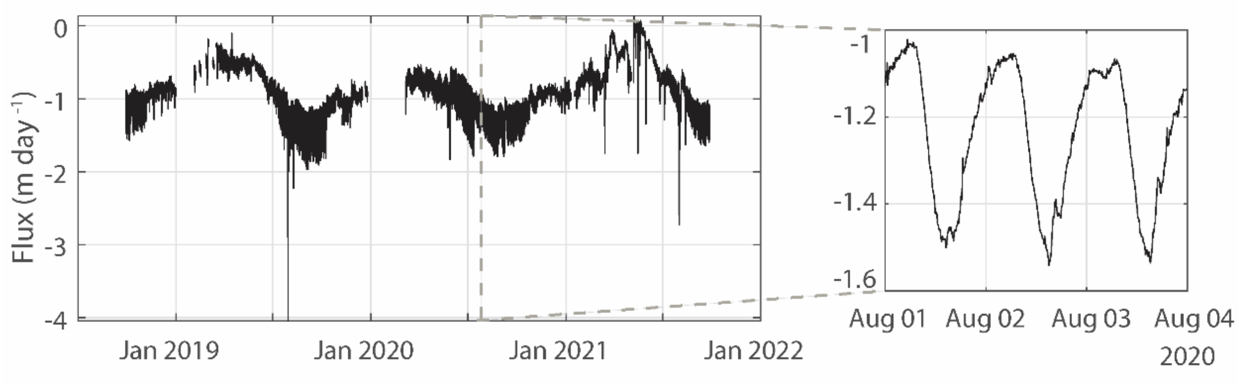

Figure 4. Example time series of streambed fluxes at site PC-B showing sub-daily, daily (seasonal), and interannual variability. Spikes in July 2019 and spring 2021 reflect high-stage events.

Daily variability was defined as variations that occur between daily median values, relative to the annual median. Daily variability was assessed with statistics calculated on subsets of daily and monthly median fluxes for each study year. Day of year plots were also constructed for several locations to qualitatively evaluate seasonal trends. Day of year plots show 7-day moving median fluxes against the day of the calendar year.

Interannual variability was defined as longer-term variations that occur between years associated with long-term trends and short-term events that alter streambed conditions. Interannual variations were assessed by calculating annual and seasonal mean and median fluxes for study years and comparing data across years. Because the number of annual observations was three (i.e., three study years), interannual variations have limited statistical confidence and are primarily described qualitatively rather than quantitatively.

2.3 Results

2.3.1 Hydraulic Properties

Hydraulic conductivity from field tests corrected for 20°C ranged from 0.15 to 140 m day^{-1}, with mean values across sites ranging from 2.6 to 89 m day^{-1} (**Table 1**). Model-calibrated hydraulic conductivities ranged from 8.64 to 69 m day^{-1} and were within ± an order of magnitude of field-based mean K for each site (**Table 2**). Values are within the expected range for sandy streambed sediments (e.g., Landon et al., 2001). Calibrated K values were used in flux calculations at the four sites with stilling wells (PC-B, PC-C, CC-A, CC-B) (Table 2). At sites where model calibration was not performed (PC-DS, PC-MU), fluxes were calculated with the mean field-estimated K value. Grain sizes were generally finer and better sorted at Cherry Creek than Plum Creek. At all but site CC-B, erosion and deposition of sediments was minimal, with <9 cm of total elevation change from the starting streambed surface elevation. A beaver dam constructed downstream from CC-B resulted in sediment buildup that is discussed in section 3.3.2. Field data for erosion and deposition are shown in ***Figure 5***.

Table 1. Estimated K values (m day-1) from falling head tests corrected for 20°C.

		East Plum Creek				*Cherry Creek*	
	Date	PC-DS[1]	PC-MU	PC-B	PC-C	CC-A	CC-B
FHTs (0-0.15m)	12-Jul-19	2.2	4.2	-	-	-	-
		-	9.0	-	-	-	-
		-	8.7	-	-	-	-
	8-Nov-19	-	-	5.3	26.5	-	-
		-	-	0.15	3.7	-	-
		-	-	-	18	-	-
	12-Aug-20	-	-	1.9	2.5	51	27
		-	-	3.1	3.6	140	0.46
		-	-	-	-	77	21
	Mean	-	7.3	2.6	10.8	89	16
	σ	-	2.7	1.8	4.9	13	15

Notes: [1]Only one falling head test conducted at PC-DS due to subsequent flooding.

Table 2. Estimated K values (m day-1) from model calibration corrected for 20°C.

		K (m day-1)		Calibration Flux Range (m day^{-1})	
	Date Range	Sample	Geo. Mean	min	max
PC-B	10/1 to 10/15/2019	41.9	18	-1.82	-0.95
	8/6 to 8/18/2020	10.8		-1.66	-1.00
	9/15 to 10/1/2020	10.1		-1.65	-1.03
	10/1 to 10/15/2020	15		-1.59	-1.00
	10/15 to 11/1/2020	19		-1.37	-0.80
	4/1 to 4/15/2021	17		-0.63	-0.10
	4/15 to 5/1/2021	26		-0.92	-0.20
PC-C	8/6 to 8/18/2020	21.5	20	-1.97	-1.03
	9/15 to 10/1/2020	24		-1.89	-1.15
	10/1 to 10/15/2020	26		-1.80	-1.02
	10/15 to 11/1/2020	30		-1.49	-0.73
	4/1 to 4/15/2021	8.8		-2.27	0.00
CC-A	5/15 to 6/1/2021	38	9	-0.11	-0.03
	8/1 to 8/15/2021	15		-0.18	-0.02
	8/15 to 9/1/2021	4		-0.11	-0.02
	9/1 to 9/15/2021	7		-0.13	-0.04
	9/15 to 10/1/2021	3		-0.13	-0.04
CC-B	9/15 to 10/1/2020	121	69	-3.17	0.02
	10/1 to 10/15/2020	104		-1.85	-0.61
	10/15 to 11/1/2020	53		-10.63	-0.58
	4/1 to 4/15/2021	49		-2.13	-1.13
	4/15 to 5/1/2021	49		-3.15	-1.58

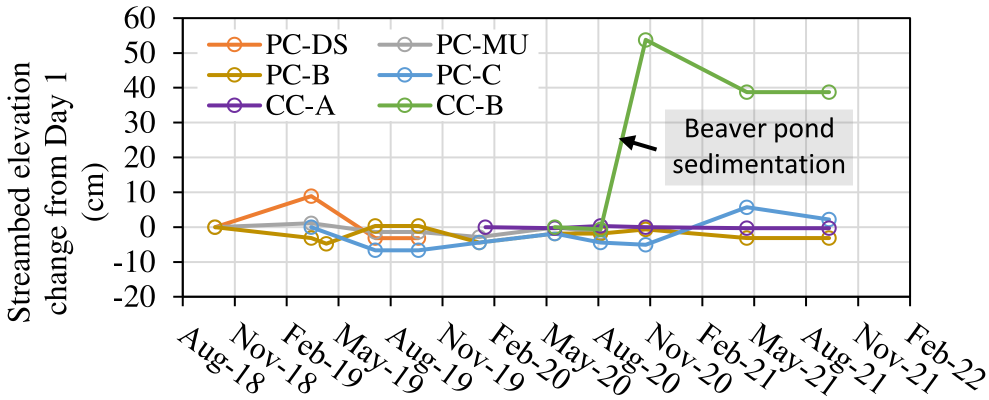

Figure 5. Field records of sediment erosion and deposition at each study site.

2.3.2 Streambed Fluxes

Summary statistics for streambed fluxes at all six monitoring stations are provided in Table 2. Between September 30, 2018, and September 30, 2021, streambed fluxes ranged from -12 to 0.62 m day^{-1} at study sites, with annual medians ranging from -4.0 to 0.23 m day^{-1} and MADs from 0.004 to 1.7 m day^{-1} (***Table 3***; ***Figure 6***). All study sites exhibited both positive and negative fluxes during the study period, indicating that transitions between losing and gaining conditions are common. While annual mean and median fluxes had the same direction and order of magnitude for individual sites, histograms of fluxes show skewed distributions (***Figure 6b***), with Fischer's skewness coefficients ranging from -1.3 to 0.97. CC-B shows a bimodal distribution, likely related to a beaver dam that was constructed nearby in October 2020. Skewness and bimodal distributions indicate that fluxes exhibit non-normal distributions through time, warranting the use of resistant statistical metrics. The magnitude of fluxes did not systematically increase or decrease with distance downstream, as has been documented by others (e.g., Hatch et al., 2010). This may be related to the relatively short length of study reaches. A similar statistical analysis and compilation were performed for the measured vertical hydraulic gradients (VHGs); those results are included in Table 4 but are not discussed herein.

Table 3. Summary statistics of streambed flux values (m day^{-1}). Daily, D, statistics (MAD, QCV) are calculated using the median flux values for each day of the water year. Sub daily, S-D, statistics (MAD, QCV), are calculated using 10-minute detrended fluxes.

						n		MAD			QCV			$\frac{MAD_{S-D}}{MAD_D}$
Site	year	Mean	Med	Min	Max	All/S-D	D	All	S-D	D	All	S-D	D	
CC-A	2	-0.0042	-0.0016	-0.21	0.17	35746	249	0.030	0.011	0.0241	37	14	38	0.46
CC-A	3	-0.044	-0.041	-0.23	0.058	52118	363	0.019	0.0065	0.019	0.96	0.34	0.96	0.35
CC-B	2	-3.5	-3.8	-9.0	0.016	8770	62	1.2	0.20	1.1	0.62	0.12	0.65	0.18
CC-B	3	-4.5	-4.0	-12	-0.34	51597	360	1.7	0.09	1.7	0.98	0.05	0.96	0.056
PC-DS	1	0.0009	0.0068	-0.098	0.056	52417	364	0.0074	0.0032	0.0046	2.6	1.3	2.2	0.70
PC-DS	2	-0.0029	-0.0039	-0.041	0.022	34620	241	0.0043	0.0010	0.0045	2.3	0.56	2.3	0.22
PC-MU	1	-0.067	-0.069	-0.33	0.096	52417	364	0.059	0.0074	0.054	1.7	0.26	1.7	0.14
PC-MU	2	0.018	0.011	-0.23	0.62	52551	365	0.048	0.0064	0.047	9.0	1.3	9.4	0.14
PC-MU	3	0.20	0.23	-0.15	0.53	51588	359	0.16	0.0044	0.16	1.6	0.04	1.6	0.027
PC-B	1	-0.89	-0.87	-4.0	-0.11	43168	306	0.28	0.045	0.28	0.63	0.12	0.65	0.16
PC-B	2	-1.0	-0.99	-1.8	-0.54	41180	291	0.18	0.065	0.18	0.36	0.16	0.36	0.36
PC-B	3	-0.79	-0.85	-2.7	0.14	50808	353	0.19	0.030	0.17	0.47	0.07	0.44	0.17
PC-C	1	-0.46	-0.45	-5.2	0.095	27781	193	0.14	0.042	0.13	0.61	0.22	0.60	0.32
PC-C	2	-0.75	-0.61	-2.1	-0.089	40245	283	0.30	0.037	0.28	1.3	0.13	1.3	0.13
PC-C	3	-2.1	-1.9	-5.0	0.40	38382	268	1.0	0.29	1.0	1.1	0.31	1.1	0.28

Table 4. Summary statistics of streambed vertical hydraulic gradients (VHGs) (-). Daily, D, statistics (MAD, QCV) are calculated using the median VHG values for each day of the water year. Sub daily, S-D, statistics (MAD, QCV) are calculated using 10-minute detrended VHGs.

						n		MAD			QCV			$\frac{MAD_{S-D}}{MAD_D}$
Site	year	Mean	Med	Min	Max	All/S-D	D	All	S-D	D	All	S-D	D	
CC-A	2	0.0003	0.00020	-0.028	0.040	35746	249	0.004	0.0016	0.004	44	16	39	0.44
CC-A	3	0.0065	0.0067	-0.0092	0.030	52118	363	0.0023	0.0010	0.0021	0.69	0.31	0.63	0.48
CC-B	2	0.056	0.061	-0.0003	0.13	8770	62	0.017	0.0022	0.0176	0.55	0.09	0.52	0.13
CC-B	3	0.082	0.083	0.0077	0.202	51597	360	0.025	0.0014	0.0250	0.59	0.04	0.60	0.06
PC-DS	1	-0.0015	-0.0046	-0.035	0.048	52417	364	0.0044	0.0020	0.0029	2.4	1.2	1.4	0.69
PC-DS	2	0.0022	0.0028	-0.012	0.022	34620	241	0.0030	0.0007	0.0032	2.3	0.54	2.3	0.22
PC-MU	1	0.012	0.012	-0.019	0.057	52417	364	0.012	0.0013	0.0119	2.1	0.28	2.0	0.11
PC-MU	2	-0.0028	-0.0019	-0.095	0.037	52551	365	0.010	0.0012	0.0091	9.7	1.3	9.3	0.13
PC-MU	3	-0.040	-0.048	-0.10	0.026	51588	359	0.024	0.0008	0.0237	1.4	0.0	1.4	0.03
PC-B	1	0.061	0.065	0.0076	0.23	43168	306	0.016	0.0022	0.0140	0.54	0.09	0.54	0.16
PC-B	2	0.069	0.070	0.037	0.12	41180	291	0.011	0.0030	0.0102	0.32	0.11	0.33	0.29
PC-B	3	0.058	0.065	-0.010	0.16	50808	353	0.017	0.0017	0.0154	0.58	0.05	0.57	0.11
PC-C	1	0.026	0.026	-0.0062	0.27	27781	193	0.006	0.0021	0.0056	0.48	0.19	0.42	0.38
PC-C	2	0.043	0.039	0.0056	0.10	40245	283	0.016	0.0016	0.0152	0.85	0.093	0.83	0.11
PC-C	3	0.12	0.12	-0.033	0.23	38382	268	0.050	0.015	0.0497	0.79	0.26	0.80	0.31

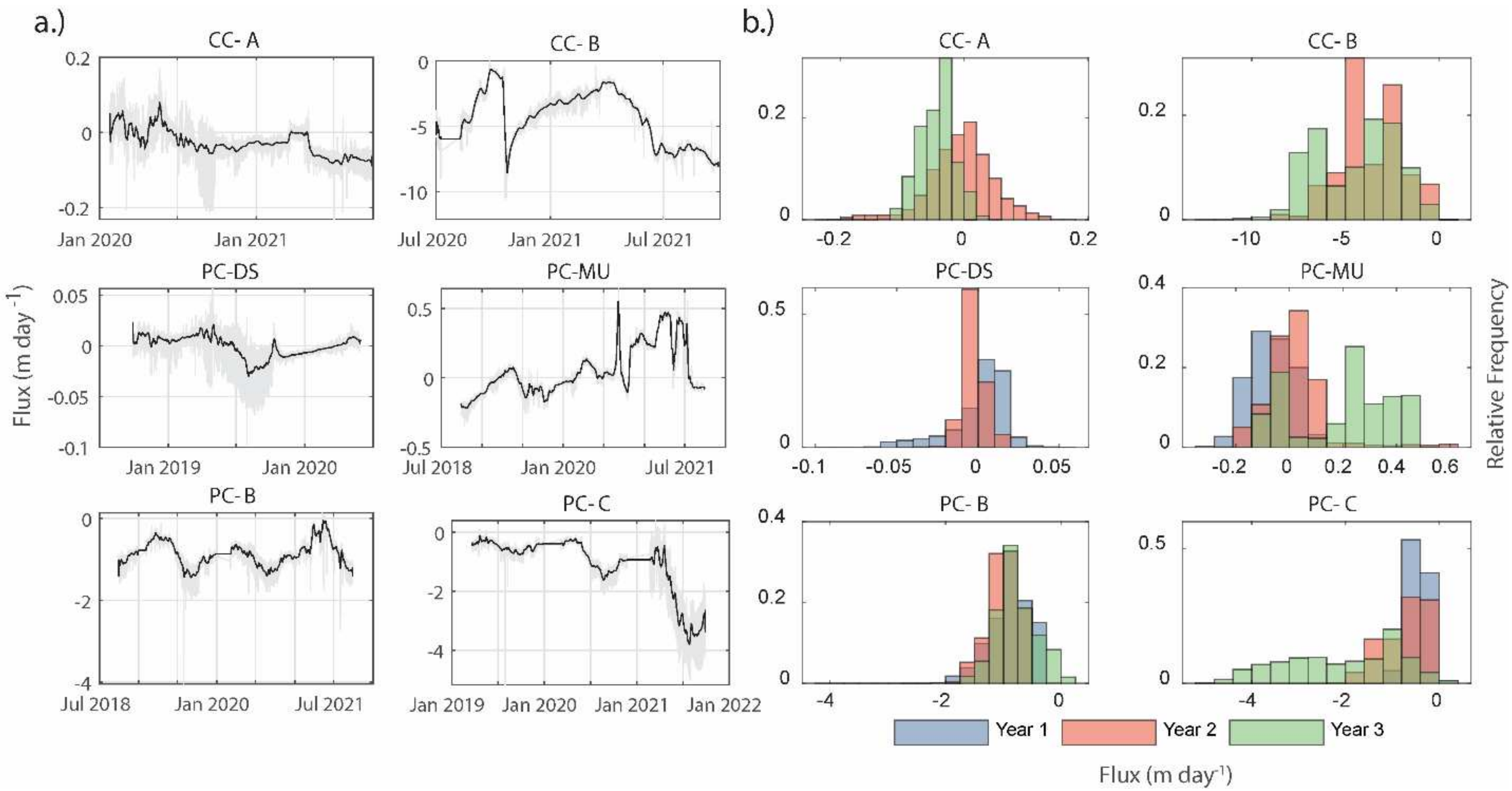

Figure 6. a.) Streambed fluxes for the study period with black line indicating 7-day moving average and gray line indicating 10-minute time series, and b.) relative frequency histograms for 10-minute time series.

2.3.3 Temporal Variability

2.3.3.1 Sub-daily

Sub-daily variations in fluxes were recorded at all study sites and primarily involved a regular daily rise and fall with minima occurring between 10:00 and 18:00, and maxima between 6:00 and 8:00. Sub-daily variations were generally greater during growing season months. ***Figure** 7* plots subsets of detrended fluxes, highlighting sub-daily patterns. The daily minimum detrended flux indicates either daily maximum stream loss or minimum groundwater discharge to streams, depending on whether the site is losing or gaining respectively. Notably, two sites with mean fluxes close to zero, PC-MU and PC-DS, exhibited sub-daily transitions from upwelling to downwelling conditions during spring months.

Median absolute deviations for yearly sets of detrended fluxes ($MAD_{S\text{-}D}$) ranged from 0.001 to 0.3 m day^{-1}. $MAD_{S\text{-}D}$ calculated for individual days ranged from <0.001 to 1.04 and exhibited seasonal trends, with greater values from July to September and lesser from December through February. $MAD_{S\text{-}D}$ was generally smaller for sites with fluxes close to zero, though $QCV_{S\text{-}D}$, which ranged from 0.04 to 14, showed the opposite relationship. This may be attributed to issues with the QCV for variables that cross zero, as discussed. However, I note that it also indicates that sub-daily variability does occur even for fluxes close to zero. For example, site PC-MU, which had median fluxes close to zero throughout the study period, recorded a daily range of fluxes between 0.007 and 0.3 m day^{-1}, with an average daily range of 0.05 m day^{-1}.

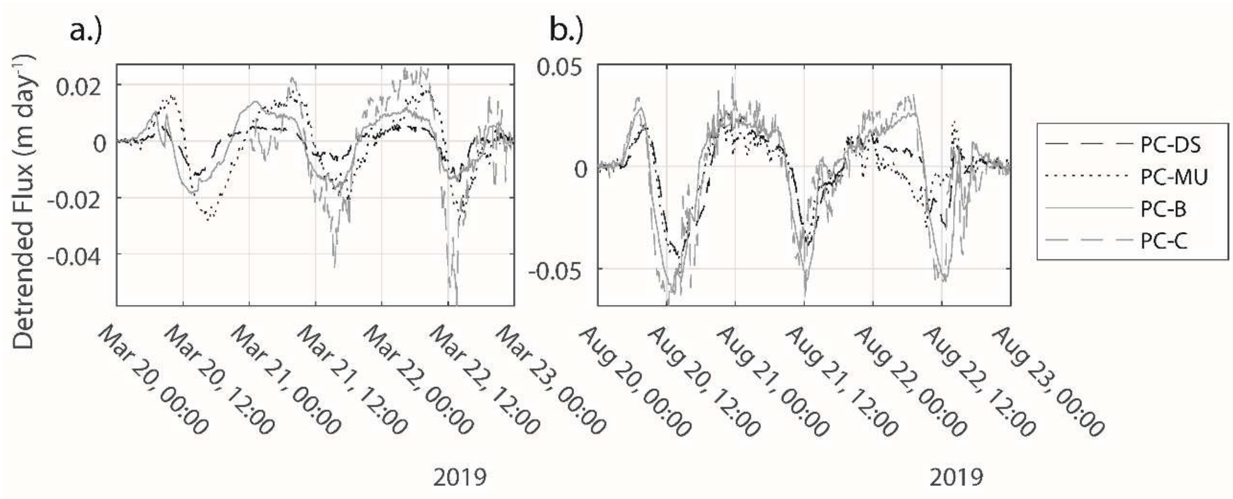

Figure 7. Detrended fluxes shown for representative 3-day periods during the (a) non-growing season and (b) growing season in 2019. Stations CC-A and CC-B were not installed yet.

2.3.3.2 Daily and Seasonal

Daily variations occurred as seasonal fluctuations around the annual median and event-based variations associated with stage change events (e.g., storm events and beaver dam construction). Median absolute deviations for yearly sets of daily median fluxes (MAD_D) ranged from 0.005 to 1.7 m day^{-1} across study sites and years, with more strongly losing sites having greater values. MAD_D ranged from 0.2x to 15x the median annual flux, with an average of 1.7x. Corresponding QCVs ranged from 0.36 to 38 and, like $QCV_{S\text{-}D}$, increased exponentially for annual median fluxes close to zero and decreased for greater magnitude fluxes. Daily statistics for MAD were greater than corresponding sub-daily statistics (based on detrended fluxes) at all sites (Table 2), indicating that daily variability, when calculated over an entire water year, was greater than sub-daily variability. For individual sites and water years, the ratio of $MAD_{S\text{-}D}$ to MAD_D ranged from 0.03 to 0.7, with an average of 0.25.

Seasonality was recorded as stronger losing and weaker gaining conditions during summer months (July-September), and weaker losing and stronger gaining conditions during winter and spring months (February – April). Day-of-year plots for fluxes (***Figure 8***) demonstrate seasonal signals across years which are apparent even when the magnitude of overall fluxes changed between years. More strongly losing sites showed stronger seasonal patterns (PC-B and PC-C) compared to weaker losing and gaining sites (PC-MU). Monthly MADs also exhibited seasonality, with summer months generally having greater values than winter months (***Figure 10***). Notably, this finding is less significant for VHGs. Seasonal mean and median fluxes are reported in ***Table 5***.

Daily variability also occurred due to events that triggered changes in stream stage. During the study period, sites PC-MU and CC-B were impacted by beaver dams that were constructed nearby on the channel, providing an unexpected opportunity to capture the impacts of dams on streambed exchange rates. The beaver dam at PC-MU was constructed in August of 2020 approximately 15m upstream from the PZ nest. The deep piezometer recorded a 0.1m rise in hydraulic head, accompanied by a 10x increase in upward streambed flux (0.06 to 0.6 m day^{-1}). Over the following month, the streambed flux magnitude decreased, eventually reaching values lower than pre-dam rates; this decline was accompanied by a decrease in both shallow and deep piezometer heads. After a month, heads and fluxes increased and remained higher relative to pre-dam rates. While the dam was not directly monitored, the decrease followed by increase in heads and flux in the weeks following dam creation indicates a possible breach and subsequent re-construction of the dam.

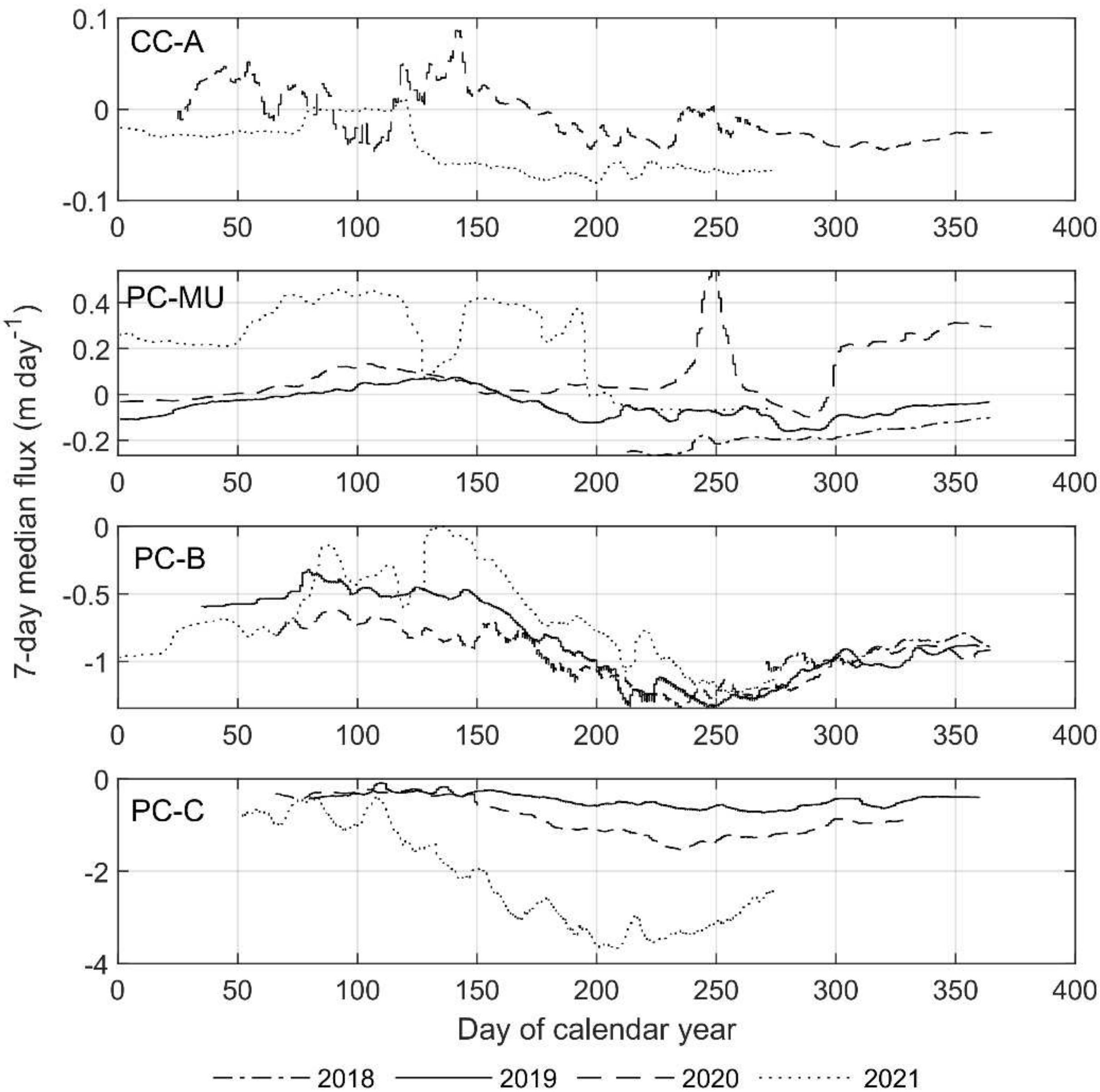

Figure 8. 7-day moving median flux plotted for day of year. Strong seasonality is recorded at sites PC-B and PC-C and weaker seasonality at CC-A and PC-MU.

Table 5. Seasonal and annual mean and median flux values (m day-1) calculated using 10-minute data sets.

		x̄			Median		
Location	Season	Y1	Y2	Y3	Y1	Y2	Y3
CC-A	Fall	-	-0.041	-0.041	-	-0.024	-0.037
CC-A	Winter	-	0.023	-0.028	-	0.026	-0.026
CC-A	Spring	-	0.011	-0.024	-	0.010	-0.011
CC-A	Summer	-	-0.025	-0.072	-	-0.015	-0.069
CC-A	**Annual**	**-**	**-0.0042**	**-0.044**	**-**	**-0.0016**	**-0.041**
CC-B	Fall	-	-0.80	-4.6	-	-0.71	-4.6
CC-B	Winter	-	-	-2.9	-	-	-3.0
CC-B	Spring	-	-	-2.7	-	-	-2.5
CC-B	Summer	-	-3.9	-6.7	-	-4.1	-6.9
CC-B	**Annual**	**-**	**-3.5**	**-4.5**	**-**	**-3.8**	**-4.0**
PC-DS	Fall	0.0052	-0.0075	-	0.0077	-0.0088	-
PC-DS	Winter	0.0071	-0.0043	-	0.0083	-0.0045	-
PC-DS	Spring	0.010	0.0043	-	0.013	0.0039	-
PC-DS	Summer	-0.013	-	-	-0.0083	-	-
PC-DS	**Annual**	**0.0009**	**-0.0029**	**-**	**0.0068**	**-0.0039**	**-**
PC-MU	Fall	-0.17	-0.090	0.14	-0.17	-0.088	0.23
PC-MU	Winter	-0.058	-0.012	0.29	-0.052	-0.020	0.26
PC-MU	Spring	0.044	0.093	0.36	0.048	0.099	0.44
PC-MU	Summer	-0.075	0.082	0.10	-0.075	0.033	-0.036
PC-MU	**Annual**	**-0.067**	**0.018**	**0.21**	**-0.069**	**0.011**	**0.24**
PC-B	Fall	-1.0	-1.1	-1.0	-0.93	-1.0	-0.99
PC-B	Winter	-0.65	-0.78	-0.81	-0.61	-0.78	-0.78
PC-B	Spring	-0.49	-0.78	-0.27	-0.49	-0.75	-0.28
PC-B	Summer	-1.1	-1.2	-0.88	-1.1	-1.2	-0.84
PC-B	**Annual**	**-0.89**	**-1.0**	**-0.79**	**-0.87**	**-0.99**	**-0.85**
PC-C	Fall	-0.76	-0.63	-1.3	-0.72	-0.56	-1.1
PC-C	Winter	-	-0.40	-0.85	-	-0.39	-0.74
PC-C	Spring	-0.313	-0.32	-1.3	-0.32	-0.30	-1.3
PC-C	Summer	-0.54	-1.2	-3.2	-0.53	-1.2	-3.2
PC-C	**Annual**	**-0.46**	**-0.75**	**-2.1**	**-0.45**	**-0.61**	**-1.9**

The beaver dam at CC-B was constructed in October of 2020 (***Figure 6*** and ***Figure 9***), approximately 3 m downstream from the PZ nest. The initiation of the dam was not observed during field visits, but the stilling well recorded a rapid stage increase of 0.45m accompanied by a 9x increase in downward fluxes over the span of a few days (-1.1 to -10.5 m day^{-1}), presumably signaling the construction of the dam. Over the following 7 months, fluxes gradually returned to rates recorded during the months preceding the dam. While the dam did not appear actively maintained, it remained in place for the duration of the study period and progressively accumulated sediment. During a site visit approximately one month after dam formation, 0.5m of accrued sediment was logged at the PZ nest. Six months after dam formation, 0.15m of the accrued sediment was eroded. Although fluxes returned to pre-dam rates, PZ heads remained elevated by about 0.5m. Beaver dams had a significant influence on fluxes at impacted sites, with up to a 9x increase in seepage (downwelling) at CC-B, and a 10x increase in groundwater discharge (upwelling) at PC-MU and impacts that persisted for several months. During the years when beaver dams were constructed nearby, both sites recorded less sub-daily variability, while CC-B recorded greater, and PC-MU recorded lesser daily variability.

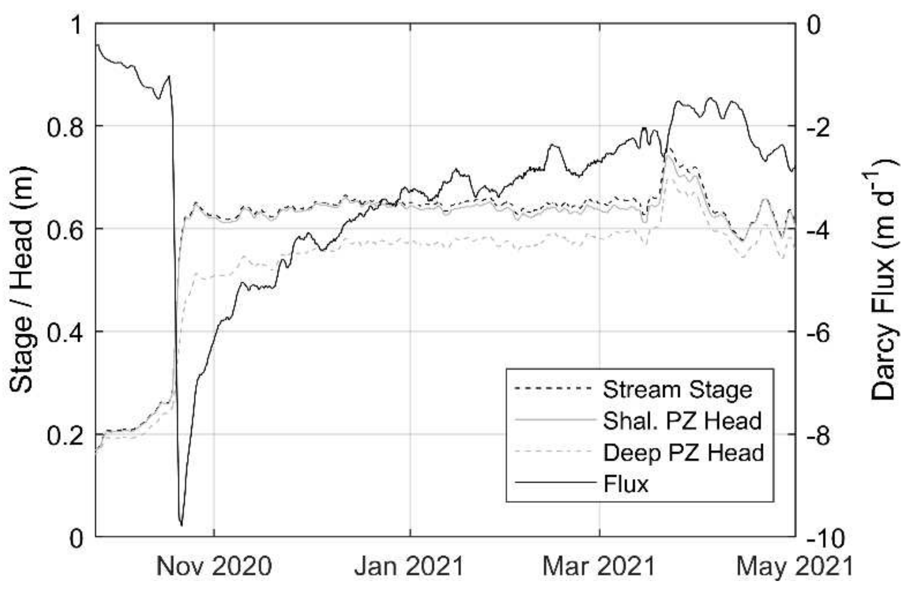

Figure 9. Groundwater head (relative to streambed datum), stream stage, and flux at site CC-B, highlighting the impacts of a beaver dam that formed 3m downstream from the nest in October 2020.

2.3.3.3 Interannual

Interannual variability was recorded as trends that persisted across years and non-trending variations between annual medians. Site PC-MU had progressively stronger upward fluxes throughout the study period, while sites PC-C, CC-A, CC-B, and PC-DS became more strongly losing. Site PC-B had greater fluxes during year 2 than during years 1 and 3. Interannual trends recorded in annual medians were generally reflected across all seasons, though winter fluxes tended to be most consistent across years. For example, 7-day moving median fluxes show site PC-C as progressively more losing during 2021 compared with 2020 (***Figure 8***). For most locations, annual and seasonal mean and median fluxes maintained the same order of magnitude across years. Exceptions to this included PC-MU, which had progressively stronger upward fluxes through time, and CC-A which had stronger seepage fluxes during year 3. The direction of interannual changes was not consistent across sites such that as some sites became more strongly losing across years, others became less strongly losing. Seasonal means and medians also showed variability across years, particularly at sites characterized by low-magnitude fluxes where it was common for seasonal medians to change 2 to 5-fold between years**.** For sites with variability between annual medians, one or two seasons typically accounted for most of the variability. For example, at PC-MU and CC-A, summer and fall fluxes had the greatest changes between years, while spring and winter fluxes stayed consistent.

2.4 Discussion

Weakly losing and gaining sites (e.g., PC-DS and PC-MU) had smaller MADs than strongly losing sites (e.g., CC-B), indicating that stronger downward fluxes were associated with greater variability. ***Figure 7*** highlights this finding with plots of MAD versus annual mean for fluxes and the absolute value of VHGs at each site showing significant negative (slope = -0.376, R^2 = 0.93, p<1e-8) and positive (slope = 0.28, R^2=0.74, p=<1e-4) relationships, respectively. Similarly, summer months were characterized by stronger losing fluxes (***Figure 8***) and greater MADs (***Figure 6***), indicating seasonal differences in the degree of variability. Other authors have also identified greater temporal variability at more strongly losing sites (Alexander and Caissie, 2003).

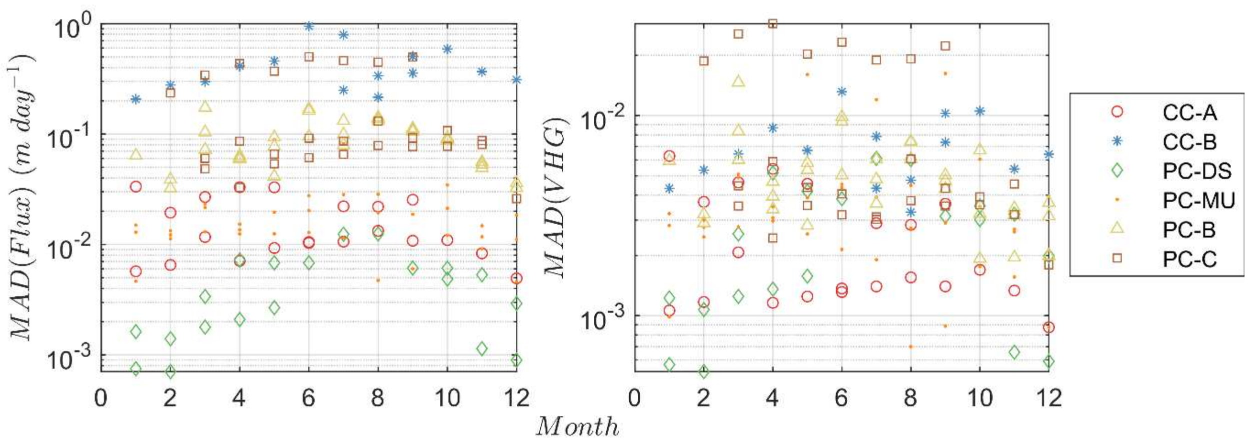

Figure 10. Monthly median absolute deviation (MAD) for flux and VHG. MAD values for each location/month were calculated using all (10-minute frequency) data; values are provided for multiple water years, where data availability permits.

Temperature related changes in K may partially explain the observed increase in variability with stronger downward fluxes. Losing streams typically have greater streambed temperature fluctuations at daily and seasonal timescales because advection drives temperature oscillations deeper into the streambed (Constantz, 2008). Because streambed K varies with temperature, and fluxes are greatly influenced by streambed K, strongly losing sites would inherently have more variability than weakly losing sites under variable temperature conditions. VHGs also exhibit increasing variability with magnitude (***Figure 11b***), but are not directly impacted by temperature related changes in K. The few sites with positive annual mean fluxes (negative VHGs) also record greater MADs, suggesting that variability may be inherently related to flux magnitude and not solely dependent on losing versus gaining conditions.

Annual median fluxes between years were generally within the same order of magnitude, whereas seasonal medians often showed variations spanning ± an order of magnitude. From this I conclude that variability within a given year may be greater than variability across years. Additionally, the relative consistency of annual medians despite seasonal variations suggests that stronger losing periods may be offset by gaining or weaker losing conditions to even out median seepage rates between years.

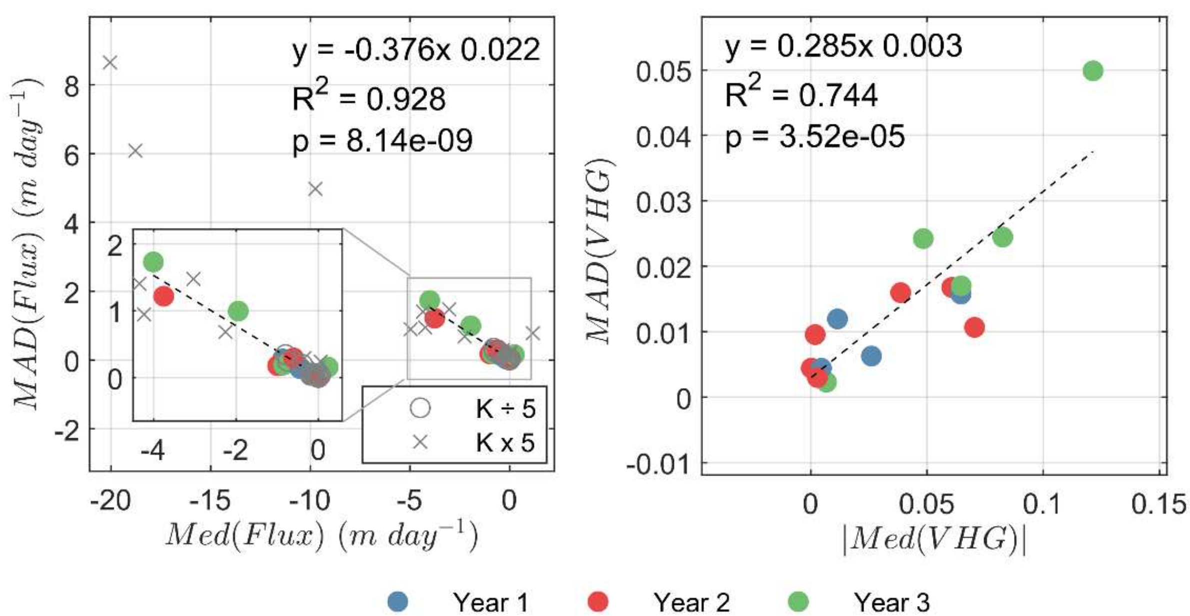

Figure 11. Median absolute deviation (MAD) versus annual median for a.) fluxes and b.) absolute value of VHGs.

2.4.1 Controls on temporal variability

Sub-daily variations in streambed flux were persistent year-round in the perennial streams that I studied. This topic deserves discussion as other authors have noted that causes for observed sub-daily variations at field sites are not fully understood (e.g., Arntzen et al., 2006;

Gribovski et al., 2010, Lautz, 2010). Here, I discuss three identified controls for sub-daily variations at our study sites: ET, temperature, and stage.

2.4.1.1 Evapotranspiration

Evapotranspiration contributes to sub-daily and seasonal variability in streambed fluxes as plants cycle through minimum and maximum groundwater uptake following diurnal and seasonal patterns related to solar energy and other factors. Phreatophytes (e.g., cottonwoods and willows) are common along riparian lowlands and floodplains where study reaches are located (Keith and Maberry, 1973) (Figure 1). ET-related effects typically cause maximum seepage losses in the late afternoon when plant water uptake peaks, primarily during growing season months (e.g., Gribovski et al., 2010; Yue et al., 2016). Daily maximum seepage losses occurred between 12:00 and 14:00 during growing season months. This timing is consistent with expected effects from ET. PZ nests located in phreatophyte zones (PC-DS and PC-MU) generally exhibited lower mean annual fluxes and temporal variability than other sites. This result is counterintuitive to expected findings because phreatophytes typically account for significant water uptake from shallow groundwater. I expect this result may be due to the upwelling conditions encountered at both PC-DS and PC-MU. Groundwater discharge may provide subsurface water that is necessary for phreatophytes to persist. That these sites record lesser temporal variability suggests either that upwelling sites are inherently less variability, or that ET is not the primary driver for recorded temporal variations.

2.4.1.2 Temperature effects

At losing-stream sites, cross correlations between seepage (negative flux) and stream temperature were used to identify lag times between maximum daily seepage and temperature.

Plots of cross-correlation coefficients for different time lags (i.e. cross-correlation functions) indiciate maximum correlations for lags of -1 to -2 hours during growing-season months, and ± 1 hour during non-growing season months (**Figure** ***12***). Maximum fluxes therefore occurred 1-2 hours before maximum stream temperatures during the growing season, corresponding to times when stream stage and shallow groundwater heads are at daily minima. During non-growing months (November-March), maximum daily seepage losses occurred between 14:00 and 18:00, within an hour of daily maximum stream temperatures. This suggests that temperature and streambed fluxes are strongly-linked year-round, and that related changes in K may be an important factor controlling sub-daily variability during winter months when ET is negligible.

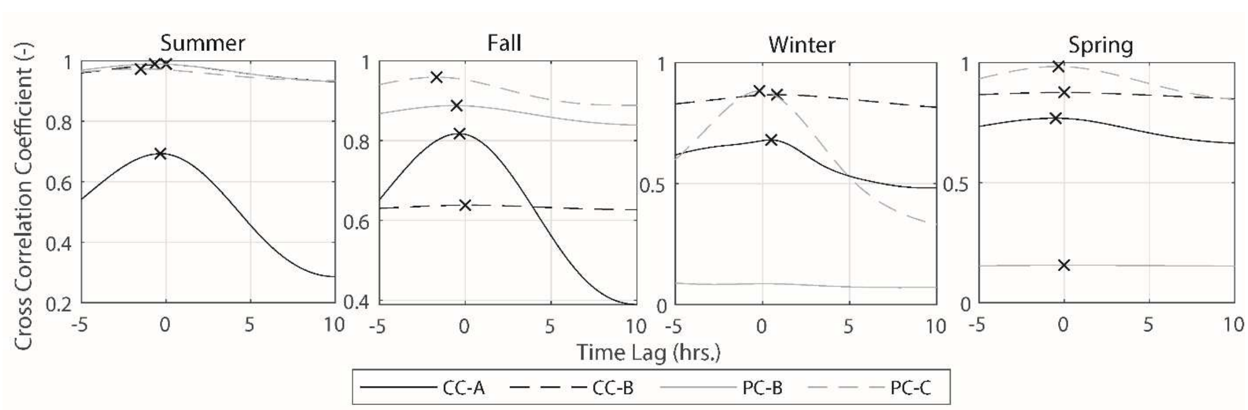

Figure 12. Cross correlation analysis of seepage (negative flux) and stream temperature using 10-minute data from summer 2020 through spring 2021. 'x' is the maximum cross correlation coefficient for lags <1 day, indicating hourly lag time between daily maximum flux.

Constantz et al. (1994) identified temperature-induced variations in K as the predominant control for sub-daily variations in seepage rates along ephemeral streams. Here, daily temperature ranges recorded in streambed sediments ranged from 0-5°C in the winter to 10-15°C in the summer, corresponding to daily changes in K of up to 10% and 30% during winter and summer months respectively. When temperature effects on K were removed from Darcy estimates, significant sub-daily variability remained, indicating that temperature induced variability only partially explains sub-daily variations in flux. Lautz (2012) similarly recorded sub-daily flux changes between 20-40% that could not be explained by temperature changes alone.

The diurnal stream temperature signal can be described as a quasi-sinusoidal wave that closely follows cycles of atmospheric temperature and solar radiation. As the stream temperature signal is transferred through the streambed via conduction and advection, it becomes dampened and lagged increasingly with depth, resulting in vertical streambed temperature profiles that vary through time and space (e.g., **Figure *13a***). Because streambed K is dependent on water temperature, K likewise exhibits spatiotemporal variations which results in variable configurations of K throughout the streambed profile, even in homogeneous systems. Spatially variable K has been shown to impact the response of fluxes to changes in stage or groundwater levels (e.g., Arntzen et al., 2006; Pryshlak et al., 2015), but effects of temporally varying K are less well understood.

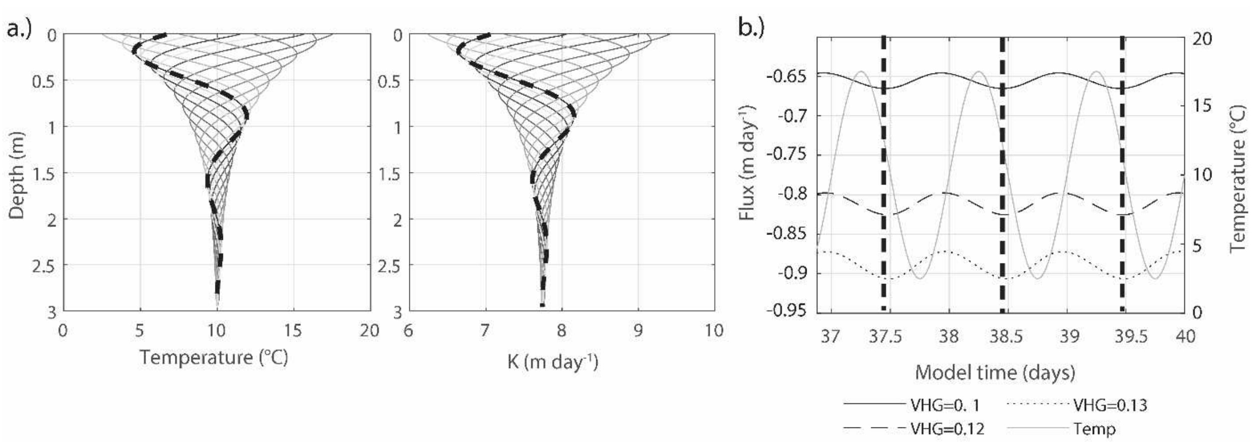

Figure 13. a.) Simulated streambed temperature and K profiles for 12 times (bihourly) during a 24-hour period and b.) simulated streambed fluxes for three different VHGs imposed by upper and lower constant head boundaries. Dashed lines represent time of maximum seepage.

As noted previously, daily temperature changes along PC and CC streams were especially large (up to 15°C), making this an opportune location to test the significance of large temperature fluctuations (and associated changes in K). Using the 1D numerical model described previously, I ran forward synthetic simulations to investigate the effects of a fluctuating temperature boundary condition on streambed fluxes. Model modifications included the assignment of upper and lower constant-head boundaries, specification of a constant temperature representing the average deep groundwater temperature at the lower boundary, and simulation of the upper temperature boundary as a sine wave with amplitudes between 2.5 and 7.5°C and mean of 10°C. Constant head boundaries enforced a target global VHG across the model domain.

Results from three simulations with varying VHG (0.1, 0.12, 0.13) are plotted in **Figure *13b***. Notably, despite constant head boundaries, model results show fluxes with sub-daily variations that include maximum downward fluxes approximately 4.8 hours after maximum upper boundary temperatures. Model results showed that sub-daily flux variations increased with increasing VHG and temperature range. Field data reveal daily maximum downward fluxes within an hour of daily maximum stream temperature during winter, and within 2 hours of maximum stream temperatures during summer (e.g., **Figure *12***), which did not align with results from this model. Notably, however, the field setting and calibration scenarios involved variable stream stage conditions, which was not considered in this analysis. With a modeled VHG of 0.1, daily flux ranges were between 0.02 and 0.03 m day^{-1}, which is substantially less than observed variations at sites with similar VHGs (e.g., 0.6-0.8 m day^{-1} at CC-B). Temperature related variations in K throughout the streambed profile are likely not a predominant cause for observed sub-daily variability. However, for other study sites with greater fluxes and VHGs, this may be an important control for explaining sub-daily flux variations.

2.4.1.3 Stage

Changes in stream stage (i.e., discharge) can drive water into and out of the streambed, causing diel signals in streambed fluxes that propagate into the aquifer (e.g., Arntzen et al., 2006; Loheide and Lundquist, 2009). Stilling wells recorded year-round sub-daily variations in stream stage, with typical daily ranges for sites between 1.5 and 4.5 cm. Arntzen et al. (2006) evaluated streambed flux variations driven by variable stage caused by hydropower generation. They identified a hysteretic relationship between river stage and VHG and connected observed effects to the increased lag in pressure propagation with depth. Using a similar methodology to Arntzen et al. (2006), I plotted VHGs vs stream stage with directional vectors connecting successive points to evaluate temporal relationships between stage and flux across seasons. Plots were constructed for three-day periods in August, October, January, and April, to consider changing dynamics throughout the year (**Figure *14*** and ***Appendix A, Figure A1***). VHGs are plotted rather than fluxes to avoid interpreting effects of temperature related changes in K.

Compared to Arntzen et al. (2006), our results show a similar hysteretic relationship between stage and VHG, particularly during growing-season months. The August and October plots record counterclockwise paths with maximum VHGs a few hours after peak stage. Note that VHG estimates are for 0.5m below the streambed so a lag is expected. Subsequently, stage begins to rise and comparatively lower VHGs are recorded for equivalent stage values. As the stream pulse propagates through the streambed, deeper groundwater heads rise while at the same time, stage begins to fall. Greater rates of ET occur as stage is falling, possibly enhancing the signal as the aquifer responds to plant-water uptake. The signal becomes less interpretable during winter months. January plots reveal a clockwise path, indicating that VHG and stage increase simultaneously. During winter months, daily stream stage changes are significantly resulting in lesser changes at depth. Additionally, during winter months, VHGs are overall less (lesser seepage). The apparent relationships between stage and VHG suggest that stage variations exert year-round controls for sub-daily flux variations.

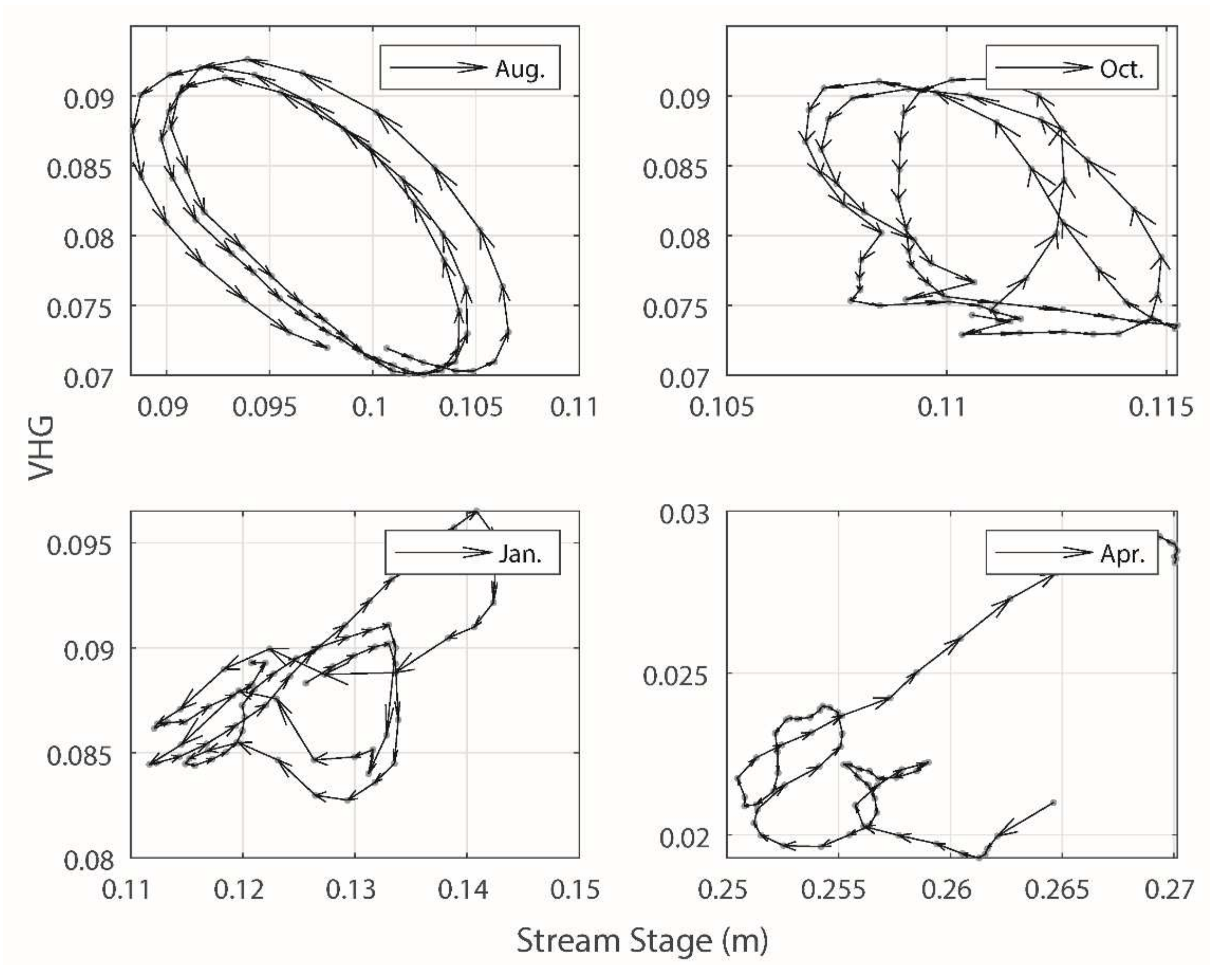

Figure 14. Comparison of vertical hydraulic gradient (VHG) and stream stage during three-day periods in August 2020, October 2020, January 2021, and April 2021 at location PC-B. Hourly vectors indicate paths through time.

2.4.1.4 Beaver Dams

The long-term nature of this study afforded the opportunity to study changing flux dynamics from beaver dam construction. Beaver dam effects imparted the greatest magnitude source of daily and interannual variability during the study period. Similar to other studies, I identified increased downwelling on the upstream side of dams, and increased upwelling on the downstream side of dams (e.g., Wade et al., 2020). Daily mean fluxes changed up- and downstream of dams by up to 1000% following dam construction, and annual mean fluxes were significantly affected during years with dams. Beaver dams may also cause reduced sub-daily variability at sites upstream of dams as indicated by lower sub-daily MADs compared to other years and sites. For streams prone to beaver dam formation where beavers construct and maintain dams, significant daily and interannual spatiotemporal flux variations should be expected. The impacts of beaver dams persisted in recorded fluxes for several months, indicating that future studies that assess the impacts of beaver dams may want to consider time-periods spanning months rather than days.

I also note that the active deposition and erosion of sediment associated with beaver dam formation complicates the application of temperature-based methods for estimating fluxes because sedimentation and erosion result in variable instrument depths relative to the streambed. While some authors have developed methods for tracking changing sensor depth with heat (e.g., Luce et al., 2013), uncertainty around sensor depths can lead to inaccurate flux estimates. Our setup consisted of welded piezometers which ensures that the distance between measurement points remains fixed through time, thereby maintaining accuracy for flux estimates in zones with sedimentation and erosion. Our setup can result in relative changes in flux measurement depths in cases of sedimentation and erosion, which could affect statistics from temporal variation analysis.

2.4.1.5 Watershed variables

To evaluate whether interannual variations correspond to changes in watershed variables, annual median fluxes and variability (MAD_D) were compared to mean annual stream discharge (USGS gage stations 06070900 and 06712000), water-year precipitation and snow totals for the headwater catchment (SNOTEL) and basin floor (Colorado Climate Center), evaporative demand (as evaporative demand drought index, EDDI; NOAA Physical Sciences Laboratory), and reference ET (CoAgMet) for the E. Plum and Cherry Creek watersheds. With only three years of data, relationships lacked statistical confidence and are therefore described qualitatively. For all watershed variables except headwater and basin precipitation and snow, no clear relationships or trends were identified. At five of the six sites, MAD_D for a given year positively correlated with headwater and valley floor precipitation. MAD_D was negatively correlated with annual basin snow; during the study period, years with greater basin precipitation typically corresponded to lesser snow. I expect that longer-term data sets and more sophisticated watershed models would provide greater clarity on observed patterns.

2.4.2 Uncertainty in streambed hydraulic conductivity

Like all geologic materials, the hydraulic conductivity of streambed sediments is spatially variable and is difficult to characterize. Because fluxes are linearly dependent on K, even minor inaccuracies in K estimates can result in significant errors in flux estimates. In this study, streambed hydraulic conductivities at each site were initially estimated by conducting falling-head tests at multiple locations surrounding the piezometer nest. Uncertainty in field-based K estimates result from heterogeneity in streambed sediments and spatial offsets between field tests and VHG calculation intervals. At four of the six monitoring sites, the conductivity was further refined through heat-transport modeling (i.e., a best-fitting K value was identified by comparing simulated and observed temperature time series). Model-calibrated K values are applicable to the subsurface interval between piezometer openings, which is consistent with the interval used to measure vertical hydraulic gradients. This consistency is a strength of the study approach; it reduces uncertainty in the reported flux values.

The potential for temporal changes in streambed K also introduces uncertainty into flux estimates. Erosion and deposition, biological activity, and entrapped gas have been known to cause changes in K which are equally difficult to characterize (e.g., Genereux et al., 2008). In this study, I used subsurface K values at four of the monitoring sites where sufficient data were available for heat-transport modeling. This subsurface interval is less affected by surficial changes associated with sediment aggradation and incision and is likely below the interval affected by biological activity (e.g., Song et al., 2010). K values were also estimated for multiple time periods, improving the temporal coverage of the estimation period.

Despite efforts to constrain the K value at each site, uncertainty persists. A sensitivity analysis was conducted to better understand how this uncertainty would impact the results and conclusions of this study. This involved recalculating fluxes and statistics for high- and low-K scenarios with K multiplied and divided by a factor of 5. The chosen sensitivity factor reflects the approximate degree of K-uncertainty attributed to temporal changes in a study by Genereux et al. (2008). Results from the sensitivity analysis show that the magnitude of the streambed flux, and therefore the median and MADs, are sensitive to the K value used in Darcy calculations. However, the relative comparison of variability (e.g., QCV and $MAD_{S\text{-}D}/MAD_D$) is not sensitive to the assumed K value. Table 3 presents a subset of the streambed flux summary statistics for base case, high-K, and low-K scenarios. Importantly, the QCV and MAD_{S-D}/MAD_D ratio for a particular year and site are not sensitive to the K value. Results of this uncertainty analysis are also shown in Figure 7. Greater mean fluxes show an increasingly greater MAD with a linear trend. Notably, the independence of this relationship from K is highlighted by the consistent finding for VHGs, with VHG and MAD_D having a positive linear relationship.

Table 3 – Summary statistics from K-uncertainty analysis

		Median			MAD			QCV			MAD_{-D}/MAD_D		
Site	Year	LK	K	HK	LK	K	HK	LK	K	HK	LK	K	HK
CC-A	2	-0.0003	-0.002	-0.008	0.012	0.06	0.3	37.9	37.4	40	0.46	0.46	0.46
CC-A	3	-0.008	-0.041	-0.21	0.0079	0.04	0.2	0.96	0.96	0.96	0.35	0.35	0.35
CC-B	2	-0.75	-3.8	-19	0.47	2.3	11.7	0.62	0.62	0.62	0.18	0.18	0.18
CC-B	3	-0.80	-4.0	-20	0.78	3.9	19.6	0.98	0.98	0.98	0.06	0.06	0.06
PC-DS	1	0.001	0.007	0.034	0.0035	0.018	0.09	2.58	2.59	2.5	0.67	0.70	0.70
PC-DS	2	-0.001	-0.004	-0.020	0.0018	0.009	0.05	2.31	2.33	2.25	0.22	0.22	0.23
PC-MU	1	-0.014	-0.069	-0.34	0.024	0.12	0.59	1.71	1.72	1.72	0.14	0.14	0.14
PC-MU	2	0.002	0.011	0.053	0.019	0.094	0.47	9.0	9.0	9.0	0.14	0.14	0.14
PC-MU	3	0.046	0.23	1.2	0.076	0.38	1.89	1.63	1.63	1.63	0.03	0.03	0.03
PC-B	1	-0.17	-0.87	-4.3	0.11	0.55	2.8	0.63	0.63	0.63	0.16	0.16	0.16
PC-B	2	-0.20	-0.99	-5.0	0.072	0.36	1.8	0.36	0.36	0.36	0.36	0.36	0.36
PC-B	3	-0.17	-0.85	-4.2	0.080	0.40	2	0.47	0.47	0.47	0.17	0.17	0.17
PC-C	1	-0.090	-0.45	-2.3	0.055	0.28	1.4	0.61	0.61	0.61	0.32	0.32	0.32
PC-C	2	-0.12	-0.61	-3.0	0.15	0.76	3.8	1.26	1.26	1.26	0.13	0.13	0.13
PC-C	3	-0.39	-1.9	-9.7	0.41	2.1	10.3	1.05	1.05	1.05	0.28	0.28	0.28

Notes: K = Original K (Table 2), HK = High K (K x 5), LK = Low K (K/5); median, MAD, and QCV presented for all data (10-minute frequency).

Further, I demonstrate that significant temporal variations occur as a result of ET, stage, and temperature even without considering potential transience in the intrinsic permeability. Significant variability was recorded in VHGs, which are independent of K estimates. The timescales for which I summarize variability have relevant implications for measuring fluxes in the field and quantifying uncertainty. At individual sites, sub-daily variability (MAD_{S-D}) spanned up to 70% of daily variability at sites (MAD_{S-D}/MAD_D), indicating sub-daily variability can be significant. The broad daily range of fluxes for site PC-MU demonstrates how sub-daily variations have implications for field measurements and data collection.

2.4.3 Implications of variability for groundwater recharge

Understanding temporal variations in streambed flux is important for predicting annual groundwater recharge rates under changing climate scenarios. Future changes in temperature and precipitation are expected to significantly impact the timing and magnitude of streamflow in Rocky Mountain region (Ray et al., 2008). Stream seepage is often the predominant mechanism of recharge for mountain-front aquifers in arid and semi-arid regions (Wilson and Guan, 2004). Understanding seasonality in recharge rates and sources is essential for predicting how rates of groundwater recharge will respond to climate change (Jasechko et al., 2014).

Our results indicate significant seasonality in seepage rates at some, but not all, locations, characterized by stronger seepage during summer months and weaker seepage during winter months. Additionally, three sites, CC-A, PC-MU, and PC-DS exhibited seasonal transitions from gaining to losing conditions, with gaining conditions typically occurring during spring months and losing during late summer through late winter. Numerical integration was performed to estimate yearly seepage recharge for 7-day moving average fluxes at the three locations that transitioned seasonally. Results show how seasonal recharge contributions during late summer and fall are offset by groundwater discharge during winter and spring months (**Figure *14***). PC-MU is losing for almost half the year despite having net groundwater discharge to streams.

Conversely, PC-DS is characterized by gaining conditions for most of the year, but this site has a net seepage loss (i.e., surface water losses exceed gains) due to a strong losing event that started in June 2019. These results demonstrate how year-round study periods are essential to characterizing annual dynamics of groundwater recharge. Future studies that assess site-specific changes in recharge due to changing climates will benefit from long-term characterizations of streambed fluxes to capture changing seasonality through time.

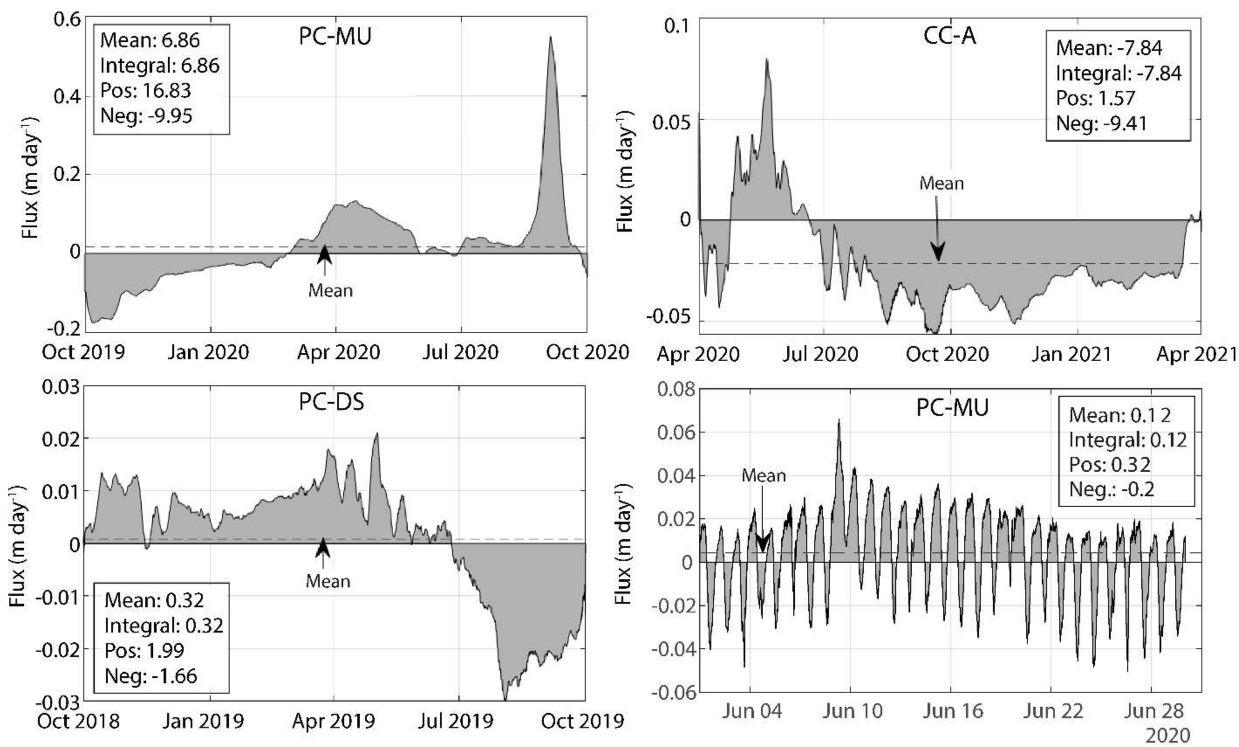

Figure 15. Influence of flux variability on groundwater recharge (first three panels) and hyporheic exchange (last panel, lower right) for representative time periods. The mean streambed flux is indicated by dashed line. Boxes list volumes of water exchanged per area of streambed (m3 m-2) during the plotted time period; value obtained using the mean flux is compared to the integrated value (shaded area). Total positive and negative volumes are also reported.

2.5 Implications of variability for hyporheic exchange

Hyporheic exchange describes the process by which stream water infiltrates the subsurface and returns to the stream over relatively short distances. Hyporheic flow paths are typically conceptualized as zones of downwelling and upwelling connected by horizontal flow paths. Our results record regular flow reversals at multiple locations, with two locations (PC-DS and PC-MU) recording daily transitions between upwelling and downwelling conditions during spring months. Others have documented seasonal and event-related transitions between losing and gaining conditions at individual sites (e.g., Essaid et al, 2008; Anibas et al., 2016), though few have documented a daily flow reversal. Despite having mean fluxes close to zero, these locations record potential for considerable rates of vertical exchange. Site PC-MU, for example, records 4-times the volume of exchanged water when calculated with sub-daily Darcy fluxes compared to the exchange volume obtained with the mean of fluxes for a representative 30-day period during June of 2020 (e.g., fourth panel in **Figure *15***). Considerations of sub-daily variability may be important when evaluating hyporheic processes, particularly along streams that are not strongly gaining or losing.

2.6 Conclusions

I monitored streambed fluxes in two perennial mountain-front streams over a 3-year study period and found significant temporal variability in fluxes at sub-daily, daily, and interannual scales. Seasonal and event related changes in flux were found to contribute the greatest overall variability, with daily variability (MAD_D values) on average 1.7 times but reaching up to 15 times the median annual flux at study sites. Seasonality was stronger at some sites than others but was consistently recorded as stronger losing and weaker gaining conditions during summer months, and weaker losing and stronger gaining conditions during winter and spring months. Sub-daily variability was also considerable; the average $MAD_{S\text{-}D}/MAD_D$ ratio across all sites/water years considered was 0.25. The magnitude and direction of mean annual fluxes stayed relatively consistent across years at our study sites. Causes for sub-daily variations were further explored, and linked to variable stage, evapotranspiration, and temperature-related changes in K. By quantifying the magnitude of temporal variations, I demonstrate how temporal changes may contribute similar magnitudes of variation and uncertainty as their more heavily researched spatial counterparts.

2.7 References

Alexander, M. D., Caissie, D. (2003). Variability and comparison of hyporheic water temperatures and seepage fluxes in a small Atlantic salmon stream. Ground Water, 41(1), 72-82.

Anibas, C., Schneidewind, U., Vandersteen, G., Joris, I., Seuntjens, P., Batelaan, O. (2016). From streambed temperature measurements to spatial-temporal flux quantification: using the LPML method to study groundwater–surface water interaction. Hydrological Processes, 30(2), 203-216.

Argerich, A., Martí, E., Sabater, F., Ribot, M. (2011). Temporal variation of hydrological exchange and hyporheic biogeochemistry in a headwater stream during autumn. Journal of the North American Benthological Society, 30(3), 635-652

Arntzen, E. V., Geist, D. R., Dresel, P. E. (2006). Effects of fluctuating river flow on groundwater/surface water mixing in the hyporheic zone of a regulated, large cobble bed river. River Research and Applications, 22(8), 937-946.

Barkmann, P. E., F. Scot Fitzgerald, L. A. Sebol, W. Curtiss, J. Pike, Annette Moore, and B. Taylor. "OF-15-10 Geology and Groundwater Resources of Douglas County." Geology and Groundwater. Open File Reports. Golden, CO: Colorado Geological Survey, 2015. https://coloradogeologicalsurvey.org/publications/geology-groundwater-resources-douglas-colorado.

Boano, F., Demaria, A., Revelli, R., Ridolfi, L. (2010). Biogeochemical zonation due to intra-meander hyporheic flow, Water Resources Research, 46(2) W02511, 10.1029/2008WR007583

Boulton, A. J., Foster, J. G. (1998). Effects of buried leaf litter and vertical hydrologic exchange on hyporheic water chemistry and fauna in a gravel-bed river in northern New South Wales. Australia. Freshwater Biology, 40(2), 229-243.

Briggs, M. A., Lautz, L. K., McKenzie, J. M., Gordon, R. P., Hare, D. K. (2012). Using high-resolution distributed temperature sensing to quantify spatial and temporal variability in vertical hyporheic flux. Water Resources Research, 48(2).

Briggs, M. A., Wang, C., Day-Lewis, F. D., Williams, K. H., Dong, W., Lane, J. W. (2019). Return flows from beaver ponds enhance floodplain-to-river metals exchange in alluvial mountain catchments. Science of the Total Environment, 685, 357-369.

Buffington, J. M., Tonina, D. (2009). Hyporheic exchange in mountain rivers II: Effects of channel morphology on mechanics, scales, and rates of exchange. Geography Compass, 3(3), 1038-1062.

Caissie, D. (1991). The importance of groundwater to fish habitat: Base flow characteristics for three Gulf Region rivers. Can. Data Rep. Fish. Aquat. Sci., (814), 30.

Caissie, D., Kurylyk, B. L., St-Hilaire, A., El-Jabi, N., MacQuarrie, K. T. (2014). Streambed temperature dynamics and corresponding heat fluxes in small streams experiencing seasonal ice cover. Journal of Hydrology, 519, 1441-1452.

Cardenas, M. B. (2010). Thermal skin effect of pipes in streambeds and its implications on groundwater flux estimation using diurnal temperature signals. Water Resources Research, 46(3).

Cardenas, M. B., Wilson, J. L., Haggerty, R. (2008). Residence time of bedform-driven hyporheic exchange. Advances in Water Resources, 31(10), 1382-1386.

Cardenas, M. B., Wilson, J. L., Zlotnik, V. A. (2004). Impact of heterogeneity, bed forms, and stream curvature on subchannel hyporheic exchange. Water Resources Research, 40(8).

Cermác, V., Rybach, L. (1982). Thermal properties: Thermal conductivity and specific heat of minerals and rocks. Landolt-Börnstein Zahlenwerte und Functionen aus Naturwissenschaften und Technik, Neue Serie, Physikalische Eigenschaften der Gesteine, 305-343.

Cheng, D., Song, J., Wang, W., Zhang, G. (2019). Influences of riverbed morphology on patterns and magnitudes of hyporheic water exchange within a natural river confluence. Journal of hydrology, 574, 75-84.

Clauser, C., and Huenges, E. (2013). Thermal Conductivity of Rocks and Minerals: AGU Reference Shelf. 105–126.

Cognac, K. E., Ronayne, M. J. (2020). Changes to inter-aquifer exchange resulting from long-term pumping: implications for bedrock groundwater recharge. Hydrogeology Journal, 28(4), 1359-1370.

Conant Jr., B. (2004). Delineating and quantifying ground water discharge zones using streambed temperatures. Groundwater, 42(2), 243-257.

Constantz, J. (1982). Temperature dependence of unsaturated hydraulic conductivity of two soils. Soil Science Society of America Journal, 46(3), 466-470.

Constantz, J., Thomas, C. L., Zellweger, G. (1994). Influence of diurnal variations in stream temperature on streamflow loss and groundwater recharge. Water Resources Research, 30(12), 3253-3264.

Constantz, J. (2008). Heat as a tracer to determine streambed water exchanges. Water Resources Research, 44(4).

Essaid, H. I., Zamora, C. M., McCarthy, K. A., Vogel, J. R., Wilson, J. T. (2008). Using heat to characterize streambed water flux variability in four stream reaches. Journal of environmental quality, 37(3), 1010-1023.

Fritz, B. G., Arntzen, E. V. (2007). Effect of rapidly changing river stage on uranium flux through the hyporheic zone. Groundwater, 45(6), 753-760.

Genereux, D. P., Leahy, S., Mitasova, H., Kennedy, C. D., Corbett, D. R. (2008). Spatial and temporal variability of streambed hydraulic conductivity in West Bear Creek, North Carolina, USA. Journal of Hydrology, 358(3-4), 332-353.

Gilmore, T.E., Genereux, D.P., Solomon, D. K., Solder, J. E., Kimball, B. A., Mitasova, H., Birgand, F. (2016). Quantifying the fate of agricultural nitrogen in an unconfined aquifer: Stream-based observations at three measurement scales. Water Resources Research, 52(3), 1961-1983.

Gooseff, M. N., Anderson, J. K., Wondzell, S. M., LaNier, J., Haggerty, R. (2006). A modelling study of hyporheic exchange pattern and the sequence, size, and spacing of stream bedforms in mountain stream networks. Oregon, USA. Hydrological Processes, 20(11), 2443-2457.

Goto, S., Yamano, M., Kinoshita, M. (2005). Thermal response of sediment with vertical fluid flow to periodic temperature variation at the surface. J. Geophys. Res., 110, B01106.

Gribovszki, Z., Szilágyi, J., Kalicz, P. (2010). Diurnal fluctuations in shallow groundwater levels and streamflow rates and their interpretation–A review. Journal of Hydrology, 385(1-4), 371-383.

Griglio, F. P., Tonina, D., Luce, C. H. (2013). Spatiotemporal variability of hyporheic exchange through a pool-riffle-pool sequence, Water Resources Research, 49, 7185–7204.

Hamdhan, I. N., Clarke, B. G. (2010). Determination of thermal conductivity of coarse and fine sand soils. Proceedings of World Geothermal Congress (1-7).

Hancock, P. J., Boulton, A. J., Humphreys, W. F. (2005). Aquifers and hyporheic zones: towards an ecological understanding of groundwater. Hydrogeology Journal, 13(1), 98-111.

Harvey, J. W., Bencala, K. E. (1993). The effect of streambed topography on surface-subsurface water exchange in mountain catchments. Water Resources Research, 29(1), 89-98.

Hatch, C. E., Fisher, A. T., Revenaugh, J. S., Constantz, J., Ruehl, C. (2006). Quantifying surface water–groundwater interactions using time series analysis of streambed thermal records: Method development. Water Resources Research, 42(10).

Hatch, C., Fisher, A. T., Ruehl, C., Stemler, G. (2010). Spatial and temporal variations in streambed hydraulic conductivity quantified with time-series thermal methods, J. Hydrol., 389, 276–288.

Hayashi M, Rosenberry, D. O. (2002) Effects of ground water exchange on the hydrology and ecology of surface water. Ground Water 40:309–316

Healy, R. W. (2010). Estimating groundwater recharge. Cambridge university press.

Healy, R. W., Ronan, A. D. (1996). Documentation of computer program VS2DH for simulation of energy transport in variably saturated porous media—Modification of the US Geological Survey's computer program VS2DT. Water-Resources Investigations Report, 96, 4230.

Healy, R.W, 1990, Simulation of solute transport in variably saturated porous media with supplemental information on modifications to the U.S. Geological Survey's computer program VS2D: U.S. Geological Survey Water-Resources Investigations Report, 90-4025, 125 p.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chap. A3, 458 p., https://doi.org/10.3133/tm4a3.

Hester, E. T., Young, K. I., Widdowson, M. A. (2013). Mixing of surface and groundwater induced by riverbed dunes: Implications for hyporheic zone definitions and pollutant reactions. Water Resources Research, 49(9), 5221-5237.

Hvorslev, M. J. (1951). Time lag and soil permeability in ground-water observations (No. 36). Waterways Experiment Station, Corps of Engineers, US Army.

Jasechko, S., Birks, S. J., Gleeson, T., Wada, Y., Fawcett, P. J., Sharp, Z. D., ... & Welker, J. M. (2014). The pronounced seasonality of global groundwater recharge. Water Resources Research, 50(11), 8845-8867.

Jones, J. B., Mulholland, P. J. (1999). Streams and ground waters. Elsevier.

Kalbus, E., Reinstorf, F., Schirmer, M. (2006). Measuring methods for groundwater–surface water interactions: a review. Hydrology and Earth System Sciences, 10(6),873-887.

Kasahara, T., Hill, A. R. (2006). Hyporheic exchange flows induced by constructed riffles and steps in lowland streams in southern Ontario, Canada. Hydrological Processes, 20(20), 4287-4305.

Käser, D. H., Binley, A., Heathwaite, A. L., Krause, S. (2009). Spatio-temporal variations of hyporheic flow in a riffle-step-pool sequence. Hydrological Processes, 23(15), 2138-2149.

Keery, J., Binley, A., Crook, N., Smith, J. W. (2007). Temporal and spatial variability of groundwater–surface water fluxes: Development and application of an analytical method using temperature time series. Journal of Hydrology, 336(1-2), 1-16.

Kennedy, C. D., Genereux, D. P., Mitasova, H., Corbett, D. R., Leahy, S. (2008). Effect of sampling density and design on estimation of streambed attributes. Journal of Hydrology, 355(1-4), 164-180.

Kennedy, C. D., Genereux, D. P., Corbett, D. R., & Mitasova, H. (2009). Spatial and temporal dynamics of coupled groundwater and nitrogen fluxes through a streambed in an agricultural watershed. Water Resources Research, 45(9).

Keith, J. R., Maberry, J. (1973). Vegetation map of the Parker Quadrangle, Arapahoe and Douglas counties, Colorado. U.S. Geological Survey Map I-770-N.

Landon, M. K., Rus, D. L., Harvey, F. E. (2001). Comparison of instream methods for measuring hydraulic conductivity in sandy streambeds. Groundwater, 39(6), 870-885.

Lappala, E. G., Healy, R. W., Weeks, E. P. (1987). Documentation of computer program VS2D to solve the equations of fluid flow in variably saturated porous media (Vol. 83, No. 4099). Department of the Interior, US Geological Survey.

Lautz, L. K. (2010). Impacts of nonideal field conditions on vertical water velocity estimates from streambed temperature time series. Water Resources Research, 46(1).

Lautz, L. K. (2012). Observing temporal patterns of vertical flux through streambed sediments using time-series analysis of temperature records. Journal of hydrology, 464, 199-215.

Lewandowski, J., Putschew, A., Schwesig, D., Neumann, C., Radke, M. (2011). Fate of organic micropollutants in the hyporheic zone of a eutrophic lowland stream: Results of a preliminary field study. Science of the Total Environment, 409(10), 1824-1835.

Livers, B., & Wohl, E. (2015). An evaluation of stream characteristics in glacial versus fluvial process domains in the Colorado Front Range. Geomorphology, 231, 72-82.

Loheide, S. P., II, Lundquist, J. D. (2009). Snowmelt-induced diel fluxes through the hyporheic zone, Water Resources Research, 45, W07404.

Luce, C. H., Tonina, D., Gariglio, F., & Applebee, R. (2013). Solutions for the diurnally forced advection-diffusion equation to estimate bulk fluid velocity and diffusivity in streambeds from temperature time series. Water Resources Research, 49(1), 488-506.

Malcolm, I. A., Middlemas, C. A., Soulsby, C., Middlemas, S. J., & Youngson, A. F. (2010). Hyporheic zone processes in a canalised agricultural stream: implications for salmonid embryo survival. Fundamental and applied limnology, 176(4), 319.

Matthai, H. F. (1969). Floods of June 1965 in South Platte River Basin, Colorado. Geological Survey Water-Supply Paper, 1850-B.

Palkovics, W. E., Petersen, G. W., Matelski, R. P. (1975). Perched water table fluctuation compared to streamflow. Soil Science Society of America Journal 39(2),343-348.

Pryshlak, T. T., Sawyer, A. H., Stonedahl, S. H., Soltanian, M. R. (2015). Multiscale hyporheic exchange through strongly heterogeneous sediments. Water Resources Research, 51(11), 9127-9140.

Rau, G. C., Cuthbert, M. O., McCallum, A. M., Halloran, L. J., Andersen, M. S. (2015). Assessing the accuracy of 1-D analytical heat tracing for estimating near-surface sediment thermal diffusivity and water flux under transient conditions. Journal of Geophysical Research: Earth Surface, 120(8), 1551-1573.

Ray, A. J., Barsugli, J. J., Averyt, K. B., Wolter, K., Hoerling, M., Doesken, N., ... & Webb, R. S. (2008). Climate change in Colorado: a synthesis to support water resources management and adaptation. Colorado Water Conservation Board Rep, 52.

Rosenberry, D. O., LaBaugh, J. W. (2008). Field techniques for estimating water fluxes between surface water and ground water. U.S. Geological Survey. 4-D2.

Rosenberry, D. O., Pitlick, J. (2009) Local-scale variability of seepage and hydraulic conductivity in a shallow gravel-bed river. Hydrological Processes, 23(23), 3306-3318.

Scanlon, B. R., Healy, R. W., Cook, P.G. (2002). Choosing appropriate techniques for quantifying groundwater recharge. Hydrogeology journal, 10(1), 18-39.

Sebok, E., Duque, C., Engesgaard, P., Boegh, E. (2015). Spatial variability in streambed hydraulic conductivity of contrasting stream morphologies: channel bend and straight channel. Hydrological Processes, 29(3), 458-472.

Sengers, J. V., Watson, J. T. R. (1986). Improved international formulations for the viscosity and thermal conductivity of water substance. Journal of physical and chemical reference data, 15(4), 1291-1314.

Shanafield, M., Cook, P. G. (2014). Transmission losses, infiltration and groundwater recharge through ephemeral and intermittent streambeds: A review of applied methods. Journal of Hydrology, 511, 518-529.

Simpson, S. C., Meixner, T. (2012). Modeling effects of floods on streambed hydraulic conductivity and groundwater-surface water interactions. Water resources research, 48(2).

Solomon, D. K., Gilmore, T. E., Solder, J. E., Kimball, B., Genereux, D. P. (2015). Evaluating an unconfined aquifer by analysis of age-dating tracers in stream water. Water Resources Research, 51(11), 8883-8899.

Song, J., Chen, X., & Cheng, C. (2010). Observation of bioturbation and hyporheic flux in streambeds. Frontiers of Environmental Science & Engineering in China, 4(3), 340-348.

Song, J., Zhang, G., Wang, W., Liu, Q., Jiang, W., Guo, W., Tang, B., Bai, H. and Dou, X. (2017). Variability in the vertical hyporheic water exchange affected by hydraulic conductivity and river morphology at a natural confluent meander bend. Hydrological Processes, 31(19), 3407-3420.

Stallman, R. W. (1965). Steady one-dimensional fluid flow in a semi-infinite porous medium with sinusoidal surface temperature. Journal of geophysical Research, 70(12), 2821-2827.

Storey, R. G., Howard, K. W., Williams, D. D. (2003). Factors controlling riffle-scale hyporheic exchange flows and their seasonal changes in a gaining stream: A three-dimensional groundwater flow model. Water Resources Research, 39.2.

Surridge, B. W., Baird, A. J., Heathwaite, A. L. (2005). Evaluating the quality of hydraulic conductivity estimates from piezometer slug tests in peat. Hydrological Processes: An International Journal, 19(6), 1227-1244.

Tonina, D., Buffington, J. M. (2009). A three-dimensional model for analyzing the effects of salmon redds on hyporheic exchange and egg pocket habitat. Canadian Journal of Fisheries and Aquatic Sciences, 66(12), 2157-2173.

U.S. Geological Survey (2022). National Water Information System data available on the World Wide Web (USGS Water Data for the Nation), accessed [January, 2022], at URL [http://waterdata.usgs.gov/nwis/].

Van Grinsven, M., Mayer, A., Huckins, C. (2012). Estimation of streambed groundwater fluxes associated with coaster brook trout spawning habitat. Groundwater, 50(3), 432-441.

Vogel, H. (1921), Das Temperaturabhangigkeitsgesetz der Viscositat von Flussigkeiten, Physikalische Zeitschrift, 22:645–646.

Wade, J., Lautz, L., Kelleher, C., Vidon, P., Davis, J., Beltran, J., Pearce, C. (2020). Beaver dam analogues drive heterogeneous groundwater–surface water interactions. Hydrological Processes, 34(26), 5340-5353.

Whiting, P. J., Pomeranets, M. (1997). A numerical study of bank storage and its contribution to streamflow. Journal of Hydrology, 202(1-4), 121-136.

Winter, T. C., Harvey, J. W., Franke, O. L., Alley, W. M. (1998). Ground water and surface water: a single resource. U.S. Geological Survey, Circ. 1139.

Wilson, J. L., Guan, H. (2004). Mountain-block hydrology and mountain-front recharge. Groundwater recharge in a desert environment: The Southwestern United States, 9, 113-137.

Yue, W., Wang, T., Franz, T. E., Chen, X. (2016). Spatiotemporal patterns of water table fluctuations and evapotranspiration induced by riparian vegetation in a semiarid area. Water Resources Research, 52(3), 1948-1960.

CHAPTER 3: GEOLOGIC CONTROLS ON THE TRANSITION TO HYDRAULIC DISCONNECTION IN MULTI-AQUIFER SYSTEMS: A NUMERICAL STUDY USING OBJECT-BASED GEOSTATISTICS AND VARIABLY SATURATED FLOW MODELING

3.1 Introduction

As a water table is lowered beneath a surface water body (e.g., stream, lake, wetland), the presence of a low-hydraulic conductivity (low-K) unit (sometimes referred to a clogging layer) can cause negative pressure to develop in the underlying, high-K aquifer. With successive water table lowering, the zone of negative pressure expands, causing the aquifer pressure to drop below the air entry pressure at the interface. At this point, the aquifer begins to desaturate, and the system enters a transitional flow regime characterized by unsaturated flow (Fox and Durnford, 2003; Brunner et al., 2011). As the water table declines, it also drives up the hydraulic gradient which increases the flux through the saturated low-K unit. At the same time, the aquifer hydraulic conductivity, which is dependent on saturation, progressively decreases until flow occurs predominantly through the low-K unit. Eventually, the flux asymptotically approaches a maximum value, and the flow regime becomes disconnected such that further lowering of the water table no longer significantly affects the seepage flux (Brunner et al., 2011).

The presence of unsaturated conditions (i.e., in transitional and disconnected flow regimes) has critical implications for water budgets. When pumping occurs, water is initially derived from aquifer storage, but over time, is increasingly derived from decreased discharge and increased recharge (i.e., capture) (Theis, 1941). When unsaturated conditions develop, the rate and volume of recharge that occurs in response to pumping is drastically reduced (Brunner et al., 2011). Models which neglect to consider potential for desaturation may overestimate recharge, particularly as pumping times increase. This can result in inaccurate estimates of water budgets and over-predictions of future water supplies. Unsaturated zones can also cause reduced pumping capacity, greater well drawdown (Fox and Durnford, 2003; Su et al., 2007), increased lag times, altered flow paths, and greater infiltration volumes during flood events; particularly where unsaturated zones are thick (>15m) (Fox and Durnford, 2003; Fleckenstein et al., 2006; Vázquez-Suñé et al., 2007; Hunt et al., 2008; Desilets et al., 2008; Frei et al., 2009). Therefore, considering the potential for unsaturated conditions can also be important when interpreting aquifer tests, contaminant transfer behavior, and aquifer response to floods.

Hydrogeologic controls for the development of unsaturated conditions in shallow regions beneath streams, lakes, and wetlands are well established (e.g., Rosenberry, 2000; Fox and Durnford 2003; Brunner et al. 2009a; 2009b; 2011; Tang et al., 2017; Schilling et al., 2017). However, few have considered the potential for unsaturated conditions to develop within or between aquifers. According to Brunner et al. (2009a), a 1D system with an aquifer underlying a clogging layer may begin to desaturate if the following conditions are met:

$$\frac{K_c}{K_a} \leq \frac{h_c}{d+h_c} \quad \text{(Eq. 3.1)}$$

where K_c and h_c are the clogging layer hydraulic conductivity (LT^{-1}) and thickness (L), K_a is the aquifer hydraulic conductivity (LT^{-1}), and d is the surface water depth (L). An analogy can be drawn between the near stream environment and a heterogeneous alluvial-bedrock aquifer system wherein low-K bedrock aquifer units represent the streambed clogging layer and saturated alluvium represents ponded water (**Figure *16***). Using the criteria in Equation 3.1, 15 m of saturated alluvium overlying a bedrock aquifer with a hydraulic conductivity contrast (high-K to low-K ratio) of 300 would generate unsaturated conditions with low-K layers as thin as 0.05 m. Significant water table drawdown would be required for aquifer pressures to drop below the air entry pressure. However, unprecedented pumping in recent decades has led to global groundwater depletion and drawdown of up to 290 m in heavily pumped aquifers, with annual drawdown rates ranging from 0.1 to 10 meters per year (Wada et al., 2010; Werner et al., 2013). As global pumping rates continue to rise, the potential for disconnection may expand to new settings where it has not previously been documented. This simplified example indicates the potential for unsaturated conditions to develop between and within heterogeneous aquifers, warranting further investigation for unsaturated potential in a realistic 3D aquifer system.

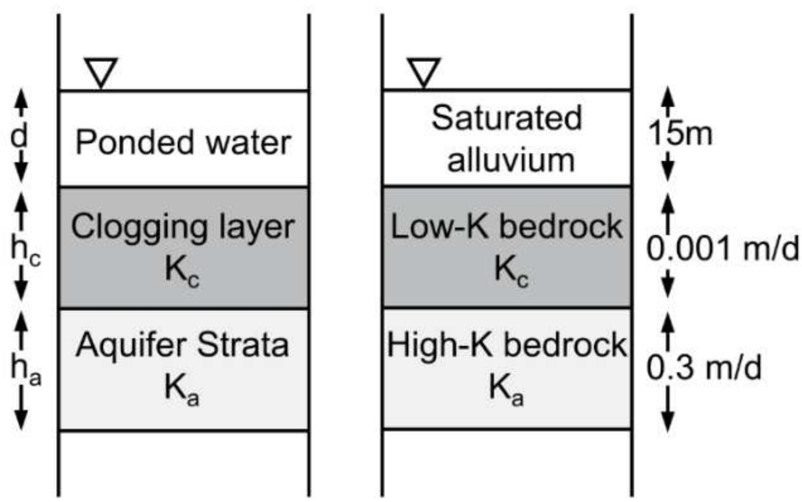

Figure 16. Illustration comparing a stream-aquifer scenario (with streambed clogging layer), left column, to a heterogeneous alluvial-bedrock aquifer system scenario, right column. Modified from Brunner et al. (2009a).

Previous authors have evaluated effects of unsaturated regions beneath surface water bodies using variably saturated flow models. Rosenberry (2000) used a two-dimensional (2D) variably saturated flow model to simulate the development of an unsaturated zone beneath a lake and found that the extent of the unsaturated zone depends on permeability contrasts, anisotropy, bed slope, and sediment thickness. Su et al. (2007) used a three-dimensional (3D) multiphase flow and transport model to simulate pumping-induced unsaturated regions beneath perennial rivers and found increased potential for unsaturated zones with increasing hydraulic conductivity contrasts and reduced pumping capacity in unsaturated zones. Shanafield et al. (2012) used a 2D variably saturated flow model to examine aquifer response to transient stream stage. Frei et al. (2009), used 3D variably saturated flow models to investigate patterns and dynamics of river-aquifer exchange in heterogeneous systems.

These and other studies indicate that geologic heterogeneity, or the distribution of low- and high-permeability subsurface materials is a key control for disconnection dynamics. Low-permeability units dictate if and where desaturation can occur (Frei et al., 2009; Brunner et al., 2011). Conversely, high-permeability units provide preferential flow paths that can dominate aquifer recharge (Swanson et al., 2006; Maples et al., 2020), river seepage (Fleckenstein et al., 2006; Frei et al., 2009), and well pumping response (Meier et al., 1998; Knudby and Carrera, 2006; Dann et al., 2008; Ronayne et al., 2008; DesRoches et al., 2013). Spatial variations in connection status and exchange flux are often attributed to patterns in subsurface heterogeneity (Brunner et al., 2011; Fleckenstein et al., 2006; Frei et al., 2009; Tang et al., 2017). Desilets et al. (2008) and Shanafield et al. (2012) describe how sharp hydraulic conductivity contrasts can alter flow paths when low-conductivity units drive lateral flow, thereby reducing vertical fluxes and leading to thicker perched zones. Geometry, connectivity, and other properties of the heterogeneity structure may also provide important controls in variably saturated settings (e.g., Bruen and Osman, 2004; Frei et al., 2009; Renard and Allard, 2013).

Heterogeneity is incorporated into groundwater flow models through aquifer parameter inputs. Because complete three-dimensional information about subsurface heterogeneity is never available, geostatistical methods are used to generate 3D realizations of aquifer properties. The choice of geostatistical methods can impact flow results, even across methods that enforce the same mean and covariance functions. For variably saturated flow problems, it is especially important that methods reproduce geologic structure and connectivity associated with depositional environments (Lee et al., 2007; Irvine et al., 2012).

In this work, I consider a heterogeneous stream-alluvial-bedrock aquifer sequence and explore the potential for unsaturated conditions within the bedrock aquifer caused by regional water table decline. I utilize geostatistical methods that reproduce realistic aquifer heterogeneity and apply Monte Carlo analysis to consider numerous realizations. I use variably saturated numerical flow models to simulate flow and saturation within the system and evaluate how aquifer exchange rates are sensitive to heterogeneity and saturation. I expand related previous work to a 3D setting. Specific objectives of this study include the following: (i) Evaluate how heterogeneity and saturation affect aquifer exchange rates in a multi-aquifer system with a lowering water table in the deeper aquifer; (ii) Determine whether unsaturated conditions and disconnected flow regimes might develop, and under what conditions; (iii) Evaluate how the connectivity structure of sandstone in the bedrock aquifer influences inter-aquifer exchange and disconnection.

3.2 Geologic setting

This study includes a numerical modeling investigation of desaturation that is based on data from a stream-alluvial-bedrock aquifer sequence within the Denver Basin of Colorado, USA. The Denver Basin bedrock aquifers occur within a bowl-shaped structure comprised of Late Cretaceous through Paleogene-age sandstones and mudstones (Raynolds, 2002). These water-bearing sedimentary rocks form a series of confined and unconfined bedrock aquifers. Where bedrock aquifers crop out at the surface, they are incised and overlain by Quaternary alluvial deposits that make up an alluvial aquifer system along which modern day tributaries of the South Platte River flow (Lindsey et al., 2021).

The Denver Basin aquifers comprise a regionally significant water resource for growing populations along Colorado's Front Range Urban Corridor. While alluvial aquifers are replenished by mountainous headwater streams, bedrock aquifers receive limited recharge from minimal outcrop extent and precipitation (Graham and VanSlyke, 2004; Paschke et al., 2011). This lack of deeper recharge coupled with long-term pumping has led to substantial groundwater depletion in the bedrock aquifers (Paschke et al., 2011; Ruybal et al., 2019). Associated water table declines are upwards of 80 m in some regions (Paschke et al., 2011; Ruybal et al., 2019), placing the bedrock water table tens of meters below the base of alluvium and creating the potential for unsaturated conditions to develop (Cognac and Ronayne, 2020). Regional flow models predict additional water table declines, localized flow reversals, and increased flow from the alluvial to bedrock aquifers (Paschke et al., 2011). However, current predictions rely on coarsely discretized numerical flow models that are unable to account for realistic heterogeneity and saturation effects. A recent higher-resolution, 2D modeling study highlights the potential for reduced connection between alluvial and bedrock aquifers (Cognac and Ronayne, 2020).

This study focuses on an area in the south-central Denver Basin where significant bedrock aquifer pumping has occurred. Nearby wells screened in the D2 Sequence within the Dawson aquifer, the uppermost bedrock aquifer, show rates of water level decline that average -1m year^{-1} (***Figure 17***). Within the study area, the upper lithified sediments are incised and overlain by an alluvial aquifer which is hydraulically connected to Cherry Creek, a tributary of the South Platte River. The bedrock aquifer sediments were deposited by fluvial systems draining the Rocky

Mountain Front Range during the Laramide orogeny (late Paleocene and Early Eocene) (Raynolds, 2002). The D2 sequence contains heterolithic arkosic strata that include coarse, multi-storied channel sandstone beds separated by overbank mudstones with trough-crossbedding common at outcrops (Raynolds, 2002). Sandstone channels comprise approximately 30-40% of the strata (Raynolds et al., 2001), and average channel thicknesses are approximately 3-4 m (Raynolds, 2002; Paschke et al., 2011). D2 strata are noted to contain relatively uniform lithofacies with limited lateral compositional and textural variability (Raynolds, 2002).

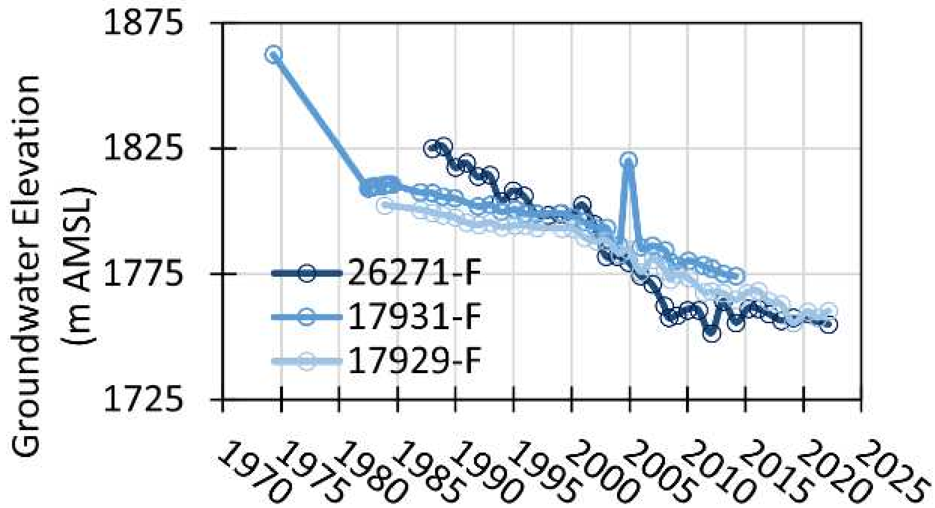

Figure 17. Historical water levels for three Dawson aquifer wells in the study area identified by permit number (CDWR, 2022).

3.3 Methods

Geostatistical and numerical flow modeling were performed to evaluate fluxes within the stream-alluvial-bedrock aquifer system under variably saturated conditions. Geologic heterogeneity is represented through facies-scale geostatistical models conditioned to geophysical log data. Multiple realizations are generated for Monte Carlo analysis to evaluate uncertainty related to geologic heterogeneity. Groundwater flow is simulated using a variably saturated flow model where hydraulic parameters are assigned using heterogeneity realizations. Simulation results are analyzed to determine the controls on inter-aquifer exchange and hydraulic disconnection.

3.3.1 Geostatistical simulation of aquifer heterogeneity

A variety of geostatistical methods have been developed to simulate geologic heterogeneity for sedimentary depositional systems (see Koltermann and Gorelick, 1996; de Marsily et al., 2005; Pyrcz and Deutsch, 2014). Fluvial deposits require methods that reproduce long-range connectivity of high-K units (i.e., channels) and low-K barriers to flow. This study employed the hierarchical, object-based geostatistical modeling code FLUVSIM (Deutsch and Tran, 2002), which successively places channel objects within a background overbank mudstone facies and allows for conditioning to observed facies within boreholes.

Conditioning data included nine shallow and deep resistivity geophysical logs spanning the upper Denver Basin D2 sediments obtained from the Colorado Division of Water Resources (CDWR, 2020) and the Colorado Geological Survey (personal communications) (***Figure 18***). Although gamma logs were available, local uranium deposits are known to influence gamma signals and hydrogeologists typically rely on resistivity data for determining sandstone and mudstone intervals (Musgrove et al., 2014). Resistivity logs were normalized using a quantile transform and a cut-off value was applied to distinguish sandstone from mudstone. Detailed geophysical log data and types are provided in ***Appendix B*** (***Table B1*** and ***Figures B1-B3***). The interpreted sand fraction at each borehole ranged from 0.28 to 0.41, with an overall fraction of 34% for the study area.

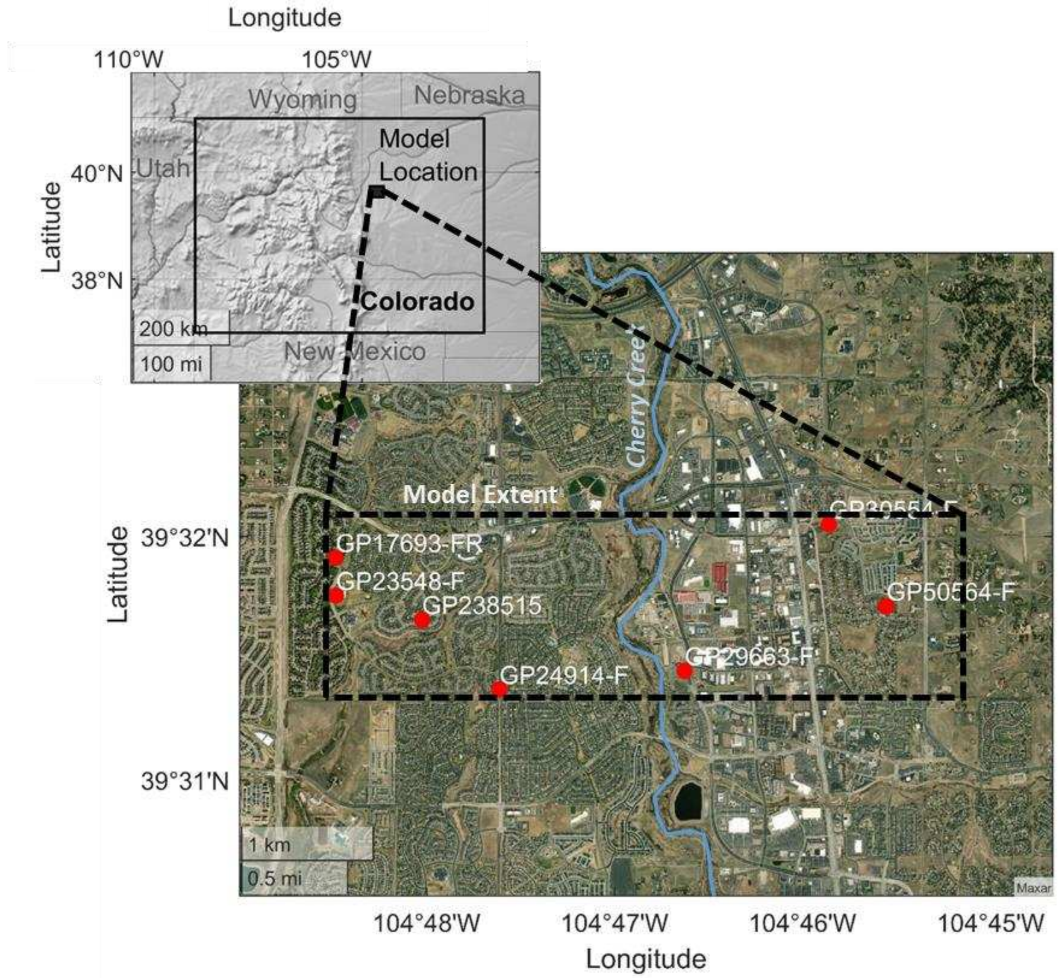

Figure 18. Location of model and geophysical logs

Facies realizations were generated for a base case sandstone (channel) fraction of 35% and for additional scenarios that assumed channel fractions of 20%, 50%, and 75%, resulting in a total of 200 realizations (50 per channel fraction). Except for the 20% set, simulations were conditioned to lithofacies from geophysical logs (i.e., sandstone or mudstone) by assigning maximum priority to conditioning during channel placement. To avoid compromising channel proportion accuracy, conditioning was turned off for the 20% simulations. FLUVSIM input parameters, including model discretization and channel geometry information, are summarized in ***Table 6***.

The alluvial aquifer was represented as a single large channel with geometry constrained by bedrock depths from boring logs. Alluvial aquifer geometry was consistent across all geostatistical realizations. An example set of facies realizations for each channel fraction is presented in **Figure *19***.

Table 6. Geostatistical and variably-saturated flow modeling input parameters

FLUVSIM Parameters	**Values**		
# rows, columns, layers	75,152,149		
Δx, Δy, Δz	27.7 m, 16.9 m, 0.41 m		
Channel			
proportions	20%, 35%, 50%, 70%		
orientation	0° (North/South)		
sinuosity (departure, length scale)	Departure 350 m; Length scale: 900 m		
Thickness	4 m ±1 m		
width to thickness ratio	200		
MODFLOW-USG Parameters	**Sandstone**	**Mudstone**	**Alluvium**
Saturated Hydraulic Conductivity (m day^{-1})	0.3	0.001	65
Specific Storage (m^{-1})	0.000017	0.000056	0.0001
Specific Yield (-)	0.18	0.15	0.36
Specific Retention (-)	0.12	0.23	0.1
van Genuchten α (m^{-1})	0.79	1.9	14.5
van Genuchten n (-)	10.4	1.31	2.68
Brooks-Corey P (-)	3.21	9.45	4.19

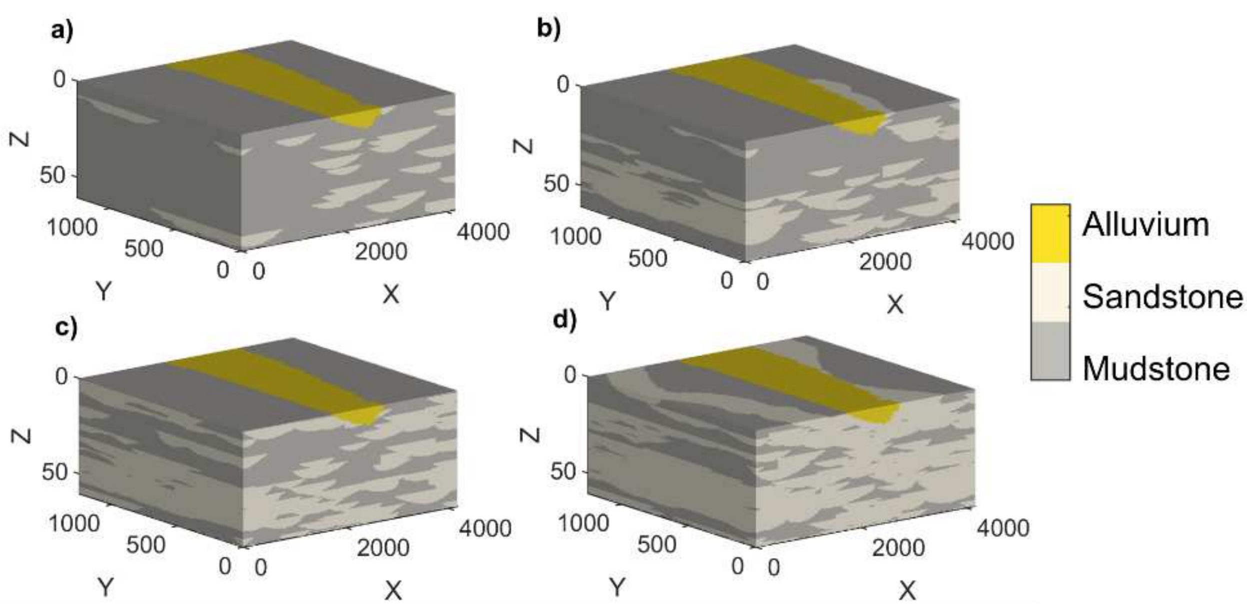

Figure 19. Simulated lithofacies for example realizations with sandstone channel fractions of a) 20%, b) 35%, c) 50%, d) 75%. Distances along each axis are reported in meters.

3.3.2 *Variably-saturated flow modeling*

Three-dimensional, variably-saturated flow (VSF) modeling was performed to investigate the potential for unsaturated conditions and disconnection within and between the alluvial and bedrock aquifers. Specifically, the model was designed to evaluate how regional, long-term drawdown in the bedrock aquifer impacts fluxes and saturation throughout the stream-alluvial-bedrock system. To evaluate the sensitivity of results to bedrock aquifer heterogeneity, the model was then run for a range of heterogeneity scenarios wherein hydraulic parameters were assigned using geostatistical realizations.

Simulations were performed using the block-centered transport process for MODFLOW-USG (MFUSG), a finite volume, unstructured grid version of MODFLOW (Panday et al., 2013; Panday, 2019). MFUSG was selected for variably saturated flow solution, robust solver capabilities, and open-access licensing. MFUSG applies a control volume finite difference scheme to approximate a solution to the governing equation for 3D variably saturated transient flow:

$$\frac{\partial}{\partial x}\left(K_x k_{rw}\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y k_{rw}\frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z k_{rw}\frac{\partial h}{\partial z}\right)-W=\emptyset\frac{\partial S_w}{\partial t}+S_w S_s\frac{\partial h}{\partial t} \qquad (3.2)$$

where K_x, K_y, and K_z are the principal components of saturated hydraulic conductivity [L T^{-1}] along the x, y, and z axes, respectively; k_{rw} is the relative permeability, a dimensionless value that ranges from 0 to 1 as a function of water saturation; h is the hydraulic head [L]; Wis a volumetric source or sink per unit volume [T^{-1}]; $\emptyset$ is the drainable porosity taken to be the specific yield, S_w is the degree of water saturation, which is a function of pressure head; S_s is specific storage [L^{-1}], and t is time [T].

Functional expressions are used to relate the relative permeability, hydraulic head, and water saturation in the solution to Equation 3.2. Effective saturation, S_e, defined as $(S_w-S_{wr})/(1-S_{wr})$ where S_{wr} is the residual saturation, is calculated in MFUSG as a function of the pressure head through a modified van Genuchten equation (van Genuchten, 1980):

$$S_e = \begin{matrix} [1+(\alpha\psi)^n]^{-m} & for\ \psi<0 \\ 1 & for\ \psi>0 \end{matrix} \qquad \text{(Eq. 3.3)}$$

where α, *n* and *m* (*m*=1-1/*n*) are the van Genuchten parameters, ψ is pressure head [L] (ψ=z-h), and z is elevation. The relative permeability term (k_{rw}) utilized in Equation 3.2 is dependent on effective saturation, $k_{rw} = S_e^P$ (Brooks and Corey, 1966) where *P* is the Brooks-Corey coefficient (Panday, 2019).

The 3D model domain includes a stream-alluvial-bedrock aquifer sequence spanning 4,686 m in the x-direction (east-west), 1,260 m in the y-direction (north-south), and 70 m vertically below the streambed. The model is oriented parallel to the direction of the stream and regional groundwater flow, which is towards the north. The domain is divided into a regularly spaced grid with 75 rows, 152 columns, and 149 layers with corresponding grid-spacing of 27.7 m, 16.9 m, 0.41 m in the x-, y-, and z-directions respectively.

The lateral and bottom edges of the model (x=0, x=4,686 and z=0 m) were assigned no-flow boundaries to reflect regional streamlines that run parallel to model edges (**Figure 20**). An assumption is made that minimal directional changes or cross gradients occur near these boundaries during flow simulations. The up- and down-gradient edges of the model (y=0 and y=1,260) represent regional groundwater head contours for the alluvial and bedrock aquifers wherein the alluvial heads are relatively stable through time and the bedrock regional water table lowers from pumping wells beyond the extent of the model domain. These boundaries were assigned using no-flow boundaries and head-dependent flux boundaries as implemented with the General Head Boundary (GHB) Package (Harbaugh et al., 2000) in MODFLOW. The external head values for alluvial aquifer GHB cells were constant through time. Boundary heads for the bedrock aquifer were successively lowered during transient simulations at a rate of 1m per year, which approximates the long-term rate of groundwater level decline recorded in Denver Basin wells (CDWR, 2022). GHB conductance values were calculated using the hydraulic conductivity assigned to each cell. The perennial stream was modeled using a head-dependent flux boundary as implemented in MODFLOW's River Package. The simulated flux into or out of river cells is proportional to an assigned stream width, stage, and streambed hydraulic conductivity, as well as the calculated head difference between the stream and adjacent aquifer (Harbaugh et al., 2000).

Aquifer hydraulic properties were assigned constant values for each simulated lithofacies (**Table 6**). Saturated hydraulic conductivity, specific storage (Ss), and specific yield (Sy) were assigned based on published values for sandstone and mudstone units of D2 sediments (Lapey 2001; Woodard et al., 2002; Paschke et al., 2011, Brown & Caldwell, 2017). Water retention parameters, including the residual water saturation (Swr) and the van Genuchten α and n, for the alluvium are averages for sand samples reported in Carsel and Parrish (1988). Within the bedrock aquifer, water retention parameter values were assigned based on published estimates for similar sedimentary rocks (van Genuchten 1980; Lapey 2001; Parajuli et al. 2017). Brooks Corey coefficients were estimated from van Genuchten parameters using equivalence relationships developed by Seytoux et al. (1996).

A flow simulation consisting of one steady-state and 14 transient stress periods was performed for each of the 200 heterogeneity scenarios. Transient stress periods spanned 5 years for a total simulation length of 70 years. Timestep lengths were automatically selected through an adaptive time-stepping algorithm and were allowed to vary between 0.1 and 200 days. Following each simulation, flow from the alluvial to bedrock aquifer was computed using the USGS program ZoneBudget-USG (Harbaugh, 1990). Unique zones were assigned to each lithofacies to calculate flow between units. Upgradient (1-5) and downgradient (71-75) rows were excluded from the analysis to avoid boundary effects.

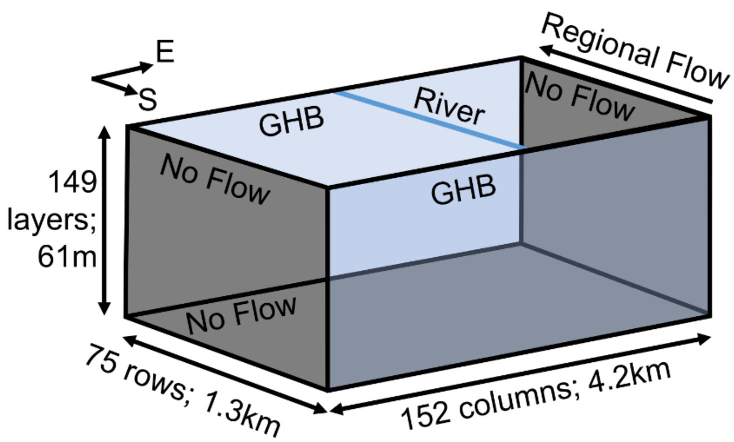

Figure 20. Model discretization and boundary conditions

3.3.3 Connectivity metrics

Sandstone channel connectivity within the bedrock aquifer was evaluated using a variety of metrics. Static connectivity describes the connectivity structure of a given parameter field as determined by the geologic architecture (King, 1990; Renard and Allard, 2013). In this case, I am interested in the connectivity between high-permeability lithofacies. Static connectivity was evaluated using binary parameter fields where sandstone and alluvium were assigned a value of 1 representing higher permeability geologic material, and mudstone cells were assigned a value of

0 for low permeability. To account for dynamic processes, I also considered saturation-dependent high-K connectivity using a thresholding scheme. Model cells with $K(\psi) > 0.01$ m day^{-1}, indicating relatively high conductivity sandstone or alluvium, were assigned a value of 1, and all other cells were assigned a value of 0. The corresponding binary field, which is based on the flow simulation results, was then analyzed to assess the high-K conductivity structure. I refer to this as the saturation-dependent or dynamic connectivity when interpreting results.

Connected component analysis was performed for the static and dynamic (processed from VSF model output) binary fields using the CONNEC3D program (Pardo-Iguzquiza and Dowd, 2003). Each connected component (CC) is a unique body of connected high-K cells wherein cells are considered connected if they intersect along a 3D grid face (Renard and Allard, 2013). For each of the 200 simulated static and dynamic fields, statistics were calculated to summarize key aspects of the CC geometry. These include: the total number of connected components (NCC), maximum number of cells for all CCs (MCC), the maximum vertical span (Z-dimension) for all CCs (ZCC), and the maximum vertical span for CCs that directly contact the alluvium (ABZCC). Saturation-dependent or dynamic conductivity statistics are indicated with 'DY' appended to the variable name (i.e., NCC_{DY}, MCC_{DY}, ZCC_{DY}, $ABZCC_{DY}$).

The alluvium is expected to be more hydraulicly connected to the bedrock aquifer where channelized sandstones are in direct contact with alluvium. Connectivity at the alluvial-bedrock aquifer interface was quantified as the percentage of bedrock aquifer cells that share a face with alluvial aquifer cells and which contain sandstone, designated $A\text{-}B_{SS\_\%}$.

Connectivity metrics were compared to the final (late-time) A-B flow rate for each simulation and multiple linear regressions were performed to quantify whether and which connectivity metrics had the strongest influence on inter-aquifer exchange and hydraulic disconnection potential.

3.4 Results

3.4.1 Pressure, Saturation, and Flux dynamics

The VSF model was used to evaluate the effects of a declining bedrock water table on hydraulic heads, saturation, and exchange flows within the stream-alluvial-bedrock aquifer system. Simulation results for a representative heterogeneity scenario are presented in Figure 21, demonstrating modeled pressure heads (ψ) and saturation as the regional water table is progressively lowered. Each panel corresponds to a simulation timestep wherein the year is equal to the magnitude of water table lowering in meters (i.e., rate = -1 m year^{-1}). The depth of the alluvium varies across the model domain and is 14 m below the model top for the example cross-sectional slice.

The first panel depicts simulation year 5 (regional water table at 5 m below the model top; 9 m above the alluvium base). At this time, the bedrock aquifer shows increasing ψ with depth and a continuous water table (i.e., ψ=0 contour) can be traced across the alluvial and bedrock aquifer. In the year 20 panel (regional water table 6 m below alluvium base), a low-pressure zone begins to form between the deeper groundwater and alluvium indicated by low ψ values surrounding the two uppermost sandstone channels underlying the alluvium. The year 30 panel (regional water table 16 m below alluvium) shows the initiation of negative pressure beneath the alluvium, indicated by ψ=0 contours within the same two channels. This marks the beginning of unsaturated flow for this simulation and indicates a transitional flow regime. Notably, the negative pressure initiates within upper sandstone channels. The remaining three panels (year 45, 55, and 70) depict an expanding negative pressure and unsaturated zone between the alluvium and deeper water table. Both ψ and saturation show percolation features that connect saturated zones wherein flow occurs through saturated mudstone between unsaturated channels. By the final timestep, a perched alluvial aquifer is separated from the deeper water table by a partially to fully unsaturated zone. Notably, saturation patterns remain irregular and are distinctly influenced by the fluvial architecture. Not all channels become unsaturated. Close inspection of Figure 4b reveals instances where connected channels are partially unsaturated in the upper portion while lower portions form saturated zones that serve as conductive pathways.

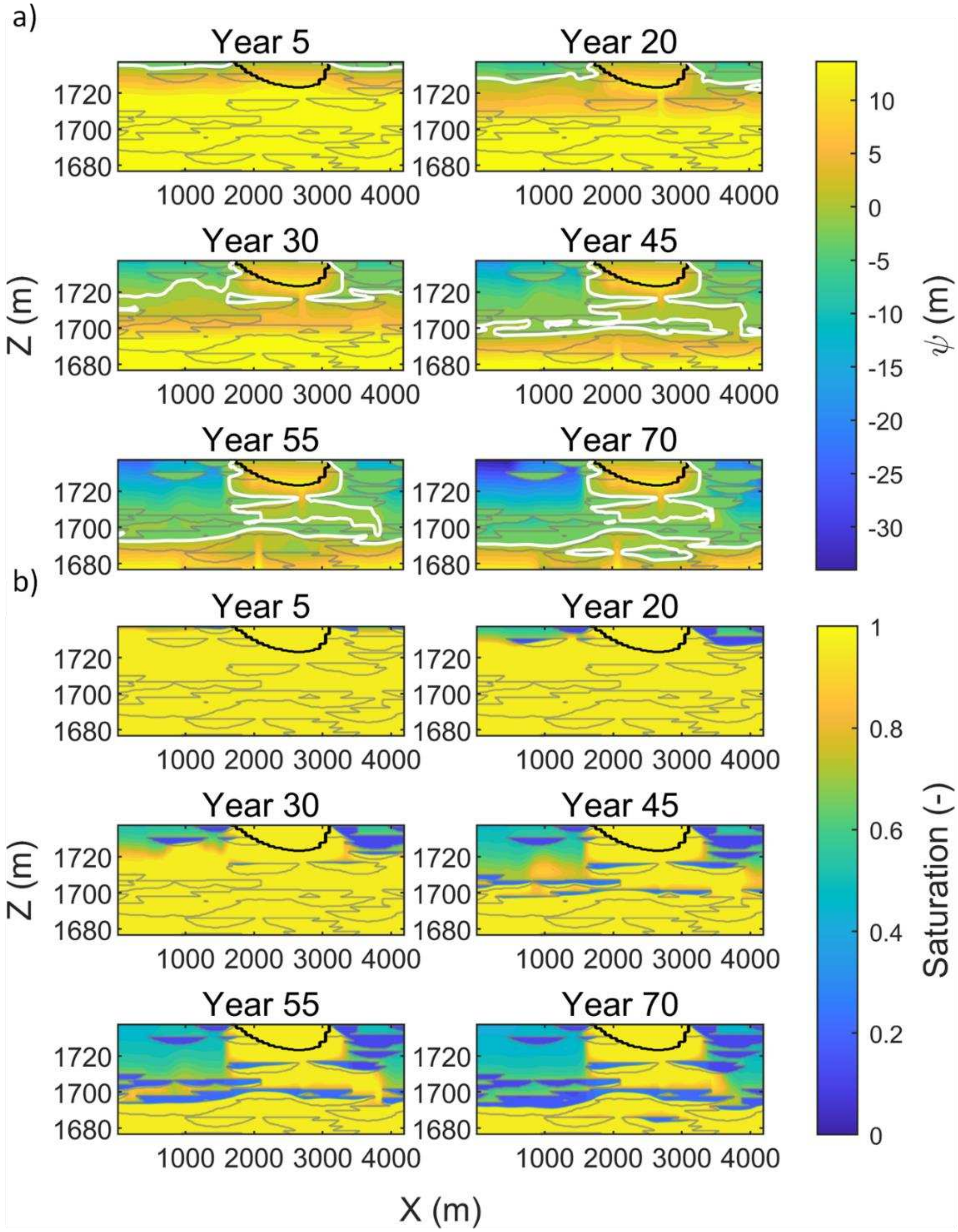

Figure 21. Pressure head and saturation results for an example realization (35% channel fraction) showing connected (Year 5 & 20), transitional (Year 30 & 45), and disconnected (Year 55 and 70) flow regimes. Results are displayed for model row 70. The black line indicates the extent of the alluvium and the white line in (a) represents pressure head = 0 m.

1.1. Inter-aquifer flow and seepage

Simulated alluvial-to-bedrock aquifer flow (A-B flow) and river inflow for the corresponding simulation are plotted in **Figure 22**. A-B flow rates range from 26 to 347 m^3 day^{-1}, and the simulated river inflow (seepage integrated over all river cells) ranges from 1624 to 2066 m^3 day^{-1}. During the transient simulation, some of the induced recharge from the stream (Δ river inflow) exits the model domain via the alluvial aquifer, rather than contributing to increased A-B flow. For both quantities, the simulated flow rates progressively increase as the regional water table is lowered. Initially, this increase follows a relatively linear trend. Around year 20, the rate of increase begins to drop off with successive water table lowering. To evaluate how resulting rates change during the simulation, the percent change between timesteps was calculated as $\Delta\{\text{A-B } Flow\} = \left\{100\ \times\ \left(Q_{AB_{ts}} - Q_{AB_{ts-1}}\right)/Q_{AB_{ts-1}}\right\}$. The percent change decreased rapidly at the beginning of the simulation and approached 0% at the simulation end for A-B flux and 0.5% for river inflow.

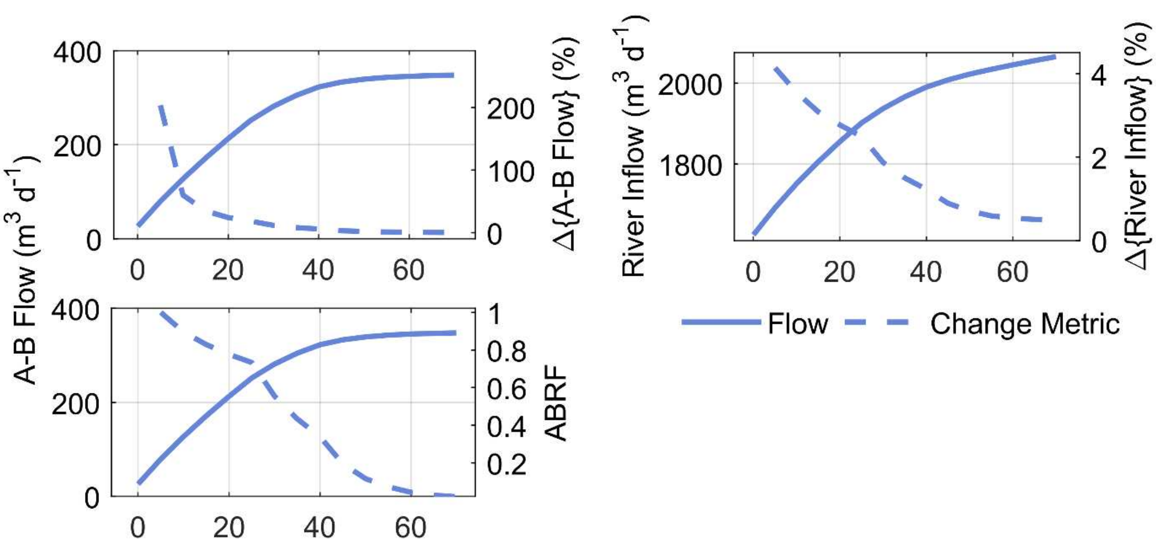

Figure 22. Simulated alluvial-to-bedrock aquifer flow (A-B flow) and river inflow with associated percent change (Δ{A-B Flow} and Δ{River Inflow}) and ABRF for example realization with 35% channel fraction.

Because unsaturated conditions alter the relationship between water table change and the resulting change in flow, it is more useful to define a metric that quantifies this relationship directly. Here, I define a normalized A-B flow response function (ABRF) that relates the change in A-B flow rate to corresponding changes in bedrock water table position:

$$ABRF = \frac{(\Delta Q_{AB}/\Delta h)^m}{(\Delta Q_{AB}/\Delta h)^1} \quad \text{(Eq. 3.4)}$$

where $(\Delta Q_{AB}/\Delta h)^m$ is the change in A-B exchange flow between successive stress periods m and m-1 divided by the change in prescribed hydraulic head value (regional water table change) for those periods, and $(\Delta QAB/\Delta h)^1$ is the corresponding ratio for the first transient stress period. For consistency, time steps at the end of each stress period are used in the calculations. The normalization by $(\Delta Q/\Delta h)^1$ results in a dimensionless metric, enabling comparisons across multiple simulations with different flow magnitudes. The value of ABRF approaches zero for a disconnected flow regime (i.e., further lowering of the regional water table does not increase the amount of A-B exchange).

For the example simulation, ABRF starts at 1 (by definition) and progressively decreases over time to a final value of 0.019. This indicates a 98% reduction in the flow response over the course of the simulation. Around simulation year 25-30, ABRF shows a distinct slope change which is related to the initiation of unsaturated conditions (see **Figure *21***). Around year 50, ABRF begins to flatten, indicating minimal change in flow with additional decline in regional water table position. For the remainder of the simulation, ABRF asymptotically approaches zero, indicating full hydraulic disconnection.

1.2. Sensitivity to sandstone channel architecture

Simulations were performed for a range of geostatistical realizations to better understand how flow, saturation, and connection status are influenced by geologic heterogeneity. A-B flow, Δ{A-B Flow}, and ABRF values for all simulations and channel fractions are plotted in ***Figure 23*** along with the ensemble mean and range. Corresponding statistics for the final timestep are summarized in ***Table** 7*.

Table 7. Summary of model results for the final timestep (Year 70) for all simulations.

		Ensemble	Minimum	Maximum	σ	CV
A-B Flow (m^3 day^{-1})	20%	313	184	501	81	0.26
	35%	344	222	508	60	0.17
	50%	650	423	1258	154	0.24
	75%	1913	919	3187	554	0.29
Δ{A-B Flow}	20%	0.53	0.27	0.97	0.15	0.28
	35%	0.32	0.085	0.55	0.10	0.31
	50%	0.25	-0.010	0.74	0.16	0.64
	75%	0.68	-0.040	2.1	0.45	0.66
ABRF	20%	0.037	0.018	0.075	0.012	0.34
	35%	0.022	0.0058	0.040	0.0074	0.34
	50%	0.016	-0.00080	0.054	0.011	0.69
	75%	0.055	-0.0027	0.19	0.040	0.74

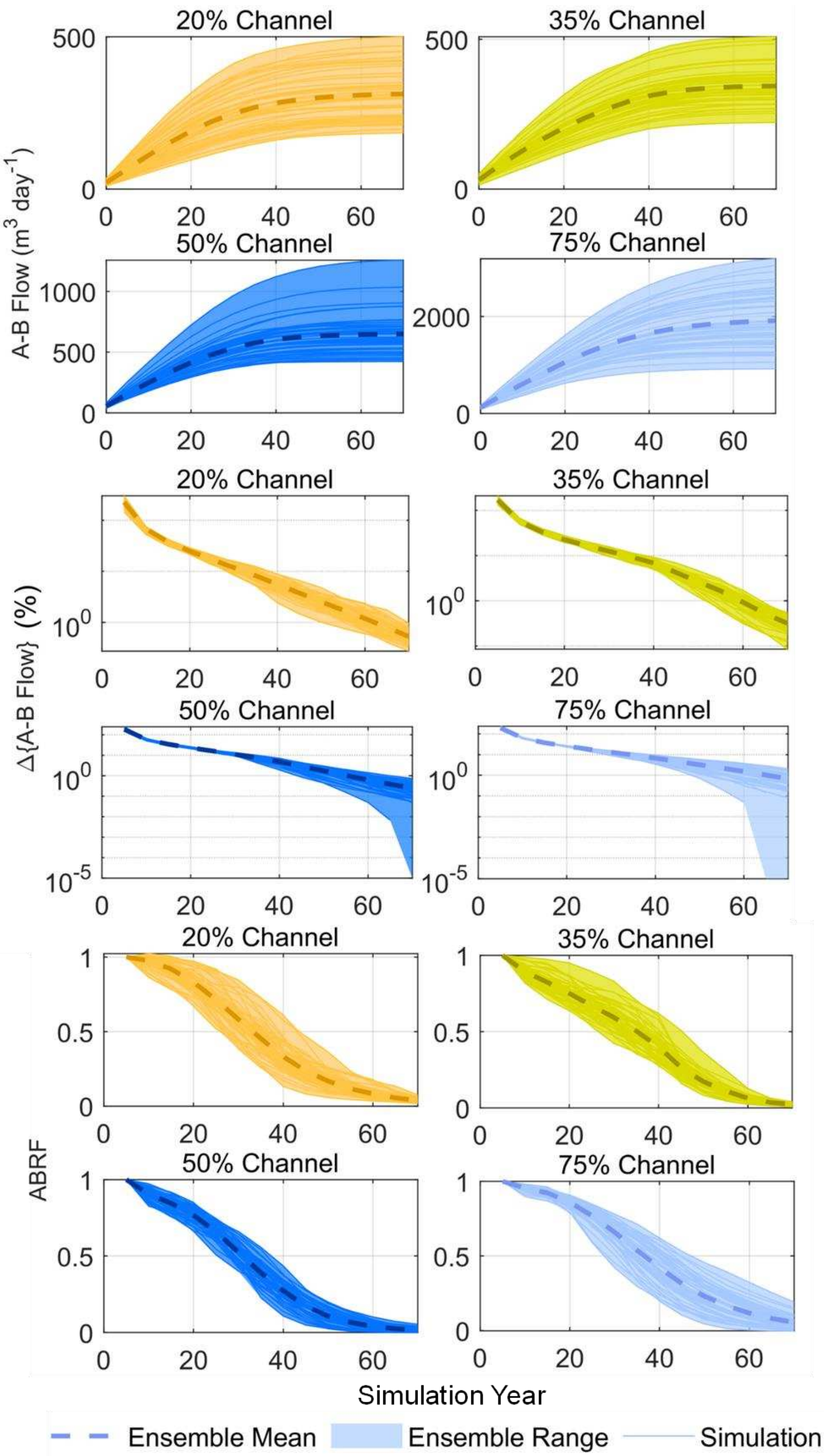

Figure 23. Simulated alluvial-to-bedrock (A-B) flow, A-B flow response function (ABRF), and percent change in A-B flow (Δ{A-B Flow}) for all realizations. Ensemble mean and range are plotted for 20%, 35%, 50% and 75% bedrock aquifer channel fractions.

3.4.2 A-B Flow

Final flow rates from the alluvial to bedrock aquifer (A-B flow) for all simulations ranged from 184 to 3187 $m^3\ day^{-1}$ with ensemble mean values of 313, 344, 650, and 1913 $m^3\ day^{-1}$ for channel fractions 20%, 35%, 50%, and 75% respectively. The standard deviation (σ) of final A-B flow rates also generally increased with increasing channel fraction except for the 20% scenarios which had a greater σ than 35%. This may be related to the 20% scenarios being unconstrained by geophysical log data. The coefficient of variation (CV) increased with increasing channel fraction, indicating variability about the ensemble mean was greater for higher bedrock sandstone fractions. For all simulations, A-B flow rates follow similar patterns through time as the example scenario described previously. Flow increased linearly at first and then starts to flatten with successive water table lowering. Some simulations show a more pronounced flattening than others, and simulations with a lower final A-B flow rate tend to have a flatter final slope.

The A-B flow rate is positively associated with the fraction of flow that occurs through sandstone channels and with river inflow (***Figure C1*** in ***Appendix C***). The magnitude of the flow rate tends to be higher when a greater proportion of the flow occurs through sandstone channels. Additionally, higher vertical groundwater flow rates, associated with a more transmissive aquifer system at depth, promote more seepage from the losing stream.

3.4.3 A-B Percent change

Final modeled values of Δ\{A-B Flow\} ranged from -0.04 to 2.1%, with ensemble mean values of 0.53%, 0.32%, 0.25%, and 0.68% for channel fractions 20%, 35%, 50%, and 75% respectively (***Table 7***). Negative percent change indicates a decrease in A-B flow rate between successive timesteps and is a strong sign that A-B flow rates have stabilized (i.e., hydraulic disconnection). The ensemble mean did not trend with changes in channel fraction. All simulations had a final Δ\{A-B Flow\} less than 1%, indicating minimal change in flow with successive lowering, except for ten of the fifty scenarios with 75% bedrock channel fraction. The σ and CV increased with increasing bedrock channel fraction, indicating greater variability for higher bedrock sandstone fractions. Like the example scenario, Δ\{A-B Flow\} decreased rapidly in the first 10 years and then asymptotically approached zero for the duration of the simulation for all realizations.

3.4.4 Alluvial-to-bedrock flow response function

Final ABRF values ranged from -0.0008 to 0.19 with ensemble mean values of 0.037, 0.022, 0.16, and 0.055 for channel fractions 20%, 35%, 50%, and 75% respectively. Again, negative values indicate a decrease in A-B flow rate between successive timesteps and point to stabilizing A-B flow and hydraulic disconnection. The ensemble mean ABRF increased slightly with increasing channel fraction except for the 20% realizations. The range for final ABRF values was relatively large for the 75% channel fraction (-0.0027 to 0.19), indicating the final flow response ranged from 0.27% to 19% of the initial flow response depending on the bedrock aquifer heterogeneity. The σ and CV generally increased with increasing channel fraction, again except for the 20% channel fractions. ABRF values followed similar trends as the example scenario, with the most significant changes occurring between 20 and 40 years and a gentle decrease for the remainder of the simulation.

3.4.5 Connectivity structure of heterogeneity

Static and dynamic connectivity metrics were evaluated as predictors of the final A-B flow rate. No single metric independently predicted the A-B flow. Exhaustive connectivity results for all metrics are included in ***Appendix C*** (***Figures C2*** and ***C3***).

Multiple linear regression (MLR) analysis was performed to evaluate whether a combination of metrics served as better predictors for the final A-B flow rate. All connectivity metrics were initially included as predictor variables and insignificant predictors (high p-value) were successively removed. Resulting MLR coefficients and model fits are presented in ***Table 8***. Combinations of connectivity metrics predict resulting flow rates with statistical significance ($p<0.05$) for all MLRs. R^2 values indicate that models explain between 52 and 88% of the variability in final A-B flow rates. The five identified strongest predictors are, in approximate order of decreasing significance, A-$B_{SS\_\%}$, ZCC_{DY}, $ABZCC_{DY}$, NCC_{DY}, and MCC_{DY}. Notably, R^2 values increase with increasing channel fraction, indicating sandstone connectivity is a stronger predictor of A-B fluxes for higher sandstone fractions.

The A-B interface sandstone fraction (A-$B_{SS\_\%}$) was the strongest predictor for final A-B flow rate for all but the 75% channel fraction MLR, with A-B flow rates increasing with increasing sandstone present at the A-B interface. This is unsurprising, because sandstone contacting the alluvium provides a high-K conduit for preferential flow to occur. For the 20% simulations, the few scenarios with no sandstone contacting the alluvium recorded the lowest A-B flow rates. Interestingly, the A-$B_{SS\_\%}$ is less important for the 75% channel simulations, suggesting that once some threshold is crossed, additional sandstone at the interface no longer increases the exchange rate.

The remaining significant predictors include dynamic connectivity metrics that consider the spatial organization of saturation-dependent hydraulic conductivity values. The ZCC_{DY} and $ABZCC_{DY}$ represent the largest vertical span of connected components (generally and specifically contacting the alluvium). Where significant, the ZCC_{DY} is negative, suggesting that final A-B flow rates are higher for lower ZCC_{DY}. A greater vertical extent of saturated, connected sandstones was expected to result in greater A-B flow rates. One possible explanation is that thinner connected sandstones prohibit the formation of unsaturated zones, thereby resulting in greater final flow rates. Conversely, the $ABZCC_{DY}$ coefficient is positive, indicating that thicker saturated sandstone bodies contacting the alluvium are correlated with higher A-B flow rates. This result is consistent with the finding that thick saturated sandstones contacting the alluvium effectively thicken the shallow saturated zone leading to greater A-B flow rates.

The NCC_{DY} has a negative coefficient for all MLRs indicating that as the number of connected components increases, the final A-B flow rate decreases. A greater number of individual connected components typically means that individual components are less connected. Therefore, models with less overall connectivity had lower A-B flow rates. MCC_{DY} has the smallest coefficient for all simulations and varies between negative and positive without a clear trend. This metric is difficult to explain given these results. One would intuitively expect greater A-B flow rates with increasing size of connected components. However, a larger sandstone body could also have a greater chance of becoming desaturated. The interpretation is generally inconclusive regarding this metric.

Table 8. Results from multiple linear regression analysis performed for each channel fraction (all model realizations considered) using significant predictor variables $A\text{-}B_{SS\_\%}$, MCC_{DY}, NCC_{DY}, ZCC_{DY}, $ABZCC_{DY}$ and response variable of A-B flow ($m^3\ day^{-1}$).

		Predictor			Model		
		Estimate	[1]SE	p-value	[2]RMSE	[3]R-squared	p-value
20%	Intercept	153	48	**2.67E-03**	56.7	0.52	4.0E-07
	$A\text{-}B_{SS\_\%}$	3	0.7	**2.78E-04**			
	MCC_{DY}	0.0010	0.0005	5.27E-02			
	NCC_{DY}	-1.1	1.03	2.80E-01			
	ZCC_{DY}	-2	1.8	2.14E-01			
	$ABZCC_D$	3	1.6	8.26E-02			
35%	Intercept	3	62	9.63E-01	39.8	0.57	4.3E-08
	$A\text{-}B_{SS\_\%}$	2.7	0.5	**4.71E-06**			
	MCC_{DY}	3.2E-05	1.7E-04	8.49E-01			
	NCC_{DY}	-0.028	0.4	9.47E-01			
	ZCC_{DY}	0.5	0.8	5.32E-01			
	$ABZCC_D$	5	0.8	**6.88E-08**			
50%	Intercept	100	124	4.24E-01	104	0.56	6.0E-08
	$A\text{-}B_{SS\_\%}$	6	1.2	**2.55E-05**			
	MCC_{DY}	0.0019	0.0005	**4.78E-04**			
	NCC_{DY}	-2.6	1.0	**1.04E-02**			
	ZCC_{DY}	-5	1.3	**8.53E-04**			
	$ABZCC_D$	1.06	0.6	7.04E-02			
75%	Intercept	403	408	3.29E-01	193	0.88	3.1E-20
	$A\text{-}B_{SS\_\%}$	-5	4	1.95E-01			
	MCC_{DY}	0.010	0.00079	**3.76E-16**			
	NCC_{DY}	-2.5	2.2	2.73E-01			
	ZCC_{DY}	-34	4	**6.26E-11**			
	$ABZCC_D$	3	2.1	1.56E-01			

[1]SE is standard error.

[2]RMSE is root mean square error

[3]R^2 is adjusted for the number of predictors in the model.

[4]Significant p-values at the 95% confidence level are bolded

A similar MLR analysis was conducted for the ABRF, and the same connectivity metrics were found to be significant predictors of ABRF. R^2 values were lower overall (0.17 to 0.59) but similarly increased for increasing channel fraction, highlighting the increasing influence of channel connectivity on disconnection dynamics with greater sandstone fraction. Results from this analysis are included in ***Appendix C*** (***Table C1***).

3.5 Discussion

This study combines geostatistical methods and variably saturated flow modeling to evaluate potential for and effects of unsaturated zones on flow between two aquifers. Results provide valuable insights for pumped aquifer dynamics and associated implications for management.

3.5.1 Disconnection Dynamics

The definition of a disconnected flow regime states that successive decreases in water table no longer substantially affect the resulting flux and that the flux asymptotically approaches a constant maximum value (Brunner et al., 2011). As Brunner et al. (2011) note, because the maximum flow rate is approached asymptotically, the only way to distinguish transitional from disconnected systems is by defining an arbitrary cutoff value where changes are considered negligible. During numerical modeling performed for this study, all scenarios developed unsaturated regions between the alluvium and bedrock, indicating the occurrence of a transitional flow regime. Further, final Δ{A-B Flow} for most simulations is very small (<1%), indicating that successive water table lowering minimally affects the resulting A-B flow rate and that the A-B flow is approaching a stable value. To distinguish transitional from disconnected flow regimes, I herein apply a practical cutoff for the Δ{A-B Flow} of 1% (i.e., the flow regime is deemed disconnected when Δ{A-B Flow} < 1%). Applying this cutoff, all but 10 of the 200 simulations result in hydraulic disconnection between the alluvial and bedrock aquifer, with the 10 transitional scenarios associated with 75% bedrock channel fractions. he ABRF for this cutoff is 0.08, indicating disconnected systems are characterized by a flow response that is reduced by at least 92%, compared to a fully connected flow regime.

Simulated flow rates from the alluvial to bedrock aquifer transition more smoothly than similar studies that modeled the effects of unsaturated conditions on near-surface infiltration and recharge. Previously reported flow rates follow a clear linear trend during connected conditions and then abruptly transition and flatten at disconnection (e.g., Brunner et al., 2011; Cognac and Ronayne, 2020). Our results show an initial linear period followed by a longer transitional phase that grades into disconnection.

I expect that several factors may explain the long transitional phase observed in models. Complex saturation and hydraulic conductivity fields are dynamic and change throughout the simulated period due to varying saturation. Given the changing nature of heterogeneity and the influence of heterogeneity on flow, it is unlikely that flow rates would stabilize before the saturation-dependent heterogeneity field also becomes stable. Additionally, existing studies have evaluated disconnection dynamics using relatively homogenous fields with a single clogging unit (e.g., Bruen and Osman, 2004; Brunner et al., 2009b), and other studies have limited the analysis to 1D and 2D model domains (e.g., Bruen and Osman, 2004; Desilets et al., 2008; Brunner et al., 2009a; Cognac and Ronayne, 2020). Increased heterogeneity has been found to significantly affect exchange rates, pressure heads, and the state of connection/disconnection (Bruen and Osman, 2004). 2D heterogeneity fields tend to have lower connectivity compared to analogous 3D fields with comparable statistics and architectures (Renard and Allard, 2013). The dynamic 3D heterogeneity described herein produces variable saturation patterns and connection status across the model, and in turn, smoother total A-B flow rates through time.

Importantly, VSF model simulations conducted in this study were transient, which is another factor that may explain the observed smooth, long-duration transitional phase. Fundamental reviews and explanations of aquifer disconnection have utilized exemplary models that run steady-state simulations (Brunner et al., 2009b; 2009b; Brunner et al., 2011). Transient models account for the lag time in changing saturation and pressure head on resulting flows. Previous studies have documented the effect of lag times on recharge rates in models with unsaturated conditions (Hunt et al., 2008). Transient simulations are assumed to produce smoother recorded flow rates.

3.5.2 Heterogeneity controls for inter-aquifer flow

Results demonstrate how geologic heterogeneity strongly influences aquifer exchange rates under scenarios of water table lowering. I plotted the saturated volumes of alluvium, sandstone, and mudstone for several realizations to further evaluate how flow and saturation relate to heterogeneity. Plots highlight how realizations with lower bedrock sandstone fractions (≤35%) tend to have significant volumes of saturated mudstone underlying the alluvium and dispersed saturated sandstone channels with minimal sandstone present at the alluvium-bedrock contact (e.g., ***Figure 24a***). The lack of sandstone at the A-B interface forces flow to occur through the low-K mudstone, limiting the magnitude of A-B flow. In contrast, for realizations with greater sandstone fractions (≥50%), saturated bedrock primarily includes sandstone, and the lack of saturated mudstone indicates flow is occurring primarily through sandstone. Multiple channels intersect the alluvium, creating lateral and vertical conduits for A-B exchange. Greater A-B flow rates tend to occur where channels are highly connected and multiple conduits span the vertical model domain. For scenarios with lower A-B flow rates, sandstone "pinch points" channeled flow into a single, narrow conduit spanning the vertical domain. While it is not clear from evaluated metrics what dictates whether a pinch point will occur, the fluvial architecture clearly dictates the geometry of the saturated zone and therefore recharge conduits. These results further explain the finding that realizations with lower sandstone fractions tend to have less sandstone at the A-B contact and lower A-B flow rates overall.

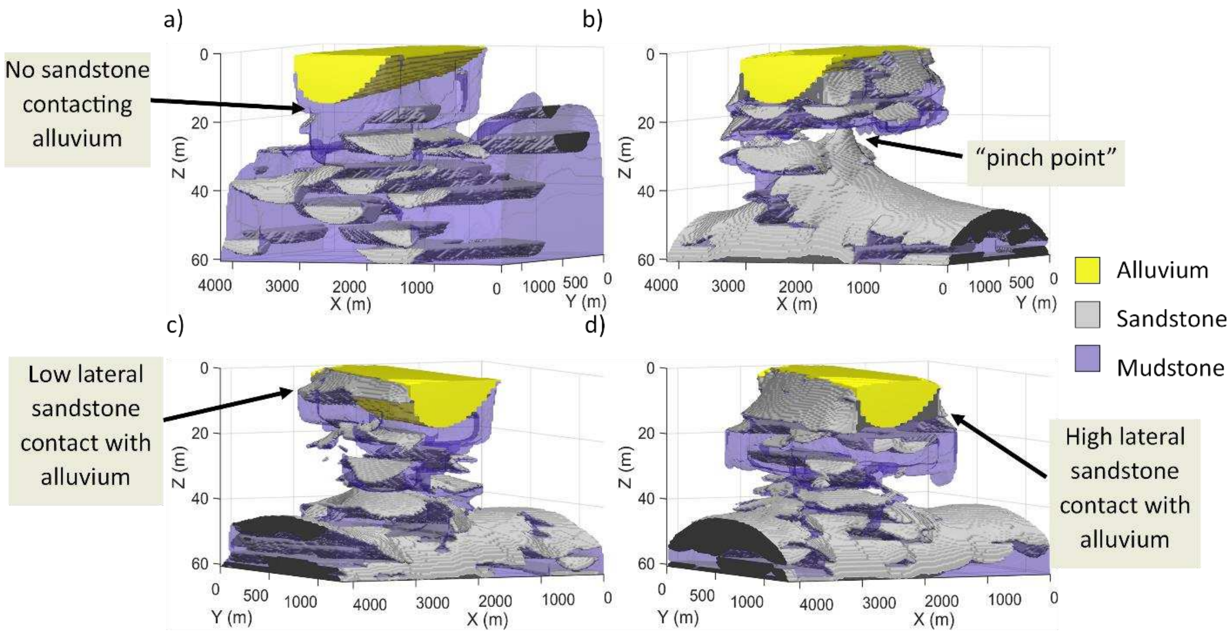

Figure 24. Saturated alluvium, sandstone, and mudstone volumes for example realizations with a) 20% and b) 75% bedrock sandstone fraction, and the c) lowest, and d) highest A-B flow rate for the 50% bedrock sandstone fraction.

Mudstone units underlying the alluvium are analogous to the clogging layer in studies of groundwater-surface water disconnection. Brunner et al. (2009b; 2011) describe how the relative thicknesses of the aquifer and clogging units influence potential for unsaturated conditions to develop. They conclude that while thicker clogging units promote desaturation, a thin aquifer layer relative to the clogging layer can prohibit unsaturated conditions from developing. In this study, I identify a similar relationship. Realizations with thick saturated mudstone units and relatively thin sandstones were less likely to desaturate (***Figure 24a***). Conversely, realizations with large, highly connected sandstone units and thin mudstones tended to produce large unsaturated zones (***Figure 24b***) and lower final A-B flow rates compared to other realizations with the same sandstone fraction.

Volume plots also reveal important dynamics related to the alluvium-bedrock contact. Realizations with more sandstone contacting the alluvium tend to have thicker and wider perched saturated zones (***Figure 24c, 7d***). Thicker perched zones enable greater pressure to form at the perched zone base, promoting greater flow through underlying units. Applying the 1D analogy in *Equation 1* and holding all else equal, a thicker saturated alluvium decreases the potential for unsaturated conditions. Therefore, when thicker perched zones cause greater pressure transfer to underlying units, there is less opportunity for negative pressure to develop. As for widening, sandstone units that contact lateral edges of the alluvium tend to remain saturated. This effectively widens the perched aquifer extent and increases the likelihood that the perched zone intersects underlying channels. Realizations with the highest A-B flow rates tended to have wider perched zones.

Connected component analysis revealed static and dynamic heterogeneity fields for sandstone fractions ≥ 50% to have connected bodies that vertically spanned the model domain and therefore realizations with ≥ 50% sandstone fraction could be expected to have different flow behavior from realizations with lower sandstone fractions. This is confirmed by a comparison of the simulation results for different channel fractions (**Table 7**). Regression analysis revealed that the ZCC_{DY} connectivity metric becomes increasingly predictive of A-B flow for 50% and 75% sandstone fractions (Table 3), and connectivity metrics overall better predict A-B flows with increasing bedrock channel fractions. These results suggest that A-B flow is increasingly controlled by the connectivity of high-K units, and in particular, the presence of vertically spanning channel bodies as bedrock channel fraction increases.

3.5.3 Definition of disconnection

Existing definitions for disconnected flow regimes directly reference a surface water body (e.g., Brunner et al., 2011). In this study, I demonstrate a scenario where the stream and shallow aquifer are connected, but an unsaturated zone develops between two aquifers deeper within the system, such that further water table lowering in the deeper aquifer no longer impacts inter-aquifer exchange rate. While others have described the development of stable stream seepage and perched aquifer conditions as deeper water tables decline (e.g., Rains et al 2006), this study specifically focuses on a stabilizing exchange rate between two aquifers. For these reasons, the definition of disconnected flow regimes should be expanded to include settings beyond the groundwater-surface water interface.

3.5.4 Management implications

When aquifers are pumped, discharge to wells is offset by induced recharge, decreased discharge, loss of storage in the aquifer, or a combination of these (Theis, 1941). The balance of terms reflects the dynamic response of the aquifer to pumping and each term is a key component of the aquifer water budget. In this study, I demonstrate how unsaturated conditions can form between two aquifers thereby limiting inter-aquifer flow and ultimately leading to a disconnected flow regime. Hydraulic disconnection causes inter-aquifer recharge induced by pumping to stabilize, thereby limiting the induced recharge to the deeper aquifer. The disconnected flow regime alters the resulting water budget, which has critical implications for groundwater management, particularly where pumping occurs over extended periods of time.

While exploratory in nature, the groundwater models in this study use realistic inputs representing the Denver Basin aquifer system. Findings suggest that as bedrock aquifer water levels in the Denver Basin decline from continued pumping, unsaturated zones may develop between the alluvial and bedrock aquifer and the potential for hydraulic disconnection may increase. As in other parts of the world, climate change is expected to increasingly stress surface water resources in Colorado (Lukas et al., 2014). As a result, groundwater will likely be increasingly utilized to meet future demands (Aeschbach-Hertig and Gleeson, 2012). In the Denver Basin and other aquifer systems experiencing long-term pumping and water level declines, it will be essential to consider dynamic processes like hydraulic disconnection may limit future rates of replenishment for an increasingly important water supply.

In this study, I demonstrate how detailed numerical modeling with explicit 3D representation of heterogeneity is essential for determining whether, where, and to what extent hydraulic disconnection will occur. Inter-aquifer flow rates that result from water table drawdown are sensitive to aspects of heterogeneity including the spatial connectivity of high-K units and high-K materials at the alluvial-bedrock contact. Regional modeling to evaluate impacts of pumping on aquifers is often limited to coarse-grid resolutions and effective upscaled parameters. While upscaled properties are often adequate for regional-scale applications, they lack necessary detail for reliably simulating processes that are strongly influenced by heterogeneity. Herein, I demonstrate how detailed modeling with explicit representation of heterogeneity is critical for determining inter-aquifer flow dynamics in heterogeneous aquifers with significant water table drawdown (i.e., long-term pumping).

As has been pointed out for stream-aquifer scenarios, disconnection between the alluvial and bedrock aquifer does not imply that A-B flow rates are unaffected by additional stresses within the system. Just as pumping near a disconnected stream can increase the extent of disconnection (Fox and Durnford, 2003), additional pumping in the bedrock aquifer could expand the spatial extent of disconnection between two aquifers. Further, just as changes in streamflow can affect seepage from disconnected streams (Sophocleous, 2002; Vazquez-Sune, 2007), changes in alluvial groundwater levels could impact A-B flow in the disconnected system.

3.6 Summary and conclusions

This study demonstrates how unsaturated conditions can develop between an alluvial and bedrock aquifer creating a scenario where further lowering of the deeper bedrock water table no longer increases alluvial-to-bedrock flow. I quantify successive changes in A-B flow that result from water table lowering and distinguish transitional from disconnected flow regimes using a threshold value of 1% change in flow between successive model stress periods. I also develop a normalized response function to quantify the changing relationship between water table lowering and the resulting change in inter-aquifer flow, which enables comparison across scenarios with different flow magnitudes. Of the 200 flow simulations, 190 reached full disconnection and 10 concluded with transitional flow regimes that were approaching disconnection.

The transition from connection to disconnection is strongly controlled by heterogeneity within the bedrock aquifer such that spatial relationships between high-K sandstone channels and low-K mudstones dictate where unsaturated regions form and the geometry of high-K flow paths. Final alluvial-to-bedrock flow rates are highly variable for models where only the locations of channelized sandstones differ, highlighting the sensitivity of aquifer exchange rates to heterogeneity. In particular, the connectivity of sandstone channels and amount of sandstone present at the alluvium-bedrock interface strongly predict resulting aquifer exchange rates.

This study extends the definition of hydraulic disconnection to a scenario involving flow between aquifers and provides a representative case study to illustrate where this may occur.

Further, I highlight how the potential for disconnection is a critical consideration for assessing aquifer response to pumping, particularly when water table drawdown is significant.

3.7 References

Aeschbach-Hertig, W., & Gleeson, T. (2012). Regional strategies for the accelerating global problem of groundwater depletion. Nature Geoscience, 5(12), 853-861.

Anderson, M.P. (1991), Universal scaling of hydraulic conductivities and dispersivities in geologic media – comment. Water Resources Research 27(6):1381–2. doi:10.1029/91WR00574

Banks, E. W., Simmons, C. T., Love, A. J., & Shand, P. (2011), Assessing spatial and temporal connectivity between surface water and groundwater in a regional catchment: Implications for regional scale water quantity and quality. *Journal of Hydrology*, 404(1-2), 30-49. doi: 10.1016/j.jhydrol.2011.04.017

Brooks, R. H., & Corey, A. T. (1966), Properties of porous media affecting fluid flow. *Journal of the irrigation and drainage division*, 92(2), 61-88. doi: 10.1061/jrcea4.0000425

Brown & Caldwell (2017), South Platte Decision Support System Alluvial Groundwater Model Update Documentation. Colorado Water Conservation Board and Division of Water Resources.

Bruen, M. P., & Osman, Y. Z. (2004), Sensitivity of stream–aquifer seepage to spatial variability of the saturated hydraulic conductivity of the aquifer. *Journal of Hydrology*, 293(1-4), 289-302. doi: 10.1016/J.JHYDROL.2004.02.003

Brunner, P., Cook, P.G., & Simmons, C. T. (2009a), Hydrogeologic controls on disconnection between surface water and groundwater. *Water Resources Research* 45, W01422. doi: 10.1029/2008WR006953. doi:10.1029/2008WR006953

Brunner, P., Simmons, C.T., & Cook, P.G. (2009b), Spatial and temporal aspects of the transition from connection to disconnection between rivers, lakes and groundwater. *Journal of Hydrology* 376: 159–169. doi:10.1016/j.jhydrol.2009. 07.023.

Brunner, P., Cook, P. G., & Simmons, C. T. (2011), Disconnected surface water and groundwater: from theory to practice. *Groundwater*, 49(4), 460-467. doi: 10.1111/j.1745-6584.2010.00752.x

Cardiff, M., Bakhos, T., Kitanidis, P. K., & Barrash, W. (2013), Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential. *Water Resources Research*, 49(9), 5395-5410. doi: 10.1002/wrcr.20356

Carsel, R. F., & Parrish, R. S. (1988), Developing joint probability distributions of soil water retention characteristics. *Water Resource Research* 24(5):755–769 doi: 10.1029/wr024i005p00755

Cognac, K. E., & Ronayne, M. J. (2020), Changes to inter-aquifer exchange resulting from long-term pumping: implications for bedrock groundwater recharge. *Hydrogeology Journal* 28(4): 1359-1370. doi:10.1007/s10040-020-02141-x

Colorado Division of Water Resources (CDWR) (2020), DWR Well Geophysical Log Database, Accessed 01/20/2020. https://dwr.state.co.us/Tools/GroundWater/GeoPhysicalLogs

Colorado Division of Water Resources (CDWR) (2022), State water level database, Statewide, Accessed 11/17/2022. https://dwr.state.co.us/Tools/GroundWater/WaterLevels

Dann, R. L., Close, M. E., Pang, L., Flintoft, M. J., & Hector, R. P. (2008), Complementary use of tracer and pumping tests to characterize a heterogeneous channelized aquifer system in New Zealand. Hydrogeology Journal, 16, 1177-1191. doi: 10.1007/s10040-008-0291-4

de Marsily, G., Delay, F., Goncalves, J., Renard, P., Teles, V., & Violette, S. (2005), Dealing with spatial heterogeneity. *Hydrogeology Journal*, 13(1), 161-183. doi: 10.1007/S10040-004-0432-3

DesRoches, A. J., Butler, K. E., & Pelkey, S. (2013), Influence of fracture anisotropy and lithological heterogeneity on wellfield response in a fluvial sandstone aquifer of the Carboniferous Moncton Subbasin, Canada. Hydrogeology Journal, 21(3), 559. doi:10.1007/s10040-012-0931-6

Desilets, S. L. E., Ferre´, T. P. A., & Troch, P. A. (2008), Effects of stream-aquifer disconnection on local flow patterns, *Water Resources Research*, 44, W09501, doi:10.1029/2007WR006782.

Deutsch, C. A., Tran, T. T. (2002), FLUVSIM: a program for object-based stochastic modeling of fluvial depositional systems. *Computers Geosciences*, 28(4), 525-535. doi: 10.1016/S0098-3004

Fleckenstein, J. H., Niswonger, R. G., & Fogg, G. E. (2006), River-aquifer interactions, geologic heterogeneity, and low-flow management. *Groundwater*, 44(6), 837-852. doi: 10.1111/j.1745-6584.2006.00190.x

Fox, G.A., & Durnford, D.S. (2003), Unsaturated hyporheic zone flow in stream/aquifer conjunctive systems. *Advances in Water Resources* 26: 989–1000. doi: 10.1016/S0309-1708(03)00087-3

Fox, G. A., & Gordji, L. (2007), Consideration for unsaturated flow beneath a streambed during alluvial well depletion. *Journal of Hydrologic Engineering*, 12(2), 139-145. doi: 10.1061/(ASCE)1084-0699(2007)12:2(139)

Frei, S., Fleckenstein, J. H., Kollet, S. J., & Maxwell, R. M. (2009), Patterns and dynamics of river–aquifer exchange with variably-saturated flow using a fully-coupled model. *Journal of Hydrology*, 375(3-4), 383-393.

Graham, G., & Vanslyke, G. (2004), Development of the regulatory framework for Denver Basin aquifers. The Mountain Geologist. doi: 10.1016/j.jhydrol.2009.06.038

Harbaugh, A.W. (1990), A computer program for calculating subregional water budgets using results from the U.S. Geological Survey modular three-dimensional ground-water flow model: U.S. Geological Survey Open-File Report 90-392, 46 p. doi:10.3133/ofr90392

Harbaugh, A. W., Banta, E. R., Hill, M. C., & McDonald, M. G. (2000), Geological Survey modular ground-water model—User guide to modularization concepts and the ground-water flow process. US Geological Survey Open-File Report 00–92. doi: 10.3133/ofr200092

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. *Groundwater*, 46(4), 551-560. doi: 10.1111/j.1745-6584.2007.00427.x

Irvine, D. J., Brunner, P., Franssen, H. J. H., & Simmons, C. T. (2012), Heterogeneous or homogeneous? Implications of simplifying heterogeneous streambeds in models of losing streams. *Journal of Hydrology*, 424, 16-23. doi: 10.1016/j.jhydrol.2011.11.051

King, P. R. (1990), The connectivity and conductivity of overlapping sand bodies. In North Sea Oil and Gas Reservoirs—II: Proceedings of the 2nd North Sea Oil and Gas Reservoirs Conference organized and hosted by the Norwegian Institute of Technology (NTH), Trondheim, Norway, May 8–11, 1989 (pp. 353-362). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-009-0791-1_30

Knudby, C., & Carrera, J. (2006), On the use of apparent hydraulic diffusivity as an indicator of connectivity. *Journal of Hydrology*, 329(3-4), 377-389. doi: 10.1016/j.jhydrol.2006.02.026

Koltermann, C. E., & Gorelick, S. M. (1996), Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research*, 32(9), 2617-2658. doi: 10.1029/96WR00025

Lapey, L. A. (2001), Hydrogeologic parameters of the Kiowa Research Core, Kiowa, Colorado. M.S. Thesis Colorado State University, Fort Collins, Colorado.

Lee, S. Y., Carle, S. F., & Fogg, G. E. (2007), Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation. *Advances in water resources*, 30(9), 1914-1932. doi: 10.1016/j.advwatres.2007.03.005

Lindsey, K. O., Barkmann, P. E., & Sebol, L. A. (2021) ON-010-02D Colorado Statewide Alluvial Aquifer (Data). Aquifers, Variable. Golden, CO: Colorado Geological Survey. https://coloradogeologicalsurvey.org/publications/colorado-alluvial-aquifer-data/.

Lukas, J., J. Barsugli, N. Doesken, I. Rangwala, and K. Wolter (2014), Climate Change in Colorado A Synthesis to Support Water Resources Management and Adaptation, Western Water Assessment, Boulder, CO.

Maples, S. R., Foglia, L., Fogg, G. E., Maxwell, R. M. (2020), Sensitivity of hydrologic and geologic parameters on recharge processes in a highly heterogeneous, semi-confined aquifer system. Hydrology and Earth System Sciences, 24(5), 2437-2456. doi: 10.5194/hess-24-2437-2020

Meier, P. M., Carrera, J., Sánchez-Vila, X. (1998), An evaluation of Jacob's method for the interpretation of pumping tests in heterogeneous formations. *Water Resources Research*, 34(5), 1011-1025. doi: 10.1029/98WR00008

Musgrove, M., Beck, J. A., Paschke, S. S., Bauch, N. J., Mashburn, S. L. (2014), Quality of Groundwater in the Denver Basin Aquifer System, Colorado, 2003–5. US Department of the Interior, US Geological Survey. doi: 10.3133/sir20145051

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 Techniques and Methods*, book 6, chap. A45, 66 p. doi: 10.5066/F7R20ZFJ

Panday, S., (2019), USG-Transport Version 1.4.0: The Block-Centered Transport Process for MODFLOW-USG, GSI Environmental, http://www.gsi-net.com/en/software/free-software/USG-Transport.html

Parajuli, K., Sadeghi, M., Jones, S. B. (2017), A binary mixing model for characterizing stony-soil water retention. Agric For Meteorol 244:1–8. doi: 10.1016/j.agrformet.2017.05.013

Pardo-Iguzquiza, E. & Dowd, P. (2003), CONNEC3D: a computer program for connectivity analysis of 3D random set models, *Computational Geosciences*, 29, 775–785. doi: 10.1016/S0098-3004(03)00028-1

Paschke, S.S. ed., (2011), Groundwater availability of the Denver Basin aquifer system, Colorado: U.S. Geological Survey Professional Paper 1770, 274 p. doi: 10.3133/pp1770

Pyrcz, M. J., & Deutsch, C. V. (2014), Geostatistical reservoir modeling. Oxford university press, USA. doi: 10.1198/tech.2005.s339

Rains, M. C., Fogg, G. E., Harter, T., Dahlgren, R. A., & Williamson, R. J. (2006), The role of perched aquifers in hydrological connectivity and biogeochemical processes in vernal pool landscapes, Central Valley, California. *Hydrological* Processes, 20(5), 1157-1175. doi: 10.1002/hyp.5937

Raynolds, R. G., Johnson, K. R., Arnold, L. R., Farnham, T. M., Fleming, R. F., Hicks, J. F., Kelley, S. A., Lapey, L. A., Nichols, D. J., Obradovich, J. D., Wilson, M. D. (2001), The Kiowa core, a continuous drill core through the Denver Basin bedrock aquifers at Kiowa, Elbert County, Colorado: *U.S. Geological Survey Open-File Report* 01-185. doi: 10.3133/ofr01185

Raynolds, R. G. (2002), Upper Cretaceous and Tertiary stratigraphy of the Denver basin, Colorado. *Rocky Mountain Geology*, 37(2), 111-134. doi: 10.2113/GSROCKY.37.2.111

Renard, P., & Allard, D. (2013), Connectivity metrics for subsurface flow and transport. Advances in Water Resources, 51, 168-196. doi: 10.1016/j.advwatres.2011.12.001

Ronayne, M. J., Gorelick, S. M., & Caers, J. (2008), Identifying discrete geologic structures that produce anomalous hydraulic response: An inverse modeling approach. *Water Resources Research*, 44(8). doi:10.1029/2007WR006635

Rosenberry, D. O. (2000), Unsaturated-zone wedge beneath a large, natural lake. *Water Resources Research*, 36(12), 3401-3409. doi:10.1029/2000WR900213

Ruybal, C. J., Hogue, T. S., & McCray, J. E. (2019), Assessment of groundwater depletion and implications for management in the Denver basin aquifer system. *JAWRA Journal of the American Water Resources Association*, 55(5), 1130-1148. doi: 10.1111/1752-1688.12755

Schilling, O. S., Irvine, D. J., Hendricks Franssen, H. J., & Brunner, P. (2017), Estimating the spatial extent of unsaturated zones in heterogeneous river-aquifer systems. *Water Resources Research*, 53(12), 10583-10602. doi: 10.1002/2017WR020409

Shanafield, M., Cook, P. G., Brunner, P., McCallum, J., & Simmons, C. T. (2012), Aquifer response to surface water transience in disconnected streams. *Water Resources Research*, 48(11). doi: 10.1029/2012WR012103

Morel-Seytoux, H. J., Meyer, P. D., Nachabe, M., Tourna, J., Van Genuchten, M. T., & Lenhard, R. J. (1996), Parameter equivalence for the Brooks-Corey and van Genuchten soil characteristics: Preserving the effective capillary drive. *Water Resources Research*, 32(5), 1251-1258. doi: 10.1029/96WR00069

Sophocleous, M. (2002), Interactions between groundwater and surface water: the state of the science. *Hydrogeology journal*, 10, 52-67. doi: 10.1007/s10040-002-0204-x

Swanson, S. K., Bahr, J. M., Bradbury, K. R., &Anderson, K. M. (2006), Evidence for preferential flow through sandstone aquifers in Southern Wisconsin. *Sedimentary Geology*, 184(3-4), 331-342. doi: 10.1016/j.sedgeo.2005.11.008

Su, G. W., Jasperse, J., Seymour, D., Constantz, J., & Zhou, Q. (2007), Analysis of pumping-induced unsaturated regions beneath a perennial river. *Water Resources Research*, 43(8). doi: doi.org/10.1029/2006WR005389

Tang, Q., Kurtz, W., Schilling, O. S., Brunner, P., Vereecken, H., & Franssen, H. J. H. (2017), The influence of riverbed heterogeneity patterns on river-aquifer exchange fluxes under different connection regimes. *Journal of Hydrology*, 554, 383-396. doi: 10.1016/j.jhydrol.2017.09.031

Theis, C. V. (1941), The effect of a well on the flow of a nearby stream. Eos, *Transactions American Geophysical Union*, 22(3), 734-738. doi: 10.1029/TR022i003p00734

van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci Soc Am J* 44: 892–898. doi: 10.2136/sssaj1980.03615995004400050002x

Vázquez-Suñé, E., Capino, B., Abarca, E., & Carrera, J. (2007), Estimation of recharge from floods in disconnected stream-aquifer systems. *Groundwater*, 45(5), 579-589. doi: 10.1111/j.1745-6584.2007.00326.x

Wada, Y., Van Beek, L. P., Van Kempen, C. M., Reckman, J. W., Vasak, S., Bierkens, & M. F. (2010), Global depletion of groundwater resources. *Geophysical research letters*, 37(20). doi: 10.1029/2010GL044571

Werner, A. D., Zhang, Q., Xue, L., Smerdon, B. D., Li, X., Zhu, X., Yu, L., & Li, L. (2013), An initial inventory and indexation of groundwater mega-depletion cases. *Water resources management*, 27, 507-533. doi: 10.1007/s11269-012-0199-6

Woodard, L., Sanford, W., & Raynolds, R. G. (2002), Stratigraphic variability of specific yield within bedrock aquifers of the Denver Basin, Colorado. *Rocky Mountain Geology*, 37(2), 229-236. doi: 10.2113/gsrocky.37.2.229

CHAPTER 4: INFLUENCE OF BEAVER DAMS ON INTER-AQUIFER EXCHANGE RATES IN A STREAM-ALLUVIAL-BEDROCK AQUIFER SYSTEM WITH PREDOMINANTLY DOWNARD FLUXES.

4.1 Introduction

Beavers are ecosystem engineers. By cutting wood and building dams, they alter hydrology, geomorphology, ecology, and related feedbacks to actively construct their habitat and generate large-scale ecological niches (Larsen et al., 2021). Because of this, beavers are increasingly being recognized for their ability to enhance water resources and offset the negative effects of climate change and drought (Bird et al., 2011; Pilliod et al., 2018; Westbrook et al., 2020; Ronnquist and Westbrook, 2021).

Prior to colonization, beaver populations ranged from 60–400 million in North America (Naiman et al., 1988). However, by the 19th century, intensive trapping and eradication reduced populations to near extinction (Baker and Hill, 2003). Over the past 100 years, conservation, reintroduction, and regulations on trapping have enabled a partial recovery, and have coincided with expanding efforts to reintroduce beavers for ecosystem restoration (Baker and Hill, 2003; Andersen et al., 2010; Pollock et al., 2017; Bailey et al., 2019). Synthetic beaver dams (i.e., beaver dam analogues) are also being used to mimic the positive effects of beavers (Wade et al., 2020; Scamardo and Wohl, 2021) and re-attract beavers to their historic habitat (Bailey et al., 2019). With rising beaver populations and growing interest in utilizing beavers for restoration, it is increasingly important to document the hydrologic effects of beavers on surface water and groundwater systems.

Particularly on low-order streams, beavers construct channel-spanning dams that promote a range of surface and near-surface hydrologic processes. Beaver dams impound water, which increases surface and groundwater storage through open-water ponds, wetlands, water table rise, increased lateral connectivity with the floodplain, and transient storage within the hyporheic zone (Lowry, 1993; Westbrook et al., 2006; Jin et al., 2009; Puttock et al., 2017; Wegener et al., 2017). By reducing stream velocity and retaining flow during storm events, beaver dams effectively reduce stormwater runoff (Bailey et al., 2019), increase water retention (Johnston & Naiman, 1987; Parker, 1986; Scamardo et al., 2022), attenuate smaller floods (Westbrook et al., 2006; Puttock et al., 2017), and in some cases increase baseflow (Majerova et al., 2015), Puttock et al., 2017; Smith et al., 2020). Beaver dams trap large volumes of sediment which adds complexity to streambed morphology, raises the streambed, and drives water vertically into the streambed and laterally into the adjacent floodplain serving to enhance and diversify hyporheic exchange (Westbrook et al., 2006; Lautz and Siegel, 2006; Briggs et al., 2013; Wade et al., 2020). The hydrologic impacts of beavers have far-reaching influence on many other riparian processes including nutrient and temperature exchange, water residence time, carbon storage, and the development of floodplain soils and vegetation (Wegener et al., 2017; Larsen et al., 2021).

While surficial and near-surface hydrologic impacts of beaver are increasingly well documented, the influence of beaver on deeper groundwater dynamics is less often considered. Increased river stage associated with natural and simulated beaver dams (i.e., <1m) have been found to raise shallow groundwater levels and hydraulic gradients extending tens of meters into the adjacent floodplain (Lowry, 1993; Westbrook et al., 2006; Bouwes et al., 2016). Increases in near-surface hydraulic heads have potential to alter deeper groundwater fluxes. For example, where regional groundwater flow is predominantly downward (e.g., in recharge zones), increased shallow groundwater head may promote stronger vertical and lateral fluxes into underlying and adjacent hydrogeologic units. Conversely, where regional groundwater flow paths are directed upward or towards the river, the influence of beaver dams on shallow and deep groundwater dynamics could be limited. This may provide an explanation for studies that have documented limited impacts to groundwater from beaver dams (and their analogues) (Feiner and Lowry, 2015; Scamardo and Wohl, 2021).

This research examines the potential for a beaver dam to influence deeper (>5-30 m) aquifer dynamics in a stream-alluvial-bedrock sequence. Data used to parameterize numerical models are from the Denver Basin. Within heavily pumped areas of the Denver Basin aquifer system, it is common for stream-alluvial-bedrock aquifer sequences to exhibit predominantly downward fluxes (Paschke et al., 2011; Cognac and Ronayne, 2020). Further, documenting sources of recharge within the Denver Basin is increasingly important as aquifers undergo depletion. Herein, I test the potential for beaver dams to increase bedrock aquifer recharge by altering hydraulic gradients at the alluvium-bedrock interface (***Figure 25***). I quantify the effects of a dam on alluvial to bedrock exchange rates using a numerical flow model. Multiple modeling scenarios are performed to test whether the influence of the dam is sensitive to the pond depth and the alluvial-bedrock contact depth. While the flow model is simple by design, inputs incorporate field observations and a realistic hydrogeologic setting, highlighting the potential for beaver dams to influence deeper aquifer dynamics.

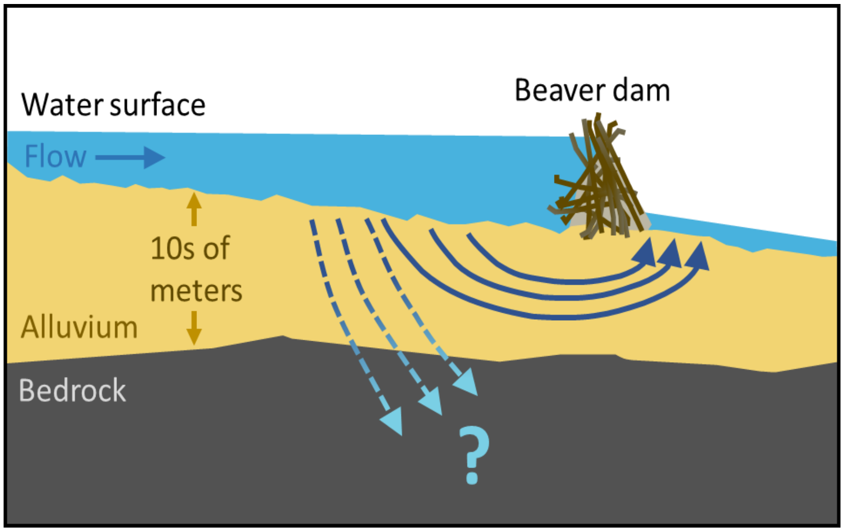

Figure 25. Conceptual figure depicting the potential influence of a beaver dam on alluvial-bedrock flow rates.

4.2 Study Area

The beaver dam represented in this study is based on a channel-spanning beaver dam that was constructed on Cherry Creek (CC) in east-central Colorado, USA during the fall of 2020 (***Figure 26***). The dam site is approximately 22 km east of the Rocky Mountain Front Range and 40 km north of the Palmer Divide, a topographic high that extends east from the Rampart Range and sources headwater tributaries of CC (Figure 27). CC is a second-order perennial stream with a meandering planform. Within the study area, CC is characterized by pool-riffle sequences, sand-and-gravel streambed, and typical channel widths of 3-5 m. The floodplain and channel are confined by quaternary alluvial terraces and incised banks. Variable discharge and intense flooding are common along CC due to convective storms. Between 1940 and 2023, the annual peak discharge along CC ranged from 34 to 9,170 cfs at a USGS stream gage located 8 km upstream (South) from the study site (U.S. Geological Survey, 2023).

Figure 26. Photographs of the beaver dam on Cherry Creek with piezometer nest visible upstream of the dam (left), and evidence of active beavers near the dam site (right).

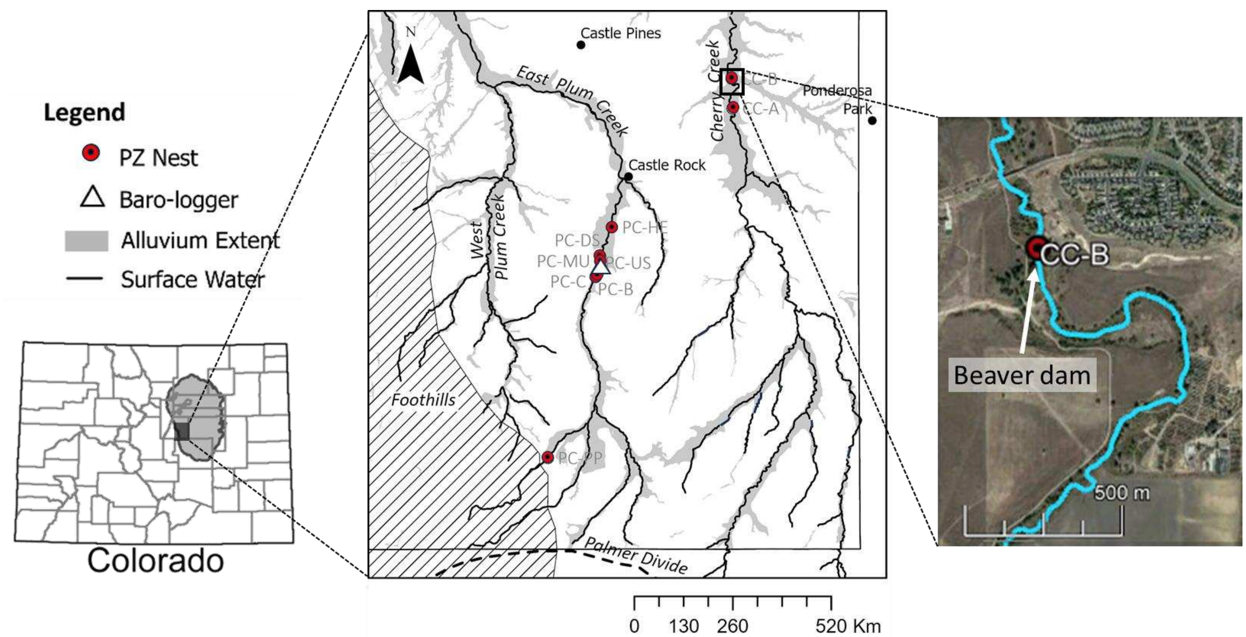

Figure 27. Location of beaver dam and nested piezometer monitoring network.

A piezometer (PZ) nest located approximately 7 m upstream of the dam captured the influence of the dam's construction and the following 7-months of stream stage, shallow groundwater levels, and hydraulic gradient. PZ data recorded a rapid increase in stage (< a few days) of 0.45 m and widening of the channel from 3 m to 5 m, accompanied by increased seepage (downwelling) to 9x the pre-dam rate (***Figure 28***). Over the following several months, seepage effects from the dam gradually returned to pre-dam conditions, however, the streambed and river stage remained elevated even 7-months later.

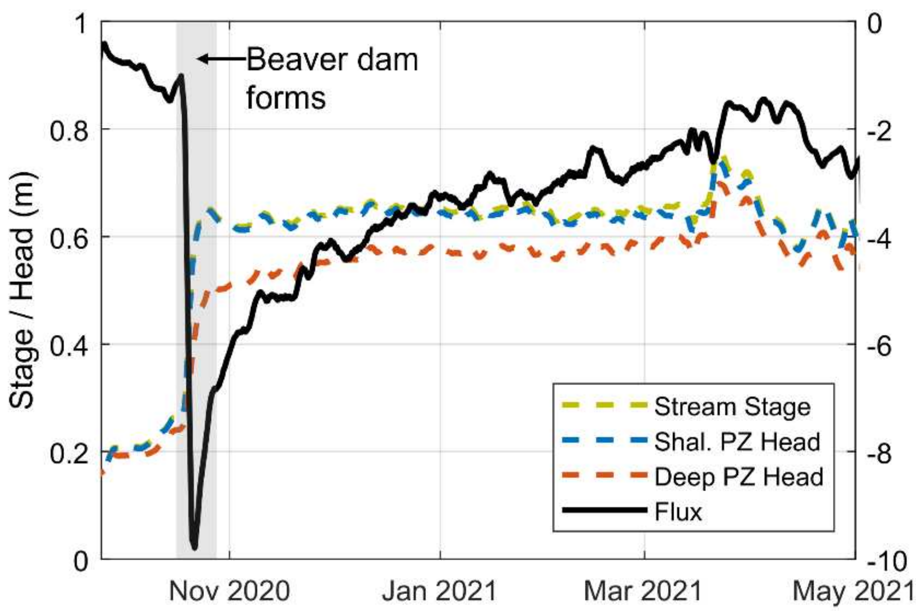

Figure 28. Effects of the fall 2020 beaver dam on stream stage, shallow and deep streambed hydraulic heads, and vertical water flux approximately 7 m upstream of the dam.

Cherry Creek overlies an alluvial aquifer system that spans the South Platte Aquifer and its tributaries. The alluvial aquifer is incised and overlain into the Denver Basin aquifer system (DBAS), a regionally significant water resource for growing populations along Colorado's Front Range urban corridor. The DBAS contains a series of confined and unconfined sedimentary bedrock aquifers that have undergone significant long-term pumping and associated head declines during recent decades (Paschke et al., 2011).

4.3 Methods

A groundwater flow model was developed to evaluate whether the presence of a beaver dam affects deeper exchange fluxes between an alluvial and bedrock aquifer. A base case scenario was developed to reflect conditions associated with the 2020 beaver dam on Cherry Creek. The influence of the dam is expected to vary with alternative beaver pond depth, alluvium-bedrock contact depths, and aquifer hydraulic properties. Therefore, I perform a sensitivity analysis to quantify the influence of the dam over a range of stream and aquifer conditions.

I simulate groundwater flow using the block-centered finite-difference code MODFLOW-2005 (Harbaugh, 2005). The model domain is discretized into 150 rows, 150 columns, and 26 layers with a regular horizontal grid spacing of $\Delta x = 13.5$ m, $\Delta y = 16.8$ m, and variable vertical spacing of $\Delta z = 1$ to 15 m (***Figure D1 in Appendix D***). The lateral extent of the alluvial aquifer was assigned using maps developed by Barkmann et al. (2015). The vertical extent of the alluvial aquifer ranges from 5 to 30 m within the region based on nearby well boring logs (CDWR, 2023). I assume an alluvial-bedrock contact depth of 10 m for the base case scenario.

External boundaries along lateral edges (X = 0 m and X = 2025 m) and the base of the model (z = 0 m) are specified as no-flow to represent regional streamlines parallel to the direction of flow and assuming minimal cross-gradients (***Figure 29***). The up- and down-gradient edges are assigned head-dependent flux boundaries as implemented in MODFLOW's General Head Boundary (GHB) Package (Harbaugh et al., 2000). These boundaries represent regional hydraulic heads external to the model and values were assigned to enforce downward fluxes between the alluvial and bedrock aquifer which are characteristic of the region. The perennial stream was modeled using a head-dependent flux boundary as implemented in MODFLOW's River Package. The simulated flux into or out of river cells is proportional to an assigned stream width, stage, and streambed hydraulic conductivity, as well as the model calculated head difference between the stream and adjacent aquifer (Harbaugh et al., 2000).

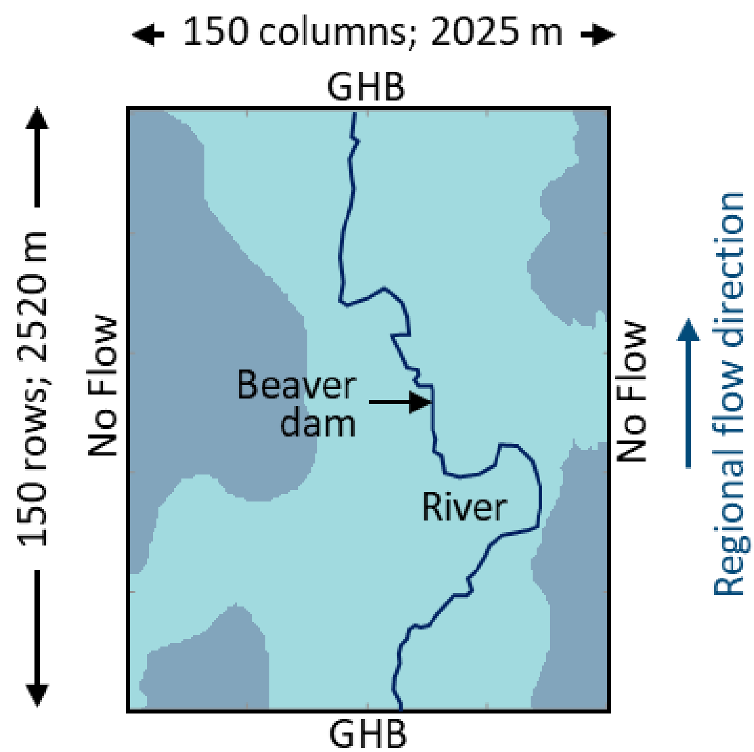

Figure 29. Model domain and boundary conditions (alluvium-bedrock contact from Barkmann et al., 2015).

Model parameters for the stream and aquifers were assigned based on published values and field measurements that were previously collected along Cherry Creek. Horizontal hydraulic conductivities were assigned constant values for alluvial and bedrock aquifers of 100 m day^{-1} and 1 m day^{-1} respectively based on values reported by Paschke et al. (2011). Vertical hydraulic was assigned using an anisotropy ratio (horizontal K to vertical K) of 10. Vertical hydraulic conductivity of the streambed was assigned a value of 10 m day^{-1} based on falling-head tests conducted within streambed sediments along Cherry Creek in the vicinity of the dam location (Cognac and Ronayne, 2023).

To quantify the influence of the beaver dam, scenarios with and without a dam ('dam' and 'no dam' respectively) were generated wherein all model conditions were identical except for the river boundary cells which were modified to reflect the influence of the dam. This included raising and widening river stage assigned to cells immediately upstream of the dam location (model row 74, column 94; Figure 29). The base case dam effects included a rise in stage of 0.45 m and increase in river width from 3 m to 5 m.

Additional scenarios were developed to test the sensitivity of results to a range of conditions. Larger magnitude stage increases due to impounded water behind the dam are expected to have a greater potential to influence deeper aquifer dynamics. In addition to the base case stage increase of 0.45m, I test stage increases of 0.25 m, 1 m, and 1.5 m to better understand how ponding depth influences results. Shallower alluvium-bedrock contact depths are expected to produce greater impacts from the dam on flow rates across the interface. In addition to the base case contact depth of 10m, I test alluvium-bedrock contact depths ranging from 5 to 30 m (***Figure 30***). To minimize the impact of variable alluvial geometry on results, I maintain a consistent lateral extent of alluvium with changes in depth. Lastly, the ratio of the alluvium to bedrock hydraulic conductivity (K_A:K_B) is expected to control the rate of exchange between the two aquifers, and therefore the influence of the dam. In addition to the base case K_A:K_B of 100, I test additional ratios of 10 and 1000.

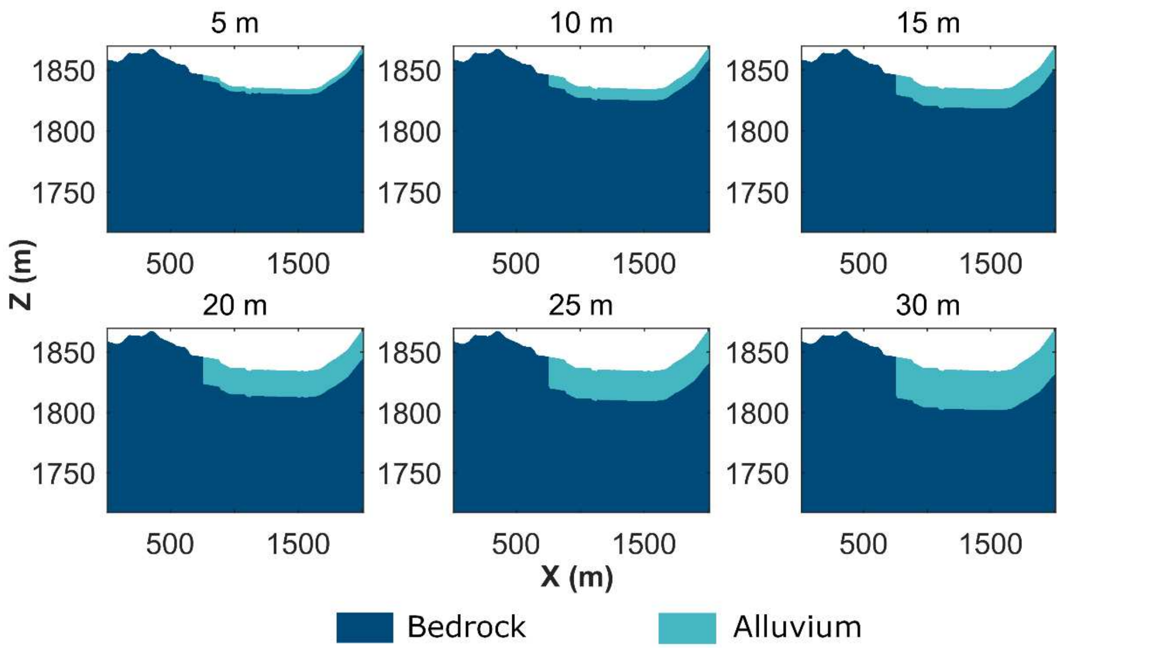

Figure 30. Alluvium-bedrock contact depths and geometries shown for the model row corresponding to the beaver dam location (row 75).

A steady-state simulation was performed for the base case and sensitivity scenarios. All model scenarios are summarized in Table 9. Flow rates between the river and aquifer units were calculated using the USGS program ZoneBudget-USG (Harbaugh, 1990). Upgradient and downgradient rows were excluded from the analysis to avoid boundary effects and to limit the analysis to the 850 m up- and down-stream of the dam.

Table 9 - Summary of model base case and sensitivity scenarios.

	Base Case	Sensitivity Scenarios
Beaver Dam	dam, no dam	dam, no dam
A-B Contact Depth (m)	10	5, 15, 20, 25, 30
Stage increase (m)	0.45	.25, 1, 1.5
Alluvial K / Bedrock K	100	10, 1000

4.4 Results

4.4.1 *Base Case*

Results document the influence of the beaver dam on flow rates throughout the stream, alluvium, and bedrock aquifer system. For the base case scenario, the addition of the beaver dam caused an increase in stream seepage to the alluvium (2307 to 2485 m^3 day^{-1}), an increase in groundwater discharge to the stream (2502 to 2648 m^3 day^{-1}), an increase in flow from the alluvial to bedrock aquifer (A-B flow) (964 to 972 m^3 day^{-1}), and an increase in flow from the bedrock to alluvial aquifer (B-A flow) (186 to 186 m^3 day^{-1}) (Table 10).

Because the exchange of water across the alluvium-bedrock contact included upwards and downwards components, the net A-B flow rate was also calculated by subtracting B-A flow from A-B flow. The net A-B flow increased from 777 to 786 m^3 day^{-1} following the addition of the beaver dam. This corresponds to an increase of 1.1%. While the percentage increase is minimal, over the course of a year, this would amount to volumetric inflow of 3,285 m^3 (~ 8.7×10^5 gallons) to the bedrock aquifer.

Table 10. Summary of flow results from the steady-state simulation for the base case scenario

	No Dam	Dam
A-B Flow (m^3 d^{-1})	964	972
B-A Flow (m^3 d^{-1})	186	186
River Seepage (m^3 d^{-1})	2307	2485
GW Discharge (m^3 d^{-1})	2502	2649
Net A-B Flow (m^3 d^{-1})	777	786

4.4.2 *Sensitivity Scenarios*

As predicted, larger magnitude stage increases corresponded to greater impacts from the beaver dam on A-B flow. The positive change in A-B flow due to the dam increased linearly with the increase in stage (***Figure 31a***). A rise in stage directly translates to an increase in the hydraulic gradient, which in hydraulically connected aquifer systems can result in a linear increase in recharge. The linear relationship between stage increase and A-B flow indicates that stage increases due to the dam produce increased A-B flow and a change in the hydraulic gradient that propagates across the A-B interface. Therefore, the dam influences deeper aquifer heads and fluxes for all scenarios tested.

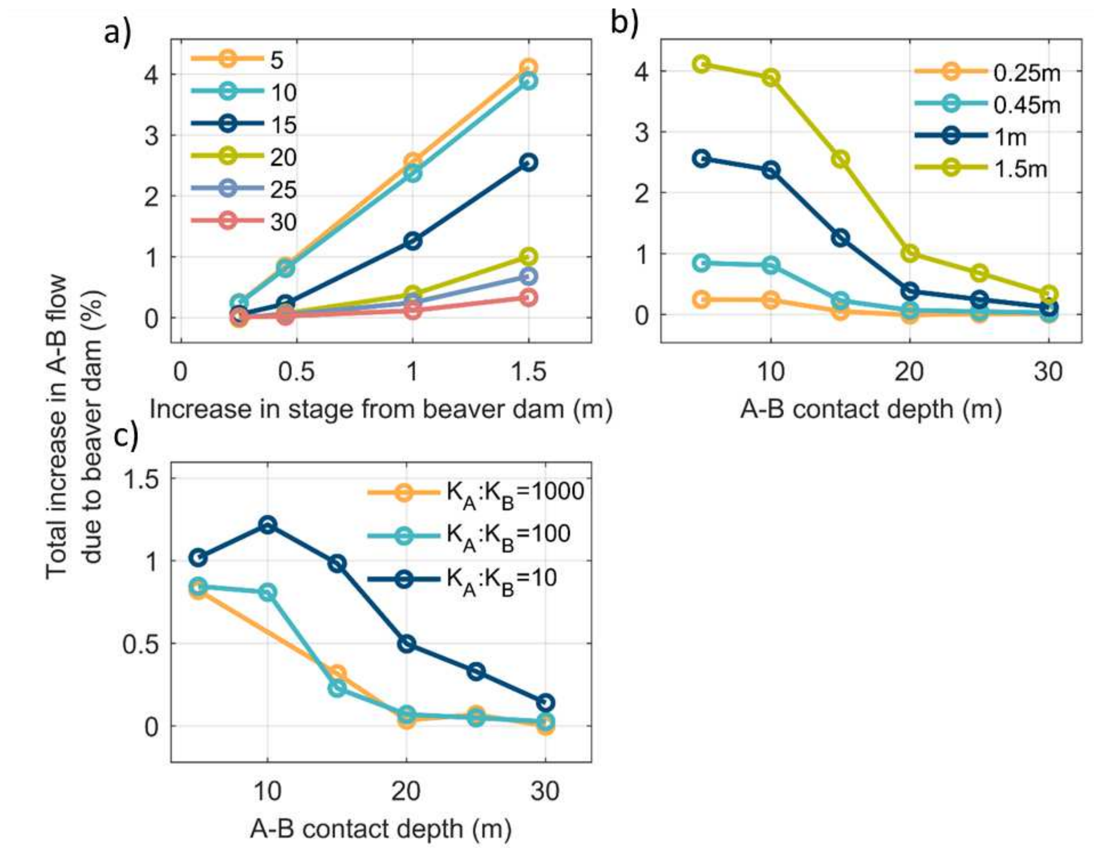

Figure 31 - Model results from the sensitivity analysis demonstrating how the total increase in A-B flow caused by the beaver dam is sensitive to a) stream stage for tested A-B contact depths b) the A-B contact depth for tested stage levels, and c) the ratio of the alluvium K (K_A) to bedrock K (K_B) for all contact depths.

The influence of the beaver dam was greatest when the alluvium-bedrock contact was shallow. In other words, as the contact position deepened, the induced increase in total A-B flow generally decreased (Figure 31b). This trend follows a reverse sigmoid curve where contact depths below 10 m appear to approach a maximum percent increase in A-B flow and contact depths greater than 20 m show increasingly negligible increase.

Interestingly, the sensitivity to stage and A-B contact depth differ for total and net A-B flow (net = total - B-A flow) (Figure 32). While the total increase in A-B flow progressively decreases with increased contact depth, the net increase in A-B flow is greater for the 10 m alluvium-bedrock compared to 5 m. This is due to a component of B-A flow offsetting the net exchange for the 5 m depth scenario. Thus, in the model, the beaver dam drives vertical flow in both the upward and downward direction across the A-B contact. This trend was consistent across all tested magnitudes of stage increase. B-A flow results indicate that B-A flow increased for shallow contact depths (<10 m) and decreased for deeper contact depths (>10m).

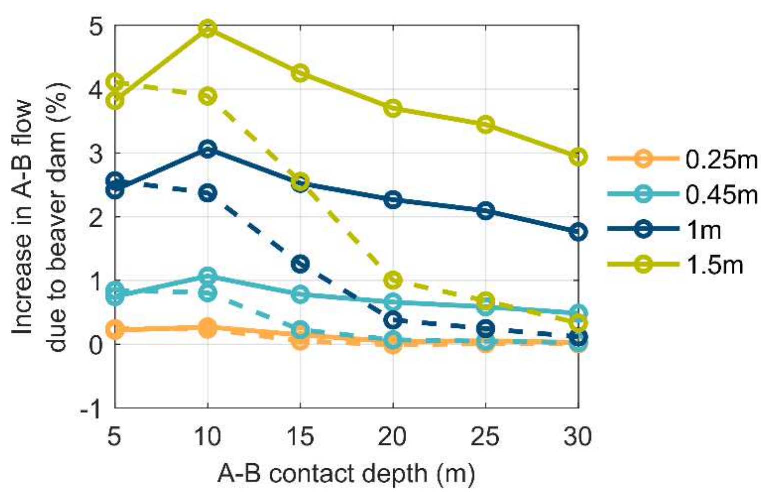

Figure 32 - Net (solid line) and total (dashed line) increase in A-B flow due to the beaver dam for tested alluvial to bedrock (A-B) contact depths.

The beaver dam had a greater impact on A-B flow rates for the scenario with the lower K_A:K_B, and a lessened impact on the scenario with greater K_A:K_B (***Figure 31c***). The transfer of water between aquifers is influenced by the hydraulic conductivity (K) of both aquifers. When groundwater flows from a more- to less-conductive aquifer, the low-K unit partially deflects flow so that it becomes more parallel to the low-K boundary. The greater hydraulic conductivity contrast (i.e., lower bedrock K compared to constant, higher alluvial K), decreases the influence of the beaver dam on deeper vertical fluxes by partially redirecting vertical flow horizontally.

4.5 Discussion

Seepage losses from streams represent an important source of groundwater recharge for many mountain-front aquifers (Wilson and Guan, 2004). This study identifies potential for beavers to drive inter-aquifer exchange been an alluvial and bedrock aquifer, a process which may constitute an important component of bedrock aquifer recharge for an increasingly depleted Denver Basin aquifer system. In many heavily pumped areas of the Denver Basin, it is common for stream-alluvial-bedrock sequences to have predominantly downward fluxes related to deep groundwater pumping (Paschke et al., 2011; Cognac and Ronayne, 2020; Cognac and Ronayne, 2023). Because of this, there may be greater potential for beaver dams to drive water deeper into the aquifer system.

Results from the simplified groundwater flow model suggest that the influence of the beavers includes increases in A-B flow ranging from 1-4% depending on the pond and alluvial depth. To assess the spatial extent of the beaver-dam influence, the dam-induced hydraulic head change was calculated throughout the aquifer as the change in hydraulic head between the scenario with and without the beaver dam (***Figure 33***). Each panel in ***Figure 33*** depicts a successively deeper alluvial-bedrock contact, with the shallowest contact scenario recording the greatest magnitude of head change within the aquifer. While relatively minor in magnitude (0-15 cm), the change in hydraulic head due to the beaver dam persists well below the alluvium base in all scenarios. For the shallowest contact depths, the head-change contours bend upward below the A-B contact and with distance from the dam. This supports the finding of greater flow from the bedrock to alluvial aquifer in shallow scenarios, as head the vertical hydraulic gradient would be altered in these areas. While not shown in ***Figure 33***, the spatial extent of the beaver dam's influence was more extensive for greater stage increases. This is consistent with the work of Wade et al. (2020) which compared field-based measurements of vertical streambed fluxes near beaver dam analogues (BDA) for induced stage increases ranging from 0.12 m to 0.4 m. They found greater downwelling upstream of the dam with the greatest stage increase (0.4 m) and nutrient signatures that indicated that the BDA induced a connection to regional groundwater flow paths. The stage differential generated by the beaver dam or BDA directly corresponds to the height of the dam, suggesting the taller dams are more likely to drive vertical flow and influence deeper aquifer dynamics.

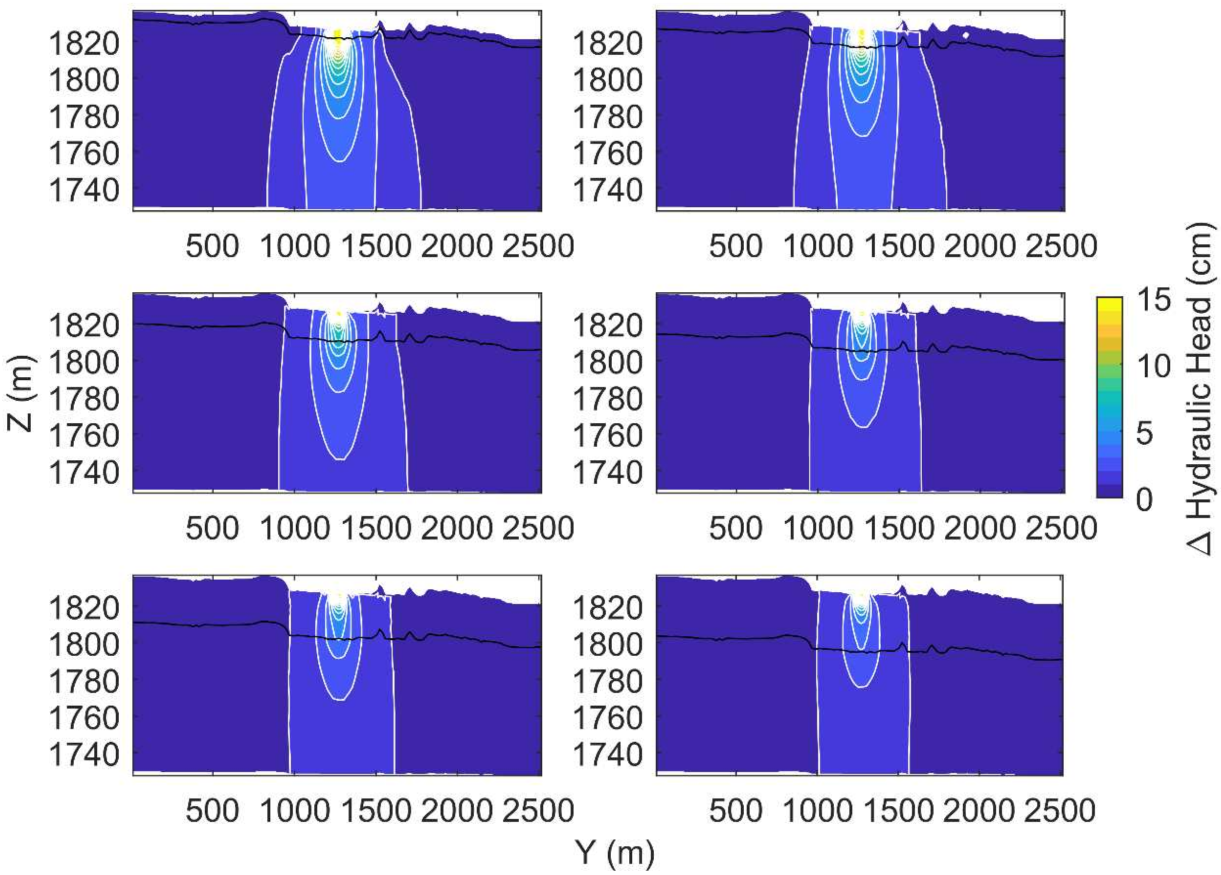

Figure 33 - Change in hydraulic head between scenarios with and without the dam along model column 94 (location of dam). Positive values indicate an increase in head following the addition of the dam. Black line indicates the base of the modeled alluvium.

The piezometer nests that captured the effects of the beaver dam on CC were part of a broader 5-year effort to monitor groundwater-surface water exchange as underlying Denver Basin aquifers undergo significant, long-term pumping and associated water level declines. During the study period (2016-2021), dozens of active beaver dams were observed on CC and other nearby streams. While a single beaver dam may have minimal impacts on shallow and deep groundwater levels, cumulative impacts of multiple dams have been found to be considerably greater (Majerova et al., 2015; Marshall et al., 2023). Dam and pond density can be upwards of ten ponds per kilometer of stream (Pollock et al., 2003; Gibson and Olden, 2014). Based on the length of the stream analyzed in mass balance results (1.6 km), our study considers a dam density of 0.6. Therefore, for systems with similar regional flow dynamics, the potential for beaver to influence inter-aquifer fluxes may be significantly greater where higher dam density exists.

This study presented a simplified groundwater flow model that has limitations related to assumptions and simplifications. Beaver dams reduce water velocity (Naiman et al., 1988; Green and Westbrook., 2009) which increases finer sediment retained by the dam, and can decrease the streambed K (Genereaux et al., 2008). I did not consider reductions in streambed K which can occur upstream of beaver dams due to the deposition of fine sediment. Beaver dams can also alter bedform dynamics (Briggs et al., 2013), which drive shallow hyporheic flow paths that return to the river. While I document an increase in flow returned to the river following the addition of the dam, I neglect streambed heterogeneity which is a major driver of hyporheic exchange, and therefore apply only a cursory representation of this process.

I also perform steady state simulations that represent a long-enduring or permanent beaver dam and average river and aquifer conditions surrounding the dam. However, beaver dams are dynamic and transient features. They are actively constructed and maintained just as they are abandoned. Large floods often damage or destroy dams, altering the lifespan of the dam and its potential impacts. Further, seasonal changes in the groundwater table and river flow would drive changes in hydraulic gradients and fluxes associated with the dam. Along Cherry Creek, groundwater levels and stream flow are highly dynamic. The recorded influence of the beaver dam on upstream vertical fluxes within the Cherry Creek streambed endured for at least seven months. However, the magnitude of the influence diminished through time. The streambed and river surface remained elevated upstream of the dam, indicating that deeper fluxes may have continued to be influenced. This is to say, the potential influence of beaver dams is transient and future research may consider a transient model that represents dynamic aspects of the dam and the duration of its influence.

Additional aspects of the model may influence results. In this study, I consider a system with predominantly downward fluxes, which are enforced in the groundwater flow model through boundary conditions (i.e., assumed downward vertical gradients were prescribed at the model edges). Additional simulations were performed wherein boundary conditions were modified to lower the magnitude of downward fluxes between the alluvial and bedrock aquifer. This resulted in a decreased effect of the dam. In the same way that stream-valley morphology determines near-surface impacts from beaver dams (Larsen et al., 2021), regional groundwater flow paths may be a critical determinant for the potential deep-aquifer impacts of beaver dams. Further research is needed to determine under what conditions specifically beaver dams are likely to affect deeper aquifer dynamics.

The depth of the alluvial-bedrock contact had a significant influence on modeled flow rates throughout the system. B-A flow, river seepage, and groundwater discharge to the river all increased with increasing contact depth. This is assumed to result from change in water moving through the model through up- and down-gradient GHB boundaries (i.e., GHB cells are assigned a greater conductance value due to higher-K, therefore thicker alluvium indicates more water moves through the model). A-B flow followed a parabolic trend with maximum rates for contact depths of 5 m and 30 m, and minimum rates for an A-B contact at 10 m. This trend followed for 'no dam' and 'dam' scenarios and is likely related to the increase in water budget related due to the deepening of a transmissive alluvium. Regardless, the net A-B flow was downward for all scenarios. A further examination of the relationships between A-B flow rates and alluvial geometry and hydraulic properties would improve the understanding of the potential for beaver dams to impact deeper aquifer fluxes. Plots of modeled flow rates for all contact depths are included in ***Appendix D, Figure D2***.

4.6 Conclusions

This study utilizes a numerical groundwater flow model to test whether a beaver dam, specifically increased stream stage and width upstream of a dam, has potential to influence deeper fluxes within a stream-alluvial-bedrock aquifer system characterized by predominantly downward fluxes. Model results document increases in alluvial to bedrock flow and changes in hydraulic head that propagate well into the aquifer (>30m) when the dam effects are simulated. However, the influence of the beaver dam is sensitive to the ponding depth, alluvial geometry, and hydraulic properties, suggesting that certain settings have greater potential for beaver to impact deeper aquifer dynamics. Results document the potential for beaver to influence deeper aquifer fluxes where regional hydraulic gradients are downward and highlight broader potential for beaver to enhance aquifer recharge to deeper aquifers.

4.7 References

Andersen, D.C., Shafroth, P.B., 2010. Beaver dams, hydrological thresholds, and controlled floods as a management tool in a desert riverine ecosystem, Bill Williams River, Arizona. Ecohydrology 3 (3), 325–338.

Barkmann, P. E., F. Scot Fitzgerald, L. A. Sebol, W. Curtiss, J. Pike, Annette Moore, and B. Taylor. "OF-15-10 Geology and Groundwater Resources of Douglas County." Geology and Groundwater. Open File Reports. Golden, CO: Colorado Geological Survey, 2015. https://coloradogeologicalsurvey.org/publications/geology-groundwater-resources-douglas-colorado.

Bailey, D. R., Dittbrenner, B. J., & Yocom, K. P. (2019). Reintegrating the North American beaver (Castor canadensis) in the urban landscape. Wiley Interdisciplinary Reviews: Water, 6(1), e1323.

Baker, B. W., and E. P. Hill. 2003. "Beaver (Castor canadensis)." In Wild Mammals of North America: Biology, Management, and Conservation, 2nd ed., edited by G.A. Feldhamer, B.C.

Thompson, and J.A. Chapman, 288–310. Baltimore, MD: The Johns Hopkins University Press.

Barkmann PE, Fitzgerald FS, Sebol LA, Curtiss W, Pike J, Moore A, Taylor B (2015) Geology and groundwater resources of Douglas County, Colorado. Colorado Geol Surv Open File Rep 15–10. https://coloradogeologicalsurvey.org/publications/geologygroundwater-resources-douglascolorado. Accessed 1 March 2018

Bird, B., O'Brien, M., Peterson, M., 2011. Beaver and Climate Change Adaptation in North America: A Simple, Cost-Effective Strategy for the National Forest System, WildEarth Guardians, Grand Canyon Trust, & the Lands Council.

Bouwes, N., Weber, N., Jordan, C. E., Saunders, W. C., Tattam, I. A., Volk, C., ... & Pollock, M. M. (2016). Ecosystem experiment reveals benefits of natural and simulated beaver dams to a threatened population of steelhead (Oncorhynchus mykiss). Scientific reports, 6(1), 28581.

Briggs, M. A., Lautz, L. K., Hare, D. K., & González-Pinzon, R. (2013). Relating hyporheic fluxes, residence times, and redox-sensitive biogeochemical processes upstream of beaver dams. Freshwater Science, 32(2), 622-641.

CDWR. (2023) Colorado Division of Water Resources Well Permit Database, Accessed August, 2023. https://dwr.state.co.us/Tools/WellPermits

Cognac, K. E., & Ronayne, M. J. (2020). Changes to inter-aquifer exchange resulting from long-term pumping: implications for bedrock groundwater recharge. Hydrogeology Journal, 28(4), 1359-1370.

Cognac, K. E., & Ronayne, M. J. (2023). Multiple timescales of streambed flux variability in two perennial mountain-front streams. Hydrological Processes, 37(3), e14840.

Feiner, K., & Lowry, C. S. (2015). Simulating the effects of a beaver dam on regional groundwater flow through a wetland. Journal of Hydrology: Regional Studies, 4, 675-685.

Genereux, D. P., Leahy, S., Mitasova, H., Kennedy, C. D., & Corbett, D. R. (2008). Spatial and temporal variability of streambed hydraulic conductivity in West Bear Creek, North Carolina, USA. Journal of Hydrology, 358(3-4), 332-353.

Gibson, P. P., & Olden, J. D. (2014). Ecology, management, and conservation implications of North American beaver (Castor canadensis) in dryland streams. Aquatic Conservation: Marine and Freshwater Ecosystems, 24(3), 391-409.

Green, K.C. and C.J. Westbrook. 2009. Changes in riparian area structure, channel hydraulics, and sediment yield following loss of beaver dams. BC Journal of Ecosystems and Management 10(1):68–79. www.forrex.org/publications/jem/ISS50/vol10_no1_art7.pdf

Harbaugh, A. W. (1990). A computer program for calculating subregional water budgets using results from the US Geological Survey modular three-dimensional finite-difference ground-water flow model (Vol. 90). Reston, VA, USA: US Geological Survey.

Harbaugh, A. W., Banta, E. R., Hill, M. C., & McDonald, M. G. (2000). Modflow-2000, the u. S. Geological survey modular ground-water model-user guide to modularization concepts and the ground-water flow process.

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, variously p.

Jin, L., Siegel, D.I., Lautz, L.K., Otz, M.H., 2009. Transient storage and downstream solute transport in nested stream reaches affected by beaver dams. Hydrol. Process. 23 (17), 2438–2449. https://doi.org/10.1002/hyp.7359.

Larsen, A., Larsen, J. R., & Lane, S. N. (2021). Dam builders and their works: Beaver influences on the structure and function of river corridor hydrology, geomorphology, biogeochemistry and ecosystems. Earth-Science Reviews, 218, 103623.

Lautz, L. K., & Siegel, D. I. (2006). Modeling surface and ground water mixing in the hyporheic zone using MODFLOW and MT3D. Advances in water resources, 29(11), 1618-1633.

Lowry, M.M., 1993. Groundwater Elevations and Temperature Adjacent to a Beaver Pond in Central Oregon (Masters Thesis).

Macfarlane, W. W., Wheaton, J. M., Bouwes, N., Jensen, M. L., Gilbert, J. T., Hough-Snee, N., & Shivik, J. A. (2017). Modeling the capacity of riverscapes to support beaver dams. Geomorphology, 277, 72-99.

Majerova, M., Neilson, B. T., Schmadel, N. M., Wheaton, J. M., & Snow, C. J. (2015). Impacts of beaver dams on hydrologic and temperature regimes in a mountain stream. Hydrology and Earth System Sciences, 19(8), 3541-3556.

Marshall, A., Zhang, X., Sawyer, A. H., Wohl, E., & Singha, K. (2023). Logjam Characteristics as Drivers of Transient Storage in Headwater Streams. Water Resources Research, 59(3), e2022WR033139.

Naiman, R.J., C.A. Johnston, and J.C. Kelly. 1988. Alteration of North American streams by beaver. Bioscience 38:753–762.

Paschke, S.S. ed., (2011), Groundwater availability of the Denver Basin aquifer system, Colorado: U.S. Geological Survey Professional Paper 1770, 274 p. doi: 10.3133/pp1770

Pilliod, D. S., Rohde, A. T., Charnley, S., Davee, R. R., Dunham, J. B., Gosnell, H., Grant, G. E., Hausner, M. B., Huntington, J. L., & Nash, C. (2017). Survey of beaver-related restoration practices in rangeland streams of the western USA. Environmental management, 61, 58-68.

Pollock, M.M., Lewallen, G.M., Woodruff, K., Jordan, C.E., Castro, J.M. (Eds.), 2017. The

Beaver Restoration Guidebook: Working with Beaver to Restore Streams, Wetlands, and Floodplains. Version 2.0. United States Fish and Wildlife Service, Portland, Oregon, 219 pp.

Puttock, A., Graham, H.A., Cunliffe, A.M., Elliott, M., Brazier, R.E., 2017. Eurasian beaver activity increases water storage, attenuates flow and mitigates diffuse pollution from intensively-managed grasslands. Sci. Total Environ. 576, 430–443.

Ronnquist, A. L., & Westbrook, C. J. (2021). Beaver dams: How structure, flow state, and landscape setting regulate water storage and release. Science of the Total Environment, 785, 147333.

Scamardo, J., & Wohl, E. (2020). Sediment storage and shallow groundwater response to beaver dam analogues in the Colorado Front Range, USA. River Research and Applications, 36(3), 398-409.

Scamardo, J. E., Marshall, S., & Wohl, E. (2022). Estimating widespread beaver dam loss: Habitat decline and surface storage loss at a regional scale. Ecosphere, 13(3), e3962.

U.S. Geological Survey (2023). National Water Information System data available on the World Wide Web (USGS Water Data for the Nation), accessed [September, 2023], at URL [http://waterdata.usgs.gov/nwis/].

Wade, J., Lautz, L., Kelleher, C., Vidon, P., Davis, J., Beltran, J., & Pearce, C. (2020). Beaver dam analogues drive heterogeneous groundwater–surface water interactions. Hydrological processes, 34(26), 5340-5353.

Wegener, P., T. Covino, and E. Wohl (2017), Beaver-mediated lateral hydrologic connectivity, fluvial carbon and nutrient flux, and aquatic ecosystem metabolism, Water Resour.

Res., 53, 4606–4623, doi:10.1002/ 2016WR019790.

Westbrook, C. J., D. J. Cooper, and B. W. Baker (2006), Beaver dams and overbank floods influence groundwater–surface water interactions of a Rocky Mountain riparian area, Water Resour. Res., 42, W06404, doi:10.1029/2005WR004560.

Westbrook, C. J., Ronnquist, A., & Bedard-Haughn, A. (2020). Hydrological functioning of a beaver dam sequence and regional dam persistence during an extreme rainstorm. Hydrological Processes, 34(18), 3726-3737.

Wilson, J. L., & Guan, H. (2004). Mountain-block hydrology and mountain-front recharge. Groundwater recharge in a desert environment: The Southwestern United States, 9, 113-137

APPENDIX A: Additional vertical hydraulic gradient plots

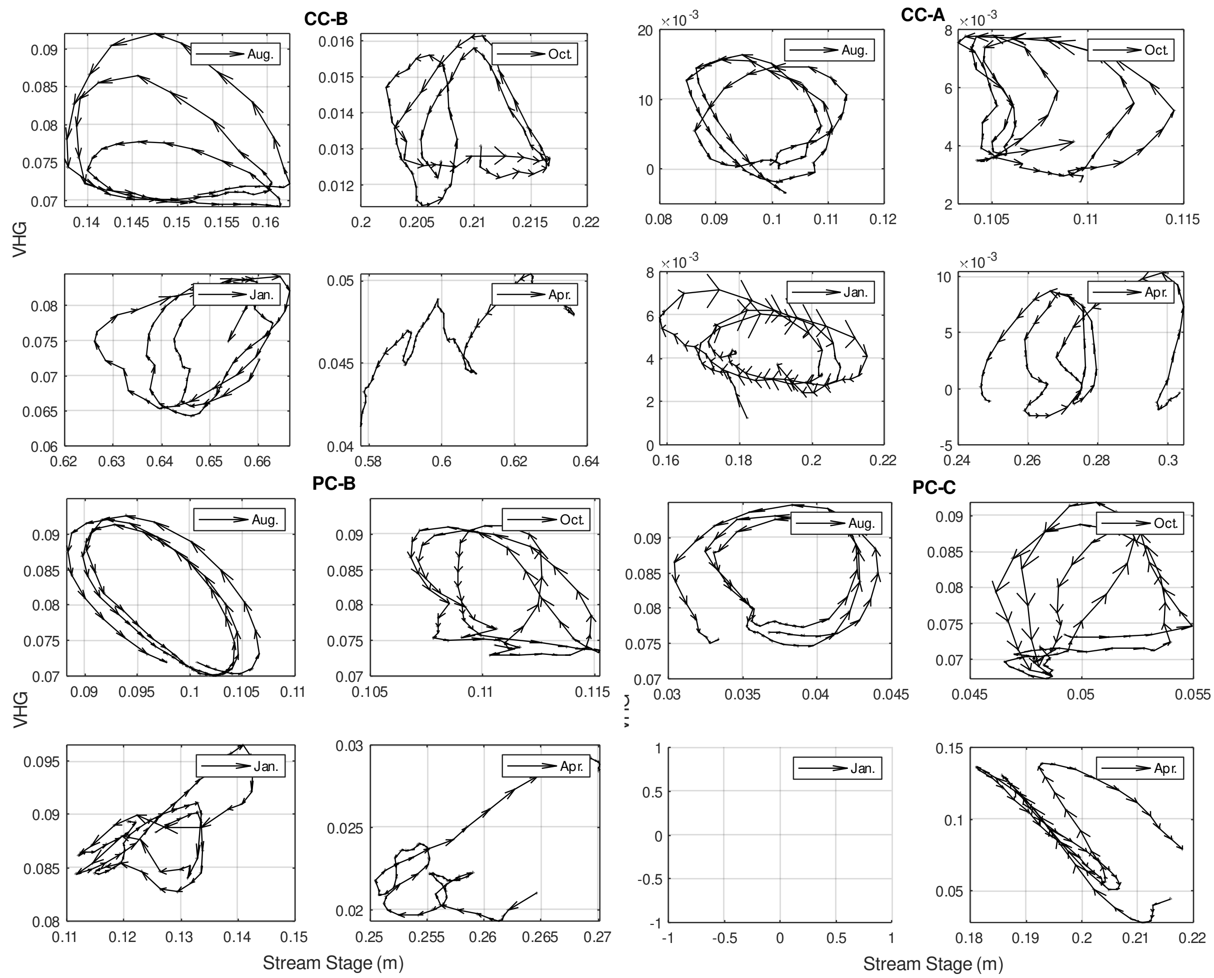

Figure A1 - Comparison of vertical hydraulic gradient (VHG) and stream stage during three-day periods in August 2020, October 2020, January 2021, and April 2021 at locations CC-B, CC-A, PC-B, and PC-C. Hourly vectors indicate paths through time.

APPENDIX B: Geophysical log data and analysis

Seven geophysical logs spanning the upper D2 sediments were used as conditioning data for geostatistical simulations (Chapter 3; Figure 18). Five of the logs contained only deep induction resistivity data, and two contained only shallow induction resistivity data (Table B1). Resistivity data were normalized using a quantile transformation to correct for differences in logging technology and penetration depth (e.g. shallow vs deep). The study area is relatively small (48 km^2) compared to the basin as a whole (18,000 km^2), therefore the assumption that bulk data exhibit stationarity within the study area (i.e., there are no trends in the mean resistivity between logs) is made. Quantile transformation involves mapping the p-quantile in a target distribution to that of a reference distribution (Figures B1 and B2) (e.g. Pyrcz and Deutsch, 2014). In this case, deep induction resistivity from well-permit 17693-FR was used as the reference distribution. Sensitivity to the choice of reference log did not significantly impact results (subsequent facies percentages were within 5% regardless of the chosen reference log). Final lithofacies are shown in Figure B3.

Table B1. Geophysical logs used for conditioning data.

DWR Permit	UTMX	UTMY	Surf. Elev. (m AMSL)	Available log type
30554-F	516493.2	4375820	1798	Shallow Induction Resistivity
238515	516497	4375533	1789	Deep Induction Resistivity
24914-F	517154	4375350	1783	Deep Induction Resistivity
17693-FR	517745	4374822	1796	Deep Induction Resistivity
29663-F	519156	4374966	1770	Deep Induction Resistivity
50564-F	520252	4376083	1783	Shallow Induction Resistivity
23548-F	520692	4375461	1812	Deep Induction Resistivity

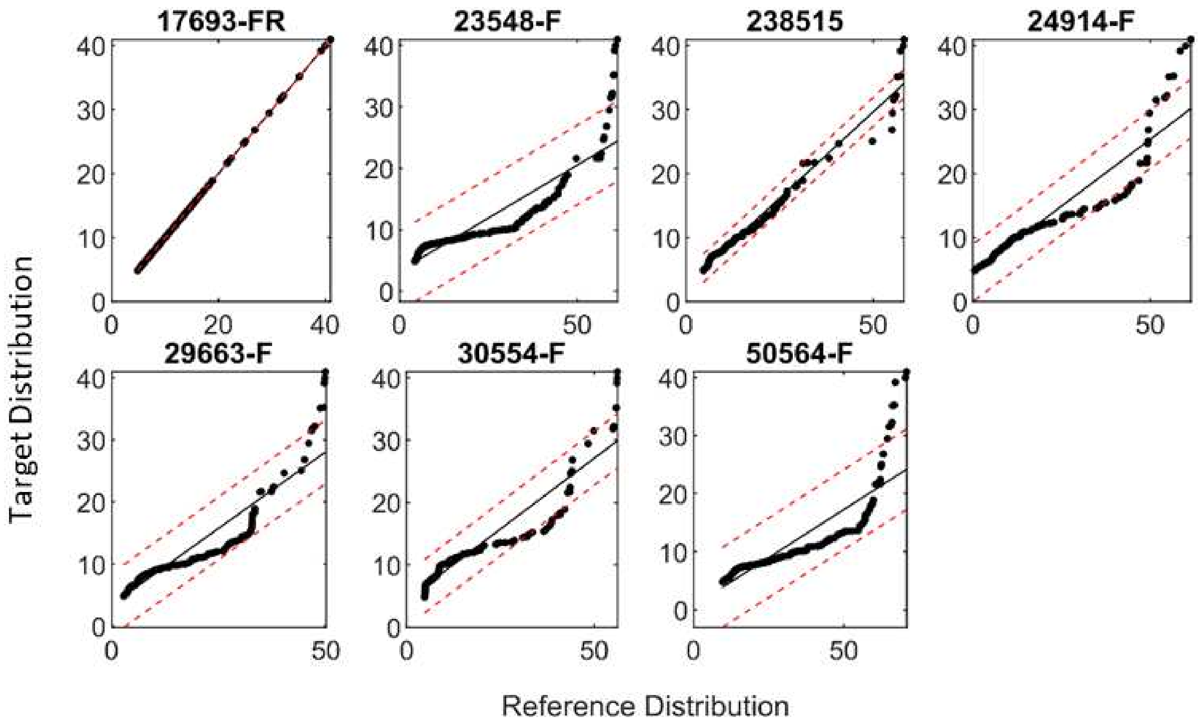

Figure B1. Quantile-quantile plots showing reference and target distributions for resistivity used for quantile transforms.

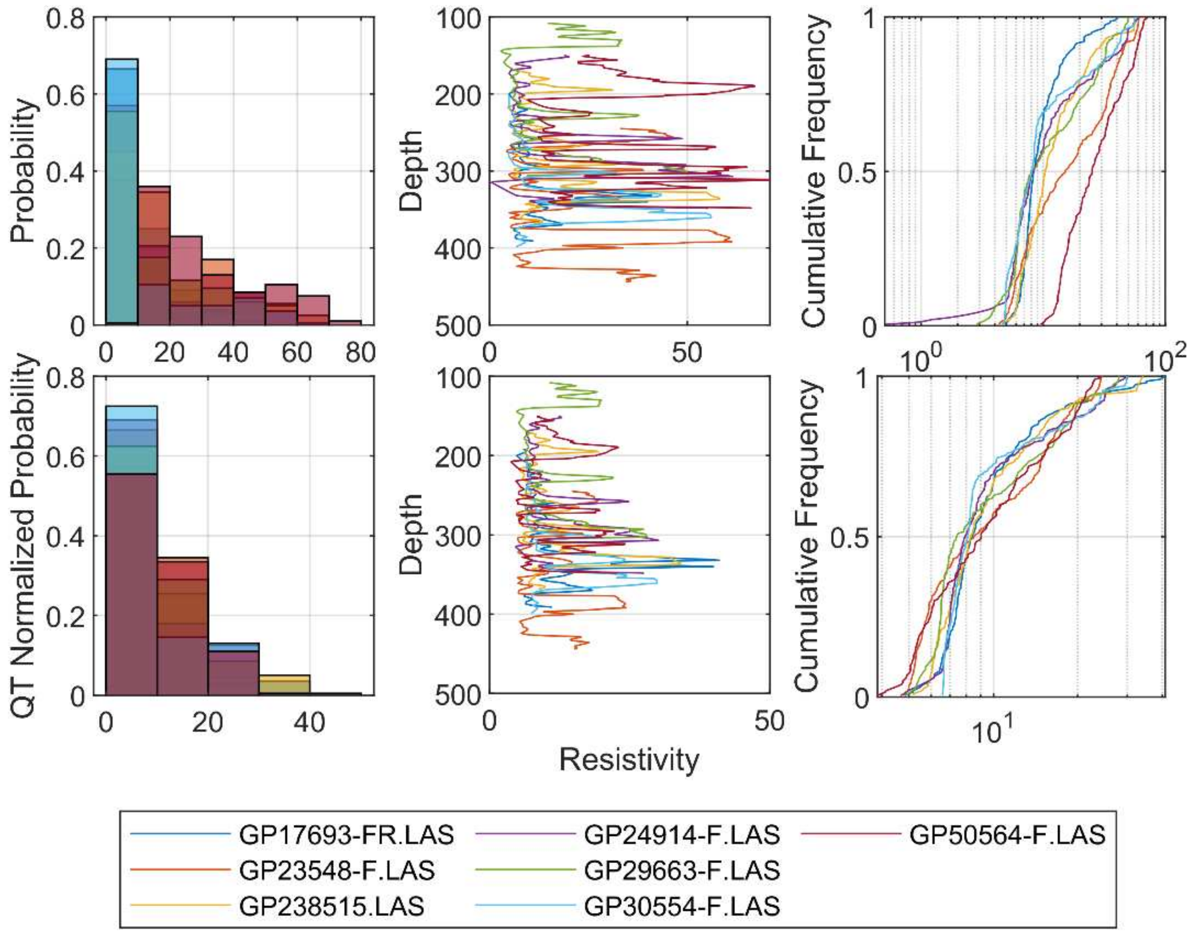

Figure B2. Log histograms, depth plots, and cumulative frequency for raw (top) and transformed (bottom) resistivity values for geophysical logs.

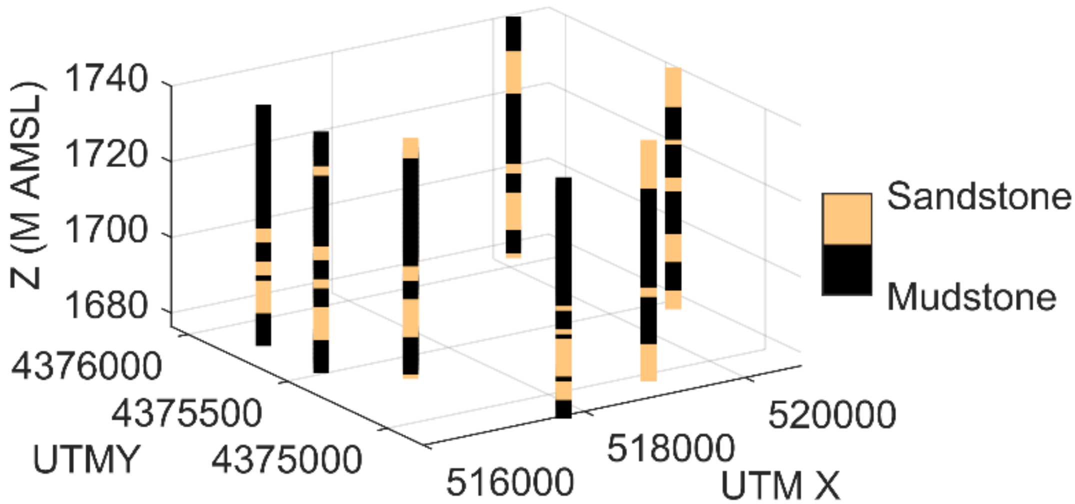

Figure B3. Lithofacies conditioning data from geophysical logs

APPENDIX C: Additional A-B flow and connectivity results

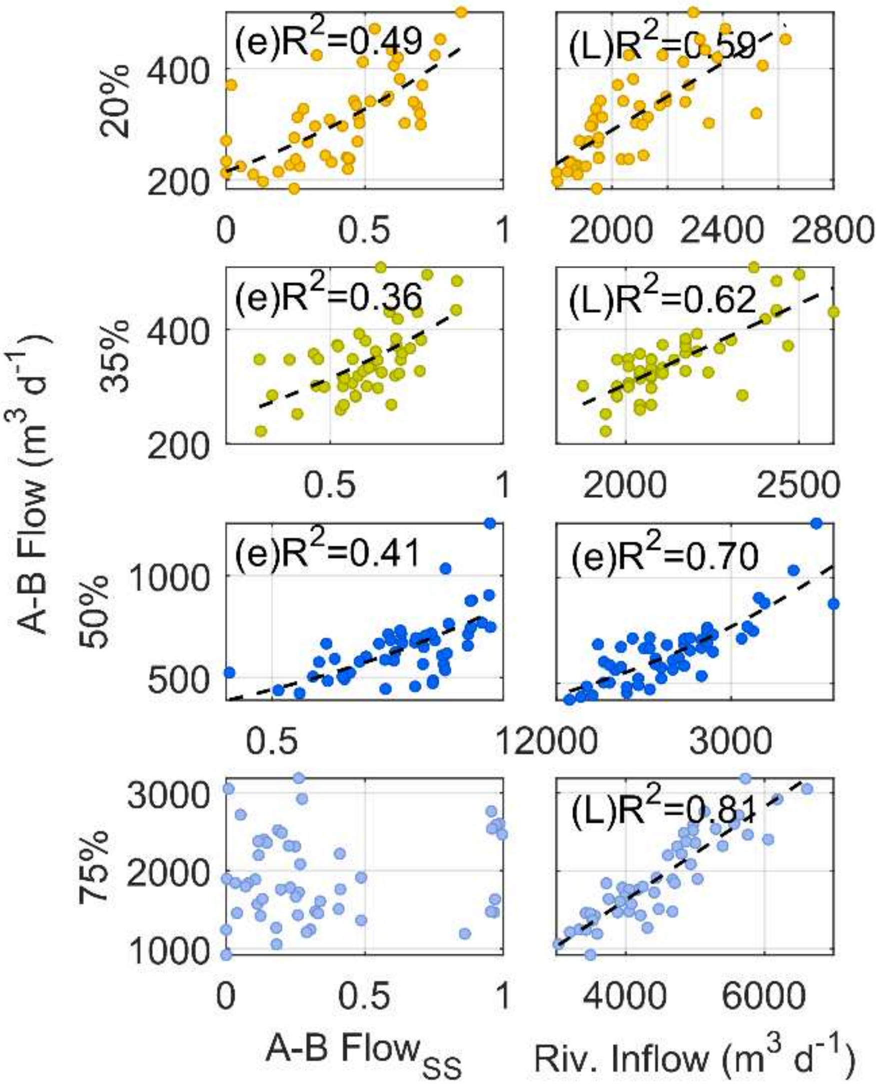

Figure C1. A-B flow rate compared to the fraction of flow that occurs through sandstone channels (A-B Flow$_{ss}$) and river inflow for simulations with 20%, 35%, 50%, and 75% bedrock aquifer sandstone fraction. Trend line and R2 are indicated as either linear (L) or exponential (e).

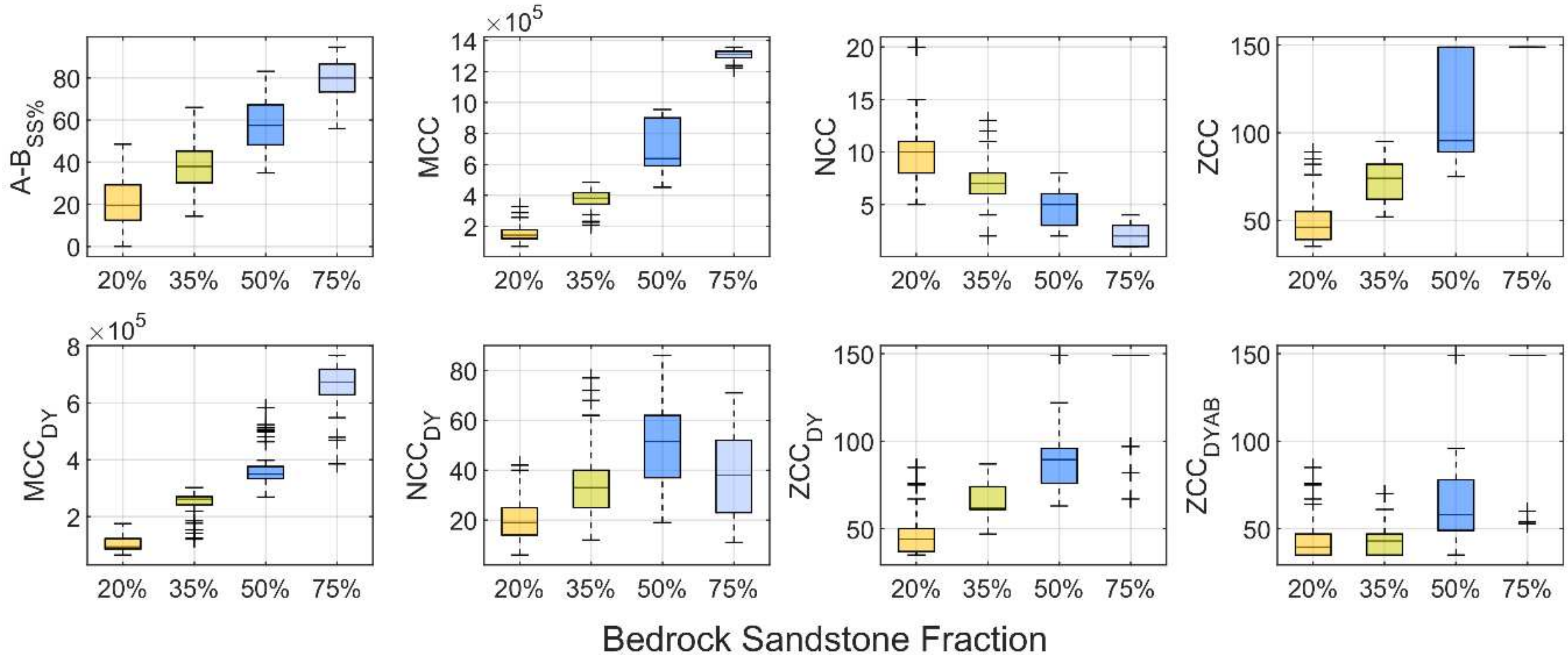

Figure C2. Boxplots of connectivity metrics for realizations with 20%, 35%, 50%, and 75% bedrock aquifer sandstone fraction.

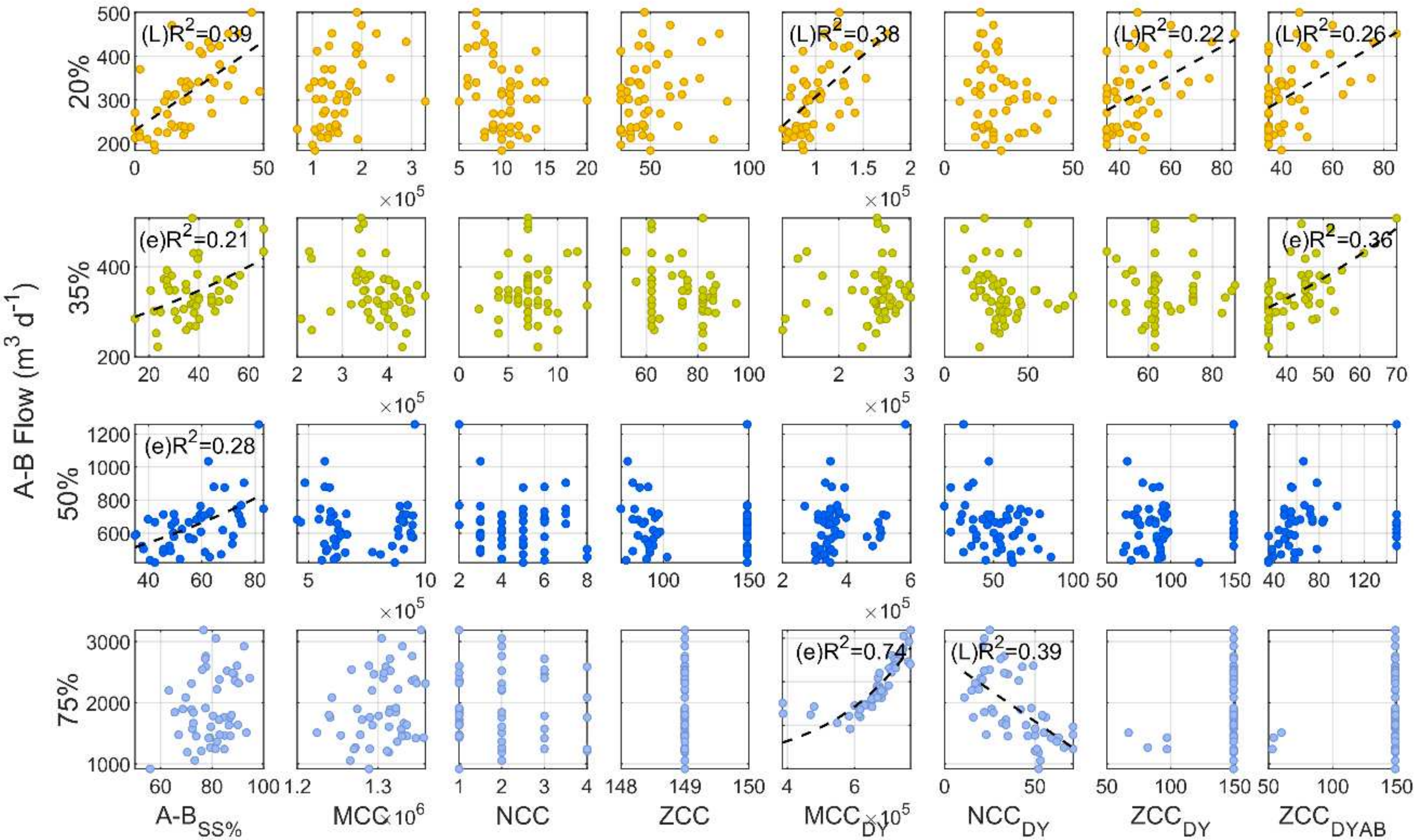

Figure C3. Final A-B flow rate compared to connectivity metrics for simulations with 20%, 35%, 50%, and 75% bedrock aquifer sandstone fraction. Trend line and R^2 are indicated as either linear (L) or exponential (e).

Table C1. Results from multiple linear regression analysis performed with for each channel fraction (all model realizations considered) using significant predictor variables A-$B_{SS\%}$, MCC_{DY}, NCC_{DY}, ZCC_{DY}, $ABZCC_{DY}$ and response variable of ABRF.

	Predictors				Model		
		Estimate	[1]SE	p-value	[2]RMSE	[3]R-squared	p-value
20%	Intercept	0.052	0.007	2.3E-	0.0115	0.17	9.3E-03
	A-BSS%	-1.7E-04	1.5E-	0.27			
	MCCDY	-2.5E-07	9.2E-	0.0080			
	ABZCC	3.4E-04	1.9E-	0.075			
35%	Intercept	0.031	0.0081	3.1E-	0.0065	0.24	1.2E-03
	A-BSS%	1.2E-04	8.4E-	0.17			
	MCCDY	-8.5E-08	2.1E-	0.0002			
	ABZCC	0	0.0	0.17			
50%	Intercept	0.0020	0.011	0.86	0.00908	0.36	6.9E-05
	A-BSS%	4.1E-04	1.1E-	4.0E-			
	MCCDY	4.3E-08	4.4E-	0.33			
	NCCDY	-2.7E-04	8.3E-	1.9E-			
	ZCCDY	-1.2E-04	1.1E-	0.30			
75%	Intercept	-0.032	0	0.56	0.0259	0.59	1.3E-08
	A-BSS%	-1.2E-04	5.5E-	0.82			
	MCCDY	0.000	1.1E-	2.1E-			
	NCCDY	-2.6E-04	0.0003	0.39			
	ZCCDY	-0.00196	5.4E-	7.0E-			
	ABZCC	0	0.0	0.85			

1SE is standard error.
2RMSE is root mean square error
3R2 is adjusted for the number of predictors in the model.
4Significant p-values are bolded

APPENDIX D: Additional beaver dam figures

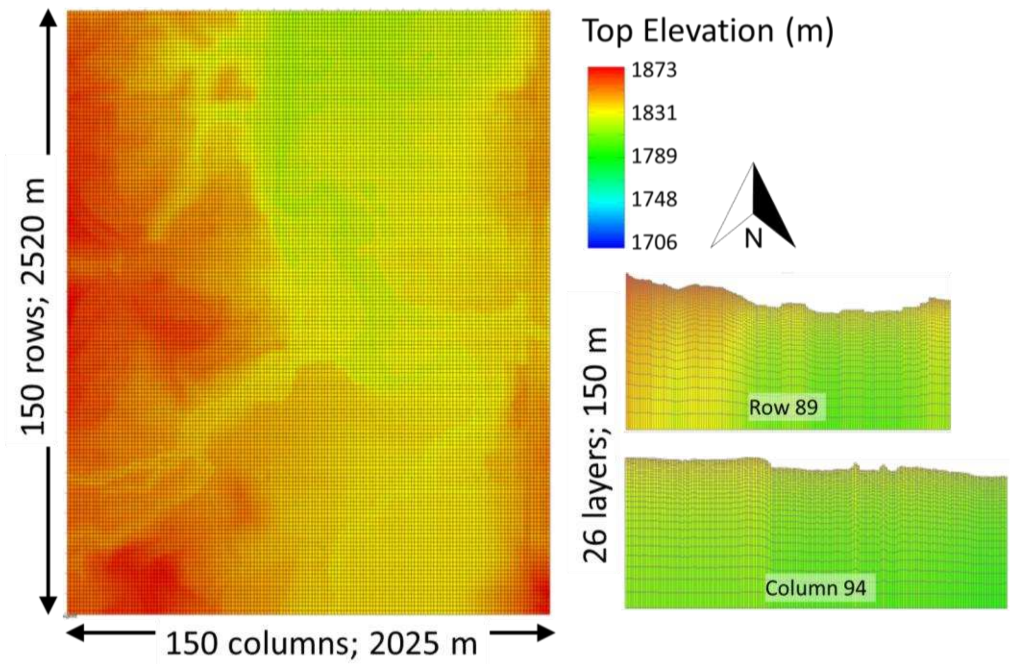

Figure D1. Model elevation and discretization.

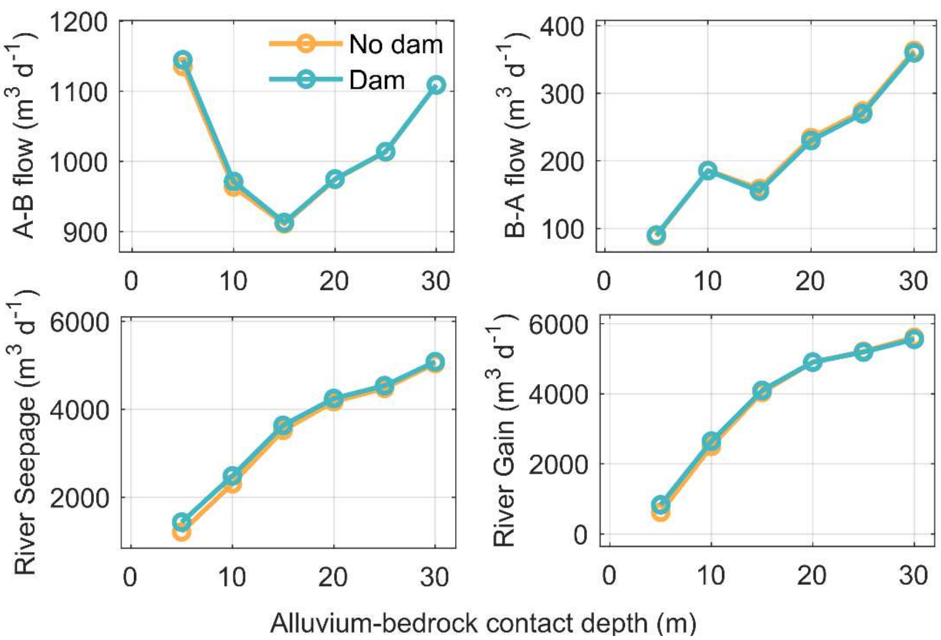

Figure D2. Alluvial to bedrock flow (A-B flow), bedrock to alluvial flow (B-A flow), river seepage, and groundwater discharge to the river (river gain) for tested alluvial-contact depths.