THESIS

EVALUATING SURFACE WATER–GROUNDWATER INTERACTIONS IN FLOODPLAINS USING SWAT+ AND GWFLOW

Submitted by

Evan Schulz

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2023

Master's Committee:

Advisor: Ryan R. Morrison
Co-advisor: Ryan T. Bailey

Ellen Wohl

ABSTRACT

EVALUATING SURFACE WATER–GROUNDWATER INTERACTIONS IN FLOODPLAINS USING SWAT+ AND GWFLOW

Floodplains are essential ecosystems that provide a variety of economic, hydrologic, and ecologic services. Within floodplains, surface water-groundwater exchange plays an important role in facilitating biogeochemical processing and can have a strong influence on hydrology through infiltration or discharge of water. These functions can be difficult to assess due to the heterogeneity of floodplains and monitoring constraints, so numerical models are useful tools to estimate fluxes, especially at a large scale. In this study, the *gwflow* module of the SWAT+ (Soil and Water Assessment Tool) ecohydrological model quantified magnitudes and spatiotemporal patterns of floodplain surface water-groundwater exchange in a mountainous watershed using an updated version of the module that directly calculated floodplain-aquifer interactions during periods of floodplain inundation. The *gwflow* module is a spatially distributed groundwater modeling subroutine within the SWAT+ code that uses a gridded network and physically based equations to calculate groundwater storage, groundwater head, and groundwater fluxes. I used SWAT+ to model an area of 7,516 km^2 in the Colorado Headwaters HUC8 watershed (14010001) and used streamflow data from USGS gages in the watershed for calibration and testing.

I evaluated model performance for scenarios with and without simulated floodplain-groundwater exchange and for three *gwflow* grid cell sizes. Models that included floodplain-groundwater interactions outperformed those without such interactions and provided valuable information about floodplain inundation and exchange rates. Furthermore, I found that smaller *gwflow* cell sizes showed similar or better performance than larger cell sizes and simulated additional information about local variations in groundwater fluxes, especially within floodplains. Finally, my analyses on the location of floodplain fluxes in the watershed showed that wider areas of floodplains, “beads,” exchanged a higher net and per area volume of water, as well as higher rates of exchange, than narrower areas, “strings.” These outputs

remained consistent across all studied cell sizes, with smaller cells simulating greater differences between bead and string floodplain regions. Study results show that floodplain surface water-groundwater exchange is a valuable process to include in hydrologic models, and model outputs could inform land conservation practices by indicating priority locations where substantial hydrologic exchange occurs.

ACKNOWLEDGEMENTS

I am incredibly grateful for the many people who have helped me pursue this graduate degree. To my advisor, Dr. Ryan Morrison, thank you for the opportunity to do research at Colorado State University. Your support and guidance over the past two years have been invaluable for my growth as a professional and individual. To my co-advisor, Dr. Ryan Bailey, thank you for teaching me the technical skills needed for this project and the many instances you've provided assistance. Dr. Ellen Wohl, thank you for joining my committee and for trusting me with work on other projects. I would also like to acknowledge the funding sources for this project, which were the Colorado State University Agricultural Experiment Station, the National Institute of Food and Agriculture of the U.S. Department of Agriculture, the Walton Family Foundation, and the National Science Foundation.

Furthermore, I would like to thank my family and friends, whose love and encouragement during my program have made an immeasurable difference in making this experience great. To the friends I've made at Colorado State University specifically, thank you for your help with coursework and research, and for pushing me to keep a work-life balance. Finally, I would like to thank my dog, Suki, for being a constant companion, source of joy, and reminder to get out and explore.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Surface Water-Groundwater Interactions in Floodplains

While there are numerous ways to define floodplains (Junk et al., 1989; Nanson & Croke, 1992), they are broadly classified as surfaces within a river corridor, adjacent to the channel, that are periodically inundated by water (Wohl, 2021). Floodplains are some of the most biologically diverse and productive landscapes on Earth (Tockner & Stanford, 2002; Ward et al., 1999), and as such, can have significant economic (Costanza et al., 1997), ecologic (Opperman et al., 2010), and hydrologic (Ward et al., 2002) impacts that extend beyond their immediate vicinity. The foremost physical function of floodplains is the attenuation of water, sediment, and nutrients, which is impacted by hydrologic exchanges, floodplain spatial extent, and floodplain heterogeneity (Wohl, 2021). Floodplain heterogeneity can vary greatly, both locally and longitudinally along a river corridor. Wider, lower gradient segments of river corridors, which Stanford et al. (1996) described as "beads," are likely to have greater diversity of habitat and storage of organic materials than their counterpart, "strings," which are narrower, steeper segments (Wohl et al., 2017). An understanding of floodplain functions requires a study of both hydrologic processes and floodplain physical properties due to the large variations in these characteristics along a river corridor.

Surface water-groundwater exchanges are of particular interest when examining floodplain hydrologic processes. Interactions with the alluvial aquifer can strongly influence hydrologic dynamics on the floodplain surface and in the river channel (Helton et al., 2014; Tonina & Buffington, 2009). This can be especially important during extreme hydrologic conditions or pollutant loading, where groundwater infiltration and discharge mitigate impacts to the ecosystem (Brunke & Gonser, 1997). Furthermore, groundwater processes can exert a strong influence on nutrient concentrations, pollutant filtering, and riparian vegetation (Stanford & Ward, 1993), which has cascading impacts to riparian ecosystems (Boulton et al., 1998). Floodplains may also provide substantial groundwater recharge; Goodrich et al. (2004) estimated that ephemeral channels, which are periodically inundated like floodplains, contribute between 15% and 40% of overall basin recharge to an aquifer in southern Arizona. Since surface water

and groundwater interactions are known to play important roles in hydrologic, ecologic, and biogeochemical processes, they should be included in any modeling efforts to understand the function of a landscape.

However, surface water-groundwater exchanges are often highly complex in floodplains due to heterogeneous hydraulic gradients and soil hydraulic conductivities (Krause et al., 2007; Woessner, 2000). The magnitude and spatial distribution of groundwater processes are determined by local controls (Cartwright et al., 2019), and can often vary with time (Andersen, 2004). While field measurements provide insight into these processes, it can be difficult to get an accurate assessment of groundwater interactions, and results can be dependent on hydrologic connectivity (Brunner et al., 2009; Martinet et al., 2009). These challenges impede our understanding of groundwater flow patterns and floodplain exchanges, and difficulties are exacerbated by anthropogenic changes that can cause novel and transient groundwater conditions, especially in floodplains.

1.2 Numerical Models

In watersheds where detailed data collection may be infeasible or inefficient, numerical models can provide insight into hydrologic, erosional, or biogeochemical processes. The incorporation of physically based equations into numerical models provides a reasonable basis for simulation accuracy, and these models are often calibrated using measured data to verify results. They are especially useful when studying interactions in complex landscapes, such as floodplains or groundwater aquifers. Models can also aid in analyzing the potential impacts of changing climate or land use, since modelers can adjust forcing variables to simulate hypothetical scenarios.

Interest in floodplain exchanges has spurred research in modeling procedures that couple surface water and groundwater processes. Efforts that include two-dimensional hydrodynamic models and floodplain subsurface flows have allowed for detailed analyses of complex flow patterns and aquifer response to floods (Bates et al., 2000; Maier et al., 2017; Saksena & Merwade, 2017). However, these studies are often limited to reach-scale systems due to high computing requirements when modeling

floodplain processes. Although they provide detailed results for smaller scales, these models are not applicable for studying processes over longer river corridors or for entire watersheds, which can be necessary to understand catchment-scale management impacts.

1.2.1 The Soil and Water Assessment Tool

The Soil and Water Assessment Tool (SWAT) is a watershed-scale modeling software that assesses a variety of properties such as nutrient cycling, water supply, and sediment dynamics (Arnold et al., 1998). SWAT discretizes watersheds into hydrologic response units (HRUs) which are lumped areas of similar soil, slope, and land use. Spatially lumped models typically require fewer input data and less calibration effort while performing as well or better than spatially distributed models (de Vente et al., 2013). However, they often cannot provide information about the location or magnitude of sources and sinks of nutrients and sediment (Arabi et al., 2006). This limits their effectiveness, especially when assessing the role of management practices on watershed water quality, which requires understanding the spatial distribution of fluxes.

Recently, Bieger et al. (2017) developed an updated version of the model, SWAT+, to better represent watershed processes using a more flexible hydrologic routing structure. While the algorithms used to calculate processes are unchanged, the change in routing structure can provide a more realistic representation of hydrologic connectivity in a watershed (Bieger et al., 2019). SWAT+ retains the same benefits of SWAT in that it is publicly available; detailed technical documentation exists for describing its modeling procedures; and it can be modified with supplemental software.

Several studies have been performed using a SWAT framework to model floodplain processes. For instance, Liechti et al. (2014) and Phiri et al. (2021) used a modified version of the SWAT reservoir unit to better model the attenuation of floods through floodplains. Sun et al. (2016, 2018) performed multiple studies on the utilization of landscape units (LUs) to better represent hydrologic and nitrate processes at both floodplain and catchment scales by improving the model simulation of surface water-groundwater exchange. Rajib et al. (2020) used SWAT to calculate forcing variables for hydrodynamic

models to map flood extents at a large scale. The versatility of the SWAT model and breadth of existing research is promising for further research using SWAT to model watershed-scale processes.

While previous studies confirm the applicability of SWAT for modeling floodplain processes at a large scale, they have not fully addressed the hydrologic interactions between surface water and groundwater in floodplains. The studies outlined above focus on surface water attenuation through floodplains, denitrification processes, or flood mapping. Integrating floodplain exchanges with groundwater is still necessary to better simulate watershed processes holistically. Floodplains do not exist without interactions with aquifers or rivers, and including these processes in a watershed model is important for accurate representation of these systems.

1.2.2 The gwflow *module for SWAT+*

SWAT and SWAT+ provide a useful modeling framework, but their default groundwater modeling procedures make several simplifying assumptions that are not physically based. These include steady state flow to streams instead of flow based on hydraulic gradients, homogeneous aquifer properties, and distinct aquifers that do not exchange flow between adjacent units. To better represent groundwater processes, Bailey et al. (2020) developed the *gwflow* module, which is a subroutine for each daily time step within the SWAT/SWAT+ modeling code. The module applies a gridded network and a control volume approach to calculate groundwater inflows and outflows, groundwater storage, and groundwater head for each cell in the grid using physically based equations. Summations of daily volumetric flow rates in a water balance equation update the storage of groundwater for each cell at each daily time step.

The *gwflow* module assumes a single-layer unconfined aquifer with a fixed volume in each cell that is bounded by bedrock at its base and the ground surface above. Each cell contains a hydraulic conductivity value in m/day (K) and specific yield value (S_y), which provides flexibility for *gwflow* to simulate heterogeneous aquifers. Boundary cells in the module are either constant-head or no-flow cells. The module uses Darcy's Law to calculate lateral flow between cells, stream-aquifer exchange, lake-

aquifer exchange, and tile drainage outflow. Simulation of stream-aquifer exchange only occurs for grid cells that intersect streams, and simulation of lake-aquifer exchange only occurs for grid cells underlying lakes. However, while the *gwflow* module includes many exchange pathways (Bailey et al., 2022), direct calculations of surface water-groundwater exchange in floodplains have not yet been included.

1.3 Study Objectives

In this study, I used an updated version of the *gwflow* module of SWAT+ to compute surface water-groundwater exchange through floodplains in a mountainous watershed in Colorado. The updated module explicitly calculates floodplain surface water-groundwater interactions during periods of simulated floodplain inundation.

The objective of this study was to quantify magnitudes and spatiotemporal patterns of floodplain surface water-groundwater exchange and to assess the importance of these exchanges at a watershed scale. These goals were achieved by: (i) evaluating SWAT+ model performance with and without the inclusion of floodplain surface water-groundwater interactions in the *gwflow* module, (ii) evaluating how grid cell resolution in the *gwflow* module impacts model outputs, and (iii) identifying the magnitude and location of surface water-groundwater exchanges in the study watershed, specifically in the context of beads and strings.

2.1 Description of Study Area

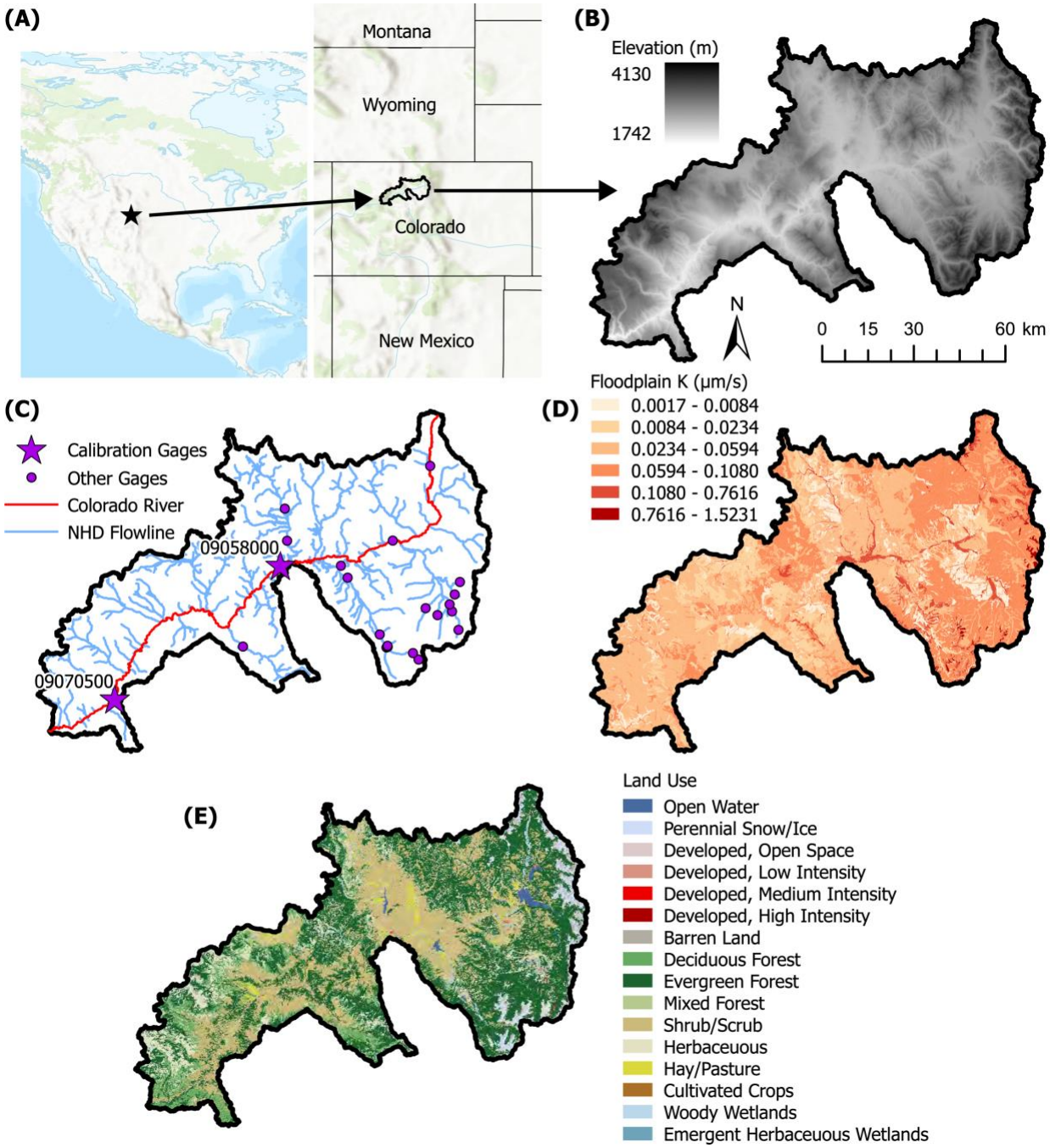

Figure 1: Location and properties of the Colorado Headwaters watershed, HUC8 14010001, showing: (A) vicinity map, (B) elevation range, (C) USGS stream gages, (D) surface soil hydraulic conductivity, and (E) land use.

I applied SWAT+ with *gwflow* to the Colorado Headwaters watershed, which is identified by the 8-digit hydrologic unit code (HUC8) 14010001 and located on the western face of the Rocky Mountains in Colorado (Figure 1). This watershed contains an area of 7,516 km^2, and land surface elevations range from approximately 1,750 m to 4,150 m.

The Colorado Headwaters watershed supplies water to regions on the eastern side of the Rocky Mountains, such as the Denver, Colorado area, through inter-basin transfers (Caldwell et al., 2012; Petsch Jr., 1985). It also contains the headwaters of the Colorado River, which is an important municipal, agricultural, and industrial water source for roughly 25 million people (Andersen et al., 2007).

2.2 Floodplain Delineation

I selected the GFPLAIN algorithm (Knox et al., 2022; Nardi et al., 2019) to delineate hydrogeomorphic floodplain areas within the study watershed and thereby identify floodplain cells. GFPLAIN implements terrain analysis techniques to extract the stream network from a digital terrain model. Each cell in the drainage network receives the maximum potential channel flow depth (h) for a recurrence interval i from the power law shown in equation (1) using contributing upstream area (A) as a scaling parameter. In this equation, a and b are dimensionless scaling parameters which were 0.0035 and 0.36, respectively, for the study watershed. The GFPLAIN algorithm then returns a gridded floodplain layer by flagging low-lying cells along river corridors and identifies the floodplain extent as the boundary of those cells that have land surface elevations lower than the corresponding maximum channel flow level. The GFPLAIN floodplain extent in this study had a recurrence interval of 100 years, 10 km^2 initiating threshold, and 30 m resolution.

$$h_i = aA^b \quad (1)$$

2.3 Updates to the *gwflow* module

As discussed in Section 1.2.2, the *gwflow* module is a subroutine in the SWAT+ code that calculates flux and storage values for every grid cell. In this study, the module included an additional

exchange pathway that explicitly calculated surface water-groundwater exchange during periods of floodplain inundation. Information related to this novel floodplain exchange term is bolded in Table 1 and Figure 2, where Table 1 provides a list of *gwflow* flux terms with their connections to other SWAT+ objects, and Figure 2 provides a conceptual overview of the *gwflow* module approach with possible inflows and outflows for each cell.

Table 1: Flux terms used in the *gwflow* module equations.

Flux in *gwflow*	Description	Connection to SWAT+
Q_{rech}	Aquifer recharge	Provided to grid cells from HRU soil profile deep percolation. Uses spatial intersections between HRUs and grid cells.
$Q_{sw \rightarrow gw}$; $Q_{gw \rightarrow sw}$	Stream seepage to aquifer; Groundwater discharge to streams	Calculated using Darcy's Law for each grid cell that geographically intersects a stream channel. Discharge to streams is provided to the corresponding channel in SWAT+.
$Q_{lake \rightarrow gw}$; $Q_{gw \rightarrow lake}$	Lake seepage to aquifer; Groundwater discharge to lakes	Calculated using Darcy's Law for each grid cell that geographically intersects a lake or reservoir. Discharge to lakes is provided to the corresponding reservoir object in SWAT+.
$Q_{fp \rightarrow gw}$; $Q_{gw \rightarrow fp}$	**Floodplain seepage to aquifer; Groundwater discharge to floodplains**	**Calculated using Darcy's Law for each grid cell that geographically intersects a floodplain. Discharge to floodplains is routed to the nearest channel in SWAT+.**
Q_{gwet}	Groundwater evapotranspiration	—
$Q_{gw \rightarrow soil}$	Groundwater transfer to the soil profile	Provided to HRUs when groundwater levels rise above the bottom of the SWAT+ soil profile.
Q_{satex}	Saturation excess flow when the water table intersects the ground surface	Calculated when the water table rises above the ground surface. Fluxes are routed to the channel object to which the surface runoff of the corresponding HRU drains.
Q_{pump}	Groundwater pumping	—
Q_{tile}	Groundwater discharge to tile drains	Calculated for grid cells that geographically include a tile drain. Fluxes are routed to the channel object to which the tile drains.
Q_{north}; Q_{south}; Q_{west}; Q_{east}	Groundwater lateral fluxes into/out of the four sides of each grid cell	—

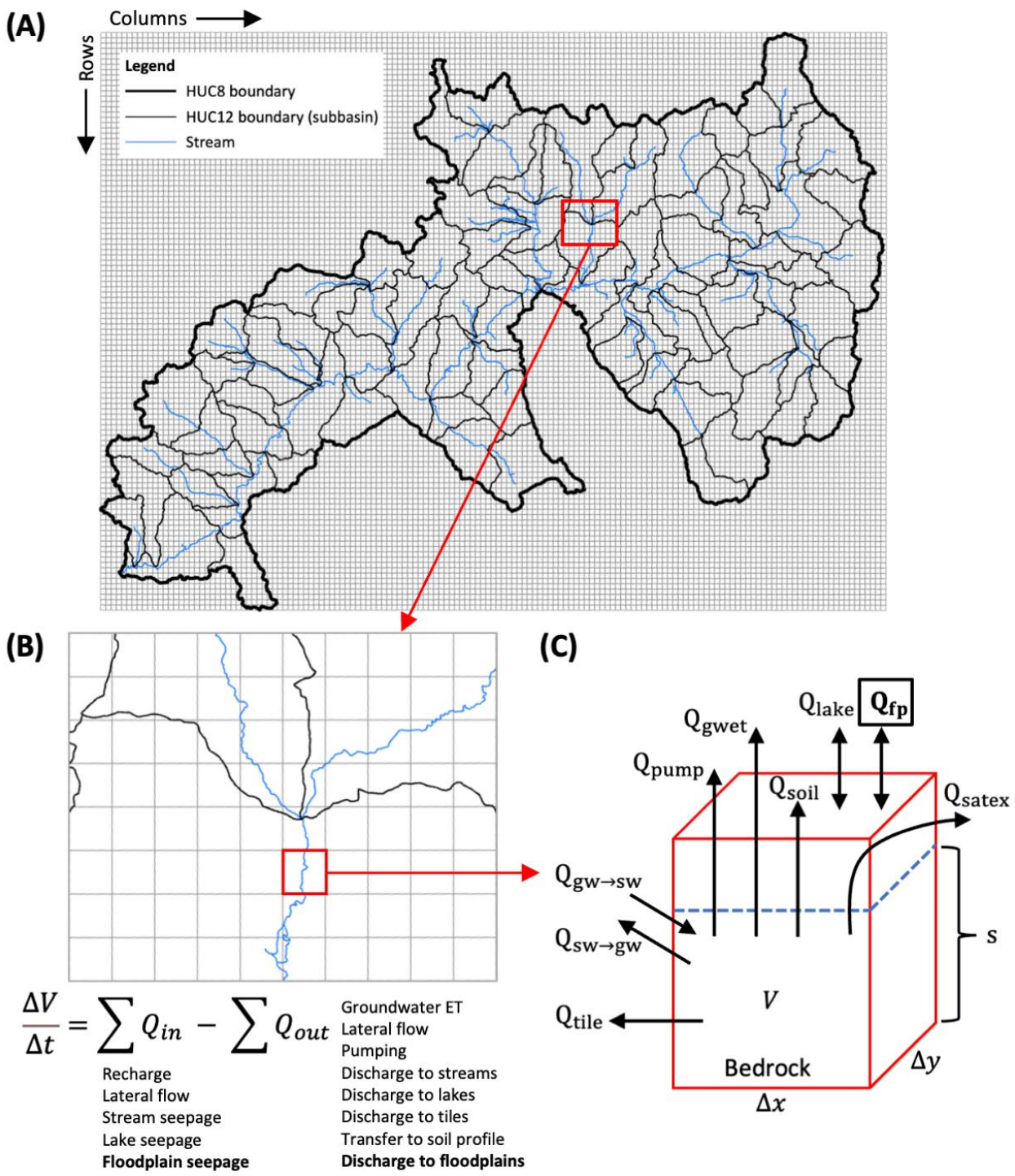

Figure 2: Layout and approach of the *gwflow* module where (A) shows the *gwflow* grid cell setup, (B) is a close-up of grid cells/streams with the general water balance equation used in *gwflow* calculations, and (C) is a conceptual schematic of the control volume approach for individual grid cells (after Bailey et al., 2023). The boxed term in (C) is a novel addition to the *gwflow* module for this study.

The updated *gwflow* module uses equation (2) to calculate floodplain exchange, where for each floodplain cell (notated by row i and column j), $Q_{fp\leftrightarrow gw\ i,j}$ is the surface water-groundwater exchange rate, $A_{fp\ i,j}$ is the area over which exchange is assumed to occur, $K_{fp\ i,j}$ is the soil hydraulic conductivity, $h_{fp\ i,j}$ is the surface water elevation, $h_{gw\ i,j}$ is the aquifer water table elevation, and z_{fp} is the distance between the ground surface and the aquifer water table (i.e., the distance infiltrating water travels through soil). Figure 3B provides a conceptual diagram of the exchange calculation.

$$Q_{fp\leftrightarrow gw\ i,j} = A_{fp\ i,j} \cdot K_{fp\ i,j} \cdot \left(\frac{h_{fp\ i,j} - h_{gw\ i,j}}{z_{fp}}\right) \quad (2)$$

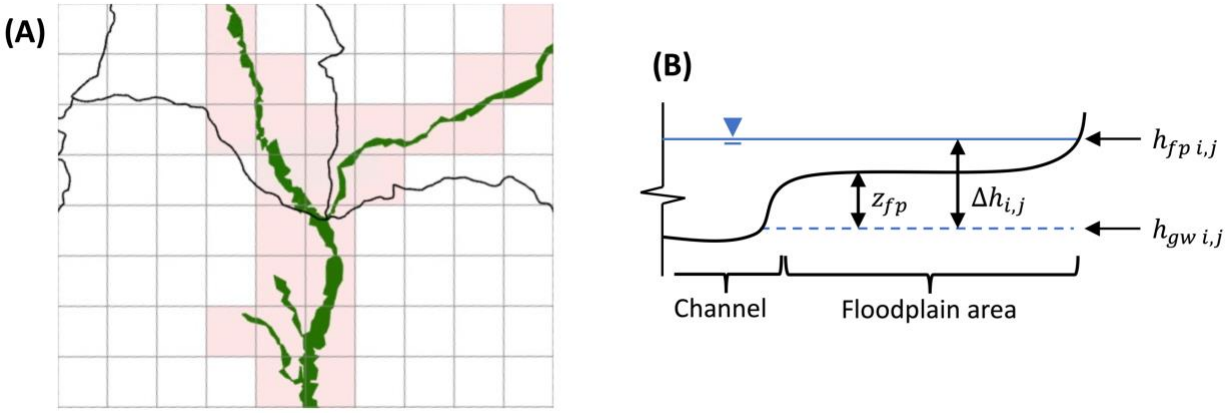

Figure 3: Schematic where (A) shows a close-up of grid cells (grey), the delineated floodplain (green), and floodplain cells (red), and (B) is a conceptual cross-section of a channel/floodplain area where surface water-groundwater exchange is calculated.

To identify periods of floodplain inundation, the *gwflow* module references the SWAT+ channel routing computations during each daily time step. SWAT+ assumes a trapezoidal channel defined by its bankfull width (width at the top of the channel), and a trapezoidal floodplain that has a bottom width equal to 5 times the channel bankfull width (Neitsch et al., 2009). The channel has sides with a 2:1 run to rise ratio while the floodplain has sides with a 4:1 run to rise ratio. When the volume of water routed to the channel exceeds its capacity, water inundates the floodplain. SWAT+ then calculates flow depth and volume as the sum of both channel and floodplain cross sections. The *gwflow* code flags when a SWAT+ channel has overbank flow and activates cells linked to that channel within the *gwflow* subroutine. The subroutine then uses the floodplain surface water elevation ($h_{fp\ i,j}$), compared to groundwater head

($h_{gw\,i,j}$) to determine $Q_{fp\leftrightarrow gw\,i,j}$. Surface water-groundwater fluxes in floodplains can either recharge the aquifer when surface water elevation is above groundwater head, or discharge to the floodplain when groundwater head is above surface water elevation.

The addition of floodplain exchanges requires a new *gwflow* input file, "gwflow.floodplain," which contains a list of cells that are within the floodplain. Each cell in the file list includes information for the area of intersection between the floodplain and the cell (A_{fp}), a hydraulic conductivity value (K_{fp}), and the ID of the connected channel. The datasets used for the creation of the *gwflow* module are shown in Table 2, with terms related to the novel floodplain exchange in bold.

Table 2: Datasets used for the construction of SWAT+ NAM models, *gwflow*, and *gwflow* with floodplain exchange (after Bailey et al., 2023).

	Dataset	Source	Data Type / Resolution
SWAT+ NAM Construction	Field boundaries	Yan and Roy (2016)	Raster / 30 m
	Crop rotation	USDA-NASS, CDL	Raster / 30 m
	Topographic slope	USGS National Elevation Dataset (Gesch et al., 2018)	Raster / 10 m
	Soil boundaries and properties	Gridded Soil Survey Geographic (Soil Survey Staff, 2014)	Raster / 10 m
	Land use, Land cover	U.S. Geological Survey, National Land Cover Data	Raster / 30 m
	Stream segments	Moore and Dewald (2016)	Vector features / —
	Lakes and reservoirs	Moore and Dewald (2016)	Vector polygons / —
	Weather	Global historical climatology network; PRISM	— / HUC12
	Water use	Dieter et al. (2018)	— / HUC8
	Discharge from facilities	Skinner and Maupin (2019)	— / HUC12
gwflow module	Geologic units	Horton (2017)	Vector polygons / —
	Tile drainage	Valayamkunnath et al. (2020)	Raster / 30 m
	Aquifer thickness	Shangguan et al. (2017)	Raster / 250 m
	Groundwater head	U.S. Geological Survey (Bailey and Alderfer, 2022)	Vector points / —
***gwflow* (floodplain)**	**Floodplain delineation**	**Knox et al. (2022)**	**ESRI Shapefile / 30 m**
	Floodplain hydraulic conductivity	**Soil Survey Geographic Database (SSURGO) (Soil Survey Staff, 2011)**	**ESRI Shapefile / —**

2.4 SWAT+ Framework

Several national scale modeling projects are currently underway to aid in simulating broad scale hydrologic, sediment, and nutrient processes (Baffaut et al., 2020; Chen et al., 2020; Cohen et al., 2018; Powers et al., 2017; Yen et al., 2016). However, these efforts are lumped at the HUC8 or HUC12 level, which only provides a general understanding for lower order streams, land use, and sediment loads at a large scale. Arnold et al. (2021) recently proposed the national agroecosystem model (NAM) due to this need for a national model capable of simulating processes for first order streams and individual fields. The NAM contains 2,121 individual SWAT+ models and encompasses field-level processes over the extent of the US, which allows it to be used as a base framework for further studies (White et al., 2022). It uses publicly available datasets, so it may be used by a variety of research groups and organizations without the need for proprietary information.

I used the Colorado Headwaters (HUC8 14010001) watershed model created in the NAM as the SWAT+ framework for this study. In alignment with the NAM's use of publicly available datasets, the goal of this study is to allow the same accessibility by utilizing only publicly available data. Table 2 outlines the datasets used for the creation of the NAM.

2.5 Model Setup

2.5.1 Study Scenarios

I considered two scenarios for this study. Both scenarios included the base SWAT+ NAM model that was only altered for parameter calibration.

1. Control Scenario: *gwflow* does not include floodplain exchanges.
2. Floodplain Scenario: *gwflow* includes floodplain surface water-groundwater exchanges.

2.5.2 Creation of gwflow *Files*

The *gwflow* module setup procedure was identical for the control and floodplain scenarios except for the addition of the "gwflow.floodplain" file in the floodplain scenario. The model simulation period

was 1 January 2000 to 31 December 2015, which encompasses both wet and dry years in the study watershed. For both scenarios, the *gwflow* module setup included three square grid cell resolutions with 250 m, 500 m, and 1000 m edge lengths.

I made some adjustments to the *gwflow* module prior to model analysis. To improve calibration efficiency, I consolidated aquifer hydraulic conductivity and specific yield values to 5 zones, which are summarized in Table 3. I analyzed soil types and used a preliminary calibration procedure to determine which zones had similar values for consolidation.

Table 3: Properties used to consolidate aquifer hydraulic conductivity zones.

Final Zone	Initial Zone	Soil Type	Initial Calibrated K (m/day)
1	1	Rhyolite, Gneiss, Granite	0.00010
2	2	Shale, Limestone	0.01034
3	3	Mafic-Hypabyssal	0.00104
2	4	N/A	0.00789
3	5	Andesite, Clastic, Siltstone	0.00100
2	6	Sandstone, Shale, Limestone	0.00903
4	7	Basalt, Volcanic	1.16548
4	8	Unconsolidated	1.05629
5	9	Unconsolidated gravel	88.9418

I determined floodplain hydraulic conductivity values from the Soil Survey Geographic (SSURGO) database (Soil Survey Staff, 2011). I consolidated these zones as well to improve calibration efficiency. From a histogram analysis of hydraulic conductivity values (Figure 4), I concluded that the values were lognormally distributed and thus fit a lognormal curve to the data to delineate 3 zones of hydraulic conductivity. I took the mean value, plus or minus one standard deviation, as the middle zone, and the remaining areas as the high and low zones. Initial *gwflow* values were the means of each zone.

I identified floodplain cells from an intersection of the *gwflow* module grid and the GFPLAIN floodplain (Section 2.2). This determined an area for each *gwflow* module grid cell ($A_{fp\ i,j}$). The maximum floodplain hydraulic conductivity value intersecting each cell determined hydraulic conductivity values ($K_{fp\ i,j}$). I selected the maximum hydraulic conductivity value under the assumption that maximum hydraulic conductivity will have the largest influence on fluxes. Finally, I used the

NHDPlus dataset (Moore & Dewald, 2016) to identify the closest channel ID for each floodplain cell. I repeated this procedure for each cell size scenario that was studied.

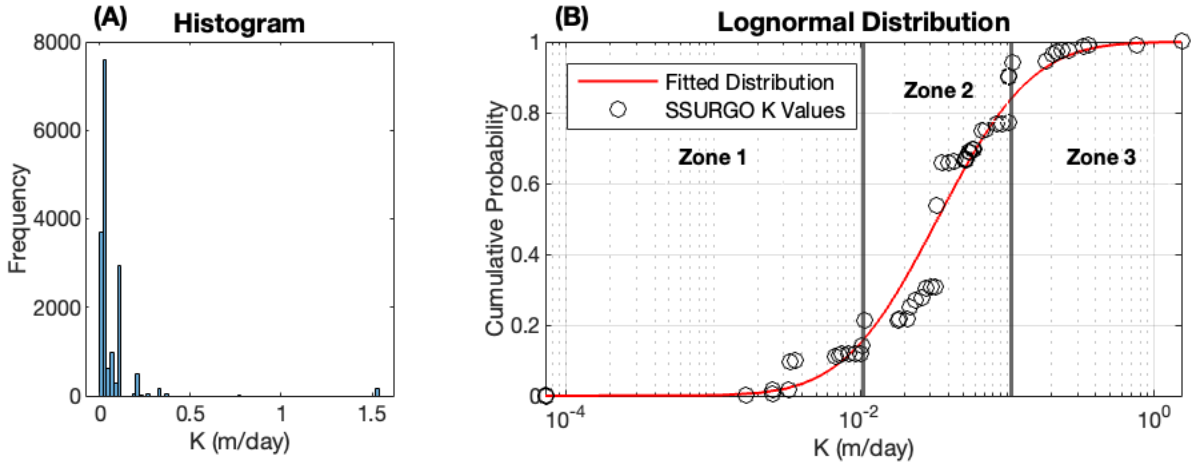

Figure 4: Histogram (A) and fitted lognormal distribution (B) for floodplain hydraulic conductivity zones.

2.5.3 Floodplain Categorization

I analyzed the GFPLAIN floodplain shapefile to identify floodplain bead and string locations, and thereby bead and string cells. I used perpendicular transects generated from centerlines of the floodplain area to estimate floodplain width. For locations where the floodplain widths (transects) were above a threshold value for bead delineation, a flag identified that transect as the location of a bead. I used a threshold value of 5 times the median stream width for each stream order, referencing widths from Downing et al. (2012). Appendix C provides a more detailed description of this procedure.

For stream orders 1, 2, and 3, the threshold widths for beads were 8.0 m, 9.5 m, and 27.5 m, respectively. Since the resolution of the GFPLAIN shapefile is 30 m, I deemed identification of beads in these lower stream orders inaccurate. Thus, I constrained further analysis of beads and strings in the watershed to stream orders 4, 5, and 6, where 6 was the maximum stream order identified in the study area. Figure 5 shows the result of intersecting *gwflow* floodplain cells with the closest transect to identify cells as representing a bead or string area.

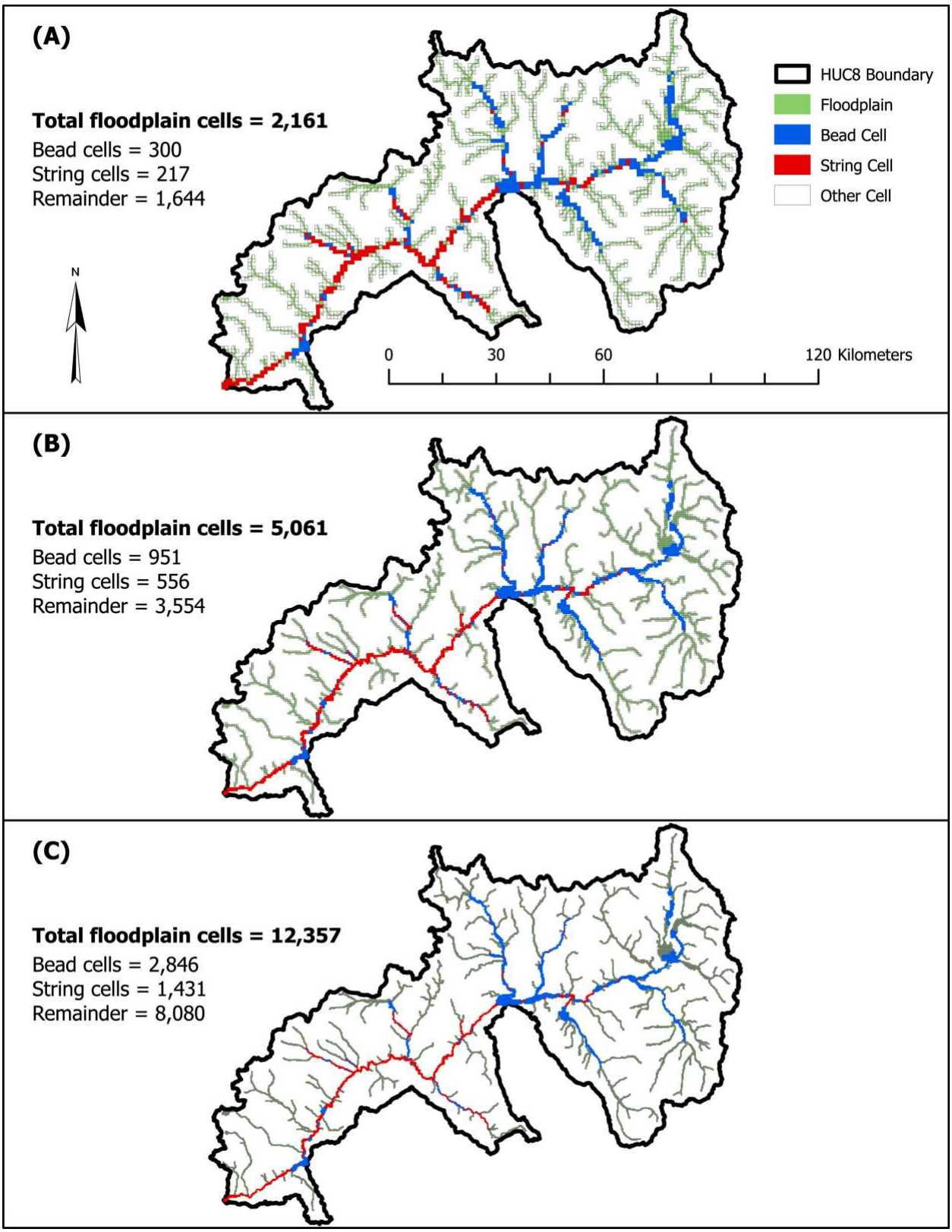

Figure 5: Floodplain cells identified as beads (blue) or strings (red) for cell sizes (A) 1000 m, (B) 500 m, and (C) 250 m.

2.6 Calibration and Testing

I limited the number of calibration targets to improve calibration efficiency and thus only used the two USGS gages located closest to the watershed outlet that had monthly average streamflow measurements for the entirety of the study period. Other gages had limited contributing area or did not include data for the entire study period. Figure 1C shows all gage locations in the watershed. The upstream calibration gage, USGS 09058000, is located near Kremmling, CO and corresponds to NHDPlus COMID 1232797 and channel ID 1970 in SWAT+. The second calibration gage, USGS 09070500, is located closer to the watershed outlet near Dotsero, CO, corresponding to NHDPlus COMID 1236065 and channel ID 2929 in SWAT+.

I used the automated model independent parameter estimation and uncertainty analysis (PEST) software (Doherty, 2010) for calibration, with a *gwflow* module grid cell size of 1000 m. I selected this cell size to decrease the runtime of each simulation, and thus the total runtime of the calibration procedure. I split the study period into 3 sections, beginning with a warm-up period from 1 January 2000 to 31 December 2002. I selected three years for the warm-up period after preliminary calibration efforts indicated that the model did not converge to realistic and stable results until around 2 to 3 years from simulation start. The calibration period was 1 January 2003 to 31 December 2011 (9 years), which captures both wet and dry years. The final 4 years (1 January 2012 to 31 December 2015) were a testing period to assess model performance without direct calibration. Thus, for each PEST iteration the simulation ran from 1 January 2000 to 31 December 2011.

The parameters included for calibration are summarized in Table 4. I selected these parameters to cover a range of watershed processes such as land surface processes (runoff generation factors, evapotranspiration factors, snowfall properties, etc.) and aquifer properties (hydraulic conductivity, specific yield, etc.). Simulation of runoff characteristics uses curve numbers, so I calibrated curve numbers for the top three land use categories for this watershed (forest, pasture, and brush, respectively). Floodplain hydraulic conductivity values were only included in the floodplain scenario. Thus, I calibrated 35 parameters for the control scenario and 38 parameters for the floodplain scenario.

Table 4: Parameters included in the calibration procedure, divided by SWAT+ parameters and *gwflow* module parameters. The bold parameters are floodplain hydraulic conductivity values which were only included in the floodplain scenario.

	Parameter	Description	Unit	Initial	Minimum	Maximum
SWAT+	CN2 FRST A	Curve number - forest	—	30.00	15.00	55.00
	CN2 FRST B	Curve number - forest	—	55.00	25.00	65.00
	CN2 FRST C	Curve number - forest	—	70.00	40.00	85.00
	CN2 FRST D	Curve number - forest	—	77.00	45.00	90.00
	CN2 PAST A	Curve number - pasture	—	39.00	25.00	55.00
	CN2 PAST B	Curve number - pasture	—	61.00	40.00	75.00
	CN2 PAST C	Curve number - pasture	—	74.00	55.00	90.00
	CN2 PAST D	Curve number - pasture	—	80.00	65.00	99.00
	CN2 BRUSH A	Curve number - brush	—	30.00	15.00	55.00
	CN2 BRUSH B	Curve number - brush	—	48.00	25.00	65.00
	CN2 BRUSH C	Curve number - brush	—	65.00	45.00	80.00
	CN2 BRUSH D	Curve number - brush	—	73.00	50.00	90.00
	EPCO	Plant uptake factor	—	1.00	0.10	1.00
	ESCO	Soil evaporation factor	—	1.00	0.10	1.00
	PERCOL	Percolation coefficient	—	0.50	0.10	1.00
	CN3 SWF	Soil water factor for CN3	—	0.95	0.10	1.00
	FALLTMP	Snowfall temperature	°C	1.00	0.01	5.00
	MELTTMP	Snow melt base temperature	°C	0.50	0.01	5.00
	MELTMX	Melt factor for snow on June 21	mm H_2O/°C-day	4.50	1.40	6.90
	MELTMN	Melt factor for snow on December 21	mm H_2O/°C-day	4.50	1.40	6.90
	TIMP	Snow pack temperature lag factor	—	1.00	0.01	1.00
	COVMX	Snow water content for 100% cover	mm H_2O	1.00	0.50	1.00
gwflow	RECH_DEL	Recharge delay	day	2.00	0.50	30.00
	$K_{aqu}1$	Aquifer K for Zone 1	m/day	0.0001	0.00001	0.001
	$K_{aqu}2$	Aquifer K for Zone 2	m/day	0.01	0.0001	0.1
	$K_{aqu}3$	Aquifer K for Zone 3	m/day	0.001	0.0009	0.01
	$K_{aqu}4$	Aquifer K for Zone 4	m/day	1	0.9	100
	$K_{aqu}5$	Aquifer K for Zone 5	m/day	100	10	5000
	S_y1	S_y for Zone 1	—	0.05	0.0005	0.4
	S_y2	S_y for Zone 2	—	0.05	0.0005	0.4
	S_y3	S_y for Zone 3	—	0.05	0.0005	0.4
	S_y4	S_y for Zone 4	—	0.25	0.0005	0.4
	S_y5	S_y for Zone 5	—	0.15	0.0005	0.4
	K_{stream}	K for stream bed	m/day	0.0001	0.0001	0.1
	d_{stream}	Thickness of stream bed	m	1	0.05	5
	$K_{fplain}1$	**Floodplain K for Zone 1**	**m/day**	**0.00529**	**0.000072**	**0.02**
	$K_{fplain}2$	**Floodplain K for Zone 2**	**m/day**	**0.0659**	**0.01**	**0.2**
	$K_{fplain}3$	**Floodplain K for Zone 3**	**m/day**	**0.81515**	**0.1**	**1.5231**

Assessment of model performance compared simulated and observed monthly average streamflow values at the gage locations. I selected Nash-Sutcliffe model efficiency coefficient (NSE), which measures model performance compared to the mean of observed time series data, and percent bias (PBIAS), which measures the tendency of the model to over- or underestimate values, as performance

metrics. Equations (3) and (4) show the NSE and PBIAS calculations, respectively, where Q_{obs}^{i} is the i^{th} observed streamflow, Q_{sim}^{i} is the i^{th} simulated streamflow, $\bar{Q}_{obs}$ is the mean of observed streamflow, and N is the total number of observations.

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(Q_{obs}^{i} - Q_{sim}^{i}\right)^2}{\sum_{i=1}^{N}\left(Q_{obs}^{i} - \bar{Q}_{obs}\right)^2} \tag{3}$$

$$\text{PBIAS} = \frac{\sum_{i=1}^{N}\left(Q_{obs}^{i} - Q_{sim}^{i}\right) \cdot 100}{\sum_{i=1}^{N}\left(Q_{obs}^{i}\right)} \tag{4}$$

I attempted to calibrate the models independently at smaller *gwflow* cell sizes, but model runtime was infeasibly long. Therefore, to assess how grid cell size impacts model results, I used the calibrated parameters from the 1000 m calibration procedure in models with smaller cell sizes (500 m and 250 m) without direct calibration. The purpose of this procedure was to assess model performance at finer spatial resolution even when computational constraints limited what cell sizes were feasible to calibrate.

3 RESULTS

3.1 Model Performance

To process intermediate outputs, I divided the PEST parameter calibration procedure into three runs for the floodplain scenario, each starting from the previous calibrated parameters, while the control scenario ran in one procedure. I ran all simulations on a desktop Intel® Core™ i7-7700 CPU @ 3.60 GHz, 64.0 GB RAM. In total, the control scenario converged after 31 optimization iterations and 2,234 model runs, and the floodplain scenario converged after 42 optimization iterations and 2,673 model runs. The calibration process completed in approximately 484 hours for the control scenario and 528 hours for the floodplain scenario.

I compared monthly average streamflow values at the calibration gages for the control and floodplain scenarios, divided by calibration and testing periods. Summarized model performance metrics by cell size, scenario, and period are shown in Table 5 for the upstream gage and Table 6 for the downstream gage. I reviewed metrics using model performance threshold values for flow suggested by Moriasi et al. (2015). Highlighted values in Table 5 and Table 6 are green for "good" model fit ($0.7 < NSE \leq 0.8$), yellow for "satisfactory" model fit ($0.50 < NSE \leq 0.70$; $\pm 10 < PBIAS < \pm 15$), and red for "not satisfactory" model fit ($PBIAS \geq \pm 15$).

Table 5: Model performance metrics for the upstream gage, USGS 09058000 (SWAT+ channel 1970).

Period	Cell Size (m)	NSE		PBIAS	
		Floodplain	Control	Floodplain (%)	Control (%)
Calibration	250	0.639	0.597	15.1	20.8
	500	0.654	0.601	16.0	21.5
	1000	0.652	0.601	18.8	26.9
Testing	250	0.562	0.571	23.0	29.5
	500	0.566	0.568	24.1	30.6
	1000	0.566	0.561	25.1	31.6

Table 6: Performance metrics for the downstream gage, USGS 09070500 (SWAT+ channel 2929).

Period	Cell Size (m)	NSE		PBIAS	
		Floodplain	Control	Floodplain (%)	Control (%)
Calibration	250	0.794	0.760	12.8	16.1
	500	0.770	0.748	13.6	16.5
	1000	0.751	0.720	20.5	25.4
Testing	250	0.740	0.707	22.0	26.8
	500	0.711	0.682	26.3	30.4
	1000	0.681	0.651	28.7	33.1

Figure 6 shows plots comparing average monthly streamflow between scenarios for the upstream USGS gage, 09058000. Figure 7 shows plots comparing average monthly streamflow between scenarios for the downstream USGS gage, 09070500. Both figures show results for the calibration grid cell size of 1000 m. The control and floodplain models both capture the overall monthly average streamflow patterns at each gage. Generally, visual assessment and model performance metrics indicate that simulated streamflow at channel 2929 is a better fit to observed data than at channel 1970.

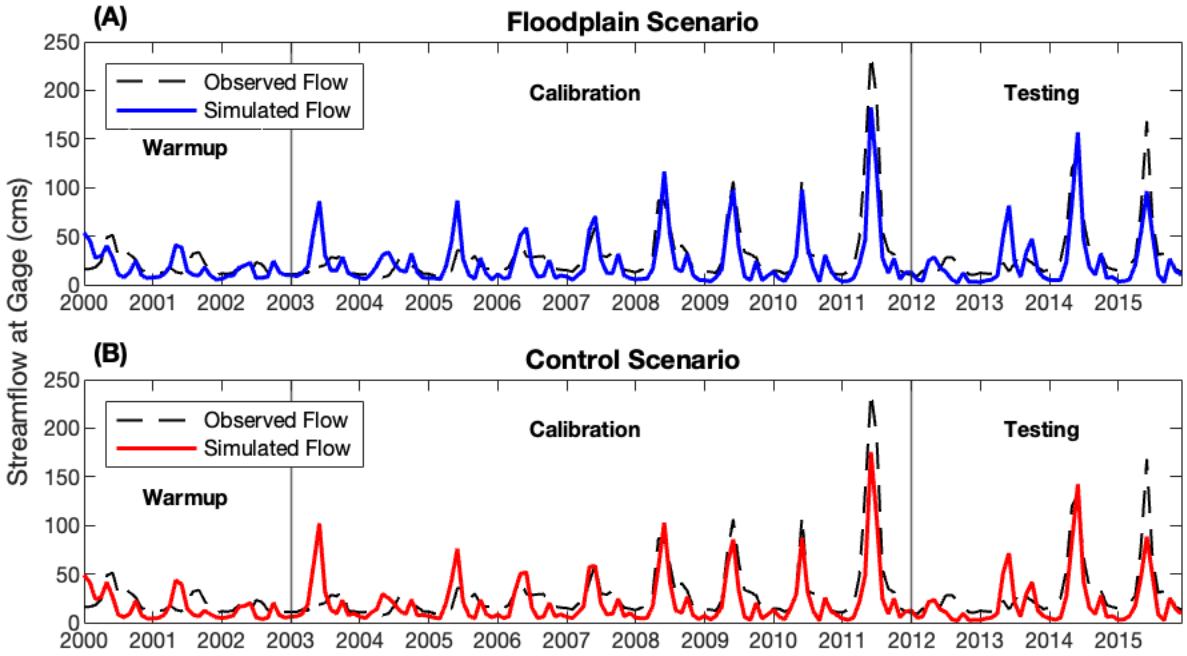

Figure 6: Monthly average streamflow at USGS gage 09058000 (SWAT+ channel 1970) for the floodplain (A) and control (B) scenarios with a *gwflow* cell resolution of 1000 m.

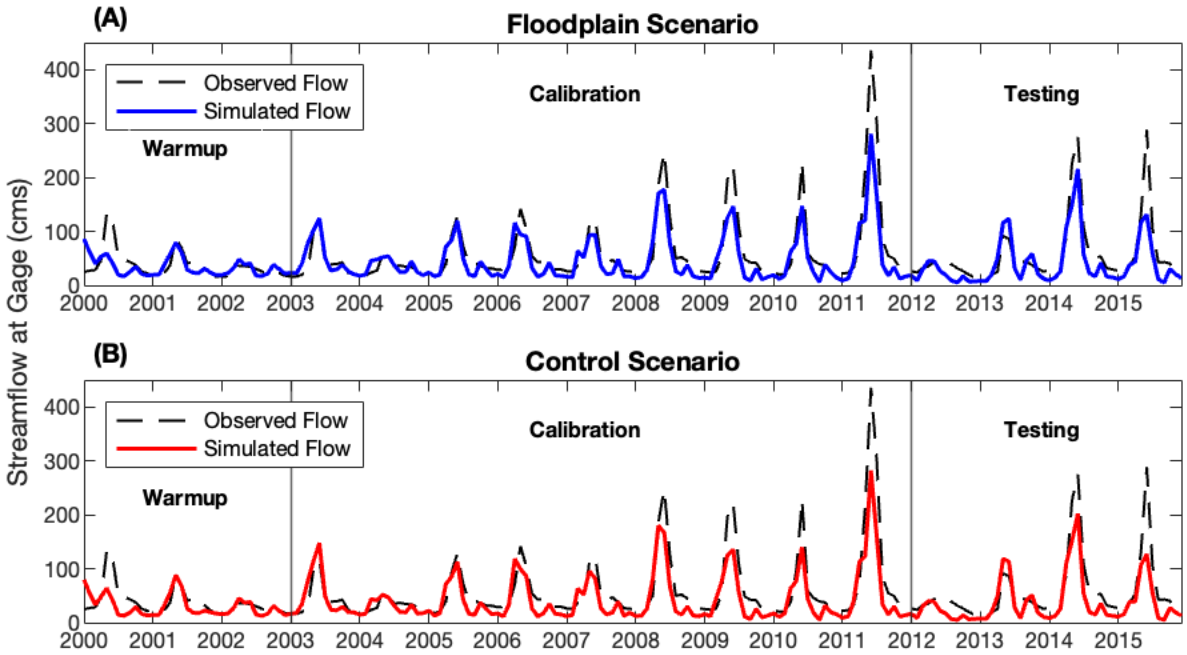

Figure 7: Monthly average streamflow at USGS gage 09070500 (SWAT+ channel 2929) for the floodplain (A) and control (B) scenarios with a *gwflow* cell resolution of 1000 m.

Metrics for the models including floodplain interactions tended to outperform their counterparts in the control models. NSE values in the calibration period were on average 0.05 and 0.03 higher in the floodplain scenario for channels 1970 and 2929, respectively. In the testing period, the scenarios had equivalent average NSE values for channel 1970 but for channel 2929 the floodplain scenario was 0.03 higher on average than the control scenario. PBIAS values were also slightly lower for the floodplain models, averaging 6.4% and 3.7% (calibration period) and 6.5% and 4.4% (testing period) lower for channels 1970 and 2929, respectively.

While the models were only calibrated using a 1000 m grid cell resolution, some smaller resolutions had improved model performance. At channel 2929, NSE values for 250 m resolution models in both scenarios were 0.04 and 0.06 higher than equivalent 1000 m resolution models. NSE values for 500 m resolution models also outperformed equivalent 1000 m resolution models at channel 2929 in both periods. At channel 1970, model performance was roughly equal across cell sizes, with the 500 m resolution tending to perform the best. Across the board, PBIAS values decreased with finer cell resolutions, although their best fit was only considered “satisfactory.”

3.2 Scenario Comparison

Calibrated parameter values and percent difference comparing the floodplain scenario to the control scenario are shown in Table 7. Percent difference was calculated from equation (5).

$$\%\ Difference = \left(\frac{floodplain - control}{control}\right) \cdot 100 \tag{5}$$

Table 7: Calibrated parameters for the floodplain and control scenarios.

	Parameter	Initial	Final - Floodplain	Final - Control	% Difference
SWAT+	CN2 FRST A	30	20.31	15.00	35.39
	CN2 FRST B	55	62.20	65.00	-4.31
	CN2 FRST C	70	56.75	42.52	33.45
	CN2 FRST D	77	51.14	45.00	13.65
	CN2 PAST A	39	38.97	55.00	-29.15
	CN2 PAST B	61	70.56	75.00	-5.92
	CN2 PAST C	74	58.01	90.00	-35.54
	CN2 PAST D	80	65.60	99.00	-33.73
	CN2 BRUSH A	30	26.62	52.93	-49.70
	CN2 BRUSH B	48	25.00	64.05	-60.97
	CN2 BRUSH C	65	61.73	45.00	37.17
	CN2 BRUSH D	73	60.47	90.00	-32.81
	EPCO	1	0.10	0.10	0.00
	ESCO	1	0.11	0.10	4.26
	PERCOL	0.5	0.97	0.94	3.71
	CN3 SWF	0.95	0.17	0.10	70.01
	FALLTMP	1	5.00	5.00	0.00
	MELTTMP	0.5	5.00	5.00	0.00
	MELTMX	4.5	3.38	3.43	-1.25
	MELTMN	4.5	1.40	1.40	-0.26
	TIMP	1	0.09	0.08	3.13
	COVMX	1	0.50	0.50	0.00
gwflow	RECH_DEL	2	5.98	3.20	87.02
	$K_{aqu}1$	0.0001	7.190E-04	8.277E-04	-13.14
	$K_{aqu}2$	0.01	1.000E-04	1.501E-04	-33.37
	$K_{aqu}3$	0.001	2.225E-03	1.824E-03	22.00
	$K_{aqu}4$	1	6.101E+00	5.857E+00	4.16
	$K_{aqu}5$	100	1.033E+02	5.746E+01	79.79
	S_y1	0.05	1.220E-01	1.349E-01	-9.54
	S_y2	0.05	6.786E-02	6.529E-02	3.92
	S_y3	0.05	1.632E-01	3.483E-01	-53.16
	S_y4	0.25	3.944E-01	3.804E-01	3.68
	S_y5	0.15	1.411E-01	1.079E-01	30.75
	K_{stream}	0.0001	2.964E-07	3.356E-07	-11.68
	d_{stream}	1	1.176E-01	6.969E-02	68.82
	$K_{fplain}1$	0.00529	5.875E-04	—	—
	$K_{fplain}2$	0.0659	9.378E-02	—	—
	$K_{fplain}3$	0.81515	3.903E-01	—	—

Final calibrated parameters varied between scenarios, with parameters in the floodplain scenario ranging from -60.97% to 87.02% of the control scenario values. The three highest differences between parameters were +87.02% for recharge delay (RECH_DEL), +79.79% for the hydraulic conductivity of aquifer zone 5 ($K_{aqu}5$), and +70.01% for the curve number soil water adjustment factor (CN3). These parameters span surface and aquifer processes.

Monthly exchange volumes in the *gwflow* module were different between the floodplain and control scenarios. While many pathways in the *gwflow* module had zero (groundwater evapotranspiration, transfer to the soil profile, discharge to tile drains) or very little (aquifer recharge, stream-aquifer, agricultural pumping, lake-aquifer) exchange volumes over the study period, monthly volumes exchanged through floodplains and saturation excess flow were more substantial. Figure 8 shows a plot of floodplain, saturation excess flow, and groundwater recharge exchange volumes from 2003-2015. Volumes were normalized by dividing by total watershed area and converting to millimeters. Simulated saturation excess flow (discharge from groundwater) was higher in the floodplain scenario than in the control scenario. This is likely due to increased water table elevation from additional aquifer recharge through floodplains.

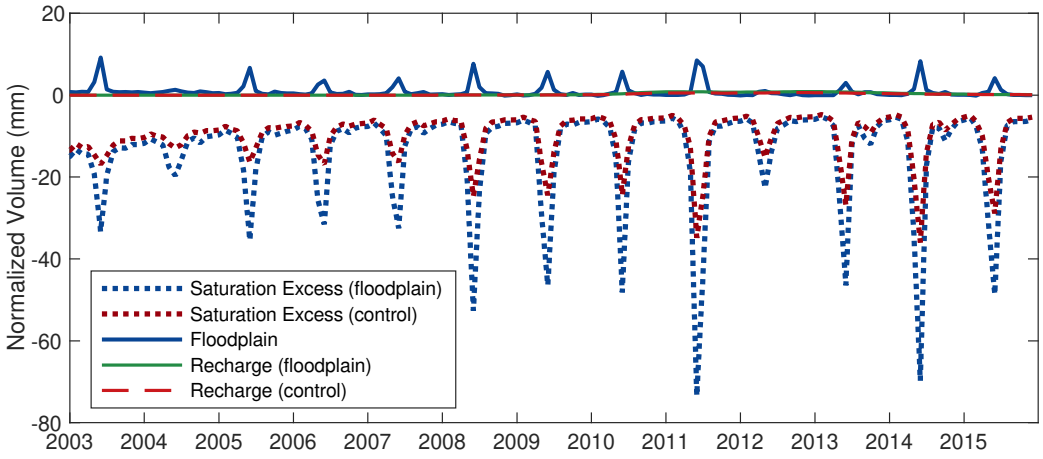

Figure 8: Cumulative monthly volume exchanged for *gwflow* processes over the study watershed at a cell resolution of 250 m, normalized by watershed area.

Monthly exchange volumes simulated for surface processes also varied between the floodplain and control scenarios. Figure 9 shows surface water (A) and shallow lateral flow (B) monthly volumes

generated from 2003-2015 for each scenario. Volumes were again normalized by the total watershed area and converted to millimeters. A higher volume of surface water runoff and shallow lateral flow was generated in the control scenario compared to the floodplain scenario. These higher volumes increase discharge to streams, whereas in the floodplain scenario, this volume of water to streams is approximately matched by the additional saturation excess flow from the *gwflow* module. Other exchange pathways in SWAT+ (e.g., snowmelt, soil percolation, surface evapotranspiration) were essentially identical between the two scenarios.

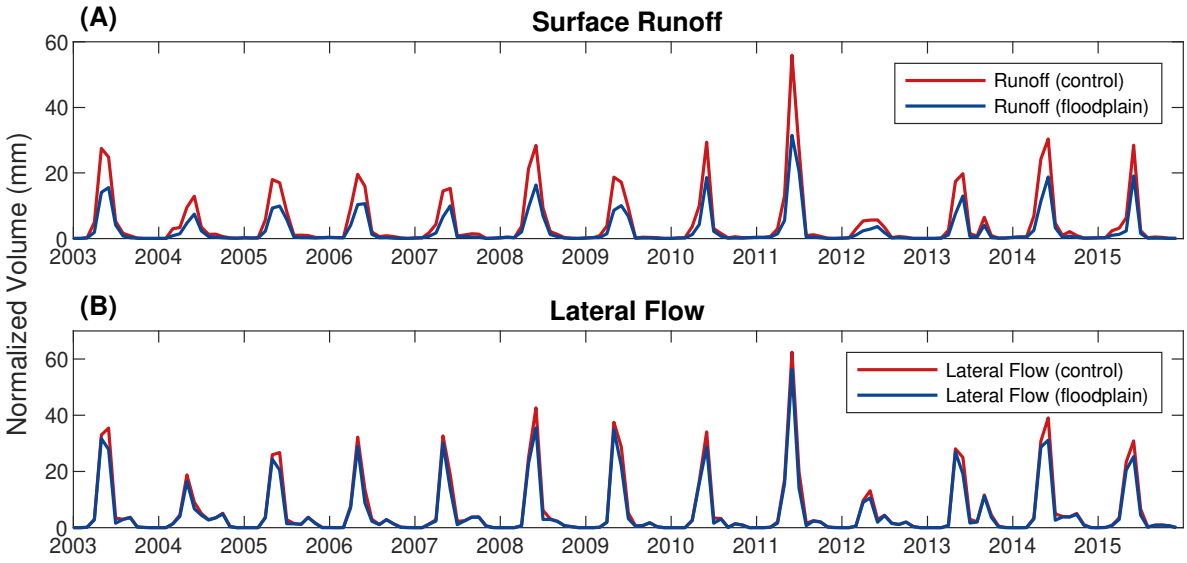

Figure 9: Cumulative monthly volume exchanged for land surface processes over the study watershed, normalized by watershed area.

Table 8 lists total exchange volumes for stream-related processes in the SWAT+ models (surface runoff, shallow lateral flow, stream-aquifer exchange, saturation excess flow, and floodplain exchange) summed from 2003-2015 and divided by watershed area to normalize values to millimeters. Each of these processes sends water directly to channel objects in SWAT+. Figure 10 shows relative percentages of volumes exchanged through these processes. The inclusion of floodplain processes increased the overall contribution of saturation excess flow by an average of 11.7% and reduced the contribution of surface runoff and shallow lateral flow by an average of 10.0% and 4.9%, respectively, compared to the control scenario.

Table 8: Sum of stream-related exchange volumes (mm) simulated from 2003-2015.

Scenario	Cell Size	Stream-Aquifer	Saturation Excess	Floodplain	Surface Runoff	Shallow Lateral Flow
Floodplain	250	39.13	-2073.59	140.81	395.21	795.46
	500	31.78	-1766.91	76.39		
	1000	20.77	-1255.75	94.97		
Control	250	73.49	-1500.88	—	709.41	931.73
	500	60.39	-1330.30	—		
	1000	40.77	-880.80	—		

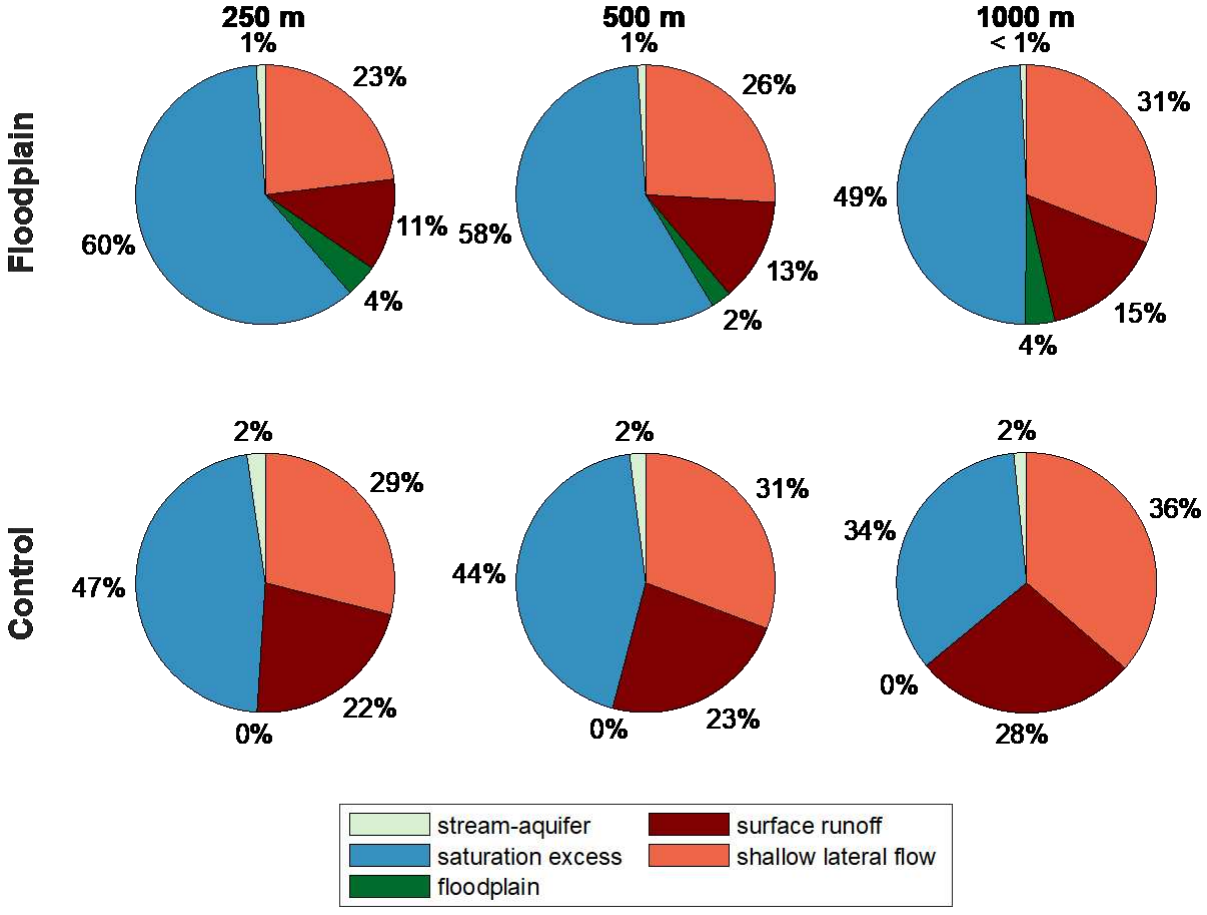

Figure 10: Relative percentage of stream-related exchange volumes for each model scenario.

Furthermore, aquifer system-response variables differed between the two scenarios. Figure 11 shows simulated head difference between the floodplain and control scenarios for 2014, the year with the largest calculated groundwater head differential. I calculated head difference using equation (6).

$$Difference = floodplain - control \quad (6)$$

I selected the head difference banding of Figure 11 to make watershed-wide patterns clear, although the full range of calculated differences was -71.57 m to 79.95 m. Regions where the floodplain scenario simulated lower groundwater head tended to be clustered around floodplain areas of higher order streams. In contrast, groundwater head was higher at upstream regions of the watershed.

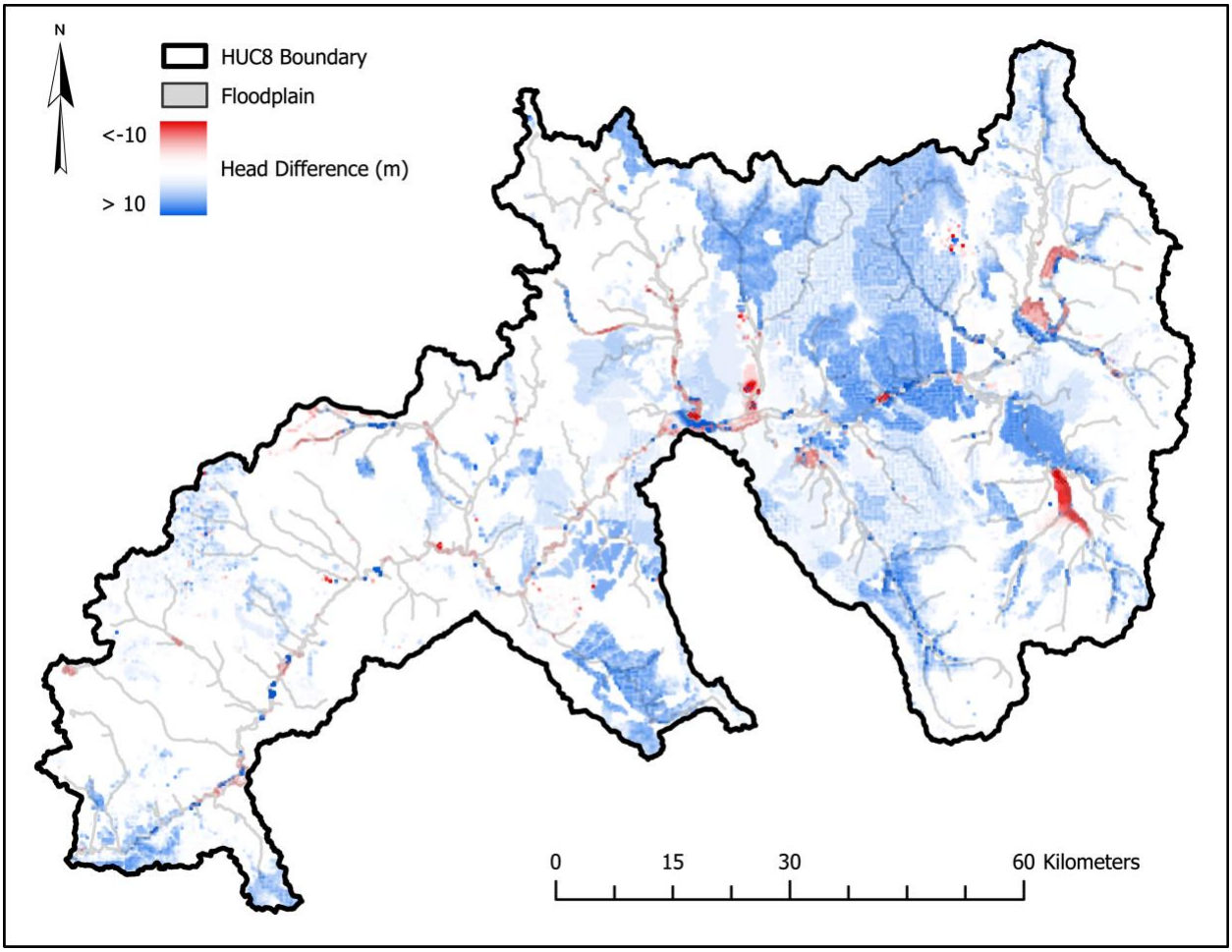

Figure 11: Simulated groundwater head difference in 2014 between the floodplain and control scenarios at a cell resolution of 250 m.

3.3 Floodplain Exchange Patterns

Since floodplain area was relatively small compared to the total watershed size, Figure 12 highlights two locations with higher simulated floodplain fluxes for the year 2014. These fluxes were normalized to a rate (m/day) by dividing average annual volumetric flux (m^3/day) by the floodplain area intersecting each cell.

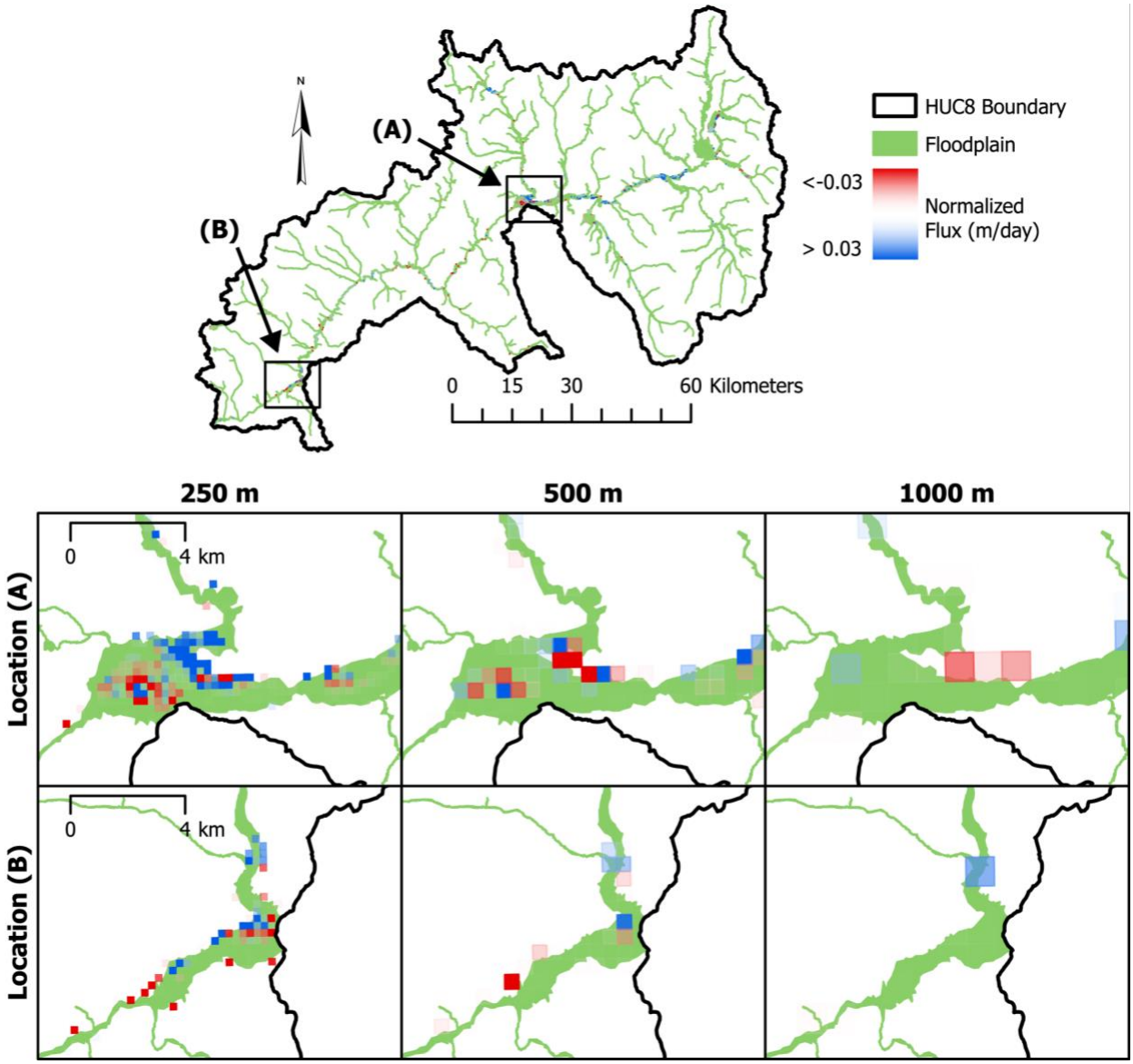

Figure 12: Normalized average annual floodplain fluxes simulated in the year 2014 at two watershed locations for each cell size.

The net volume of water exchanged through floodplains was highest at the smallest grid resolution. Net floodplain volume exchanged from 2003-2015, normalized by total watershed area, was 140.8 mm, 76.4 mm, and 95.0 mm for cell sizes of 250 m, 500 m, and 1000 m, respectively. While general locations of aquifer discharge (negative) and recharge (positive) are captured at coarse cell sizes, there is much more detail about flux magnitude and direction at the 250 m cell size.

3.4 Categorization of Floodplain Exchange

As discussed in Section 2.5.3, I delineated floodplain cells into two categories: beads or strings. Table 9 summarizes the number of bead and string cells for each cell size and stream order. Using this categorization, I created datasets of average annual flux values through bead or string cells. To compare between cell sizes, I divided average annual flux (m^3/day) by the floodplain area in each cell. If a floodplain cell had no simulated fluxes over the entire period 2003-2015, I removed it from the dataset.

Table 9: Quantification of the number of bead and string cells in the study watershed by stream order.

Cell Size (m)	Stream Order	Floodplain Cells	Bead Cells	String Cells	% Bead
250	4	1993	1690	303	84.8%
	5	568	477	91	84.0%
	6	1716	679	1037	39.6%
500	4	720	594	126	82.5%
	5	196	157	39	80.1%
	6	591	200	391	33.8%
1000	4	258	200	58	77.5%
	5	59	42	17	71.2%
	6	200	58	142	29.0%

A Kolmogorov-Smirnov test of normality on each dataset (6 total) indicated that none were normally distributed (p-values on the order of 10^{-299} or less). A non-parametric Mann-Whitney U-test comparing bead and string datasets at each cell size resulted in p-values of 0.0067, 0.0369, and 0.4395 for cell sizes 250 m, 500 m, and 1000 m, respectively. Thus, for a 5% significance level ($\alpha = 0.05$), the null hypothesis that the two datasets are derived from the same population is rejected for cell sizes 250 m and 500 m (i.e., the populations are different), while the 1000 m datasets fail to reject the null hypothesis (i.e.,

the data do not show the populations are different). Table 10 lists dataset statistics separated by beads and strings for each cell size.

Table 10: Statistics for bead and string flux datasets from 2003-2015.

		Mean (mm/day)	Median (mm/day)	Lower Quartile (mm/day)	Upper Quartile (mm/day)	Maximum (mm/day)	Minimum (mm/day)
250 m	Bead	3.954	0.000	-0.286	1.284	387.200	-288.880
	String	0.044	0.000	-0.265	0.257	101.410	-121.910
500 m	Bead	1.008	0.000	-0.463	0.559	355.540	-227.480
	String	0.568	0.000	-0.360	0.470	72.736	-93.963
1000 m	Bead	2.720	0.000	-0.257	1.265	59.822	-25.629
	String	0.932	0.000	-0.417	0.914	44.063	-94.023

The boxplots in Figure 13 provide a visual representation of the bead and string datasets. Due to many outliers at all cell resolutions, I limited the extent of the boxplots. The percent of each dataset classified as outliers is listed on relevant plots, and the minimum and maximum values for each case are included in Table 10. The average lower and upper quartiles of bead cells for all cell sizes were -0.335 mm/day and 1.036 mm/day, respectively. For string cells, the average lower and upper quartiles for all cell sizes were -0.347 mm/day and 0.547 mm/day, respectively. While the spread of data is large, most floodplain fluxes were clustered around -0.5 mm/day to 1 mm/day, where bead cells had a larger interquartile range. Median values for beads and strings at each cell size were equal at 0 mm/day while the mean values had larger differences. The difference between mean bead and mean string flux values was most pronounced for the 250 m grid size (3.910 mm/day) compared to the difference at 500 m (0.440 mm/day) or 1000 m (1.789 mm/day) grid sizes.

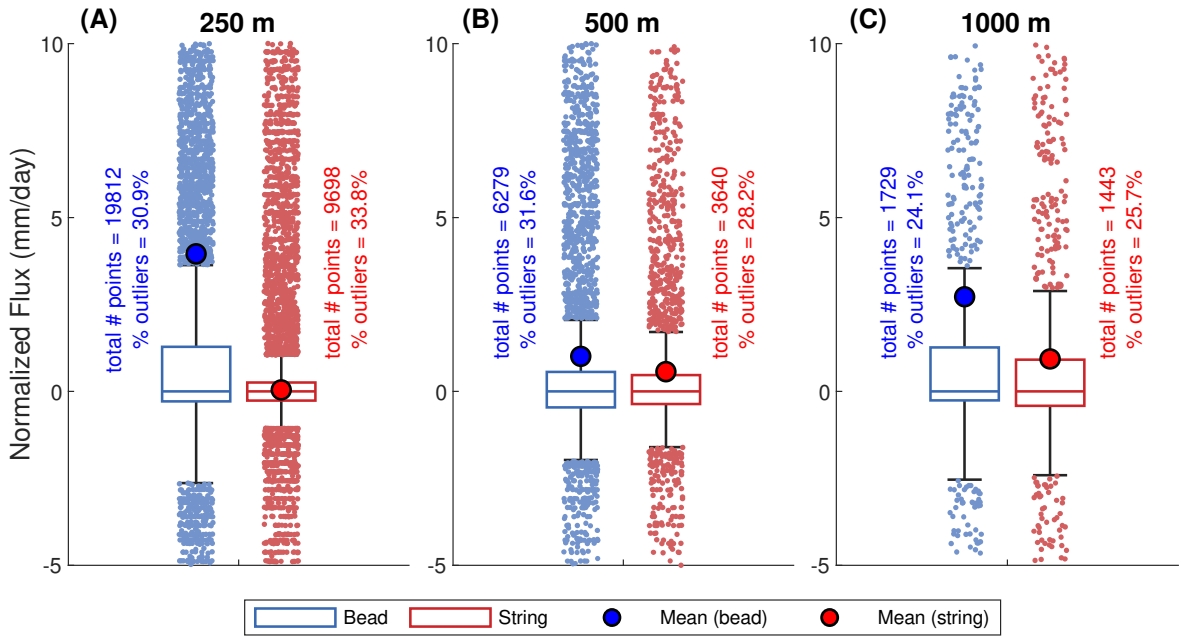

Figure 13: Normalized average annual floodplain flux by bead and string cells simulated from 2003-2015.

I calculated net water volume exchanged for bead and string cells from average annual flux. For each cell size, I multiplied the non-normalized volumetric flux (m^3/day) at every grid cell by the number of days in each year to calculate a total volume exchanged per year per grid cell. Total volume exchanged by bead and string cells per year was the sum of these volumes by category. I also summed total recharge (positive values) and total discharge (negative values) for each category per year. Figure 14 shows plots of these values.

In addition, I summed total exchange volumes over the period 2003-2015 for bead and string cells. To account for higher bead areas compared to string areas, I normalized the data by dividing total exchange volume by the area of each floodplain category. The bead areas were 98.285 km^2, 95.934 km^2, and 82.882 km^2, compared to string areas of 28.526 km^2, 28.090 km^2, and 25.298 km^2 for 250 m, 500 m, and 1000 m resolutions, respectively. Table 11 summarizes total volumes and normalized volumes at each resolution. I calculated the percent difference between bead and string volumes using equation (7).

$$\%\ Difference = \left(\frac{bead - string}{string}\right) \cdot 100 \quad (7)$$

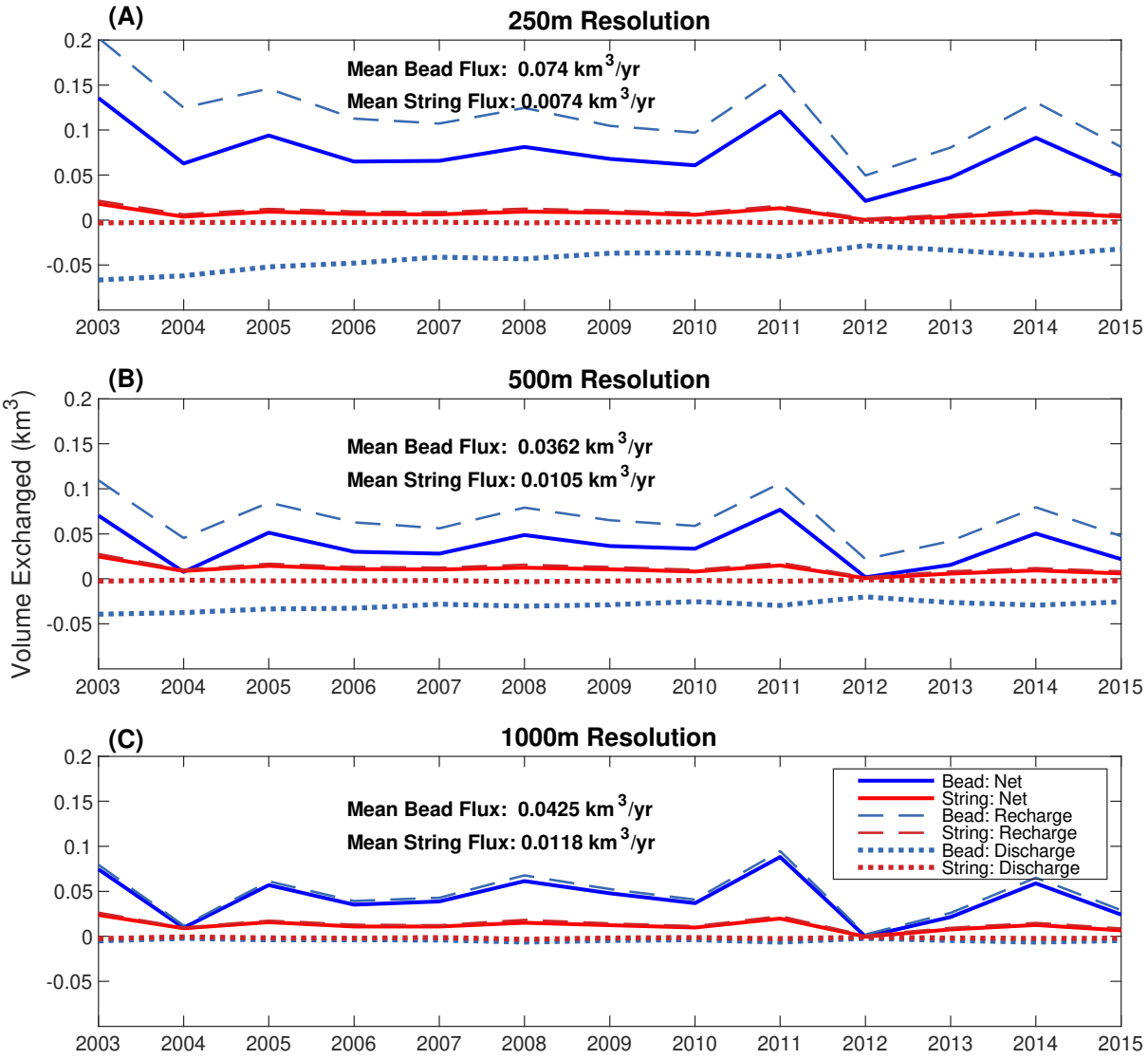

Figure 14: Total simulated recharge, discharge, and net volume exchanged by bead and string cells from 2003-2015 for varied cell resolutions.

Table 11: Total and normalized exchange volumes summed over 2003-2015 by beads and strings.

	Total Exchange (km^3)			Normalized Exchange (m)		
Cell Size (m)	Bead	String	% Difference	Bead	String	% Difference
250	0.9625	0.0962	900.0	9.7925	3.3739	190.2
500	0.4711	0.1365	245.1	4.9101	4.8599	1.033
1000	0.5523	0.1531	260.7	6.6635	6.0525	10.10

Bead cells simulated a substantially higher total exchange volume compared to string cells for all cell resolutions. Normalization of exchange volume by area reduced the percent differences, but still indicated higher exchange through bead cells for every resolution. Total volume exchanged through bead cells was also higher in each individual year than for string cells (Figure 14). Total simulated exchange volume through bead regions increased with reduced cell size, while the volume through string regions decreased. This could be due to the higher percentage of bead cells relative to string cells calculated at finer resolutions (Table 9).

4 DISCUSSION

4.1 Inclusion of Floodplain-Groundwater Interactions

When compared to measured streamflow, models including surface water-groundwater exchange through floodplains performed as well or better than control models for all cell resolutions. Floodplain scenario models also provided valuable information about floodplain inundation and exchange rates in the study watershed. From these results, I conclude that the addition of floodplain-groundwater exchange in the *gwflow* module of SWAT+ improves the model's representation of watersheds since it has similar performance while including another realistic hydrologic exchange pathway.

Furthermore, smaller *gwflow* cell sizes showed similar or better performance than larger cell sizes and better captured local patterns of floodplain-groundwater recharge or discharge (Figure 12). These smaller cell sizes also simulated a higher total and normalized volume exchanged through floodplains (Table 11). Since floodplains are complex landscapes that have heterogeneous characteristics, these smaller cells likely provided superior representation of landscape variation, which led to higher simulated exchange volumes. Thus, I conclude that spatial refinement of the *gwflow* module grid is important for improving model performance and for understanding where floodplain-groundwater exchange occurs in a watershed. I expect that direct calibration at smaller cell sizes would further improve performance and representation of exchange patterns.

The volume of water exchanged between groundwater and floodplain surface water was non-negligible, especially at smaller cell sizes where simulated exchange was the highest. Understanding how these individual hydrologic exchanges contribute to streamflow is essential for accurately simulating a watershed. Floodplains exchanged more volume with groundwater than through stream-aquifer interactions and were the largest source of groundwater recharge in the model (Figure 8). Quantification of these processes in the field or further studies are necessary to fully validate these results; however, considering model performance metrics and the physically based approach of the *gwflow* module, these floodplain processes seem to be a substantial hydrologic exchange pathway in the study watershed.

4.2 Quantification and Categorization of Floodplain Fluxes

Many studies have assessed the function of floodplain beads compared to strings. Beads in old-growth forests have higher wood loads than in strings (Livers & Wohl, 2016; Polvi & Wohl, 2013), which in turn leads to higher spatial heterogeneity (Wohl et al., 2017). Spatial heterogeneity improves the availability and diversity of habitat (Wohl et al., 2017), benefits biological productivity and species richness (Bellmore & Baxter, 2014; Hauer et al., 2016; Hood & Larson, 2014), and increases organic carbon retention (Wohl et al., 2012). These studies emphasize the importance of bead regions, although they focus on surface processes which are often easier to assess visually or through field studies. Groundwater processes, especially flux rates, are much more difficult to assess as they are not as easily measured or visualized. Moreover, quantifying processes at a watershed scale is often infeasible due to access constraints and the amount of data required.

This study introduces a novel approach for quantifying surface water-groundwater interactions through bead and string regions. My analyses on the location of floodplain fluxes within the watershed showed that beads had a higher net and per area volume of surface water-groundwater exchange than strings. These outputs remained consistent across all studied cell resolutions, with smaller resolutions simulating greater differences between bead and string floodplain regions. I also found that flux rates through bead regions were higher than for string regions, and these results were statistically significant for the two smallest cell sizes. While this quantification focuses on surface water-groundwater exchange, which is only one exchange pathway present in the complex mosaic of floodplain processes, it confirms existing field studies and our conceptual understanding of how beads function.

4.3 Implications and Recommendations

The study watershed is part of the larger Colorado River basin, which is one of the most regulated hydrologic systems in the world (Graf, 1985). Impacts to this headwater system can have cascading effects downstream, so understanding the processes that occur within the study watershed are important for managing and protecting water resources. This is especially important as the river network faces

climate-related challenges (Bair et al., 2019) that include increasing human consumption, diminishing flows (Christensen et al., 2004), and declines in riparian ecosystem health (Sankey et al., 2015).

While it is known that floodplain beads are important for protecting river system resilience and ecosystem function (Hauer et al., 2016; Wohl et al., 2017), land use changes and reductions in forest cover often impact bead function (Wohl & Beckman, 2014). Land management and conservation practices should prioritize these locations, as they are key locations that improve ecosystem health. These study results further confirm the importance of floodplain beads, specifically in the context of surface water-groundwater exchange. Through this modeling effort, I showed that floodplain beads had disproportionately higher volumes and flux rates of hydrologic exchange than strings. Thus, in addition to their improved surface heterogeneity, storage, and biological productivity, I also conclude that beads are important locations for groundwater interactions.

4.4 Study Limitations

There were several limitations to this study. I only performed calibration for models with a grid cell size of 1000 m, which was the smallest cell size feasible due to time and computing constraints. The calibration procedure for the floodplain scenario had more optimization iterations and model runs than the control scenario, which may have skewed results in favor of the floodplain scenario despite equivalent convergence parameters and similar end results. Furthermore, I did not calibrate models to groundwater well data due to limited well locations. While there were wells within the watershed that had measured groundwater head during the study period, they were clustered closely together and did not seem representative of groundwater head for the overall watershed. Some analysis was still performed to assess model fit (Appendix B), but none of the models were directly calibrated to well data.

The finest cell resolution (250 m) was coarse when considering the complexity of floodplain processes. For this study, this cell size was the smallest resolution that provided reasonable model run times. The GFPLAIN resolution also limited analysis, as it was produced at a resolution of 30 m, which was not sufficient for analysis of stream orders 1-3 in the watershed. Additionally, the threshold

floodplain width for the bead and string categorization impacted the location and number of cells selected for each category. Moreover, the SWAT+ assumption of a constant floodplain width (5 times channel width) may have impacted results as water depth and water volume calculations are not based on the variable floodplain widths present in this study watershed.

Finally, these SWAT+ and *gwflow* models were applied to a montane region in the western US which is not representative of the many variations in climate, land use, soil, and aquifer properties that exist worldwide. The results presented here should not be extrapolated to other regions without further study. The analysis on bead and string locations is also limited by this study location, and the patterns noted here require additional verification at other study sites.

5 CONCLUSIONS

In this study, I found that SWAT+ models with *gwflow* for the Colorado Headwaters watershed were able to reasonably simulate observed streamflow at two USGS gage locations. Models that included floodplain exchanges in *gwflow* were able to meet or exceed model performance metrics compared to the control scenario. In the floodplain scenario, simulated streamflow contributions from saturation excess flow increased, and contributions from surface runoff and shallow lateral flow decreased compared to the control scenario. Exchange through floodplains provided net recharge to the groundwater aquifer and was the most substantial source of groundwater recharge in the floodplain models.

Smaller *gwflow* grid cell sizes improved model performance at the downstream gage and had equivalent performance at the upstream gage. Finer grid resolutions also provided more information on the magnitude, direction, and location of floodplain exchanges throughout the watershed, and simulated higher total floodplain exchange volumes.

Analyses of floodplain fluxes indicated that bead regions of the floodplain exchanged a higher volume of water, both net and by area, than string regions. Mean values of normalized flux were also higher in bead regions than in string regions. Thus, model results suggest that efforts to conserve or protect floodplain surface water-groundwater exchanges in this watershed should focus on wider floodplain regions (beads) to preserve a larger proportion of hydrologic interactions.

For an analysis of lower stream orders, a finer resolution floodplain delineation procedure could provide enough detail to analyze changes in floodplain width and thus improve bead and string categorization at that scale. However, these analyses would likely require proprietary datasets and a smaller model scope to allow for reasonable run times. Analyses of lower order streams should be coupled with an equivalent reduction in *gwflow* grid cell size to ensure that the model is able to simulate the finer resolution.

To verify performance for various conditions, further studies should apply the updated *gwflow* module in SWAT+ to watersheds with different climate, surface/subsurface properties, and sizes.

Calibration should be performed at groundwater wells where data are available and could be improved with more computing power to calibrate smaller cell resolutions directly, or to decrease cell size further. To better capture spatial variations, additional aquifer and floodplain soil property zones could be included in future *gwflow* module setups, and a more detailed SWAT+ modeling framework could be created. A sensitivity analysis on the threshold width determining bead locations is necessary to understand how that value influences results. Finally, ground truthing of simulated exchange processes would provide valuable insight into model accuracy.

REFERENCES

Andersen, D. C., Cooper, D. J., & Northcott, K. (2007). Dams, floodplain land use, and riparian forest conservation in the semiarid Upper Colorado River Basin, USA. *Environmental Management*, *40*(3), 453–475. https://doi.org/10.1007/s00267-006-0294-7

Andersen, H. E. (2004). Hydrology and nitrogen balance of a seasonally inundated Danish floodplain wetland. *Hydrological Processes*, *18*(3), 415–434. https://doi.org/10.1002/hyp.1277

Arabi, M., Govindaraju, R. S., Hantush, M. M., & Engel, B. A. (2006). Role of Watershed Subdivision on Modeling the Effectiveness of Best Management Practices with Swat1. *JAWRA Journal of the American Water Resources Association*, *42*(2), 513–528. https://doi.org/10.1111/j.1752-1688.2006.tb03854.x

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large Area Hydrologic Modeling and Assessment Part I: Model Development1. *JAWRA Journal of the American Water Resources Association*, *34*(1), 73–89. https://doi.org/10.1111/j.1752-1688.1998.tb05961.x

Arnold, J. G., White, M. J., Allen, P. M., Gassman, P. W., & Bieger, K. (2021). Conceptual Framework of Connectivity for a National Agroecosystem Model Based on Transport Processes and Management Practices. *JAWRA Journal of the American Water Resources Association*, *57*(1), 154–169. https://doi.org/10.1111/1752-1688.12890

Baffaut, C., Baker, J. M., Biederman, J. A., Bosch, D. D., Brooks, E. S., Buda, A. R., Demaria, E. M., Elias, E. H., Flerchinger, G. N., Goodrich, D. C., Hamilton, S. K., Hardegree, S. P., Harmel, R. D., Hoover, D. L., King, K. W., Kleinman, P. J., Liebig, M. A., McCarty, G. W., Moglen, G. E., … Yasarer, L. M. W. (2020). Comparative analysis of water budgets across the U.S. long-term agroecosystem research network. *Journal of Hydrology*, *588*, 125021. https://doi.org/10.1016/j.jhydrol.2020.125021

Bailey, R., & Alderfer, C. (2022). Groundwater Data in Unconfined Aquifers—Conterminous United States. *Figshare*. https://doi.org/10.6084/m9.figshare.c.5 918738.v2

Bailey, R. T., Abbas, S., Arnold, J., White, M., Gao, J., & Čerkasova, N. (2023). Augmenting the National agroecosystem model with physically based spatially distributed groundwater modeling. *Environmental Modelling & Software*, *160*, 105589. https://doi.org/10.1016/j.envsoft.2022.105589

Bailey, R. T., Bieger, K., Arnold, J. G., & Bosch, D. D. (2020). A New Physically-Based Spatially-Distributed Groundwater Flow Module for SWAT+. *Hydrology*, *7*(4), Article 4. https://doi.org/10.3390/hydrology7040075

Bailey, R. T., Bieger, K., Flores, L., & Tomer, M. (2022). Evaluating the contribution of subsurface drainage to watershed water yield using SWAT+ with groundwater modeling. *Science of The Total Environment*, *802*, 149962. https://doi.org/10.1016/j.scitotenv.2021.149962

Bair, L. S., Yackulic, C. B., Schmidt, J. C., Perry, D. M., Kirchhoff, C. J., Chief, K., & Colombi, B. J. (2019). Incorporating social-ecological considerations into basin-wide responses to climate change in the Colorado River Basin. *Current Opinion in Environmental Sustainability*, *37*, 14–19. https://doi.org/10.1016/j.cosust.2019.04.002

Bates, P. D., Stewart, M. D., Desitter, A., Anderson, M. G., Renaud, J. P., & Smith, J. A. (2000). Numerical simulation of floodplain hydrology. *Water Resources Research*, *36*(9), 2517–2529. https://doi.org/10.1029/2000WR900102

Bellmore, J. R., & Baxter, C. V. (2014). Effects of Geomorphic Process Domains on River Ecosystems: A Comparison of Floodplain and Confined Valley Segments. *River Research and Applications*, *30*(5), 617–630. https://doi.org/10.1002/rra.2672

Bieger, K., Arnold, J. G., Rathjens, H., White, M. J., Bosch, D. D., & Allen, P. M. (2019). Representing the Connectivity of Upland Areas to Floodplains and Streams in SWAT+. *JAWRA Journal of the American Water Resources Association*, *55*(3), 578–590. https://doi.org/10.1111/1752-1688.12728

Bieger, K., Arnold, J. G., Rathjens, H., White, M. J., Bosch, D. D., Allen, P. M., Volk, M., & Srinivasan, R. (2017). Introduction to SWAT+, A Completely Restructured Version of the Soil and Water

Assessment Tool. *JAWRA Journal of the American Water Resources Association*, *53*(1), 115–130. https://doi.org/10.1111/1752-1688.12482

Boulton, A. J., Findlay, S., Marmonier, P., Stanley, E. H., & Valett, H. M. (1998). The Functional Significance of the Hyporheic Zone in Streams and Rivers. *Annual Review of Ecology and Systematics*, *29*, 59–81.

Brunke, M., & Gonser, T. (1997). The ecological significance of exchange processes between rivers and groundwater. *Freshwater Biology*, *37*(1), 1–33. https://doi.org/10.1046/j.1365-2427.1997.00143.x

Brunner, P., Simmons, C. T., & Cook, P. G. (2009). Spatial and temporal aspects of the transition from connection to disconnection between rivers, lakes and groundwater. *Journal of Hydrology*, *376*(1), 159–169. https://doi.org/10.1016/j.jhydrol.2009.07.023

Caldwell, P. V., Sun, G., McNulty, S. G., Cohen, E. C., & Moore Myers, J. A. (2012). Impacts of impervious cover, water withdrawals, and climate change on river flows in the conterminous US. *Hydrology and Earth System Sciences*, *16*(8), 2839–2857. https://doi.org/10.5194/hess-16-2839-2012

Cartwright, I., Werner, A. D., & Woods, J. A. (2019). Using geochemistry to discern the patterns and timescales of groundwater recharge and mixing on floodplains in semi-arid regions. *Journal of Hydrology*, *570*, 612–622. https://doi.org/10.1016/j.jhydrol.2019.01.023

Chen, M., Gassman, P. W., Srinivasan, R., Cui, Y., & Arritt, R. (2020). Analysis of alternative climate datasets and evapotranspiration methods for the Upper Mississippi River Basin using SWAT within HAWQS. *Science of The Total Environment*, *720*, 137562. https://doi.org/10.1016/j.scitotenv.2020.137562

Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P., & Palmer, R. N. (2004). The Effects of Climate Change on the Hydrology and Water Resources of the Colorado River Basin. *Climatic Change*, *62*(1), 337–363. https://doi.org/10.1023/B:CLIM.0000013684.13621.1f

Cohen, S., Praskievicz, S., & Maidment, D. R. (2018). Featured Collection Introduction: National Water Model. *JAWRA Journal of the American Water Resources Association*, *54*(4), 767–769. https://doi.org/10.1111/1752-1688.12664

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton, P., & van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, *387*(6630), Article 6630. https://doi.org/10.1038/387253a0

de Vente, J., Poesen, J., Verstraeten, G., Govers, G., Vanmaercke, M., Van Rompaey, A., Arabkhedri, M., & Boix-Fayos, C. (2013). Predicting soil erosion and sediment yield at regional scales: Where do we stand? *Earth-Science Reviews*, *127*, 16–29. https://doi.org/10.1016/j.earscirev.2013.08.014

Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., Barber, N. L., & Linsey, K. S. (2018). Estimated Use of Water in the United States in 2015. In *Circular* (p. 76).

Doherty, J. (2010). *PEST User-Manual: Model-independent parameter estimation*. Watermark Numerical Computing.

Downing, J. A., Cole, J. J., Duarte, C. M., Middelburg, J. J., Melack, J. M., Prairie, Y. T., Kortelainen, P., Striegl, R. G., McDowell, W. H., & Tranvik, L. J. (2012). Global abundance and size distribution of streams and rivers. *Inland Waters*, *2*(4), 229–236. https://doi.org/10.5268/IW-2.4.502

Gesch, D. B., Evans, G. A., Oimoen, M. J., & Arundel, S. (2018). *The National Elevation Dataset*. U.S. Geological Survey.

Goodrich, D. C., Williams, D. G., Unkrich, C. L., Hogan, J. F., Scott, R. L., & Hultine, K. R. (2004). Comparison of methods to estimate ephemeral channel recharge, Walnut Gulch, San Pedro River Basin, Arizona. In *Groundwater recharge in a desert environment: The southwestern United States* (pp. 77–99). American Geophysical Union.

Graf, W. L. (1985). *The Colorado River: Instability and basin management* (Resource Publication No. 84/2). Association of American Geographers.

Hauer, R. F., Locke, H., Dreitz, V. J., Hebblewhite, M., Lowe, W. H., Muhlfeld, C. C., Nelson, C. R., Proctor, M. F., & Rood, S. B. (2016). Gravel-bed river floodplains are the ecological nexus of glaciated mountain landscapes. *Science Advances*, *2*, e1600026. https://doi.org/10.1126/sciadv.1600026

Helton, A. M., Poole, G. C., Payn, R. A., Izurieta, C., & Stanford, J. A. (2014). Relative influences of the river channel, floodplain surface, and alluvial aquifer on simulated hydrologic residence time in a montane river floodplain. *Geomorphology*, *205*, 17–26. https://doi.org/10.1016/j.geomorph.2012.01.004

Hood, G. A., & Larson, D. G. (2014). Beaver-Created Habitat Heterogeneity Influences Aquatic Invertebrate Assemblages in Boreal Canada. *Wetlands*, *34*, 19–29. https://doi.org/10.1007/s13157-013-0476-z

Horton, J. D. (2017). *The state geologic map compilation (SGMC) geodatabase of the conterminous United States: US geological survey data release*. US Geological Survey.

Junk, W. J., Bayley, P. B., & Sparks, R. E. (1989). The flood pulse concept in river-floodplain systems. *Canadian Special Publications Fisheries and Aquatic Sciences*, *106*, 110–127.

Knox, R. L., Morrison, R. R., & Wohl, E. E. (2022). Identification of Artificial Levees in the Contiguous United States. *Water Resources Research*, *58*(4), e2021WR031308. https://doi.org/10.1029/2021WR031308

Krause, S., Bronstert, A., & Zehe, E. (2007). Groundwater–surface water interactions in a North German lowland floodplain – Implications for the river discharge dynamics and riparian water balance. *Journal of Hydrology*, *347*(3), 404–417. https://doi.org/10.1016/j.jhydrol.2007.09.028

Liechti, T. C., Matos, J. P., Ferràs Segura, D., Boillat, J.-L., & Schleiss, A. J. (2014). Hydrological modelling of the Zambezi River Basin taking into account floodplain behaviour by a modified reservoir approach. *International Journal of River Basin Management*, *12*(1), 29–41. https://doi.org/10.1080/15715124.2014.880707

Livers, B., & Wohl, E. (2016). Sources and interpretation of channel complexity in forested subalpine streams of the Southern Rocky Mountains. *Water Resources Research*, *52*, 3910–3929. https://doi.org/10.1002/2015WR018306

Maier, N., Breuer, L., & Kraft, P. (2017). Prediction and uncertainty analysis of a parsimonious floodplain surface water-groundwater interaction model. *Water Resources Research*, *53*(9), 7678–7695. https://doi.org/10.1002/2017WR020749

Martinet, M. C., Vivoni, E. R., Cleverly, J. R., Thibault, J. R., Schuetz, J. F., & Dahm, C. N. (2009). On groundwater fluctuations, evapotranspiration, and understory removal in riparian corridors. *Water Resources Research*, *45*(5). https://doi.org/10.1029/2008WR007152

Moore, R. B., & Dewald, T. G. (2016). The Road to NHDPlus—Advancements in Digital Stream Networks and Associated Catchments. *JAWRA Journal of the American Water Resources Association*, *52*(4), 890–900.

Moriasi, D. N., Gitau, M. W., Pai, N., & Daggupati, P. (2015). *Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria*. https://doi.org/10.13031/trans.58.10715

Nanson, G. C., & Croke, J. C. (1992). A genetic classification of floodplains. *Geomorphology*, *4*(6), 459–486. https://doi.org/10.1016/0169-555X(92)90039-Q

Nardi, F. (2020). *GFPLAIN* [Python]. https://github.com/fnardi/GFPLAIN (Original work published 2018)

Nardi, F., Annis, A., Di Baldassarre, G., Vivoni, E. R., & Grimaldi, S. (2019). GFPLAIN250m, a global high-resolution dataset of Earth's floodplains. *Scientific Data*, *6*(1), Article 1. https://doi.org/10.1038/sdata.2018.309

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., & Williams, J. R. (2009). Soil and water assessment tool theoretical documentation version 2009. *Texas Water Resources Institute*.

Opperman, J. J., Luster, R., McKenney, B. A., Roberts, M., & Meadows, A. W. (2010). Ecologically Functional Floodplains: Connectivity, Flow Regime, and Scale. *JAWRA Journal of the American Water Resources Association*, *46*(2), 211–226. https://doi.org/10.1111/j.1752-1688.2010.00426.x

Petsch Jr., H. E. (1985). *Inventory of inter-basin transfers of water in the western conterminous United States* (Open-File Report No. 85–166). US Geological Survey.

Phiri, W. K., Vanzo, D., Banda, K., Nyirenda, E., & Nyambe, I. A. (2021). A pseudo-reservoir concept in SWAT model for the simulation of an alluvial floodplain in a complex tropical river system. *Journal of Hydrology: Regional Studies*, *33*, 100770. https://doi.org/10.1016/j.ejrh.2020.100770

Polvi, L. E., & Wohl, E. (2013). Biotic Drivers of Stream Planform: Implications for Understanding the Past and Restoring the Future. *BioScience*, *63*(6), 439–452. https://doi.org/10.1525/bio.2013.63.6.6

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., Cooper, D. J., Coen, J. L., Ahmadov, R., Peckham, S. E., Grell, G. G., Michalakes, J., Trahan, S., Benjamin, S. G., Alexander, C. R., Dimego, G. J., Wang, W., Schwartz, C. S., Romine, G. S., … Duda, M. G. (2017). The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions in: Volume Issue (). *Bulletin of the American Meteorological Society*, *98*(8). https://doi.org/10.1175/BAMS-D-15-00308.1

Rajib, A., Liu, Z., Merwade, V., Tavakoly, A. A., & Follum, M. L. (2020). Towards a large-scale locally relevant flood inundation modeling framework using SWAT and LISFLOOD-FP. *Journal of Hydrology*, *581*, 124406. https://doi.org/10.1016/j.jhydrol.2019.124406

Saksena, S., & Merwade, V. (2017). *Integrated Modeling of Surface-Subsurface Processes to Understand River-Floodplain Hydrodynamics in the Upper Wabash River Basin*. 60–68. https://doi.org/10.1061/9780784480595.006

Sankey, J. B., Ralston, B. E., Grams, P. E., Schmidt, J. C., & Cagney, L. E. (2015). Riparian vegetation, Colorado River, and climate: Five decades of spatiotemporal dynamics in the Grand Canyon with river regulation. *Journal of Geophysical Research: Biogeosciences*, *120*(8), 1532–1547. https://doi.org/10.1002/2015JG002991

Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., & Dai, Y. (2017). Mapping the global depth to bedrock for land surface modeling. *Journal of Advances in Modeling Earth Systems*, *9*(1), 65–88. https://doi.org/10.1002/2016MS000686

Skinner, K. D., & Maupin, M. A. (2019). Point-source nutrient loads to streams of the conterminous United States, 2012. In *Data Series* (p. 24). https://doi.org/10.3133/ds1101

Soil Survey Staff. (2014). *Gridded Soil Survey Geographic (gSSURGO) Database for the Conterminous United States*. United States Department of Agriculture.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. (2011). *Soil Survey Geographic (SSURGO) Database* [Data set]. https://sdmdataaccess.sc.egov.usda.gov

Stanford, J. A., & Ward, J. V. (1993). An Ecosystem Perspective of Alluvial Rivers: Connectivity and the Hyporheic Corridor. *Journal of the North American Benthological Society*, *12*(1), 48–60. https://doi.org/10.2307/1467685

Stanford, J. A., Ward, J. V., Liss, W. J., Frissell, C. A., Williams, R. N., Lichatowich, J. A., & Coutant, C. C. (1996). A General Protocol for Restoration of Regulated Rivers. *Regulated Rivers: Research & Management*, *12*(4–5), 391–413. https://doi.org/10.1002/(SICI)1099-1646(199607)12:4/5<391::AID-RRR436>3.0.CO;2-4

Sun, X., Bernard-Jannin, L., Grusson, Y., Sauvage, S., Arnold, J., Srinivasan, R., & Sánchez Pérez, J. M. (2018). Using SWAT-LUD Model to Estimate the Influence of Water Exchange and Shallow Aquifer Denitrification on Water and Nitrate Flux. *Water*, *10*(4), Article 4. https://doi.org/10.3390/w10040528

Sun, X., Bernard-Jannin, L., Sauvage, S., Garneau, C., Arnold, J. G., Srinivasan, R., & Sánchez-Pérez, J. M. (2016). Assessment of the denitrification process in alluvial wetlands at floodplain scale using the SWAT model. *Ecological Engineering*, *103*, 344–358. https://doi.org/10.1016/j.ecoleng.2016.06.098

Tockner, K., & Stanford, J. (2002). Review of: Riverine Flood Plains: Present State and Future Trends. *Environmental Conservation*, *29*(3), 308–330.

Tonina, D., & Buffington, J. M. (2009). Hyporheic Exchange in Mountain Rivers I: Mechanics and Environmental Effects. *Geography Compass*, *3*(3), 1063–1086. https://doi.org/10.1111/j.1749-8198.2009.00226.x

Valayamkunnath, P., Barlage, M., Chen, F., Gochis, D. J., & Franz, K. J. (2020). Mapping of 30-meter resolution tile-drained croplands using a geospatial modeling approach. *Scientific Data*, *7*(1), Article 1. https://doi.org/10.1038/s41597-020-00596-x

Ward, J. V., Tockner, K., Arscott, D. B., & Claret, C. (2002). Riverine landscape diversity. *Freshwater Biology*, *47*(4), 517–539.

Ward, J. V., Tockner, K., & Schiemer, F. (1999). Biodiversity of floodplain river ecosystems: Ecotones and connectivity1. *Regulated Rivers: Research & Management*, *15*(1–3), 125–139. https://doi.org/10.1002/(SICI)1099-1646(199901/06)15:1/3<125::AID-RRR523>3.0.CO;2-E

White, M. J., Arnold, J. G., Bieger, K., Allen, P. M., Gao, J., Čerkasova, N., Gambone, M., Park, S., Bosch, D. D., Yen, H., & Osorio, J. M. (2022). Development of a Field Scale SWAT+ Modeling Framework for the Contiguous U.S. *JAWRA Journal of the American Water Resources Association*, *n/a*(n/a). https://doi.org/10.1111/1752-1688.13056

Woessner, W. W. (2000). Stream and Fluvial Plain Ground Water Interactions: Rescaling Hydrogeologic Thought. *Groundwater*, *38*(3), 423–429. https://doi.org/10.1111/j.1745-6584.2000.tb00228.x

Wohl, E. (2021). An Integrative Conceptualization of Floodplain Storage. *Reviews of Geophysics*, *59*(2). https://doi.org/10.1029/2020RG000724

Wohl, E., & Beckman, N. D. (2014). Leaky rivers: Implications of the loss of longitudinal fluvial disconnectivity in headwater streams. *Geomorphology*, *205*, 27–35. https://doi.org/10.1016/j.geomorph.2011.10.022

Wohl, E., Dwire, K., Sutfin, N., Polvi, L., & Bazan, R. (2012). Mechanisms of carbon storage in mountainous headwater rivers. *Nature Communications*, *3*(1), Article 1. https://doi.org/10.1038/ncomms2274

Wohl, E., Lininger, K. B., & Scott, D. N. (2017). River beads as a conceptual framework for building carbon storage and resilience to extreme climate events into river management. *Biogeochemistry*, *141*(3), 365–383. https://doi.org/10.1007/s10533-017-0397-7

Yan, L., & Roy, D. P. (2016). Conterminous United States crop field size quantification from multi-temporal Landsat data. *Remote Sensing of Environment*, *172*, 67–86. https://doi.org/10.1016/j.rse.2015.10.034

Yen, H., Daggupati, P., White, M. J., Srinivasan, R., Gossel, A., Wells, D., & Arnold, J. G. (2016). Application of Large-Scale, Multi-Resolution Watershed Modeling Framework Using the Hydrologic and Water Quality System (HAWQS). *Water*, *8*(4), Article 4. https://doi.org/10.3390/w8040164

APPENDIX A: ADDITIONAL BACKGROUND

Download links for models and datasets:

- SWAT+ is available for download at https://swat.tamu.edu/software/plus/.
- The *gwflow* module is available for download at https://swat.tamu.edu/software/plus/gwflow/.
- The GFPLAIN algorithm is available online at https://github.com/fnardi/GFPLAIN.

***gwflow* module Procedure:**

After each daily time step of the SWAT+ simulation, the *gwflow* module subroutine calculates storage in every grid cell. The module uses the following water balance equation (A.1), where $n+1$ is the current day, n is the previous day, i is the cell row index, j is the cell column index, ΔV is equal to $V_{i,j}^{n+1} - V_{i,j}^{n}$, and Δt is equal to $t^{n+1} - t^{n}$. Rearranging these terms results in equation (A.2).

$$\frac{\Delta V_{i,j}^{n+1}}{\Delta t} = sources_{i,j}^{n} - sinks_{i,j}^{n} \pm lateral\ flows_{i,j}^{n} \quad \text{(A.1)}$$

$$V_{i,j}^{n+1} = V_{i,j}^{n} + \left(sources_{i,j}^{n} - sinks_{i,j}^{n} \pm lateral\ flows_{i,j}^{n}\right) \cdot (t^{n+1} - t^{n}) \quad \text{(A.2)}$$

The source, sinks, and lateral flows are volumetric fluxes (m^3/day) and are further clarified in equations (A.3, A.4, and A.5). Table 1 defines the terms in each equation. Equations (A.3) and (A.4) represent fluxes into and out of the aquifer, respectively. The surface water-groundwater exchange term in equation (A.4) for floodplains, $Q_{gw \to fp}$, was a novel addition for this study. Equation (A.5) represents groundwater fluxes into/out of the four sides of each grid cell. These lateral fluxes allow for the spatial connectivity of the groundwater aquifer system.

$$sources = Q_{rech} + Q_{sw \to gw} + Q_{lake \to gw} + Q_{fp \to gw} \quad \text{(A.3)}$$

$$sinks = Q_{gwet} + Q_{\mathrm{gw \to sw}} + Q_{\mathrm{gw \to soil}} + Q_{gw \to lake} + Q_{satex} + Q_{pump} + Q_{tile} + Q_{gw \to fp} \quad \text{(A.4)}$$

$$lateral\ flows = \pm Q_{north} \pm Q_{south} \pm Q_{west} \pm Q_{east} \quad \text{(A.5)}$$

The *gwflow* module uses equation (A.6) to calculate groundwater volume in each cell using cell dimensions and saturated thickness, s, which is the distance from the bedrock to the water table in meters. Since Δs is equivalent to the change in groundwater head, Δh, groundwater head is determined from equation (A.7).

$$V_{i,j}^{n+1} = \left(\Delta x \cdot \Delta y \cdot s_{i,j}^{n+1}\right) \cdot S_{y\,i,j} \tag{A.6}$$

$$h_{i,j}^{n+1} = h_{i,j}^{n} + \left(sources_{i,j}^{n} - sinks_{i,j}^{n} \pm lateral\ flows_{i,j}^{n}\right) \cdot \left(\frac{\Delta t}{S_{y\,i,j} \cdot \Delta x \cdot \Delta y}\right) \tag{A.7}$$

The Python script developed by Bailey et al. (2023) to automate the creation of *gwflow* module input files uses publicly available datasets to set up the *gwflow* grid and populate each cell with relevant properties. Each grid cell contains values for aquifer thickness, ground surface elevation, aquifer properties (hydraulic conductivity, specific yield), initial groundwater head, a flag for boundary cells, a flag for the presence of tile drains, a flag for the presence of lakes/reservoirs, a flag for the presence of rivers, values for HRUs (cultivated fields) that intersect, and values for USGS wells that intersect. I used this script to create most of the input files needed to run SWAT+ with the *gwflow* module in this study.

APPENDIX B: ADDITIONAL ANALYSES

Groundwater Wells:

USGS groundwater wells are located within the study watershed, but I did not use them as calibration targets due to limited data and time constraints. However, this section analyzes model fit at several wells. I excluded results at the 1000 m grid cell size due to errors in model outputs and selected the three USGS wells with the most observed data from 2003-2015 for comparison. Figure B.1 shows the well locations. Figure B.2 compares simulated aquifer head values for the three wells. I assessed model performance using mean absolute percent error (MAPE) to account for the relatively high simulated aquifer head due to the watershed elevation. MAPE was calculated using equation (B.1) and values for each well and cell size are shown in Table B.1.

$$\text{MAPE} = \frac{100\%}{N} \cdot \sum_{i=1}^{N} \left| \frac{obs - sim}{obs} \right| \tag{B.1}$$

Table B.1: MAPE for groundwater wells in the study watershed.

		USGS 395929105510300	USGS 395648105485300	USGS 395605105473100
250 m	Floodplain	9.506	6.472	1.404
	Control	9.485	6.471	1.370
500 m	Floodplain	0.121	0.727	0.173
	Control	0.117	0.506	0.184

There was little variation of groundwater head at each well location between the floodplain and control scenarios. Models with 500 m grid cells performed much better than the equivalent models with 250 m grid cells. The 500 m resolution had a close fit to the observed data without any calibration efforts.

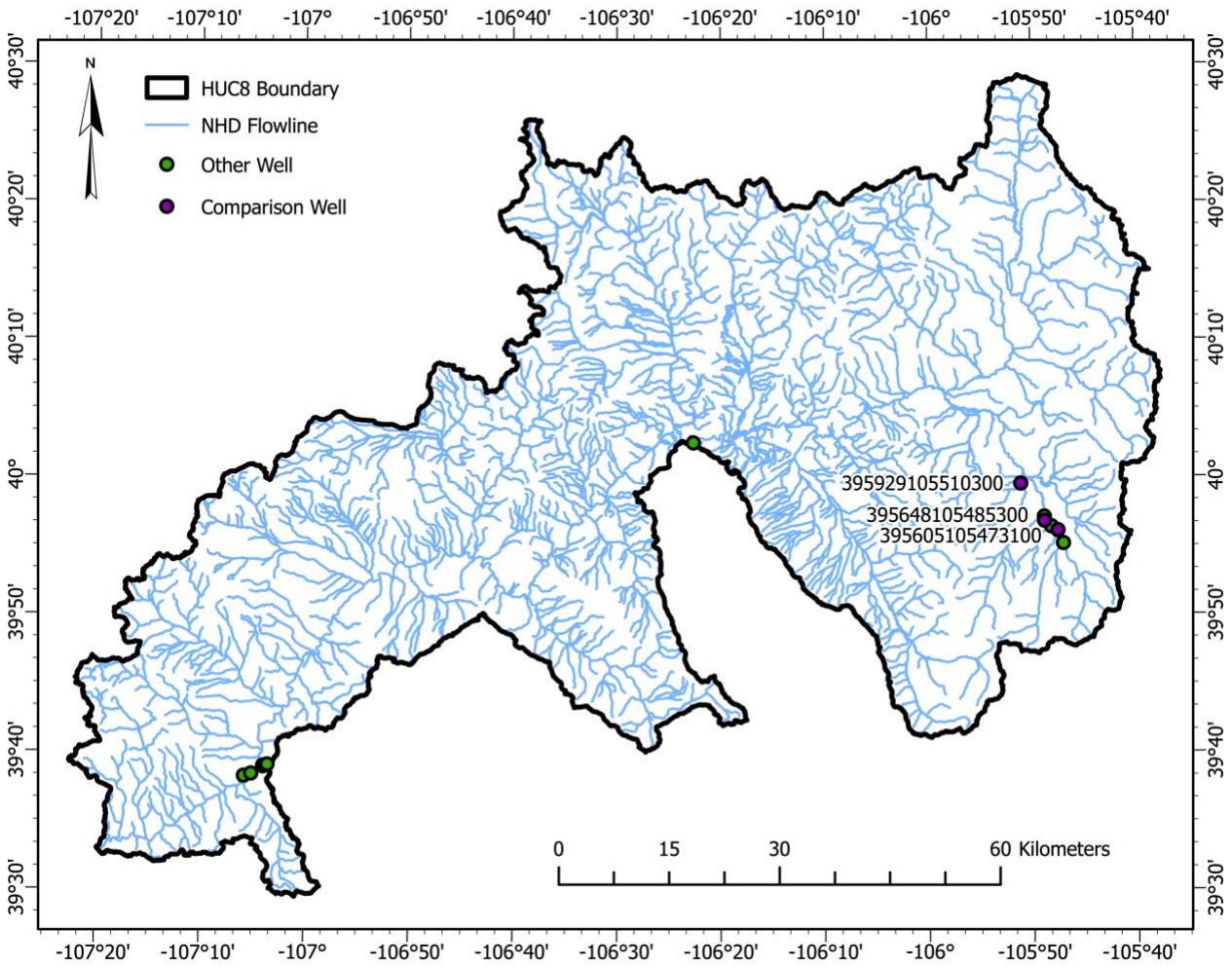

Figure B.1: Groundwater well locations in the Colorado Headwaters watershed.

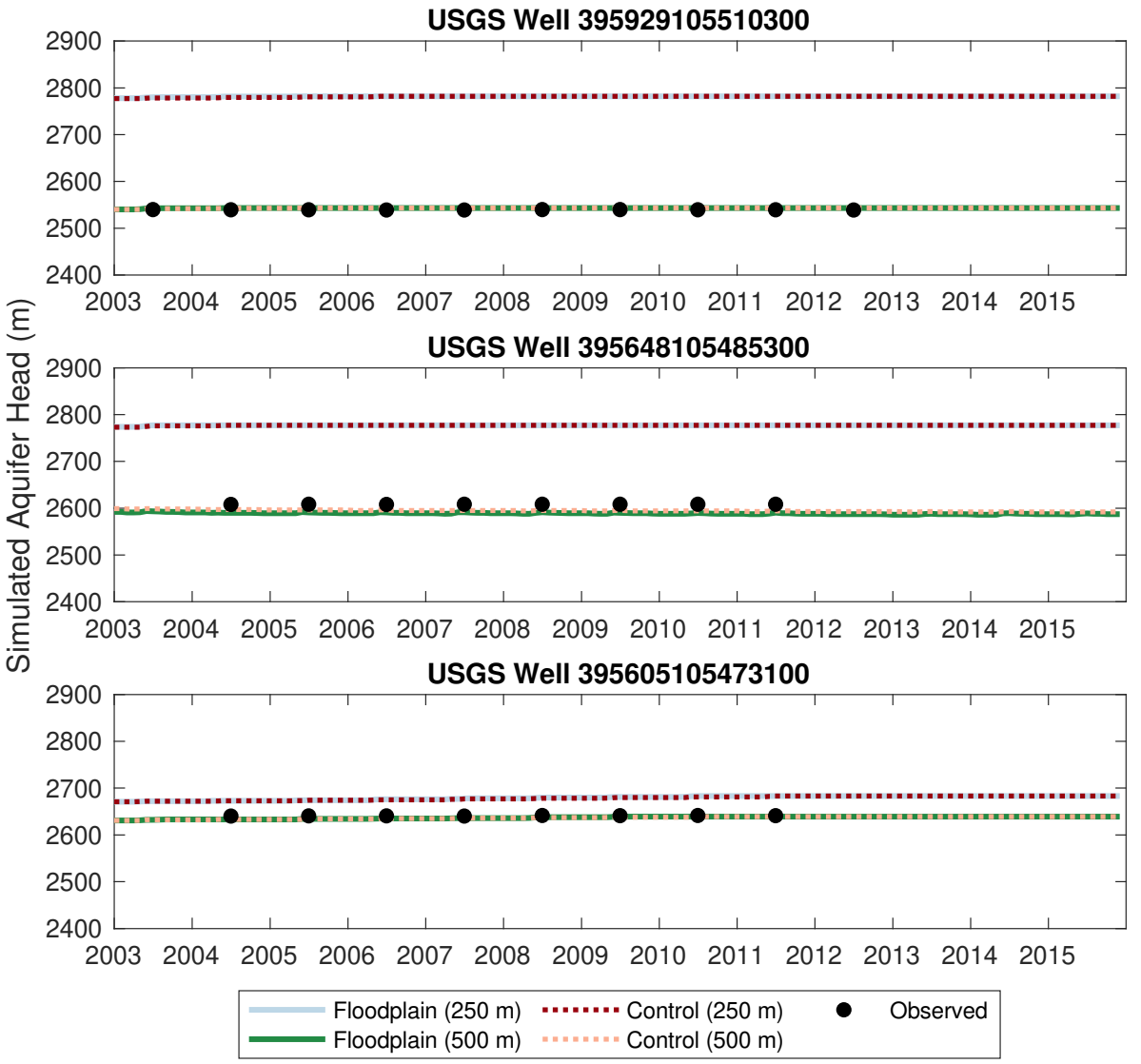

Figure B.2: Simulated aquifer head at three USGS wells in the study watershed.

Bead and String Analysis:

I analyzed the number of days per year that beads and strings were inundated in SWAT+. Figure B.3 shows boxplots of annual inundation counts over 2003-2015 for beads and strings at each cell size. Table B.2 shows summary statistics for days inundated per year. The median and mean number of days that cells were inundated were higher for string regions than for bead regions at every cell resolution.

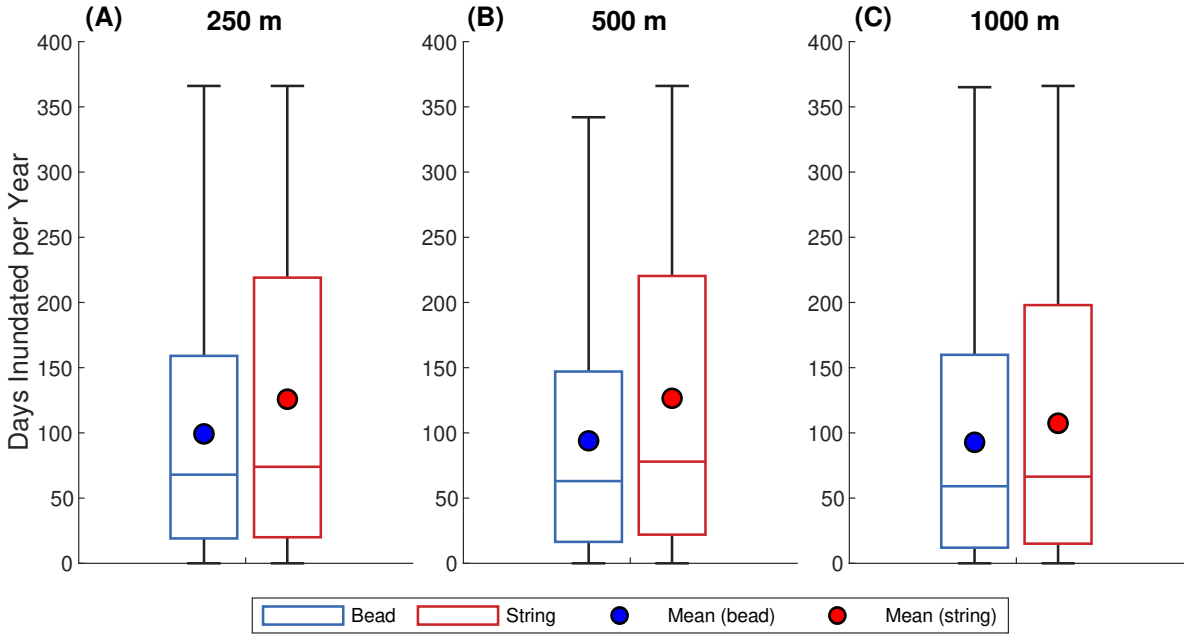

Figure B.3: Annual floodplain cell inundation count by beads and strings.

Table B.2: Summary statistics for bead and string inundation count.

		Mean	Median	Maximum	Minimum
250 m	Bead	99.23	68	366	0
	String	125.82	74	366	0
500 m	Bead	93.92	63	366	0
	String	126.50	78	366	0
1000 m	Bead	92.77	59	365	0
	String	107.40	67	366	0

To assess the magnitude of floodplain fluxes, I analyzed the absolute value of normalized fluxes in bead and string cells. Figure B.4 shows boxplots of these values and Table B.3 shows summary statistics. Mean flux values for bead cells were higher than for strings at every cell size.

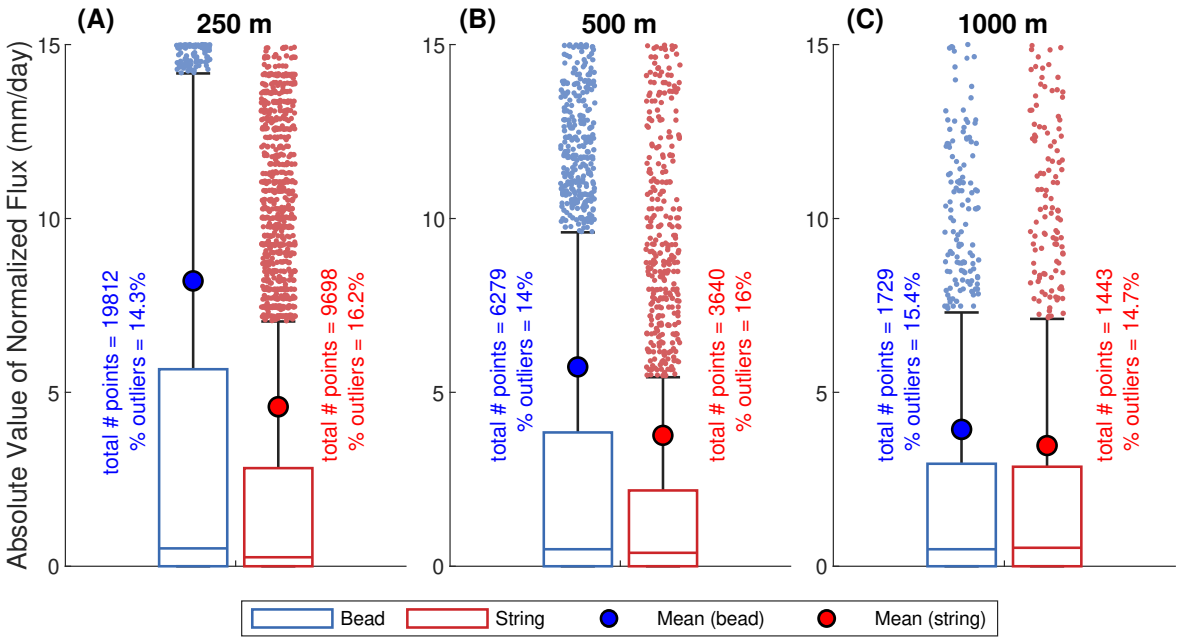

Figure B.4: Average annual absolute value of normalized floodplain fluxes by beads and strings.

Table B.3: Summary statistics for the absolute value of bead and string annual average fluxes.

		Mean (mm/day)	Median (mm/day)	Maximum (mm/day)	Minimum (mm/day)
250 m	Bead	8.204	0.514	387.200	0.000
	String	4.586	0.260	121.906	0.000
500 m	Bead	5.731	0.488	355.542	0.000
	String	3.762	0.387	93.963	0.000
1000 m	Bead	3.934	0.490	59.822	0.000
	String	3.470	0.530	94.023	0.000

As discussed in Section 3.4, the volume of water exchanged through bead and string cells was normalized by the relevant bead or string area in each year. Figure B.5 shows normalized volume exchanged per year for each cell size.

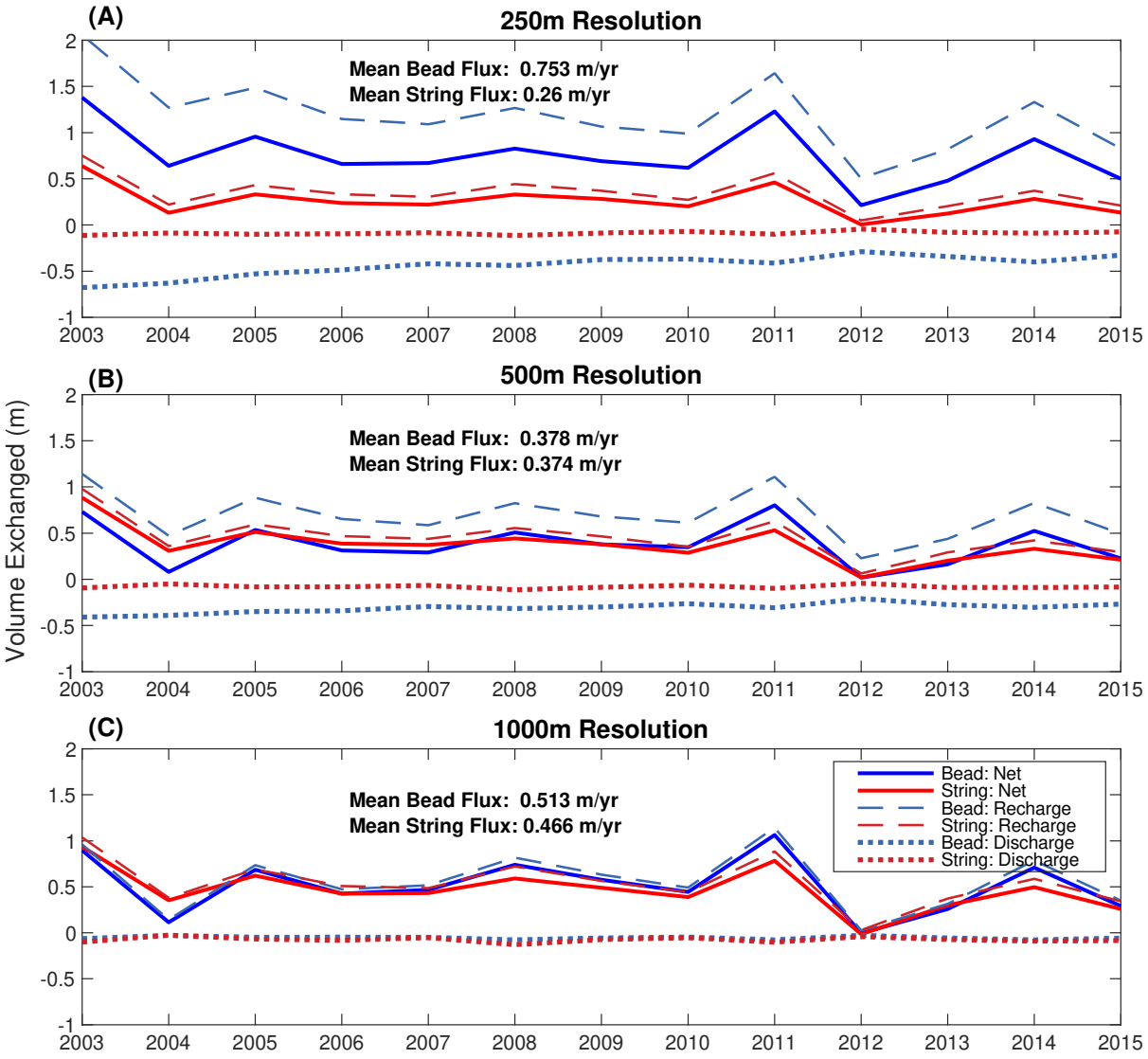

Figure B.5: Total yearly volume exchanged in beads or strings, normalized by relevant area.

The volume of water exchanged through bead and string cells varied with cell resolution. Plots of total volume (Figure B.6) and normalized volume (Figure B.7) exchanged through bead and string cells for every cell resolution are shown below.

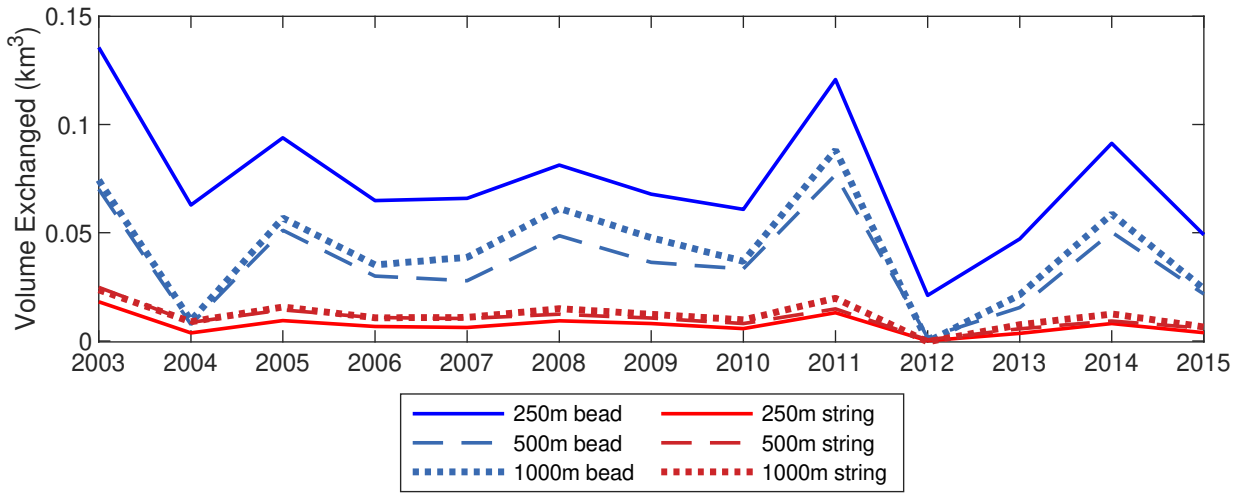

Figure B.6: Total volume exchanged in beads or strings for each cell size.

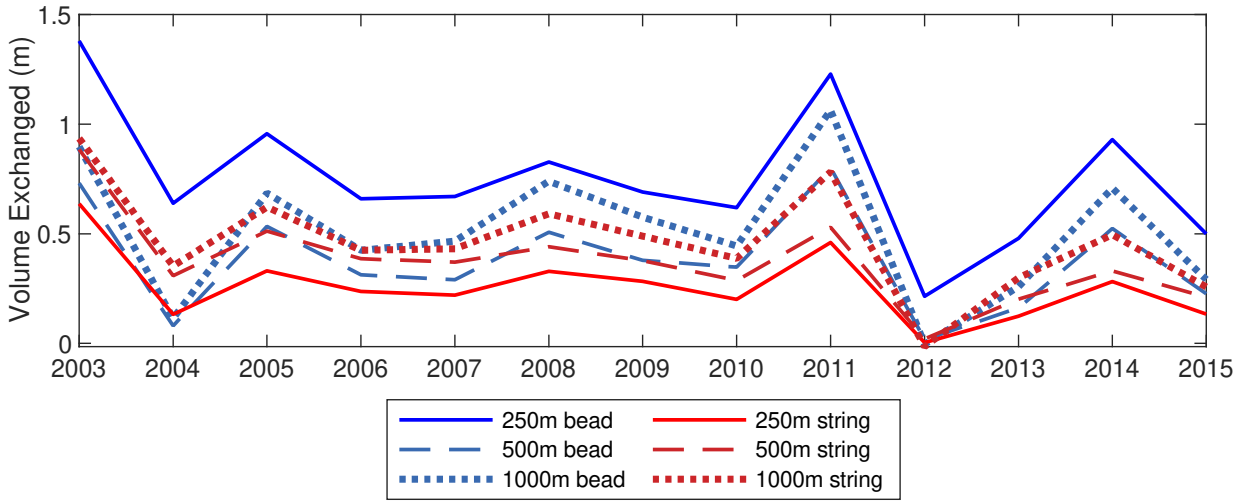

Figure B.7: Total volume exchanged in beads or strings for each cell size, normalized by relevant area.

APPENDIX C: BEAD/STRING DELINEATION

This section describes the methodology used to determine floodplain beads/strings. Identification of floodplain beads and strings was performed in ArcGIS Pro. This process was done for the GFPLAIN shapefile of floodplain area within the Colorado Headwaters watershed (HUC8 14010001).

1. Use the "Buffer" tool to expand the floodplain shapefile by 50 m. Since the GFPLAIN algorithm returns a gridded dataset, the buffering process is necessary to reduce errors in centerline creation over narrow regions of the floodplain (Figure C.2).
2. Use the "Polygon to Centerline" tool next to create a centerline through the buffered floodplain shapefile (Figure C.3). These centerlines will be used as the baseline for creating transects later in this process, which determine floodplain width.
3. Use the "Split Line at Vertices" tool to break up the floodplain centerlines. Then, use the "Spatial Join" tool to add a stream order value to each segment of the centerlines by merging segments based on a distance (500 m) and taking the maximum stream order from the NHDPlus flowlines. The split line tool was used to ensure the centerlines took values from the closest NHDPlus flowline to each small segment. Finally, merge the centerlines back together using the "Dissolve" tool (Figure C.4).
4. Manually adjust the centerlines. The main purpose of this step is to verify that the automatically created centerlines represent the center of the floodplain area. It is also important to reduce the number of centerline intersections, as this can cause overlap issues with transects later in the procedure. Verification of centerline stream orders compared to NHDPlus flowlines should also be performed (Figure C.5).
5. Merge the centerlines again using the "Dissolve" tool. Then create transects along the merged centerlines using the "Generate Transects" tool using a 250 m spacing, which is the same resolution as the smallest cell size used in this analysis. Generate transects separately for each

stream order. The width of transects generated for each stream order in this study are summarized in Table C.1. This may vary based on the study watershed, in this case, widths were determined based on a preliminary assessment of maximum floodplain width for each stream order (Figure C.6).

6. Clip transects to the floodplain area using the "Clip by Layer" tool and the GFPLAIN floodplain shapefile as the clipping layer (Figure C.7). To account for multipart objects, use the "Multipart to Singlepart" tool. Some transects extend into floodplain areas they do not originate from, so use the "Select by Location" tool to weed out transects that do not intersect a centerline of the same stream order. Some transects may not extend to the floodplain boundary, so perform a visual assessment of transects and manually extend those that do not reach the boundary (Figure C.8).

7. For the bead analysis, a floodplain width of 5 times stream width was assumed to be the threshold value for beads. Median stream width was assumed from the dataset provided by Downing et al., (2012). Table C.1 summarizes the values of median stream width and threshold bead width for each stream order.

8. The bead identification procedure was performed in ArcGIS Pro, using the code in Figure C.9. A flag of "1" was used for transects above the threshold width (beads), and a flag of "-1" was used for transects below the threshold width (strings). Figure C.10 shows the result of this analysis. Figure C.11 shows the transects limited to the stream orders analyzed in this study.

Table C.1: Summary of initial transect length, median stream width, and threshold bead width for each stream order.

Stream Order	Initial Transect Length (m)	Median Stream Width (Downing et al., 2012)	Threshold Width for Bead (m)
1	500	1.6	8.0
2	1000	1.9	9.5
3	1000	5.5	27.5
4	2000	11.0	55.0
5	2000	47.5	237.5
6	2000	99.0	495.0

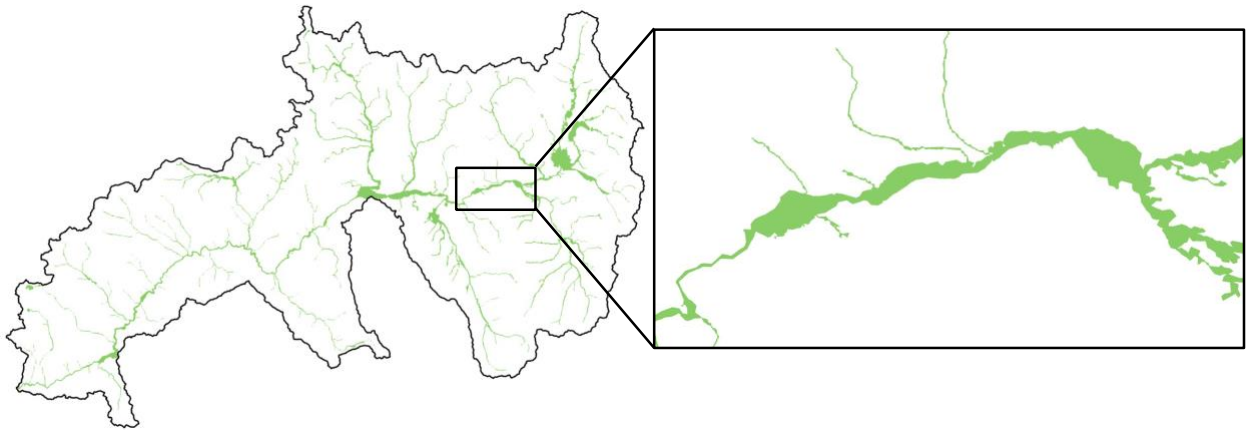

Figure C.1: GFPLAIN floodplain area in the study watershed with close-up of example area shown for this explanation.

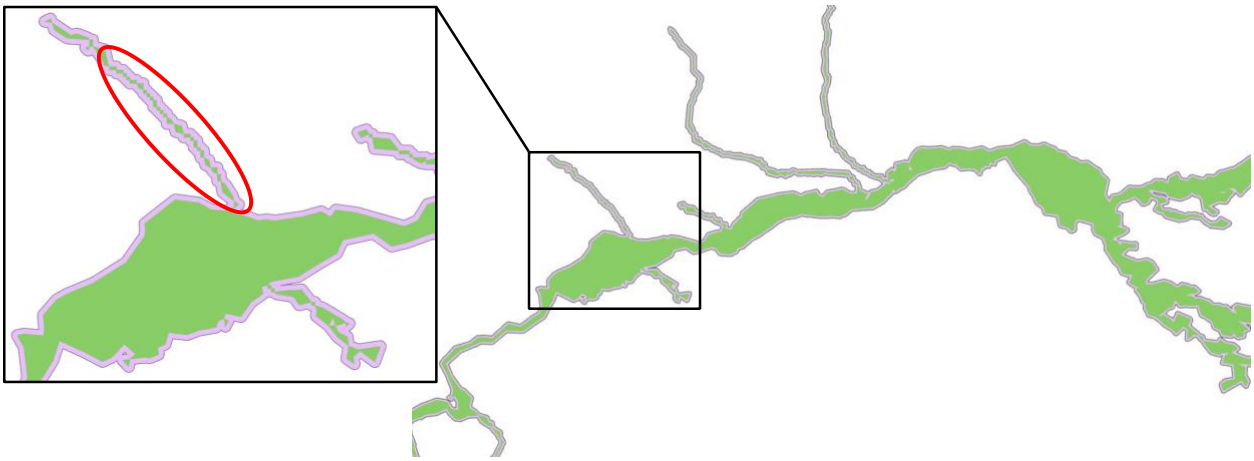

Figure C.2: Buffering of floodplain area (purple). The red circled area highlights a small GFPLAIN floodplain area that is most impacted by the buffering process.

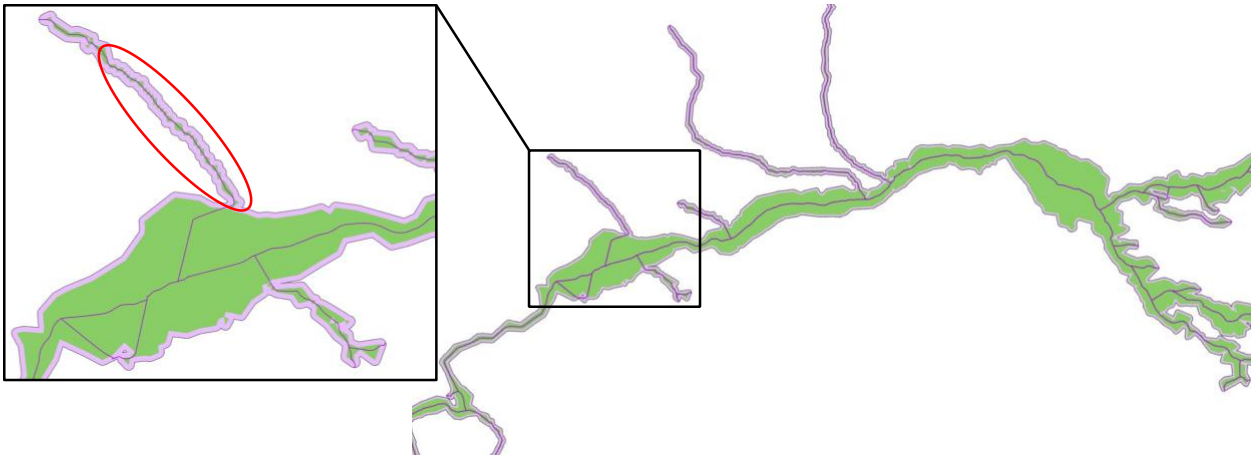

Figure C.3: Initial centerline creation from buffered floodplain. The red circled area indicates where the centerline creation is most impacted from buffering in the previous step.

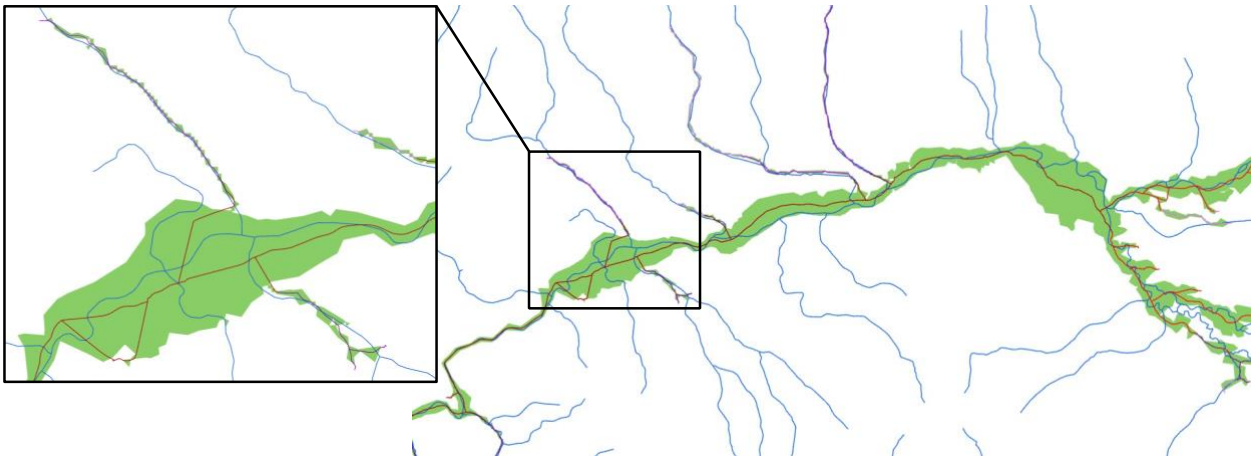

Figure C.4: Centerlines and NHDPlus flowlines. Each centerline segment is given the stream order value of the maximum NHDPlus stream order within 500 m of the segment.

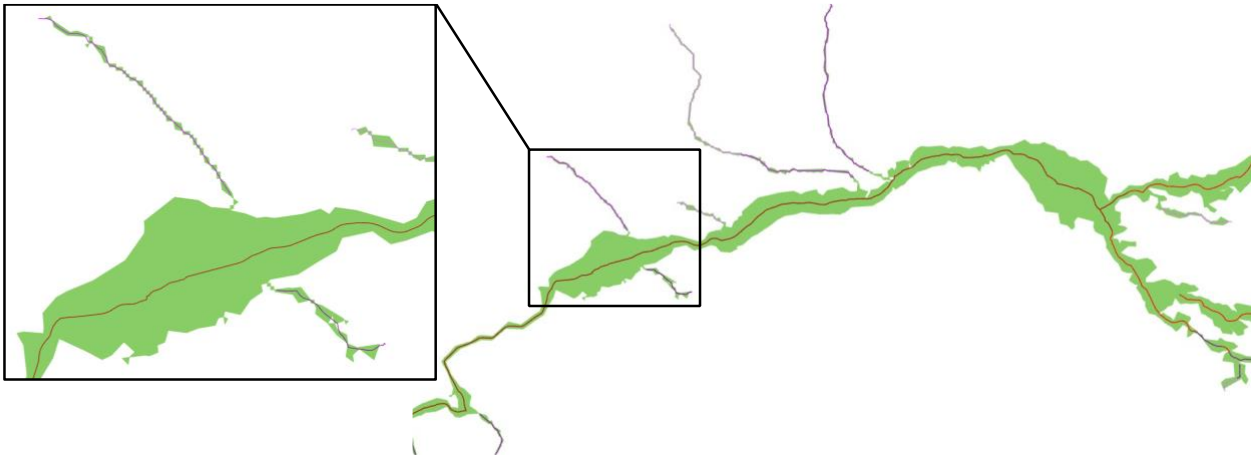

Figure C.5: Manually adjusted centerlines, where centerlines were adjusted to be more representative of the actual floodplain center, and some intersections were removed.

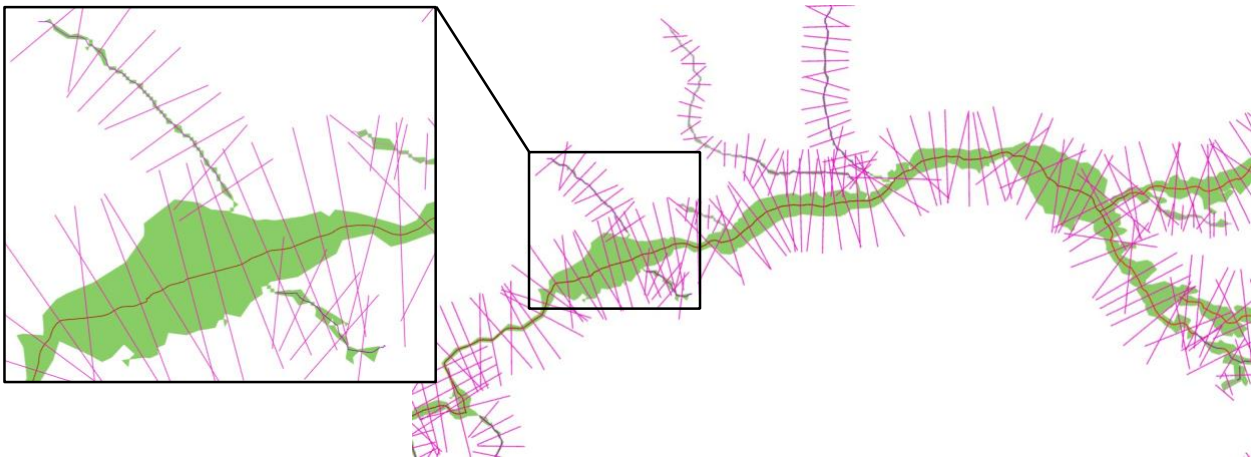

Figure C.6: Initial generation of transects with widths summarized in Table C.1.

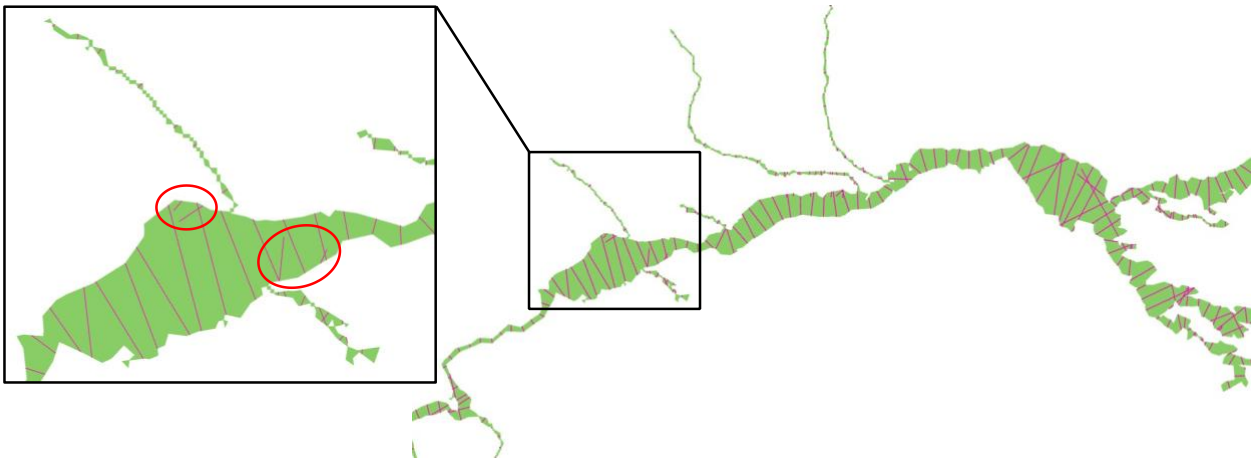

Figure C.7: Transects clipped to the floodplain boundary. The red circled areas indicate where transects extend into areas where they are not representative of actual floodplain width.

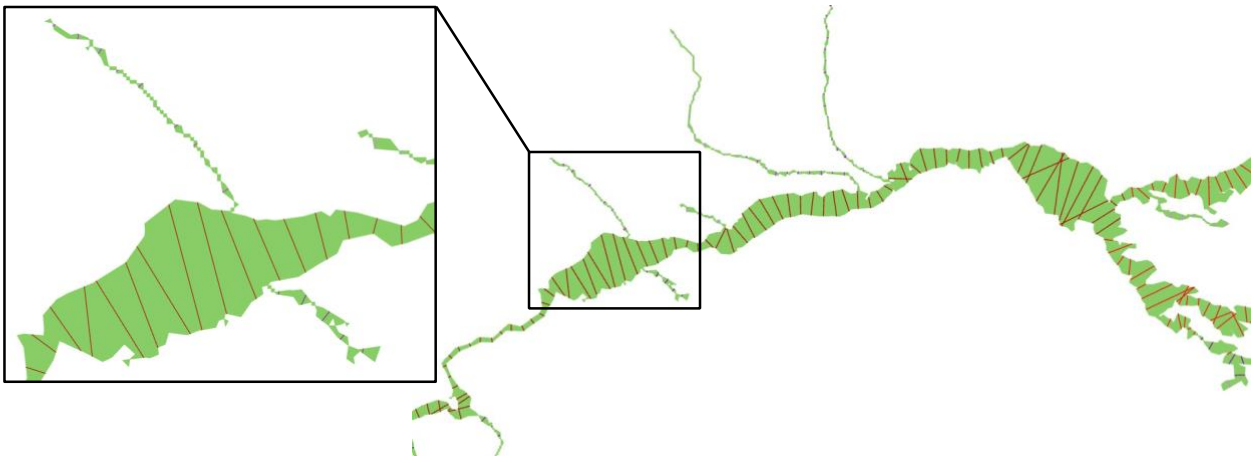

Figure C.8: Result of filtering transects by removing those that do not intersect a centerline of the same stream order.

```
beadCalc(!Shape_Length!,!StreamOrder!)
```

Code Block

```
def beadCalc(Shape_Length,StreamOrder):
    if (StreamOrder==1 and Shape_Length>8):
        return 1
    elif (StreamOrder==2 and Shape_Length>9.5):
        return 1
    elif (StreamOrder==3 and Shape_Length>27.5):
        return 1
    elif (StreamOrder==4 and Shape_Length>55):
        return 1
    elif (StreamOrder==5 and Shape_Length>237.5):
        return 1
    elif (StreamOrder==6 and Shape_Length>495):
        return 1
    else:
        return -1
```

Figure C.9: Code used to insert a flag indicating bead or string into the transect shapefile.

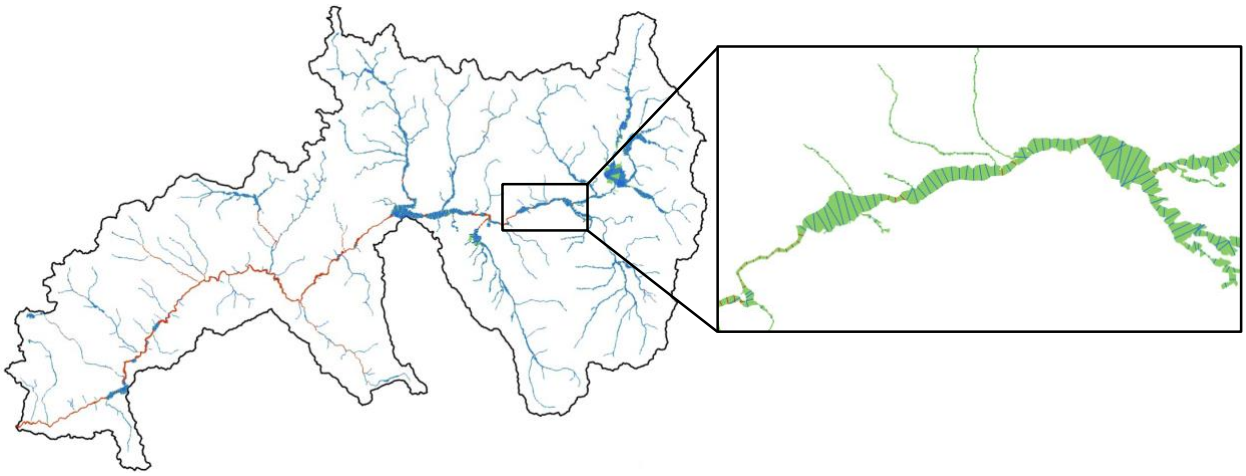

Figure C.10: Result of bead identification process, where blue transects show where lengths are over the bead threshold, and red transects show where lengths are under the threshold.

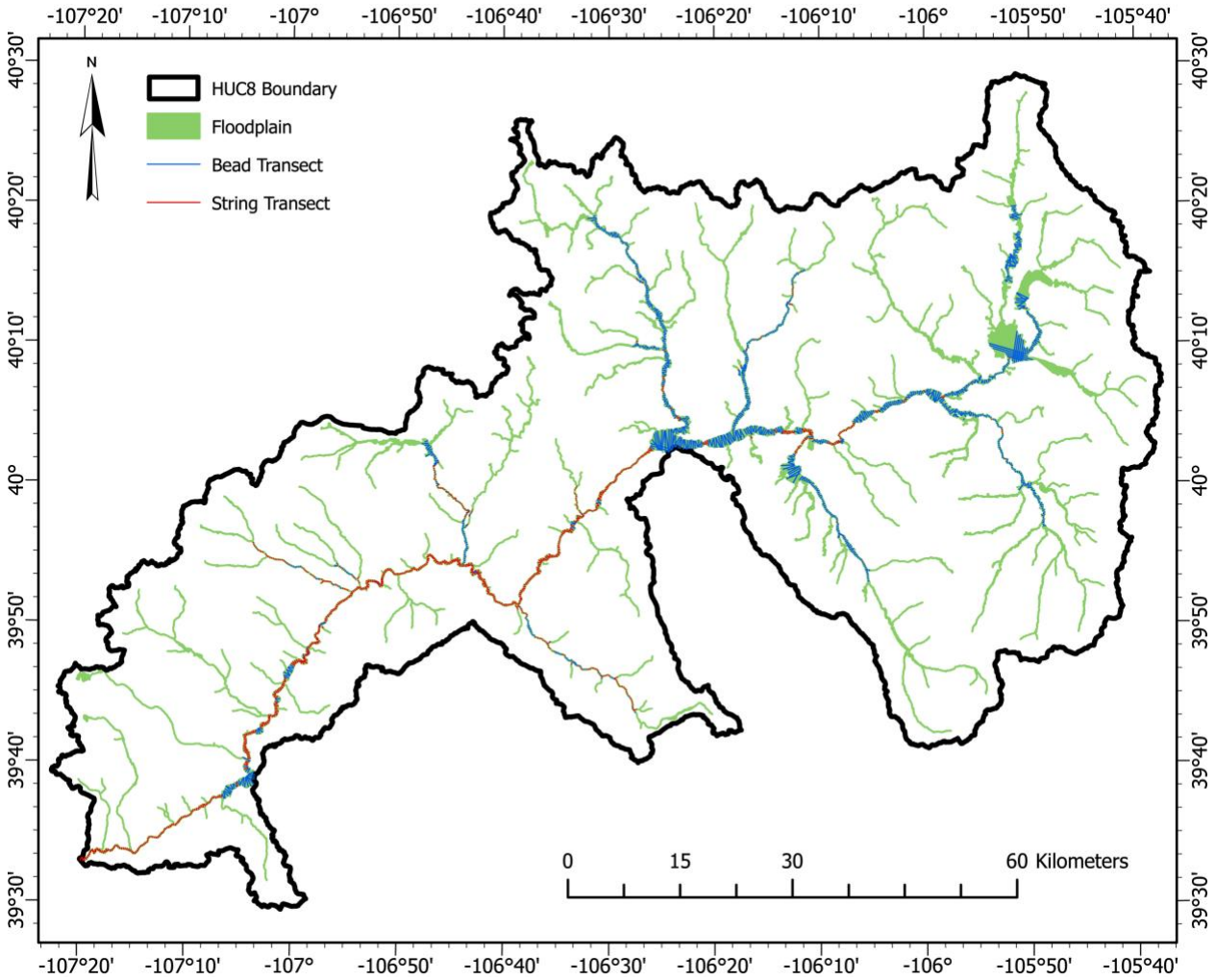

Figure C.11: Transects identifying beads and strings in the study watershed for stream orders 4-6.

APPENDIX D: POST-PROCESSING SCRIPTS

Python post-processing script to convert SWAT+ outputs to ArcGIS shapefiles:

```
import arcpy
import re
import os
import shutil
import subprocess
from arcpy import env
from arcpy.sa import *

## This script was developed to post process SWAT text file outputs into ArcGIS shapefiles.
## It was used for a case where several gwflow cell sizes and two scenarios (floodplain exchange included/excluded) were tested.
## The study years were 2000-2015 and the code specifically finds files for 2000 onwards.
## Relevant adjustments to the code should be made for alternate situations.
##
##
## Place this script in a folder (e.g. "processing_fplain", "processing_cntrl").
## The folder must also include "cell_sizes.txt" and "output_files.txt" files outlining which sizes/files to process.
##
## Make note of the file paths. These should be updated to local files. There are two file paths to update.
## Noted locations of file paths with "## ----------..."
## Also update flag for "fplain" or "cntrl" cases.
##
## The naming convention of SWAT folders here was "HUC8number_scenario_cellSize" (scenario = "fplain" or "cntrl").
## SWAT folders must follow this naming convention for the code to work properly.
## Folder created from "gwflow_create_NAM.py" must also be in the format "gwflow_gisCellSize".
##
## This process creates folders "gwflow_output..." for text files and "gwflow_analysis..." for shapefiles for each cell size.
## The two scenarios (floodplain and control) should be processed in separate folders.


# flag for floodplain or original processing.
# UPDATE FOR RELEVANT CASE.
processing_type = "fplain"

# placeholder for case where multiple HUC8 watersheds may be analyzed
huc8_target = "14010001"

year_list =
["2000","2001","2002","2003","2004","2005","2006","2007","2008","2009","2010","2011","2012","2013","2014","2015"]
```

```
# open list of cell sizes to analyze
with open("cell_sizes.txt") as cell_sizes:
    # begin loop for each cell size in the list
    for size in cell_sizes:

        # update the following path to the location of your SWAT folders
        ## ------------------------------------------------------------------------------------
        SWAT_path = "C:\Users\eyschulz\Documents\SWAT_runs\SWAT_calibrated\%s_%s_%s_cmpr1"
% (huc8_target, processing_type, size.strip())
        ## ------------------------------------------------------------------------------------

        print("start loop for processing size: %s" % size.strip())
        # open list of files to process
        with open("output_files.txt") as output_files:
            # begin loop for each output file in the list
            for file_name in output_files:
                # reset variables
                years = []
                rows = []
                old_year = " "
                # open output text file
                with open(SWAT_path + "\\" + file_name.strip()) as f:
                    # create a folder for yearly text files
                    folder_name = file_name.strip()
                    current_path = os.getcwd()
                    print("current directory: %s" % current_path)
                    newpath = "%s\gwflow_output%s\%s" % (current_path, size.strip(), folder_name)
                    print("creating folder for file %s" % folder_name)
                    if not os.path.exists(newpath):
                        os.makedirs(newpath)
                    # start loop to read annual flux values
                    for line in f: # read each line
                        if "Frequency" in line: # case for fp cell count
                            continue
                        elif "# of days" in line: # case for fp cell count
                            continue
                        elif "Annual" in line: # skip first line
                            continue
                        elif "Initial" in line: # case for gwflow_state_head
                            old_year = line # header
                            continue
                        elif ":       20" in line: # line that labels the year of study
                            if any(rows): # create new list if a year has already been recorded
                                years.append({
                                    "year":old_year,
                                    "rows":rows
                                })
                            rows = []
                            old_year = line # header
                            continue
                        elif line.strip() in year_list: # line that labels the year of study, case for fp cell count
```

```
                if any(rows): # create new list if a year has already been recorded
                    years.append({
                        "year":old_year,
                        "rows":rows
                    })
                rows = []
                old_year = line # header
                continue
            elif "Total" in line: # case for listing total fluxes
                break
            elif line == " ": # skip lines that separate years
                continue
            columns = re.split("\s+",line.strip())
            rows.extend(columns)
        # include the last year of the file
        years.append({
            "year":old_year,
            "rows":rows
            })
        # create folder labels for organized output files
        output_name = folder_name.split("gwflow_",1)[1]
        year_counter = 0
        # write output files into folders with text files for each year
        for year in years:
            print_name = year.get("year")
            if "Initial" in print_name:
                print("saving %s" % print_name.strip())
                file_name = "%s_initial.txt" % output_name
                save_name = os.path.join(newpath, file_name)
                with open(save_name,"w") as f:
                    f.write(year.get("year") + "\n")
                    counter = 1
                    for row in year.get("rows")[0:]:
                        f.write("%d\t%s\n" % (counter,row))
                        counter += 1
                continue
            print("saving %s" % print_name.strip())
            file_name = "%s_20%02d.txt" % (output_name,year_counter)
            save_name = os.path.join(newpath, file_name)
            with open(save_name,"w") as f:
                f.write(year.get("year") + "\n")
                counter = 1
                for row in year.get("rows")[0:]:
                    f.write("%d\t%s\n" % (counter,row))
                    counter += 1
            year_counter += 1
    print("end loop for processing size: %s" % size)

# convert processed flux files to ArcGIS shapefiles
arcpy.env.overwriteOutput=True
# open list of cell sizes to analyze
```

```
with open("cell_sizes.txt") as f:
    for cell_size in f:
        size = cell_size.strip()
        print("\n\nbegin ArcGIS routine for cell size: %s" % size)

        # Update the following file path to folder that contains grids from "gwflow_create_NAM.py"
        ## ----------------------------------------------------------------------------------------
        grid_folder = "C:\Users\eyschulz\Documents\SWAT_setup\gwflow_create_files\gwflow_gis%s\%s"
% (size, huc8_target)
        ## ----------------------------------------------------------------------------------------

        # folder with processed outputs from first loop in script
        output_folder = "%s\gwflow_output%s" % (current_path, size)
        print("saving .shp files to %s" % current_path)
        # the grid8 file to add columns
        grid_file = grid_folder + "\grid8_tile.shp"
        # create folder to store GIS files and set as the workspace environment
        analysis_folder = "gwflow_analysis%s" % size
        arcpy.CreateFolder_management(current_path, analysis_folder)
        local_folder = current_path + "\\" + analysis_folder
        env.workspace = local_folder
        # open list of outputs to analyze
        with open("output_files.txt") as outputs:
            for line in outputs:
                print("Creating shapefile for: %s" % line.strip())
                # name shapefile based on flux name and size
                item_name = "grid_%s_%s.shp" % (line.strip().split("gwflow_",1)[1], cell_size.strip())
                arcpy.management.CopyFeatures(grid_file,local_folder + "\\" + item_name)
                # get list of flux years to copy as columns
                DIR = output_folder + "\\" + line.strip()
                year_list = os.listdir(DIR)
                year_list.sort()
                # check to confirm that the list is in order of years (for gwhead, initial at end)
                print("confirm sorted year list: %s" % year_list)
                # copy fluxes in each year to shapefile
                year_counter = 0
                for year in year_list:
                    if year_counter == 16: # read initial gwhead state values, skips for fluxes
                        print("reading initial state values")
                        values = []
                        temp_str = "initial"
                        # create list initial
                        with open(DIR + "\\" + year) as flux_file:
                            input_fluxes = flux_file.readlines()[2:-1]
                            for flux in input_fluxes:
                                calc = flux.split("\t")[1]
                                values.append(float(calc.strip()))
                        # check max value and number of cells
                        print(temp_str + " max value: %s" % max(values))
                        print("confirm number of cells: %s" % len(values))
                        # add field to shapefile for year
```

```
                    arcpy.AddField_management(item_name, temp_str, "FLOAT", 15)
                    # insert cursor to copy list items into shapefile
                    cur = arcpy.UpdateCursor(item_name)
                    counter = 0
                    for row in cur:
                        row.setValue(temp_str, values[counter])
                        cur.updateRow(row)
                        counter += 1
                    del cur
                    break
                print("reading year 20%02d" % year_counter)
                values = []
                temp_str = "year_20%02d" % year_counter
                # create list for each year
                with open(DIR + "\\" + year) as flux_file:

                    if year == "fp_count_cell.txt": # don't cut out last line for fp cell count
                        input_fluxes = flux_file.readlines()[2:]
                        for flux in input_fluxes:
                            calc = flux.split("\t")[1]
                            values.append(float(calc.strip()))

                    input_fluxes = flux_file.readlines()[2:-1]
                    for flux in input_fluxes:
                        calc = flux.split("\t")[1]
                        values.append(float(calc.strip()))
                # check max value and number of cells
                print(temp_str + " max value: %s" % max(values))
                print("confirm number of cells: %s" % len(values))
                # add field to shapefile for year
                arcpy.AddField_management(item_name, temp_str, "FLOAT", 15)
                # insert cursor to copy list items into shapefile
                cur = arcpy.UpdateCursor(item_name)
                counter = 0
                for row in cur:
                    row.setValue(temp_str, values[counter])
                    cur.updateRow(row)
                    counter += 1
                del cur
                year_counter += 1
            # delete fields from original grid8_tile.shp file
            delete_fields =
["thick_m\n","gridcode\n","K_mday\n","Sy\n","zone\n","boundary\n","gridcode_1\n","gridcode_2\n"]
            for field in delete_fields:
                arcpy.DeleteField_management(item_name, field.strip())
    print("end ArcGIS routine for cell size: %s\n" % size)
```