12-2024

Quantitative Evaluation of Baseflow Separation Methods Using an Integrated Hydrologic Model: A Case Study in a Snow-Dominated Watershed

Jihad Othman
Utah State University

QUANTITATIVE EVALUATION OF BASEFLOW SEPARATION METHODS USING AN INTEGRATED HYDROLOGIC MODEL: A CASE STUDY IN A SNOW-DOMINATED WATERSHED

by

Jihad Othman

A plan B thesis submitted in partial fulfillment
of the requirement for the degree

of

MASTER OF SCIENCE

in

Civil and Environmental Engineering

Approved:

_____	_____
Pin Shuai, Ph.D.	Bethany Neilson, Ph.D.
Major Professor	Committee Member

_____
Wei Zhang, Ph.D.
Committee Member

UTAH STATE UNIVERSITY
Logan, Utah

2024

ABSTRACT

Quantitative Evaluation of Baseflow Separation Methods Using an Integrated Hydrologic Model: A Case Study in a Snow-Dominated Watershed

by

Jihad Othman, Master of Science

Utah State University, 2024

Major Professor: Dr. Pin Shuai
Department: Civil and Environmental Engineering

Baseflow, commonly referred to as the groundwater contribution to streamflow, constitutes approximately 50% of streamflow in mountainous regions of the Western United States. Accurately quantifying the amount of baseflow is critical for water management and decision-making, as it significantly impacts stream water quality, low flow availability, and ecosystem health. Traditionally, baseflow has been estimated using conceptual and automated baseflow separation methods, which are known to be both arbitrary and ambiguous, posing a challenge to validate them. In this study, we developed an integrated hydrologic model that seamlessly integrated the exchange between surface and subsurface flows to physically quantify the baseflow component in a snow dominated catchment—Coal Creek Watershed (CO, USA). The simulated baseflow and streamflow from the numerical model were then used as a controlled experiment to evaluate the performance of four commonly used baseflow separation methods, including the Pettyjohn and Henning (PH) graphical, the United Kingdom Institute of Hydrology (UKIH) graphical, the Eckhardt digital filter, and conductance mass balance (CMB) methods. Simulated baseflow has an average

baseflow index (BFI) of around 53% with a higher BFI in dry years versus that in wet years. In comparison to the numerical baseflow, both UKIH graphical and Eckhardt digital filter methods performed relatively well with high modified Kling-Gupta Efficiency (mKGE) (0.72 and 0.68, respectively) and Nash-Sutcliffe Efficiency (NSE) (0.58 and 0.7, respectively) values. However, UKIH graphical method performed poorer than the Eckhardt digital filter method in average and dry years when stream hydrographs resemble unimodal peaks, which are typical in snow-dominated catchments. Additionally, the Eckhardt digital filter showed better performance in matching the temporal dynamics of baseflow with a smoother hydrograph. Both the PH graphical and CMB methods did not perform satisfactorily in estimating baseflow with both mKGE and NSE values less than 0.3. Among them, the PH graphical method has consistently overestimated baseflow with an average BFI of 85%, whereas the CMB method has consistently underestimated baseflow with an average BFI of 24%. Overall, the Eckhardt digital filter method is promising for automated baseflow separation in snow-dominated catchments, though the relative importance of baseflow contribution to streamflow may be underestimated in dry years. Our findings suggest that integrated hydrologic models, when calibrated, provide a quantitative way to evaluate and improve existing baseflow separation methods. Caution should be exercised when applying automated baseflow separation methods in snow-dominated catchments, and future investigations are warranted to thoroughly evaluate these methods in catchments with diverse hydroclimate conditions.

ACKNOWLEDGMENTS

First and foremost, I would like to extend my warmest thanks to my supervisor, Dr. Pin Shuai, for supporting me through all the stages of my project. I am also deeply grateful to the committee members: Dr. Bethany Neilson, for her guidance throughout my master's program, including courses, research, and writing, and Dr. Wei Zhang, for taking the time to review my dissertation.

A special thanks to the Utah Water Research Lab (UWRL) for the grant that funded this research.

I am most indebted to my parents, sisters, and brothers for their unwavering patience, support, encouragement, and endless inspiration throughout my academic journey and the process of researching and writing this thesis. Finally, I am deeply grateful to my network of friends for their support and encouragement over the last two years.

Jihad Othman

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Baseflow, as defined by Hall (1968) is the portion of streamflow that comes from the groundwater discharge. Baseflow also represents relatively stable streamflow between storm events (Koskelo et al. 2012) and low flows during dry periods (Figure 1). Baseflow is generated from precipitation that infiltrates into the subsurface and later discharges into streams in the form of groundwater discharge. In this study, the terms baseflow and groundwater discharge are used interchangeably and refer to the same subsurface source without differentiation of shallow and deep groundwater.

Baseflow is critical to sustaining streamflow during dry periods and contributes to overall streamflow during periods of high flow (Hall, 1968; Koskelo et al., 2012). In the mountainous watersheds of the Western United States, it is estimated that over 50% of streamflow originates as baseflow (Rossman & Zlotnik, 2013). Any changes in baseflow availability will directly impact ecosystem health. Baseflow has different ion concentration, pH, and temperature than the stream water and plays a significant role in determining stream water quality (Reynolds et al., 1986), thus affecting the living organisms that form ecological diversity (Li and Yue, 2016). Additionally, as a component of the basinwide water balance (Famiglietti,2014; McNutt, 2014), baseflow is considered in water management practices, including irrigation, human consumption, and flood control (Lott & Stewart, 2013; Tan et al., 2020). Given the importance of baseflow, accurately estimating and separating baseflow from total streamflow is important (S. K. Singh et al., 2019). This process ensures a reliable assessment of water quality and quantity and supports water resource management in regions facing variable climatic conditions and increasing water demands (Eckhardt, 2008).

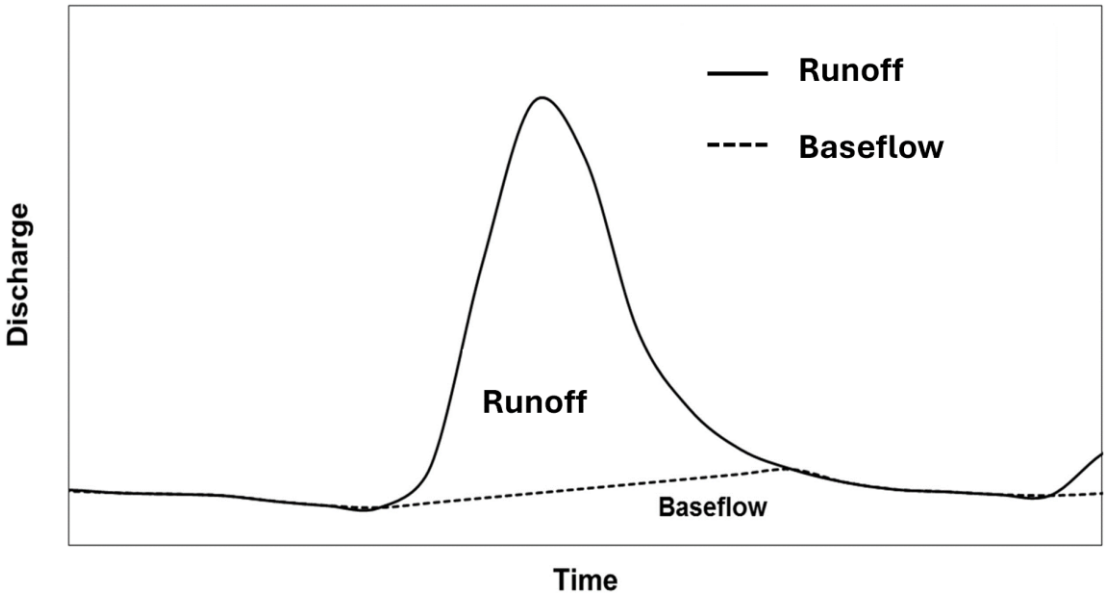


Figure 1: Baseflow (dashed line) shown as part of the streamflow (solid line) is illustrated using a straight-line graphical separation method. Note that during the low flow period, baseflow overlaps with streamflow. Adopted and modified from Bosch et al. (2017)

Numerous automated hydrograph-based and tracer-based methods have been developed and applied to effectively estimate and separate baseflow, enhancing our ability to manage and protect water resources efficiently (Table 1). The tracer-based method is used in conjunction with water balance calculations when tracer concentrations are known, assuming a defined relationship between the tracer concentration and the water source (Cook et al., 2008). In contrast, the hydrograph-based method does not use any tracer but employs conceptual algorithms to separate baseflow from streamflow (Eckhardt, 2008).

Table 1: A comparison of the baseflow separation methods used in this study

Baseflow separation methods	**Hydrograph-based**		**Tracer-based**	**Numerical**
	Graphical	**Digital Filter**	**Mass Balance**	
Methods used in this study	Pettyjohn & Henning (1979): Fixed interval Sliding	Eckhardt (2005)	Conductance Mass Balance (CMB) Miller et al. (2015)	Integrated Hydrological models Delottier et al. (2022) Partington et al. (2012)

	interval Local Minimum			
Description	Conceptual method based on hydrograph only	Conceptual method based on hydrograph only	Conceptual method based on hydrograph and specific conductance	Quantitative method based on model outputs
Software/Package (if available)	HYSEP	-	-	ATS (Advanced Terrestrial Simulator)
Required inputs	Streamflow Catchment area	Streamflow Catchment area Maximum baseflow index (BFI_{max}) Recession constant	Streamflow Continuous specific conductance measurements	Meteorological forcing data Land cover Surface and subsurface characteristics

There are two techniques derived from the hydrograph-based method: the Graphical and Digital Filter. The graphical method involves dividing the hydrograph into events of constant lengths of days and identifying the low flow point of each event using three different approaches, which are discussed in the methodology section. The digital filter method partitions continuously the hydrograph into two components, direct runoff and baseflow (Eckhardt, 2005).

One of the commonly used graphical methods is developed by Pettyjohn and Henning (1979), which uses the catchment area to determine the event length of days. There are three different algorithms included in the method, namely fixed interval, sliding interval, and local minimum (LocMin), to determine the selections of the lowflow points. The connecting line of these specific low points forms the baseflow. Before 1991, this process involved manually identifying and connecting the lowflow points (Sloto & Crouse, 1996). In

1991, White & Sloto developed HYSEP, a computer program that automates this process using various algorithms (Sloto & Crouse, 1996). The results from the White & Sloto graphical method using HYSEP closely matched those obtained from manual hydrograph separation at 307 streamflow measurement stations in Pennsylvania, USA. Since then, numerous methods have been developed, each using different algorithms for more refined baseflow separation. USGS has advanced this field by developing software such as the USGS GW Toolbox that include the HYSEP program (Barlow et al., 2022; Sloto & Crouse, 1996). Another commonly used graphical method is the smoothed minima method developed by the UK Institute of Hydrology (UKIH) in 1980. The UKIH method is considered easy to apply, as it determines baseflow by identifying and interpolating turning points derived from daily streamflow data (Gustard et al., 1992; Piggott et al., 2005; Zhang et al., 20017).

The digital filter method involves analyzing the signal in streamflow to separate the baseflow, which is assumed to be the low-frequency component of the streamflow (Guzman et al., 2015; Eckhardt, 2008; Xie et al., 2020). The Digital Filter method assumes that the outflow from the aquifer, which constitutes the baseflow, is proportional to the aquifer's storage (Eckhardt, 2005, 2008). There are multiple digital filter methods, each with a unique equation form, but they all assume that baseflow is the slow-responding, low-frequency component of streamflow, whereas quick flow is the fast-responding, high-frequency component of precipitation event response (Xie et al., 2020). The earliest of these is the LH method (Lyne and Hollick, 1979) which uses a single-parameter filtering algorithm, β, based on signal analysis and processing to separate baseflow (Yang et al., 2021). However, β has a low physical connection with the catchment characteristics (Xie et al., 2020). Chapman (1991) criticized the LH method for estimating unreasonable baseflow levels with some peak flow during low flow periods, where no peak flow should be present. Chapman updated the LH method's equation to keep the baseflow constant during non-peak flow periods, thus

preventing baseflow from decreasing during drought periods (Xie et al., 2020). Instead of using β, they introduced the filter parameter α (exponential recession) to consider the baseflow exponential recession (Eckhardt, 2008; Xie et al., 2020). Later, Chapman and Maxwell (1996) further refined the method to account for the previous day's baseflow in the calculation, maintaining a single-parameter filtering algorithm. Eckhardt (2005) derived a two-parameter recursive digital filter method with linear reservoir assumption. In addition to the α, Eckhardt introduced a new parameter, BFI_{max}, representing the maximum value of the baseflow index, defined as the long-term ratio of baseflow to total streamflow (Yang et al., 2021).

Several studies have compared these methods with baseflow estimated using other methods as controls. Gonzales et al. (2009) compared the digital filter method developed by Eckhardt (2005) to a dissolved-silica-based method with six other baseflow separation methods, concluding that the digital filter method was the most convenient. Xie et al. (2020) compared five digital methods, including all the above-mentioned, and four graphical ones in 1815 catchments across the CONUS, finding that the Eckhardt method was the best among them. On the other hand, Bhardwaj et al. (2024) applied the Eckhardt digital method in the Himalayan River basin, Northern India, along with other baseflow separation methods. Their evaluation using statistical and graphical indicators revealed that the Eckhardt digital method was not as effective, lacking versatility and reliability due to limitations in parameter availability.

Hydrograph and digital filter based conceptual methods are beneficial because they are easily automated, replicable, and require minimal data. However, despite their ease of use, these automated conceptual models for baseflow separation possess significant limitations: (1) They rely on assumptions that overlook surface, subsurface, and climate

characteristics specific to the basin, focusing mainly on discharge and basin area (Eckhardt, 2005; Huyck et al., 2005). For example, the basin area is the only input required for applying the Pettyjon and Henning (PH) graphical method. (2) Digital methods use simple assumptions that may introduce errors in baseflow estimation. For example, the recession constant is estimated by assuming a linear reservoir between groundwater storage and baseflow, which is found to be non-linear (Huyck et al., 2005). (3) The empirical parameters used in digital filter methods typically have a low physical connection to the basin characteristics (Xie et al., 2020). For example, the BFI_{max} used in Eckhardt (2005) digital filter method has only three suggested values (i.e., 0.25, 0.5 and 0.8) that correspond to basins that fall into three simple categories (i.e., perennial streams with porous aquifers, ephemeral streams with porous aquifers, and perennial streams with hard rocks). Therefore, the choice of BFI_{max} value often relies on a subjective interpretation of the hydrogeological condition of the study site (Eckhardt, 2008). Furthermore, the hydrograph-based method is often challenging to apply in snowmelt-dominated watersheds (Miller et al., 2015). This phenomenon is primarily attributed to the infrequency of quick flow events (e.g., rainfall-driven runoff events lasting less than five days) in snowmelt-dominated watersheds, where runoff is predominantly driven by prolonged snowmelt events (e.g., lasting more than a month). Snowmelt-driven hydrographs typically exhibit a unimodal peak discharge with gradual increases and decreases in flow rates and minimal variations in discharge outside the snowmelt runoff period (Julander and Clayton, 2018). Consequently, methods that utilize short intervals (e.g., 5-day blocks in the UKIH method) to identify minimum streamflow points for baseflow separation become less effective.

The tracer-based approach is a widely used method for streamflow separation that relies on mass balance equations, utilizing chemical concentrations in the stream, stream discharge, and the end member variables (Stewart et al., 2007). The end member is the

chemical concentrations of the source water. The two end-member components used to separate baseflow from the streamflow are runoff and baseflow end members (Caine, 1989; Miller et al., 2014, 2015). Different tracers including stable isotope tracers such as deuterium and oxygen, and major ions such as calcium, silica, and specific conductance were used to quantify the runoff and baseflow (Dinçer et al., 1970; Miller et al., 2014; Pinder & Jones, 1969). According to Gonzales et al. (2009), the accuracy of the tracer-based approach is affected by the type of tracer for a given catchment and its heterogeneity. For example, stable isotopes provide information about the residence time and ratio of contribution of water resources and, thus, are considered one of the most accurate chemical tracers (Kendall & Caldwell, 1998). However, the high cost associated with sample analysis is one of the main limitations of using the isotopic tracers. In contrast, specific conductance-based tracer method has become popular due to its relatively low cost and effectiveness in separating baseflow from total streamflow (Cassie et al., 1996).

The conductivity mass balance (CMB) method uses the difference in specific conductance (SC) between runoff and baseflow to separate the streamflow (Hayashi et al., 2004; Stewart et al., 2007). The CMB method is often used to separate baseflow because SC measurements are relatively easy and inexpensive to obtain and are readily available in many USGS gages (Miller et al., 2015). Additionally, the CMB method has been demonstrated successfully for baseflow estimation in snow-dominated watersheds in the Upper Colorado River Basin (Miller et al., 2014). To use this method effectively, continuous, high-frequency SC data is preferred with information about the SC of both runoff and baseflow end members (Lott & Stewart, 2013; Stewart et al., 2007). However, long-term, continuous, high-frequency (e.g., daily) SC data is often lacking in most watersheds (Miller et al., 2015). One way to estimate daily SC values is to analyze the relationship between stream discharge and SC, which generally shows an inverse relationship with a power function (Lott & Stewart, 2013;

Miller et al., 2014). Others, such as Miller et al. (2015), estimated the daily SC concentrations using a regression analysis approach. They used the discrete SC measurements in a regression model to estimate the continuous SC concentration for snow-dominated catchments. The regression-derived was described to provide an accurate baseflow estimation. Then, the estimated SC was used in the CMB method to estimate the baseflow and compare it with a graphical method developed by Wahl and Wahl (1988). They concluded that the CMB method showed better baseflow estimates than the graphical method at 12 snowmelt-dominated sites.

The CMB method can be challenging to apply in catchments with multiple sources of groundwater but is more applicable in small catchments with limited groundwater sources (Miller et al., 2014). Rumsey et al. (2015) used this approach with regression-based estimates to find daily SC values in the headwaters of the Upper Colorado River Basin. Their results showed that baseflow made up about half of the streamflow each year. However, the lack of high-frequency SC data can sometimes limit its broader use (Matsubayashi et al., 1993). Additionally, determining the end-members SC for large-scale catchments is more challenging due to high spatial variation, affecting the accuracy of this method (Sanford et al., 2011).

To understand the impact of these limitations on the baseflow estimation, (Mau & Winter, 1997), suggested validating baseflow estimations using independent methods that account for different forcing functions and hydrological conditions, such as numerical methods. These methods, which are physics-based, address many of these limitations by directly simulating baseflow from model outputs (Partington et al., 2012). However, many of those numerical models do not explicitly simulate the exchange between surface water and

groundwater and groundwater reservoir is often simplified, resulting in poor representation of baseflow (Delottier et al., 2022).

In surface hydrologic models, groundwater flow is primarily solved using simple groundwater equations such as Darcy's law (V. P. Singh, 2018). For example, Lee et al. (2014) discussed the Soil and Water Assessment Tool (SWAT), which contributes to baseflow estimation but does not directly simulate it. SWAT can simulate streamflow, surface runoff, aquifer percolation, and groundwater re-evaporation, and then these variables can be used by other models and methods, for example, conceptual methods or the integration of these hydrological processes, to estimate the baseflow (Lee et al., 2014). Consequently, this approach could not capture the dynamic nature of low flows inherent in baseflow (Partington et al., 2012).

In subsurface hydrologic models, also known as groundwater models, the conductance concept or other techniques are used to generate baseflow. However, neither surface nor subsurface models alone accurately assess the integrated exchange between surface and subsurface water, leading to incomplete estimations of baseflow dynamics (Staudinger et al., 2019). An example of a subsurface model is MODFLOW, which can simulate saturated groundwater and unsaturated flow using Darcy's law, though it often employs simplified surface flow boundaries (e.g., constant recharge) (Brunner et al., 2010).

According to Staudinger (2019), the development of surface-subsurface, called integrated hydrologic models including models such as HydroGeoSphere (HGS) (Therrien et al., 2006), MODHMS (Hydrogeologic Inc., 2006) and Parflow (Kollet & Maxwell, 2006), has filled this gap by quantitatively assessing the integrated exchange between surface and subsurface waters. These models typically solve a 2-D overland flow equation, a 3-D Richards groundwater governing equation, as well as land surface energy balance such as

evapotranspiration and snowmelt. They often simulate spatial and temporal dynamics of surface and subsurface flow and their exchanges at fine spatiotemporal resolutions (e.g., 10s meter and daily), enabling the quantification of exchange fluxes at m to 10s m scale and sub-daily frequency. Therefore, baseflow can be directly estimated across the watershed by integrating the groundwater discharge across the land surface throughout the year. Additionally, integrated hydrologic models are valuable for assessing simpler conceptual models as they eliminate the need to assume a functional relationship between baseflow and streamflow or rely on basic empirical relationships (Partington et al., 2012). Thus, until new observational techniques and conceptual hydrographical models are developed to address these gaps, numerical models remain the most reliable method for directly quantifying baseflow dynamics under varying hydrologic conditions for catchments with diverse characteristics (Partington et al., 2012).

Partington et al. (2012) applied the HGS, a physically based, integrated hydrologic model, to conduct spatial and temporal surface and subsurface simulations. However, HGS was not used alone to determine the baseflow contribution to streamflow in catchments because of the high surface and groundwater exchange, where both losing and gaining areas vary along the catchment (Partington et al., 2011). To address this, Partington (2011) developed a hydraulic mixing cell (HMC) module in combination with HGS to demonstrate spatiotemporal exchange, thus allowing for effective baseflow separation. HMC is a tracking method for all streamflow generation mechanisms, enabling the tracking of groundwater discharge directly to the stream while considering travel time and surface-subsurface exchange, both in and out. This method helps parse streamflow into runoff and baseflow (Partington et al., 2011). Because this combined method relies on quantitative analysis to theoretically examine baseflow dynamics, the resultant baseflow was used by Partington et al. (2012) to compare the performance of other conceptual methods for automated

baseflow separation. These methods included the graphical method using different packages such as HYSEP (Sloto & Crouse, 1996), PART (Rutledge, 1998) and BFLLOW (Arnold and Allen, 1999), and Eckhardt digital filter method (Eckhardt, 2005). Automated baseflow separation methods appear to be affected by the complexity of baseflow dynamics, as they could not perform satisfactorily, and no single separation method proved to be clearly superior to the others. It is worth noting that the numerical method used as the control followed the tilted V-catchment geometry employed by Panday & Huyakorn (2004) to represent the catchment. This assumption might have also impacted the control baseflow results.

Delottier et al., 2022 reviewed the applications of HGS to assess the applicability of integrated models in both small and large catchments. They found that while the model had been used to directly simulate baseflow in small catchments, it had not been applied to regional-scale catchments. Applying HGS to regional-scale catchments is computationally very expensive and time-consuming (Ledoux et al., 1984). To avoid long simulation times, Delottier used a surface water mass balance module to obtain spatial and temporal infiltration, which was then used as an input to the HGS. The model performance in simulating baseflow was acceptable by the authors and they suggested that comparing the obtained baseflow with conceptual models should be explored in future research.

Nonetheless, very few studies have simulated baseflow using integrated hydrological models, and even fewer have compared the baseflow results to those from automated conceptual methods. Therefore, there is a need for an assessment of commonly used baseflow separation methods, particularly in snow-dominated catchments.

This project aims to quantitatively evaluate commonly used baseflow separation methods to simulate baseflow using an integrated hydrologic model in a snow-dominated, mountainous headwater watershed. The hydrologic model is developed and calibrated

against field observations. Then, the simulated baseflow from the model is used as a numerical control experiment to validate the estimated baseflow from the graphical, digital filter, and conductance mass balance methods. This work provides physically based recommendations to establish which commonly used baseflow separation method is appropriate for snow-dominated, mountainous watersheds.

2 STUDY AREA

The Coal Creek Watershed, our study site, is located in West-Central Colorado within the larger East Taylor Watershed (Figure 2). Located near Crested Butte, Coal Creek falls under the Hydrologic Unit Code-12 (HUC12) 140200010204, which is part of the larger East Taylor Watershed (HUC 14020001). The watershed covers an area of 53.2 km² and is characterized by its mountainous terrain, with elevations ranging from approximately 2706 meters above sea level (m.a.s.l.) at the eastern boundary to 3770 m.a.s.l. at the western boundary. Described as a high alpine (Köppen and Geiger, 1930), snowmelt-dominated catchment, the area experiences a warm summer and humid continental climate. According to the Daymet dataset (Thornton et al., 2021), the watershed receives about 850 mm of precipitation annually, of which 530 mm falls as snow. This contributes to the total discharge measured daily by the USGS station number 09111250 at the watershed outlet. Vegetation in the watershed is diverse, comprising 62.6% evergreen forest and 20.5% shrub, along with other less dominant types.

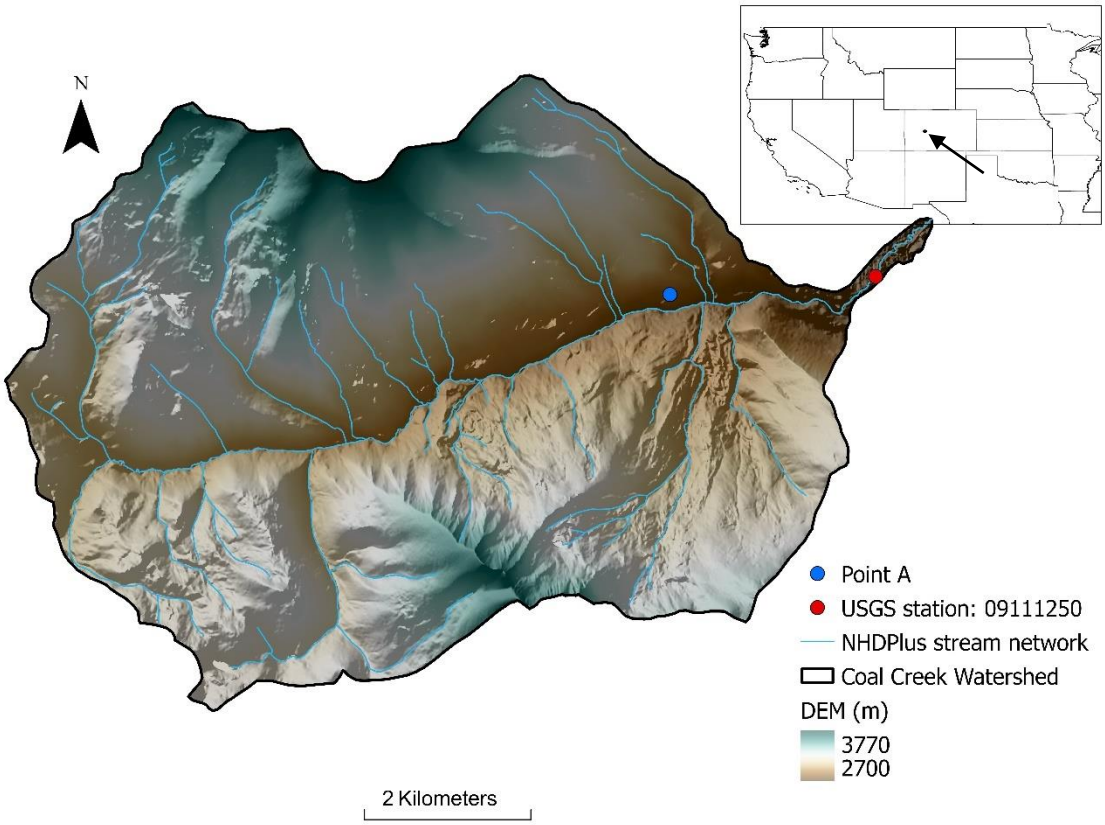


Figure 2: Coal Creek Watershed map in the context of the USA. The red point represents the USGS station 09111250, which measures the output flow of the watershed through the river network and is used in the study for numerical method validation. The river network and the Digital Elevation Model (DEM) were obtained from the National Hydrography Dataset Plus (NHDPlus).

3 METHODOLOGY

3.1 Numerical Method

3.1.1 ATS Model Setup

ATS (Advanced Terrestrial Simulator) is a fully distributed integrated model that simulates the flow of water in both surface and subsurface environments (Coon et al., 2019). It employs a two-dimensional diffusion wave approach, based on the Saint-Venant equation, to represent surface water movement, and a three-dimensional approach, using Richards equation, for subsurface flow. The Richards equation is described below:

$$\frac{\partial}{\partial t}(\emptyset s) + \nabla . q = 0, \qquad \text{Equation 1}$$

with

$$q = -\frac{1}{\mu} k_r \kappa (\nabla p + \rho g), \qquad \text{Equation 2}$$

Where ϕ is the effective porosity (-), s is the saturation (-), q is the Darcy flux (m s^{-1}), μ is the dynamic viscosity (Pa s^{-1}), k_r is the relative permeability (–), κ is the saturated hydraulic permeability (m^2), p is the water pressure (Pa), and g is the gravitational constant (m s^{-2}).

The diffusive wave approximation to overland flow is described as

$$\frac{\partial h}{\partial t} + \nabla \cdot (hv) = Q_w + Q_{ss}, \qquad \text{Equation 3}$$

With

$$v = -\frac{h^{\frac{2}{3}}}{n. \max(\in, \sqrt{\nabla z})} \nabla (z + h), \qquad \text{Equation 4}$$

where *h* is the depth of ponded water (m), v is the surface flow velocity (m s^{-1}), Q_w are all external source/sink terms (m s^{-1}), Q_{ss} is the exchange flux between surface and subsurface systems (m s^{-1}), n is Manning's coefficient (s $m^{-1/3}$), z is surface elevation (m), and ∈ is a small positive regularization to keep the equations non-singular in places with zero bed slope (m).

The model incorporates the Priestley–Taylor equation to account for evapotranspiration (ET) from sources such as snow sublimation and vegetation transpiration, integrating these factors into the surface-subsurface simulations. The original ATS model for Coal Creek was developed by Shuai et al., 2022 (For details regarding model setup, please refer to this paper). Horizontally, the model mesh consists of 171,769 grid cells with an area

ranging from 5,000 m^2 to 50,000 m^2. Vertically, there are 19 subsurface layers, representing a total depth of 28 m with varying thickness. The top layer is about 5 cm thick, and the bottom layer is 2 m thick. The model was spin up first then used in transient simulations driven by the meteorological forcing dataset.

For this study, we have extended the simulation period from 2019 to 2022. In total, the model covers eight years ranging from October 1, 2014, to October 1, 2022, with daily outputs. The first water year was excluded from the results to allow the model to stabilize. The model simulated variables such as groundwater discharge (assumed to be baseflow), streamflow, and soil saturation. The model has previously been calibrated against streamflow for Coal Creek from October 1, 2016, to October 1, 2019 (Jiang et al., 2023) .

The streamflow of the period of study is compared with the observation data and the model performance is evaluated using the modified Kling-Gupta Efficiency (KGE) and standard Nash-Sutcliffe Efficiency (NSE) (Kling et al., 2012). The modified KGE metric is represented in the equation:

$$KGE = 1 - \sqrt{(r-1)^2 + (\gamma-1)^2 + (\beta-1)^2} \qquad \text{Equation 5}$$

With

$$r = \frac{cov(S,O)}{\sigma_s \, \sigma_O} \qquad \text{Equation 6}$$

$$\gamma = \frac{\sigma_s/\mu_s}{\sigma_O/\mu_O} \qquad \text{Equation 7}$$

$$\beta = \frac{\mu_s}{\mu_o} \quad \text{Equation 8}$$

where S and O represent simulated and observed values, respectively, r is the correlation coefficient, γ is the variability ratio, β is the bias ratio, $cov(S, O)$ is the covariance between simulated and observed values, σ is the standard deviation, and μ is the mean.

The modified KGE is different from the original one by the ability to detect the generated bias due to the input effect on the variability indicator (Gupta et al., 2009; Kling et al., 2012). The KGE breaks down the model performance into three components: correlation (r), variability (γ), and bias (β). The correlation component evaluates the timing accuracy of streamflow predictions, reflecting temporal dynamics, while the variability and bias components assess the flow duration curve, indicating the magnitude of streamflow. The KGE score ranges from negative infinity, indicating the poorest model performance, to 1, signifying a perfect model where all three components reach unity. This representation of the KGE also applies to that of the NSE achieving perfect performance at the value of 1.

3.1.2 Calculating baseflow using the numerical model simulated variables

This section briefly explains the approach with the required assumptions to calculate baseflow. The hydrologic processes and variables that form the input and output of the surface water balance equations are precipitation, infiltration, runoff (both surface and subsurface), baseflow, and evapotranspiration (Dingman, 2002; Tarboton, 2003).

$$input = output \quad \text{Equation 9}$$

$$surface\ water\ net = rainfall + canopy\ drainage + snowmelt - evapotranspiration - runoff - infiltration$$ Equation 10

ATS model simulates the exchange fluxes across the river, which can be used to calculate the baseflow. Positive fluxes indicate exfiltration (water that leaves the subsurface and flows onto the river) whereas negative fluxes indicate infiltration (water enters the subsurface)

$$net\ exchange = total\ exfiltration - total\ infiltration$$ Equation 11

Integrating all positive fluxes across the surface can quantify total exfiltration (infiltration) at each timestep. Total infiltration consists of net infiltration and re-infiltration:

$$total\ infiltration = net\ infiltration + re\_infiltration$$ Equation 12

Net infiltration refers to the process by which water from precipitation or snowmelt penetrates the surface and enters the subsurface. Re-infiltration, on the other hand, describes the process where water that has already infiltrated at one location exfiltrates and then infiltrates again at a different location. For example, rainfall that infiltrates at an upstream location may exfiltrate into the stream at a downstream location, and then reinfiltrate into the subsurface (Figure 3). Similarly, total exfiltration is the sum of net exfiltration and re-exfiltration

$$total\ exfiltration = net\ exfiltration\ (e.g., baseflow) + re\_exfiltration$$ Equation 13

Net exfiltration (also known as baseflow) refers to groundwater that emerges onto the surface and contributes to streamflow. Re-exfiltration describes the process where water that has exfiltrated at one location infiltrates again and subsequently exfiltrates at a different location. Baseflow is calculated using the following equation:

$$baseflow = total\ exfiltration - re\_exfiltration \quad \text{Equation 14}$$

Re-exfiltration is assumed to be equivalent to re-infiltration for the daily calculations of baseflow. This assumption is considered reasonable under the premise that the infiltrated water is exfiltrated on the same day without any delay. Consequently, this timing assumption between infiltration and exfiltration is applied consistently across all baseflow calculations in this project.

$$re\_exfiltration = re\_infiltration \quad \text{Equation 15}$$

The subsurface water balance is shown in the following equations:

$$\Delta S = total\ infiltrration - transpiration - total\ exfiltration \quad \text{Equation 16}$$

Substituting values from Equation 12,Equation 13, and Equation 15 into Equation 16 to obtain:

$$\Delta S = net\ infiltrration - transpiration - baseflow \quad \text{Equation 17}$$

During dry periods, when there is no snowmelt and no rainfall, infiltration (e.g., net infiltration) and runoff are assumed to be zero. Using Equation 16 and the assumption of zero net infiltration during dry periods, the dry period baseflow is calculated as follows:

$$baseflow = -transpiration - \Delta S \qquad \text{Equation 18}$$

Now that the baseflow during dry periods is calculated, it is important to note that it is not continuous for the entire year. To calculate the continuous baseflow for the entire year, both total exfiltration and re-exfiltration must be known (see Equation 14). While total exfiltration is directly simulated by the model, re-exfiltration can be calculated as follows:

1. Finding the discontinuous re-exfiltration (dry): the model simulates daily total infiltration that could be used directly in Equation 12. After substituting Equation 15 in Equation 12, only net infiltration is still required to find the re-exfiltration. Assuming that snowmelt and rainfall are the only sources of infiltration during dry periods, the net infiltration is zero during dry periods. Thus, Equation 12 during dry periods is as follows:

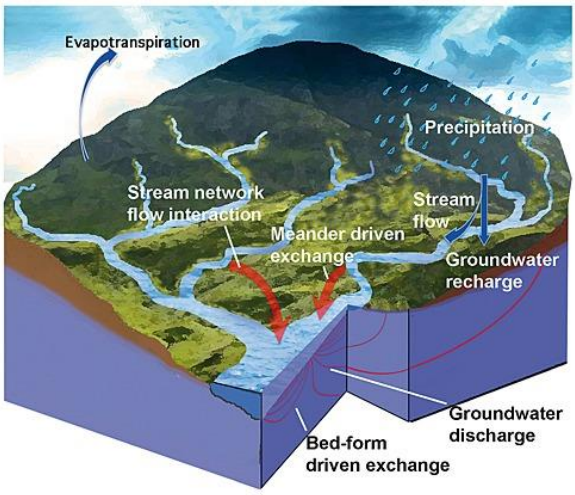
$$\begin{aligned} total\ infiltration &= actual\ infiltration + re\_infiltration \\ &= re\_exfiltration\ (dry) \end{aligned} \qquad \text{Equation 19}$$

   Thus, the discontinuous re-exfiltration of dry days is now calculated
2. Estimating the continuous re-exfiltration: the relationship between the simulated continuous total exfiltration and discontinuous re-exfiltration (dry) is extracted. Then, according to this relationship and using the continuous total exfiltration data, the daily continuous re-exfiltration is interpolated.
3. Estimating baseflow: Subtract the estimated re-infiltration from the total exfiltration to obtain the final baseflow (see Equation 14). The baseflow is checked if it exceeded the streamflow, and if so, the percentage of exceedance is calculated as below:

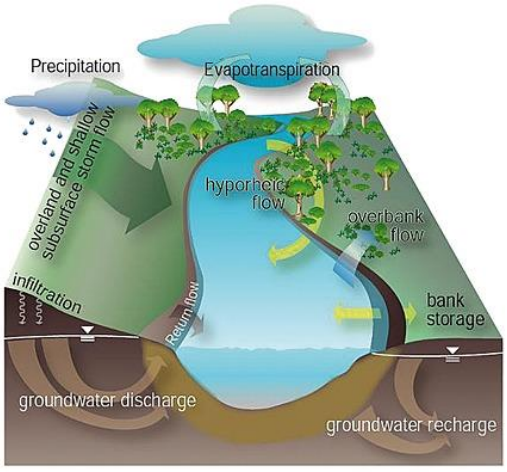
$$\%\ of\ exceedance = \frac{(estimated\ baseflow - streamflow)}{streamflow} \times 100$$ Equation 20

If baseflow exceeds the streamflow, the dry period baseflow calculated in Equation 18 replaces the exceeded days. replaces the exceeded days. If there is no dry baseflow for the exceedance days, the baseflow is removed, and those days are left blank.

a) Watershed or Basin scale



b) River Corridor scale



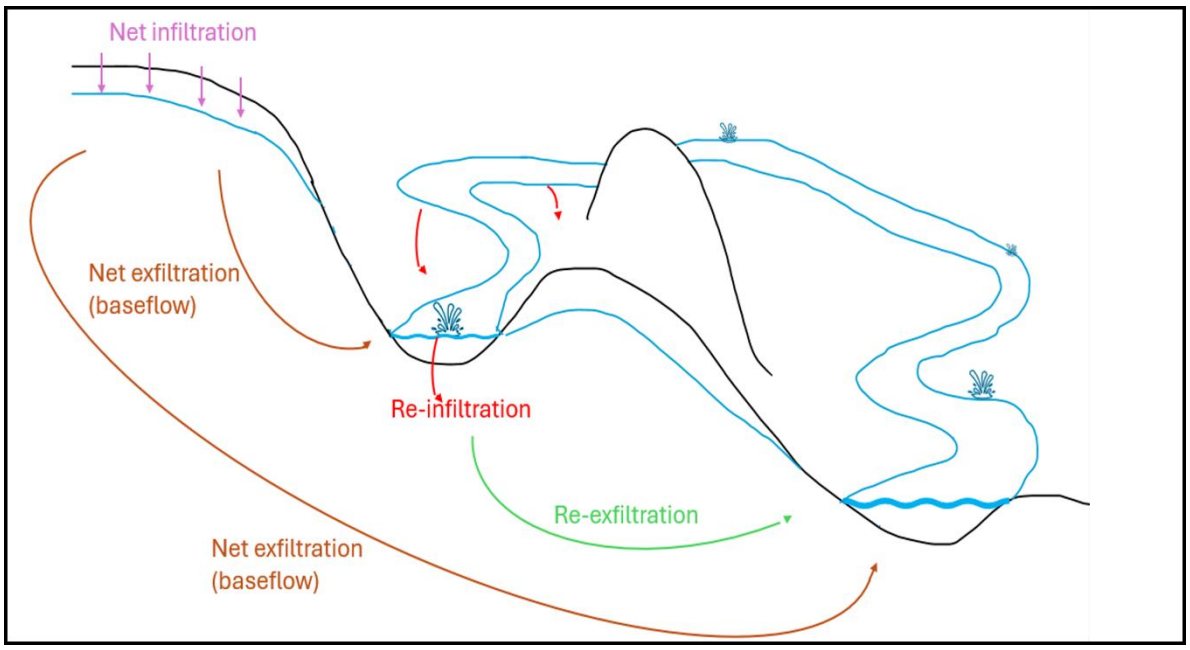


Figure 3: The exchange between the surface and subsurface along the riverbed and edges through the bed-form-driven and meander-driven exchanges. The upper figures are adapted from Boano et al. (2014). The water infiltrates and is assumed to exfiltrate in the same day without any delay. Thus, the re-infiltration is assumed to be equivalent to the re-exfiltration.

3.2 PH-Graphical Method

The PH approach is a graphical method that was developed by Pettyjohn & Henning (1979) (Table 1). The PH graphical approach can be implemented using three algorithms that systematically connect low points into a line. Each algorithm is considered a separate method: the fixed interval, sliding interval, and local minimum methods (Pettyjohn and Henning, 1979). Sloto & Crouse (1996) explained the methodology of the three approaches starting with the common requirement of these methods such as the daily streamflow that is simulated by ATS model and the duration of the surface runoff which is calculated from the empirical relation:

$$N = A^{0.2} \qquad \text{Equation 21}$$

Where N is the number of days after the runoff stops and A is the drainage area in square miles. The 2N values are then calculated. The 2N value should be changed to the closest odd integer, ensuring it remains within the range of a minimum of 3 and a maximum of 11.

The fixed-interval method works by detecting the lowest discharge within each interval of 2N* days. At this lowest point, a straight horizontal line is plotted, starting at the beginning of the 0th day of the 2N* interval and ending at the 2Nth day (Figure 4). The baseflow is represented by the line connecting these lines.

The sliding-interval method involves calculating the interval width as 2N*-1. The median day of this new interval is then assigned the lowest discharge recorded within that interval (2N*-1), creating a point with the median day as the x-axis coordinate and the lowest discharge as the y-axis coordinate. The interval is then shifted forward by one day, and the new median day is assigned the corresponding lowest discharge, forming a new point with the new median day being the previous day plus one. This process is repeated for the entire period of interest. All the obtained points are then connected by straight lines to form the baseflow hydrograph, which represents the baseflow component of the total streamflow.

The local minimum method also involves calculating the 2N*-1 width. The process starts by checking if any of the days within a 0.5(2N*-1) width to the right and 0.5(2N*-1) width to the left of the assigned day have a lower discharge than that of the assigned day. If so, the newly assigned day is the day that has the lowest discharge within the total width of (2N*-1). This process is repeated for the newly assigned day. If there is no lower discharge within this interval, this day and its discharge are assigned as a local minimum. All the local minimums are then connected by straight lines to represent the baseflow. This method can be visualized as drawing straight lines between the lowest points on the hydrograph to represent the baseflow.

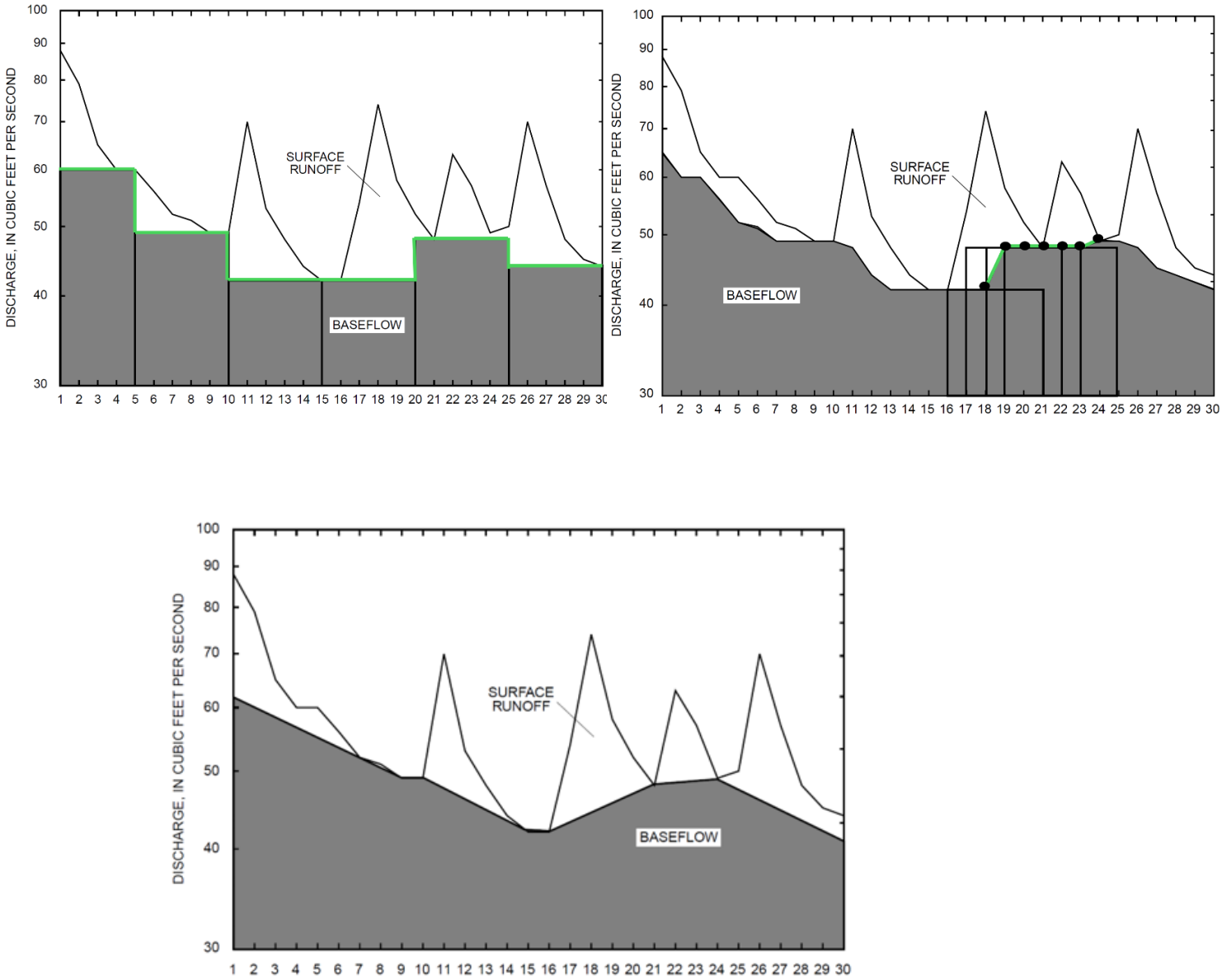


Figure 4: Three approaches of PH graphical method (The 2N* used in these examples is 5 days): (a)The fixed-interval method. The baseflow is represented by the line connecting the upper edges of the black rectangles. These horizontal lines indicate the lowest discharge within each fixed 2N*-day interval, forming the baseflow component of the hydrograph. (b) The sliding interval method where one point of the coordinates of the lowest discharge and median day of the 2N*-1 interval. All the points are connected in green. (c) The local minimum method. Adapted from Sloto & Crouse (1996)

3.3 UKIH Graphical Method

The UKIH method is a smoothed minima method that estimates the baseflow by identifying the turning points within the hydrograph. The turning points are the days of

corresponding streamflow values assumed to be the entire baseflow (Gustard et al., 1992; Piggott et al., 2005; Zhang et al., 2017). The steps required to find the turning points are described by Aksoy et al. (2009), Zhang et al. (2017), and Xie et al. (2020) as follows: (1) partitioning the hydrograph into 5-day consecutive segments without any gaps or overlapping and note the days that have the minimum streamflow in the entire segment. These are the minimum points that have streamflow values of Q_1, Q_2, Q_3, …, Q_t, where t is the order of the block, (2) grouping each of three consecutive minimum points as (Q_1, Q_2, Q_3), (Q_2, Q_3, Q_4), …, (Q_{t-1}, Q_t, Q_{t+1}), (2) selecting the groups that have $Q_t \leq 1.11\ Q_{t-1}$ and $Q_t \geq 1.11\ Q_{t+1}$. The central points of the selected groups are the turning points. Once all turning points are found, the turning points are linearly interpolated as a baseflow line. The final baseflow hydrograph is plotted by restricting the exceedance of the baseflow line against the streamflow hydrograph.

3.4 Eckhardt Digital Filter Method

Eckhardt (2005) introduced a special technique for separating baseflow by modifying existing digital filter techniques, known as the recursive digital filtering method, that partitions streamflow into baseflow and runoff:

$$y_k = f_k + b_k \qquad \text{Equation 22}$$

Where y is the total streamflow,

f is the direct runoff,

b is the baseflow,

k is the time step,

With the assumption of the linear reservoir, they followed the following algorithm:

$$b_k = \frac{(1 - BFI_{max})ab_{k-1} + (1 - a)BFI_{max}\, y_k}{1 - aBFI_{max}} \qquad \text{Equation 23}$$

Where a is the recession constant and BFI_{max} is the maximum value of the baseflow index

According to Eckhardt (2005), the main parameters required to apply this method are the recession constant and the maximum value of the baseflow index (BFI_{max}). The recession constant was calculated using different methods. Xie et al. (2020) used regression analysis to determine the recession constant for 1815 catchments. In our study, we conducted a similar recession analysis and derived a recession constant of 0.988. This value has been utilized in the application of the Eckhardt baseflow separation method.

As BFI_{max} cannot be directly measured, Eckhardt (2005) analyzed results from various watersheds with differing hydrogeological characteristics and identified three reference values: BFI_{max} = 0.8 for perennial streams with porous aquifers, BFI_{max} = 0.5 for ephemeral streams with porous aquifers, and BFI_{max} = 0.25 for perennial streams with hard rock aquifers. These are approximate values for BFI_{max} and not the exact values for different watersheds. For this study, without field investigation, the BFI_{max} was determined using a backward filtering technique (Collischonn and Fan, 2013; Xie et al., 2020). The method incorporates the recession constant and streamflow to filter the baseflow as follows:

$$b_{k-1} = \frac{b_k}{a}\ where\ (b_k \leq y_k) \qquad \text{Equation 24}$$

BFI_{max} is calculated by dividing the maximum possible baseflow by the total streamflow. The details of the backward filtering operation are thoroughly described in Collischonn and Fan (2013).

3.5 Conductance Mass Balance (CMB) Method

The CMB method requires continuous streamflow data and continuous specific conductance data (Stewart et al., 2007). In locations where only discrete measurements of SC are available, the power function that relates the SC to stream discharge can be used in estimating the continuous SC from the discrete SC data (Lott & Stewart, 2013). After estimating the continuous SC data, it can be used to determine the baseflow using the following equation:

$$Q_{BF} = Q\left[\frac{SC - SC_{RO}}{SC_{BF} - SC_{RO}}\right] \quad \text{Equation 25}$$

In this equation, Q_{BF} is the estimated daily baseflow in m^3/s, SC_{RO} is the specific conductance of the runoff end member in µS/cm, and SC_{BF} is the specific conductance of the baseflow end member in µS/cm. Miller et al. (2014) suggested that identifying the end members involves recognizing that the runoff end member, with its low SC, is best measured during snowmelt while assuming baseflow is minimal at peak runoff. The baseflow end member, which has high SC, is measured upstream, assuming that the runoff upstream is minimal, during the summer when baseflow is dominant.

4 RESULTS AND DISCUSSION

4.1 Numerical Method

The results of the model simulating the discharge of the Coal Creek watershed are presented in three plots (Figure 5). The top plot shows the time series comparison between observed discharge (obs_Q) and simulated discharge (simu_Q) over the period from 2016 to 2022. The middle plot displays the flow duration curve, which illustrates the exceedance probability of discharge rates. The bottom right plot is a scatter plot comparing observed

discharge with simulated discharge, along with performance metrics: R-square = 0.69, NSE (Nash-Sutcliffe Efficiency) = 0.68, and mKGE (modified Kling-Gupta Efficiency) = 0.76.

The model tends to underestimate discharge during wet years, such as in 2019, and overestimate discharge during dry years, such as in 2021. Additionally, the model consistently overestimates low flow conditions. To avoid uncertainties introduced by differences between the simulated and observed streamflow, the simulated streamflow is assumed to be the truth and is used in the baseflow estimation for other methods. This assumption is reasonable because the study focuses on the performance of baseflow separation methods, provided there is no significant change in the streamflow pattern. Despite these gaps, the model performs well in matching the peaks during normal years, indicating its robustness in capturing the overall discharge patterns.

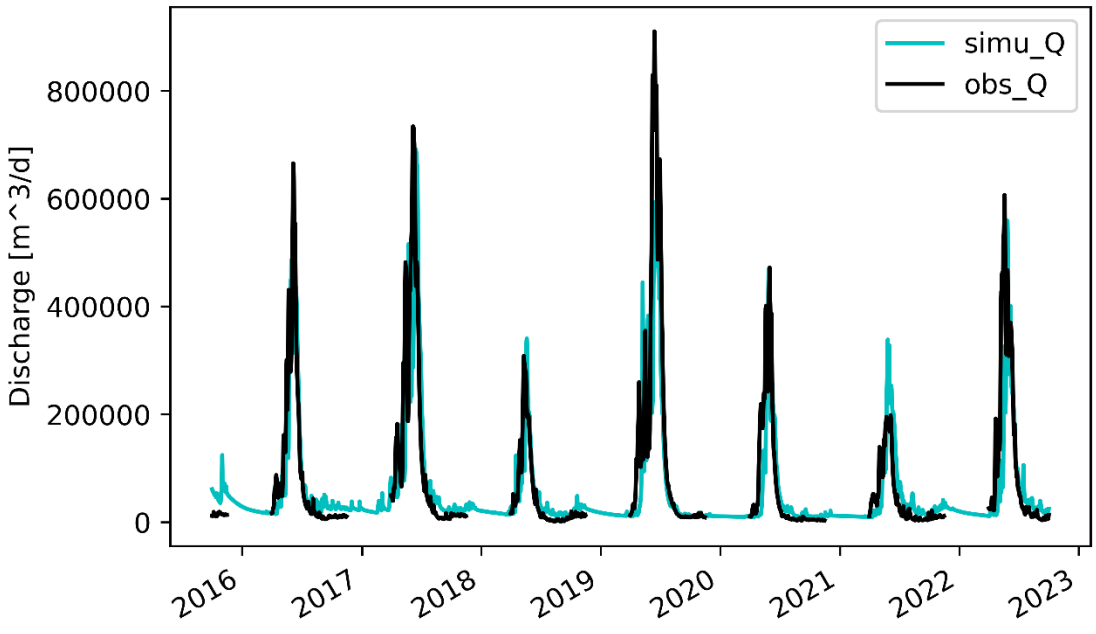
The exceedance probability shown in the flow duration curve in Figure 5 indicates that the average discharge of high flow periods from the model is slightly underestimated. However, the exceedance probability of low flow is higher for the simulated data as compared to the observed data, further supporting that the model overestimates low flows.

In general, the comparison between observed and simulated discharge in the time series plot indicates that the model captures the seasonal patterns and peak flows reasonably well. The performance metrics further support this, with R^2 equals 0.69, NSE = 0.68, and mKGE = 0.76, all indicating a good level of agreement between observed and simulated data (Kling et al., 2012; Nash & Sutcliffe, 1970). Thus, the model performance is reasonable, and its results can be used to evaluate the automated separation methods.

To understand the characteristics of the aquifer, the surface and subsurface hydrologic processes, and the relationship between the streamflow and baseflow, more analysis was done on other surface and subsurface variables that the baseflow is sensitive to.

The model simulates surface ponded depth, groundwater table, and aquifer saturation both spatially and temporally. At point A, Figure 6 shows the seasonal change in the water table. The increase in the water table in 2016 occurs at the same time as the high surface ponded depth and subsurface saturation (Figure 7). The water table decreases with the decrease in the surface ponded depth and the subsurface saturation.

The water table in the water year 2021 reached the lowest recorded during the study period (Figure 6). When examining the streamflow for this water year, the average is 0.45 m^3/s compared to the total average which is 0.68 m^3/s, classifying it as a dry year. The variations in streamflow and groundwater levels correlate well; high groundwater levels and saturation lead to high streamflow, resulting in higher low flows during summer.



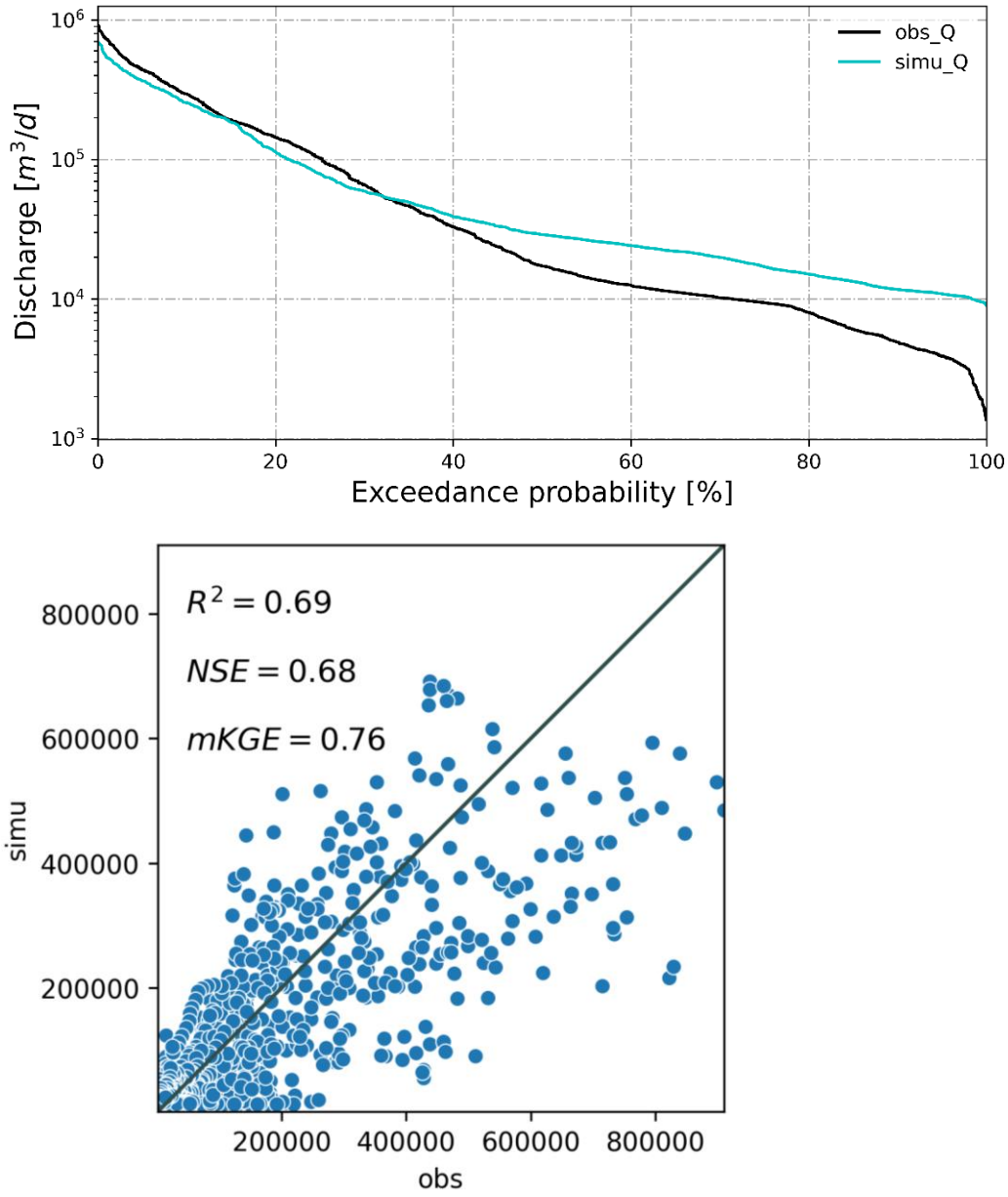
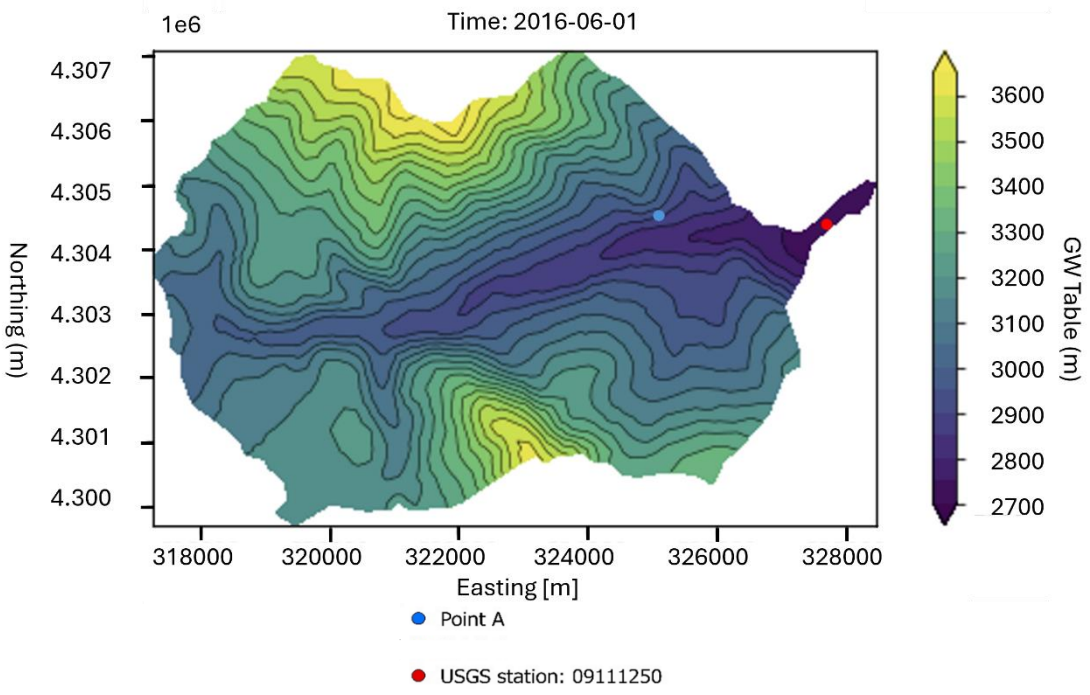


Figure 5: Comparison of observed discharge (obs_Q) and simulated discharge (simu_Q) for the Coal Creek watershed from 2016 to 2022. The top plot shows the time series of observed and simulated discharge, highlighting the model's ability to capture seasonal patterns and peak flows. The lower left plot displays the exceedance probability of discharge rates and indicates the model's performance during high flow periods (e.g., snowmelt events) while overestimating low flow conditions. The bottom right scatter plot compares observed and simulated discharge, with performance metrics R-square = 0.69, NSE = 0.68, and mKGE = 0.76, demonstrating a good level of agreement between observed and simulated data

The surface saturation maps of the entire watershed (see Figure 7) show changes in surface saturation before (May 05, 2016), during (June 15, 2016), and after the snowmelt period (July 01, 2016) in an average year, 2016. During the snowmelt, saturation reached its

maximum in most parts of the watershed, resulting in saturation excess overland flow. This runoff peak generated is closely related to variables such as the antecedent soil moisture and evapotranspiration (Godsey et al., 2014) impacting surface water infiltration and subsequent groundwater discharge, eventually, impacting the baseflow. Thus, the numerical model simulates all these discussed variables to integrate the baseflow.



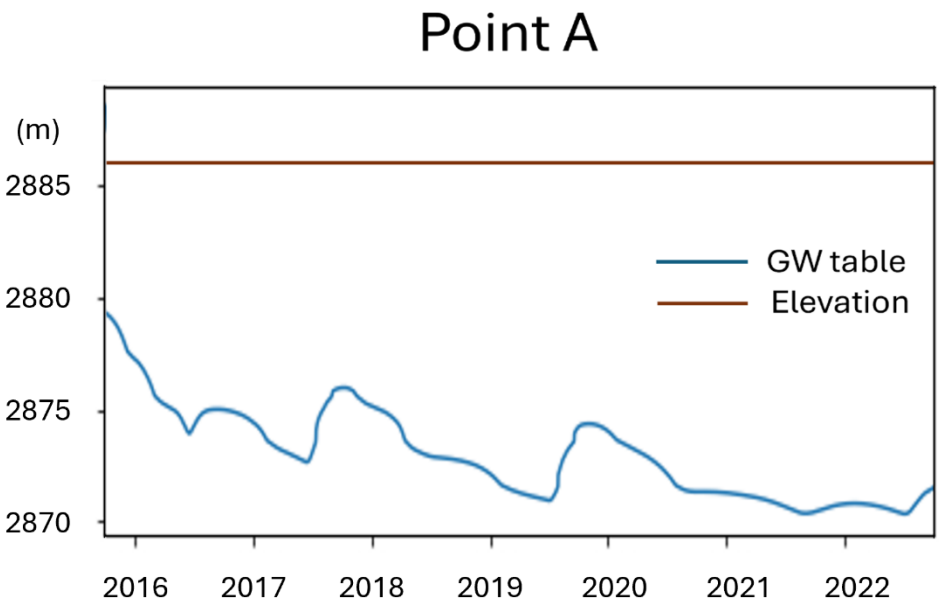


Figure 6: The groundwater level contours with point A in blue and the catchment USGS output station in red in the upper figure on June 01, 2016. The groundwater level graphs at point A in the lower figure.

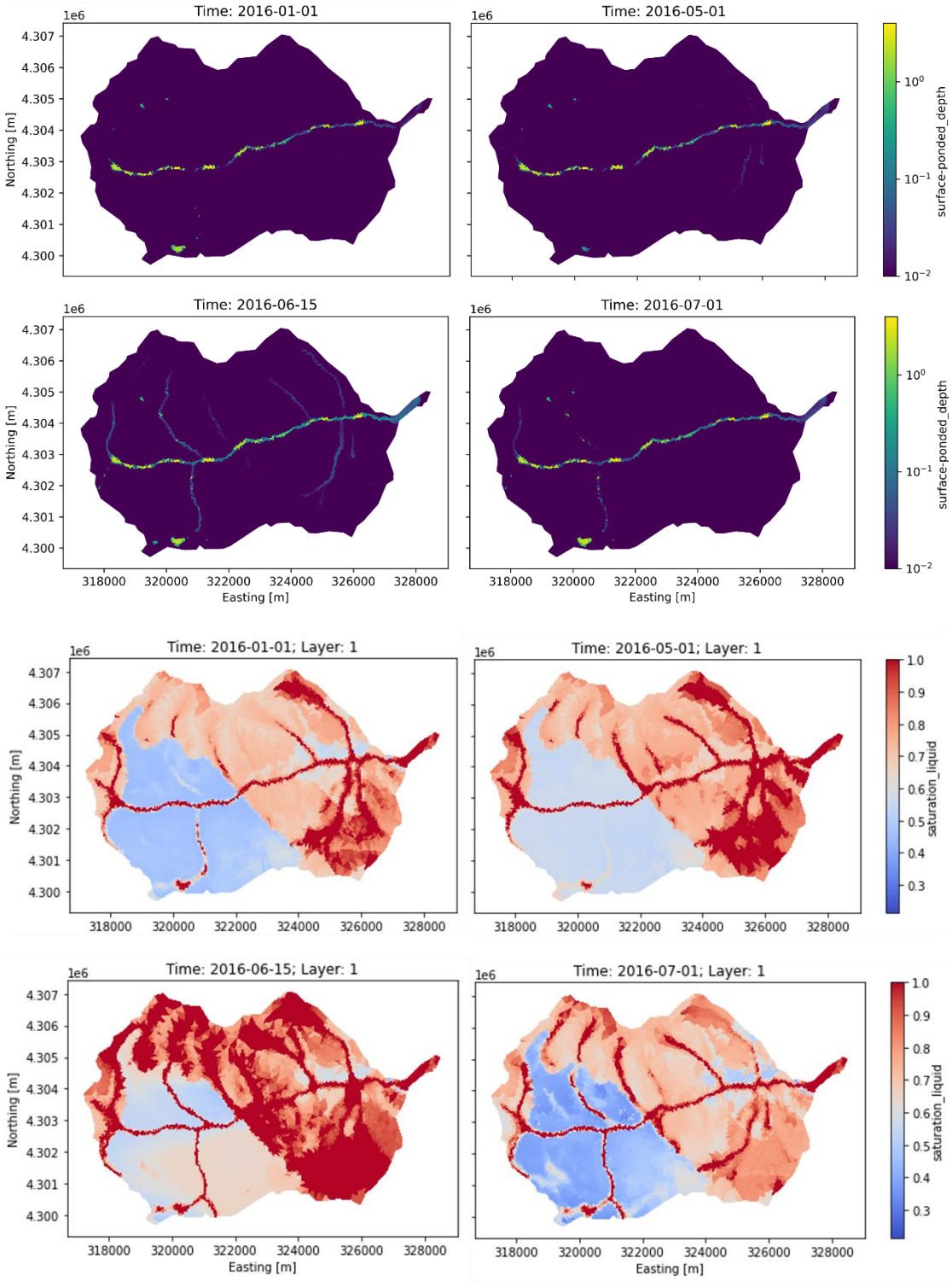


Figure 7: Upper figure: the ponded depth of the surface. Lower figure: maps of the surface saturation of the top 5 cm soil layer for the entire watershed in 2016 in January, just before the snowmelt on the first day of May, during the snowmelt on June 15, and at the end of the snowmelt on July 7th.

4.1.1 Calculating the baseflow from the ATS model

The baseflow is calculated as per Equation 19 using the simulated variables and approach discussed earlier in the methodology. Starting with dry periods, the baseflow obtained is equivalent to the total streamflow (Figure 8), based on the assumption of zero infiltration and no time lags between infiltration and exfiltration. This is also the streamflow used in the recession analysis. Figure 9 shows that during dry periods, the simulated total infiltration is not zero, indicating the existence of a re-infiltration process.

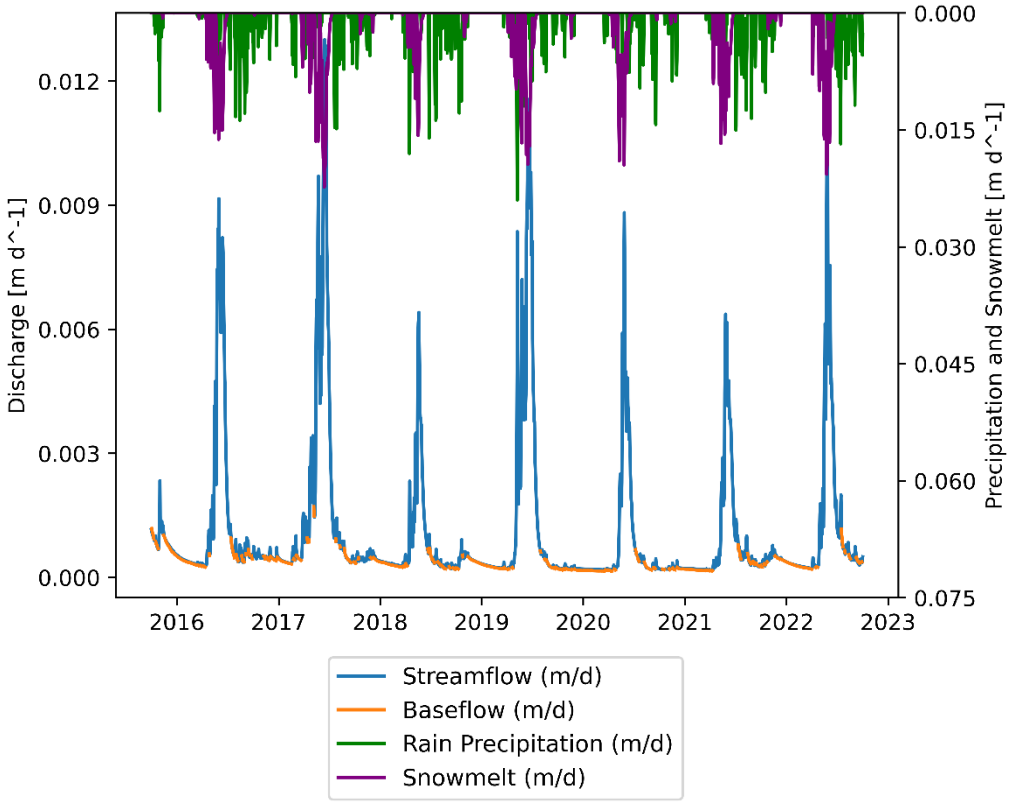


Figure 8: Dry baseflow discharge during the dry season of the year when rainfall and snowmelt are zero

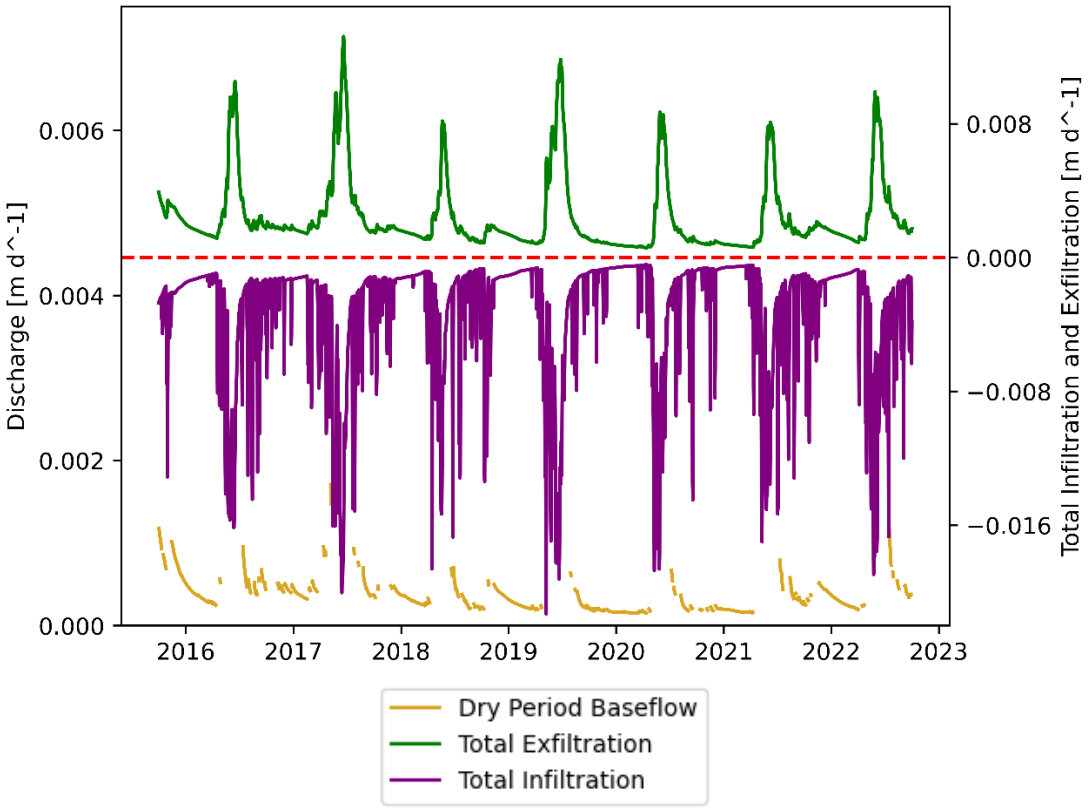


Figure 9: The dry period baseflow is shown in this plot on the left axis. The right axis shows the total infiltration (negative) and the total exfiltration (positive). The net infiltration is supposed to be zero during dry periods (when there is a dry period baseflow). The total infiltration during the dry periods is not zero, meaning that this is not the net infiltration, but the total including the re-infiltration

The continuous baseflow was calculated following the approach in calculating baseflow using the numerical model simulated variables. Starting with the interpolation of the continuous re-exfiltration, Figure 10 illustrates the linear relationship between the total exfiltration and re-exfiltration. Although some points of high re-exfiltration and total exfiltration did not fit within the linear interpolation, the R^2 value is approximately 0.99, indicating a strong correlation. It is worth noting that the discontinuous re-exfiltration data used to determine the linear correlation are of low value because they represent dry periods

only, which might affect the extrapolation of re-exfiltration during high flow periods.

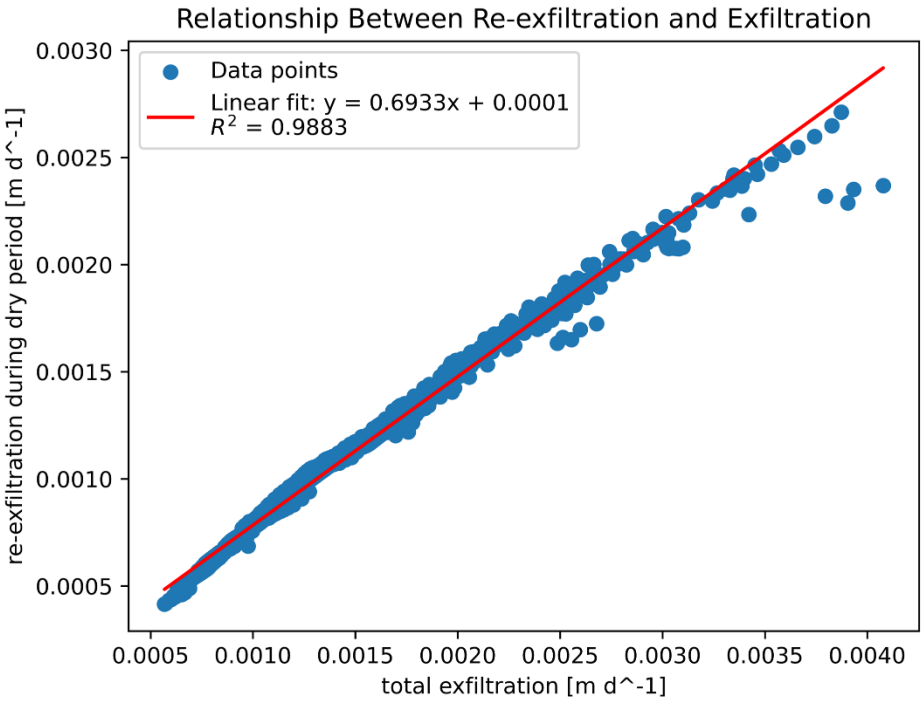


Figure 10: Relationship between re-infiltration and total exfiltration during the dry period. The linear fit equation demonstrates a high correlation (R^2 = 0.9883) between the two variables, with only minor deviations at higher exfiltration values.

The obtained baseflow is shown in Figure 11Figure 12. There were 560 days of 2555 total days (percentage of 22%) where the calculated baseflow was slightly higher than the streamflow with an average exceedance of 6.5%. To determine the timing and extent of exceedance, Table 2 shows that May is the only month without any exceedance, while July, August, and September have high average exceedance. According to Figure 12, the exceedance is not substantial when compared to the average streamflow. However, when compared to the streamflow on the same day, it appears exaggerated because the streamflow is already low, making any exceedance more noticeable. Additionally, from July to September, the streamflow exhibits many small fluctuations that should also be reflected in the re-infiltration but cannot be accurately estimated using the interpolation method. These small errors in baseflow estimation may not be noticeable during high flow periods because the significant difference between streamflow and baseflow means any extra

discharge added to the baseflow will not exceed the streamflow. However, during the lower flow periods in July, August, and September, these errors become more apparent. In other low-flow months, the baseflow does not significantly exceed the streamflow because the streamflow is relatively stable, and interpolation works effectively in these cases. Another caveat about this approach is that there may be extrapolation errors as the data used for developing the linear regression model only contains data during dry periods without any high flows.

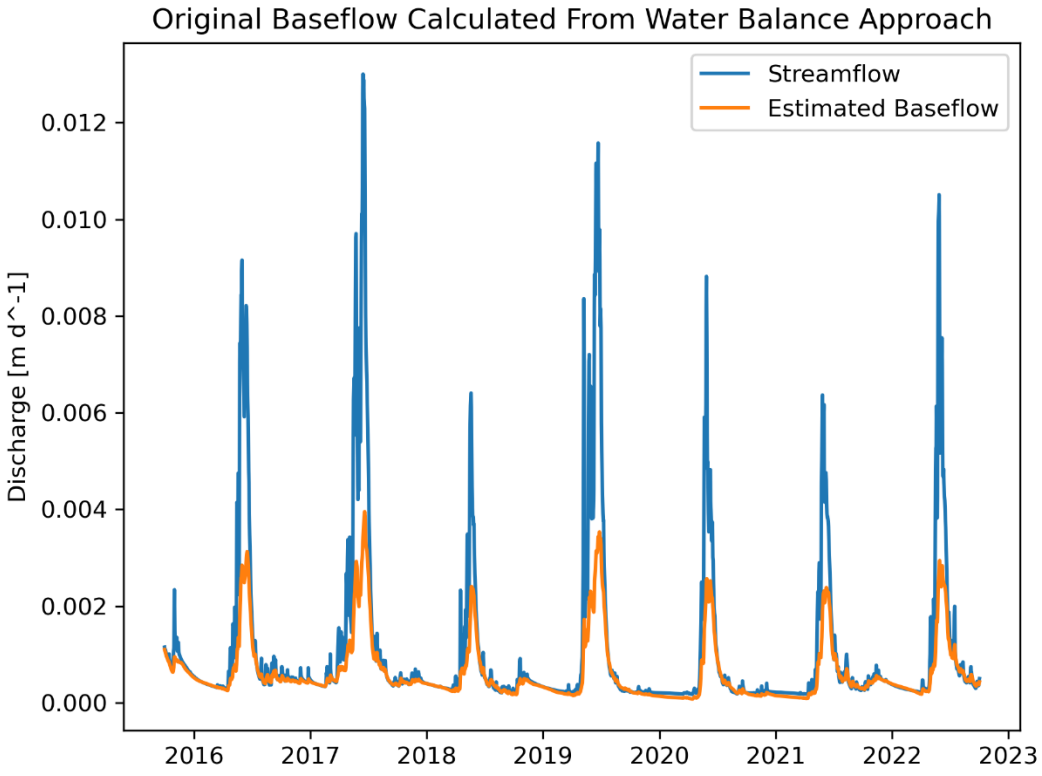


Figure 11: The baseflow calculated using water balance and re-infiltration interpolation. The baseflow exceeds the streamflow at percentage of 21.5% of the days (560 days during the period from 2015 to 2022).

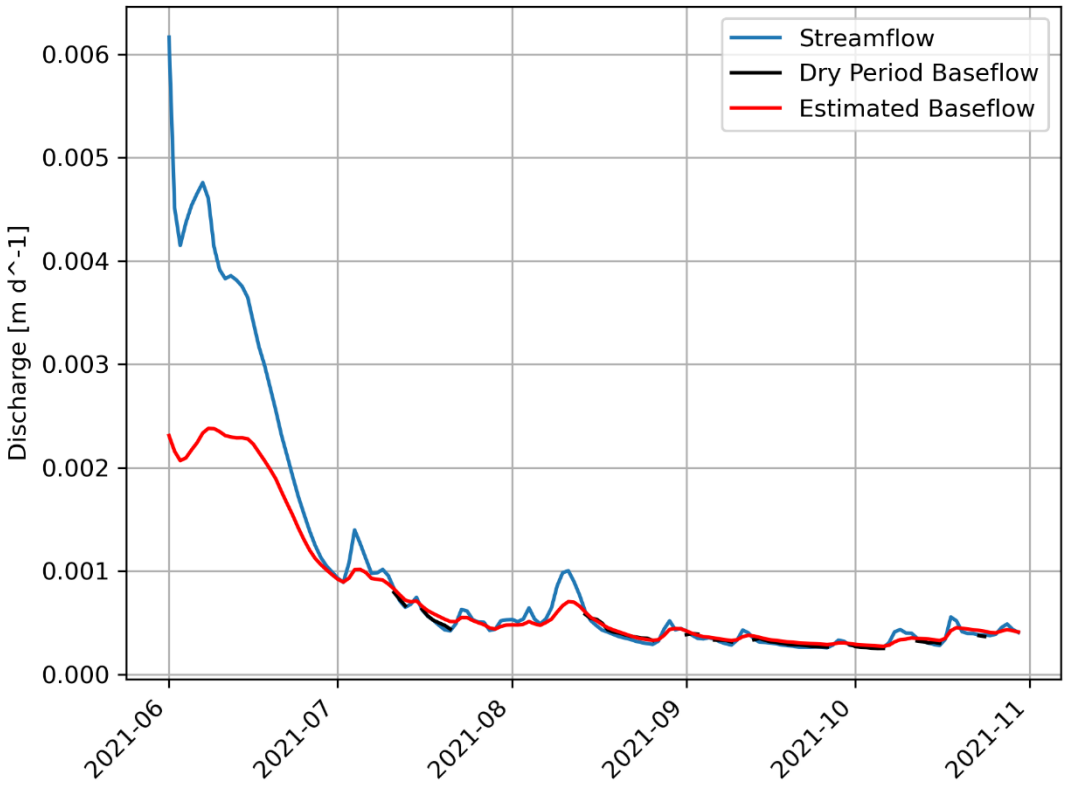


Figure 12: Zoom-in plots of the month showed significant baseflow exceedance.

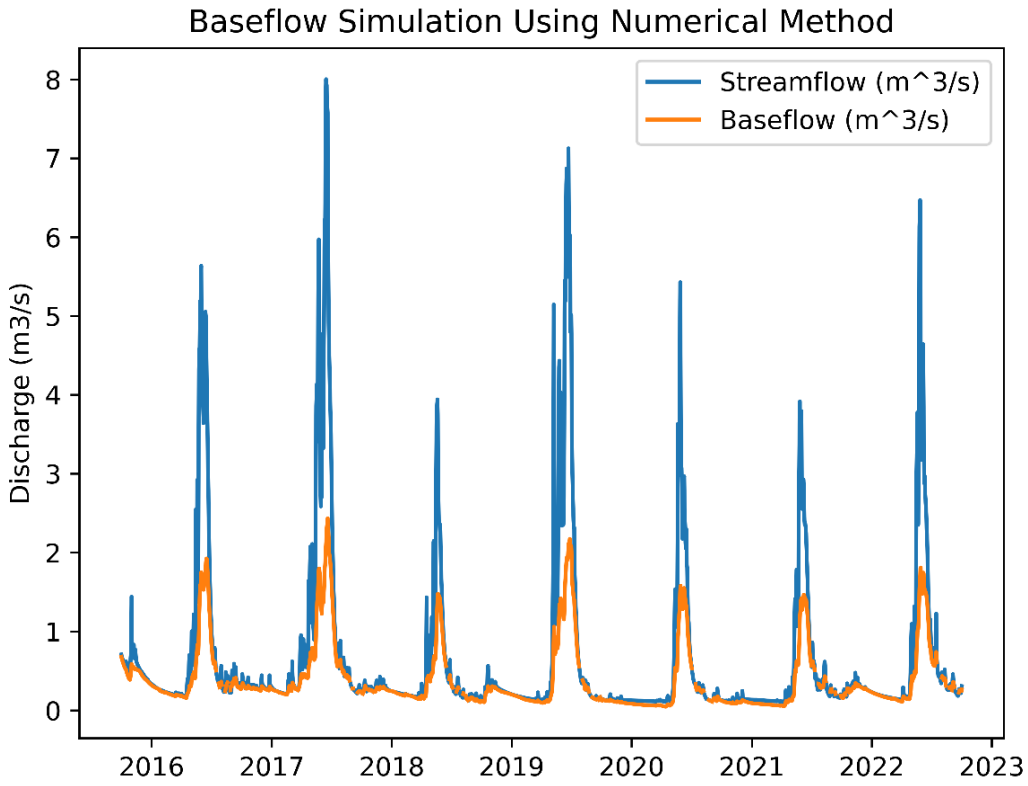
Table 2: The exceedance average of baseflow for each month of the year in percentage for the entire period (2015-2022). The statistics of exceedance such as the minimum, the maximum, and the SD.

month	average exceedance (%)	min exceedance (%)	max exceedance (%)	SD exceedance
1	2.9	0.0	7.2	2.6
2	2.9	0.0	7.2	2.0
3	2.7	0.2	6.1	1.8
6	6.5	0.3	9.7	2.8
7	7.4	0.2	23.9	5.6
8	11.0	0.0	29.9	6.5
9	9.1	0.2	24.1	6.4
10	6.1	0.2	16.8	4.6
11	1.4	0.8	2.0	0.9
12	1.1	0.0	2.7	0.9

To address this issue, the days of exceedance were replaced by the dry-period baseflow or completely removed if it was not a dry period. Consequently, small gaps appeared in the final baseflow, as shown in Figure 13. These gaps, however, do not significantly alter the overall shape of the baseflow curve. We will use this baseflow as control points in comparison with

other baseflow separation methods. It is worth noting that the final baseflow is converted to m^3/s to maintain consistency with other methods, facilitating the eventual comparison of different methods.

Baseflow variation between different water years is shown in Figure 14 and Figure 15. In terms of baseflow peak timing, the baseflow in 2017 (wet year) reached the peak around July 01, however, the highest baseflow in 2018 (dry year) occurred around June 01 (Figure 14). Also, the length of the baseflow peak is longer in 2017 than in 2018. This variation in peak timing as per dry, wet, and average years is significant and consistent for most of the years. However, in terms of baseflow magnitude compared to the total streamflow, e.g., BFI, the BFI of wet years is slightly below the average for the entire period of study and slightly higher than the average in dry years (Figure 15). The minimum BFI, 46.8%, was recorded in 2019 while the maximum BFI, 61%, was recorded in 2018. The variation didn't exceed the ratio of ±0.15 of the average BFI, 53.24%. This means that the baseflow is mostly stable.



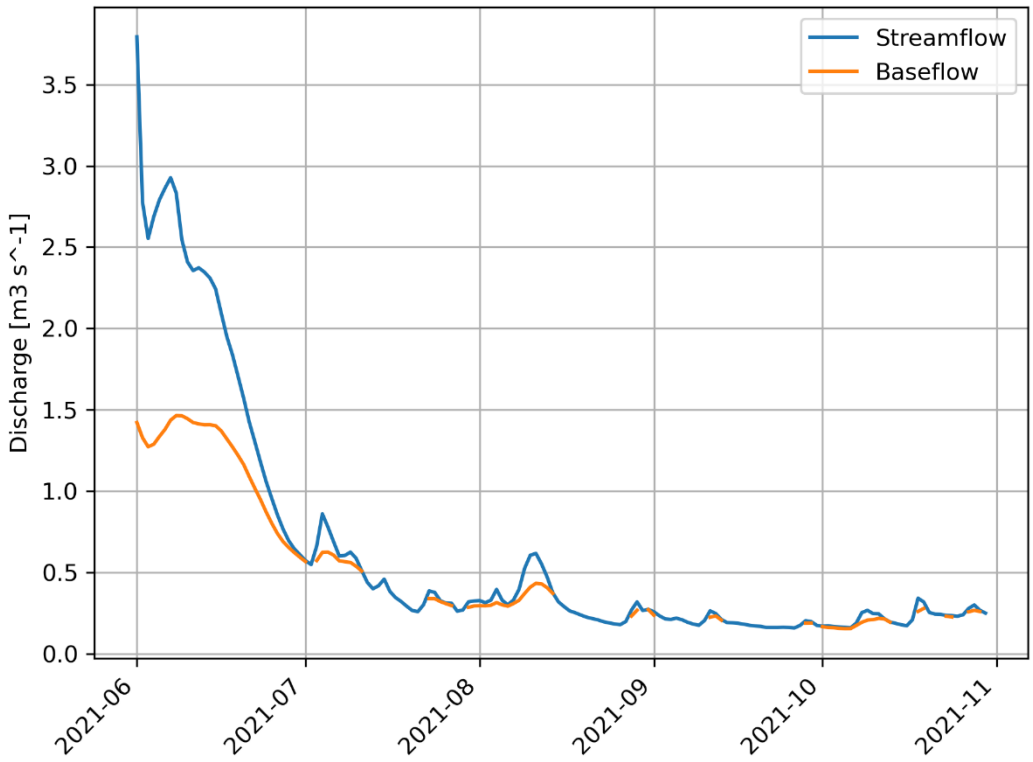
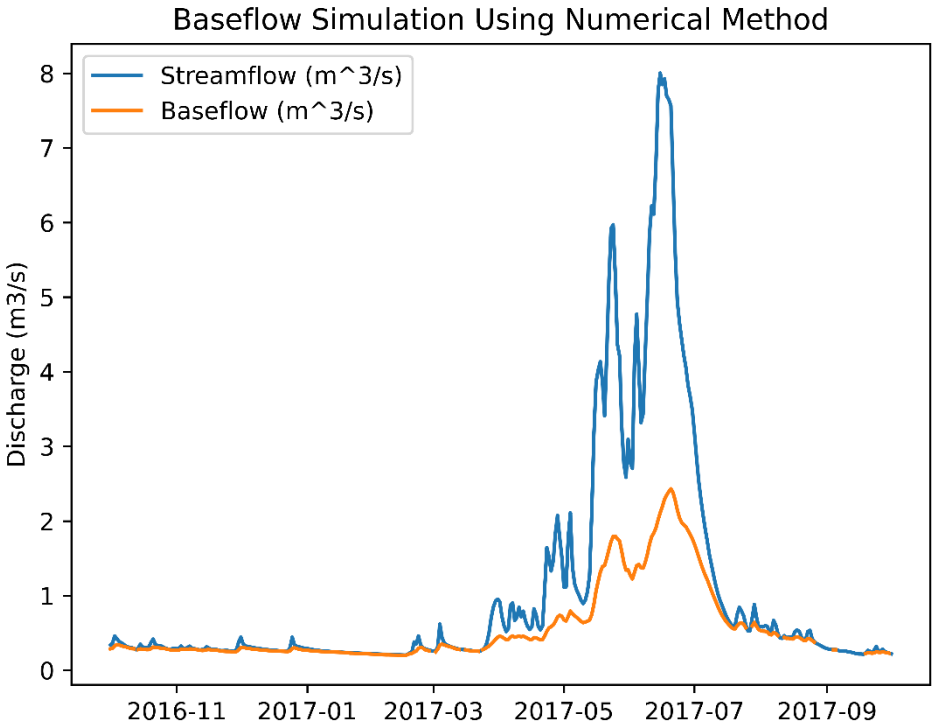


Figure 13: Baseflow of Coal Creek calculated using the results and variables of the numerical method. The upper graph shows the entire period of study from 2015 to 2022. The lower graph: zoom in to the period between July 2021 and November 2021. Short gaps appear in the curve due to the interpolation of the re-infiltration, resulting in a small margin of error in baseflow calculations that slightly exceeded the streamflow.



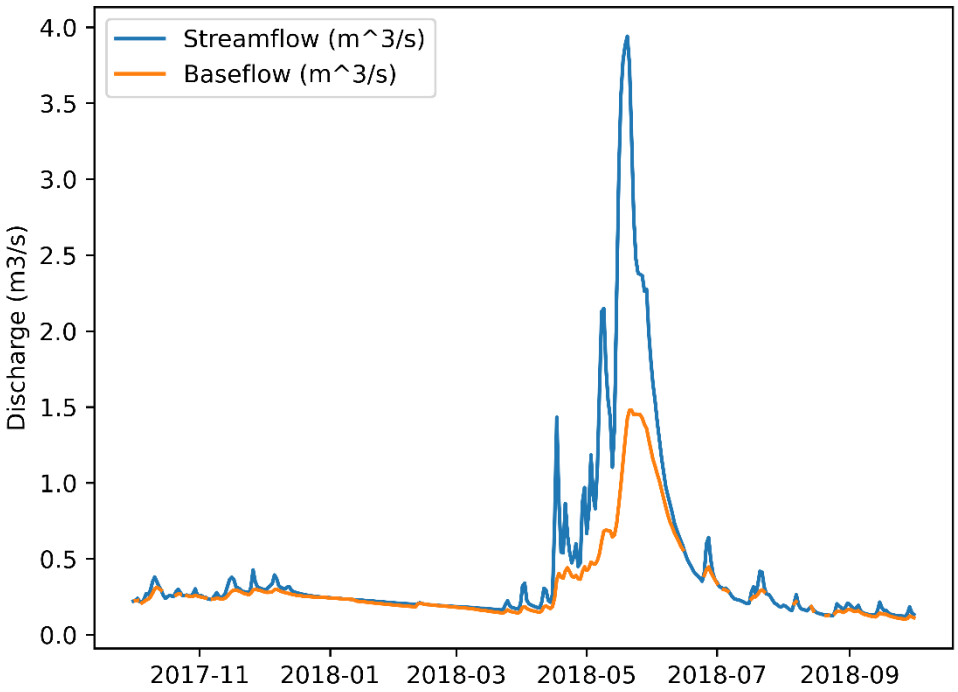


Figure 14: Comparing water years 2017 (wet year) and 2018 (dry year) baseflow. The peak is reached in dry years earlier than that of the wet year.

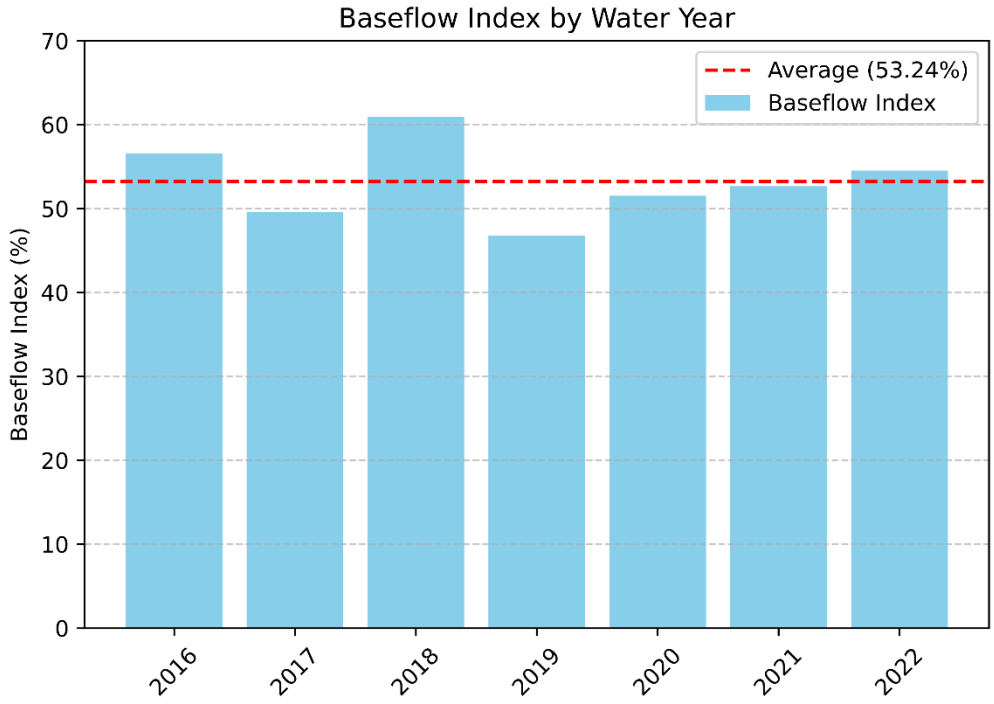


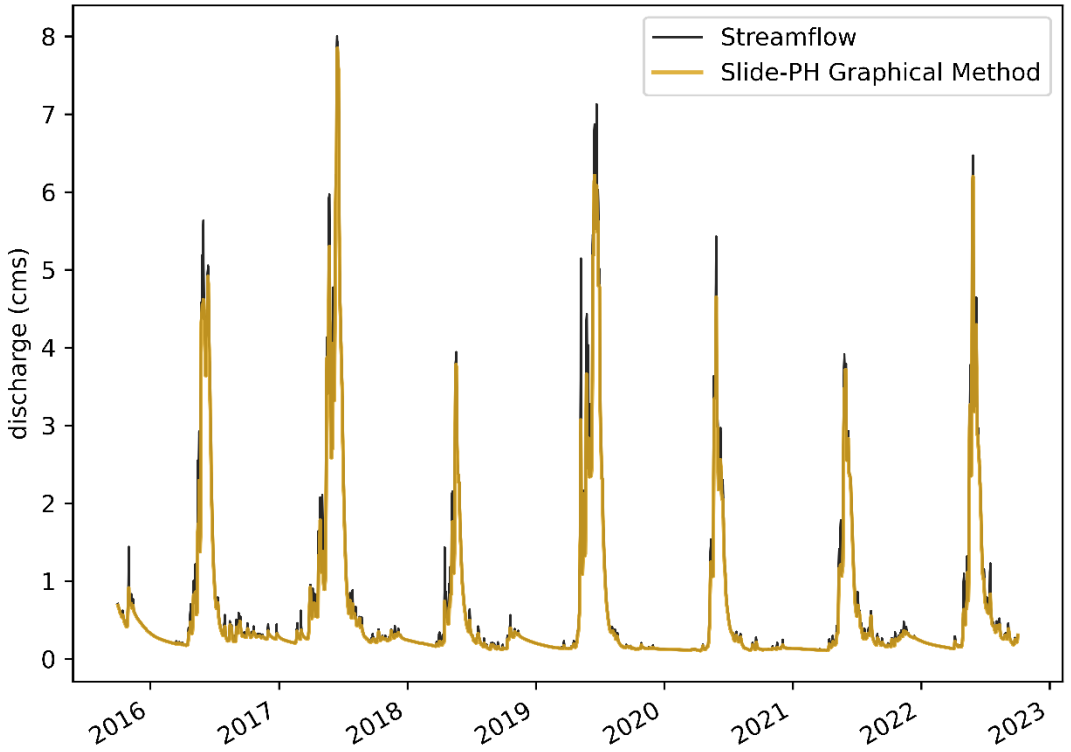
Figure *15*: BFI percentage of all water years from 2016 to 2022 compared to the average.

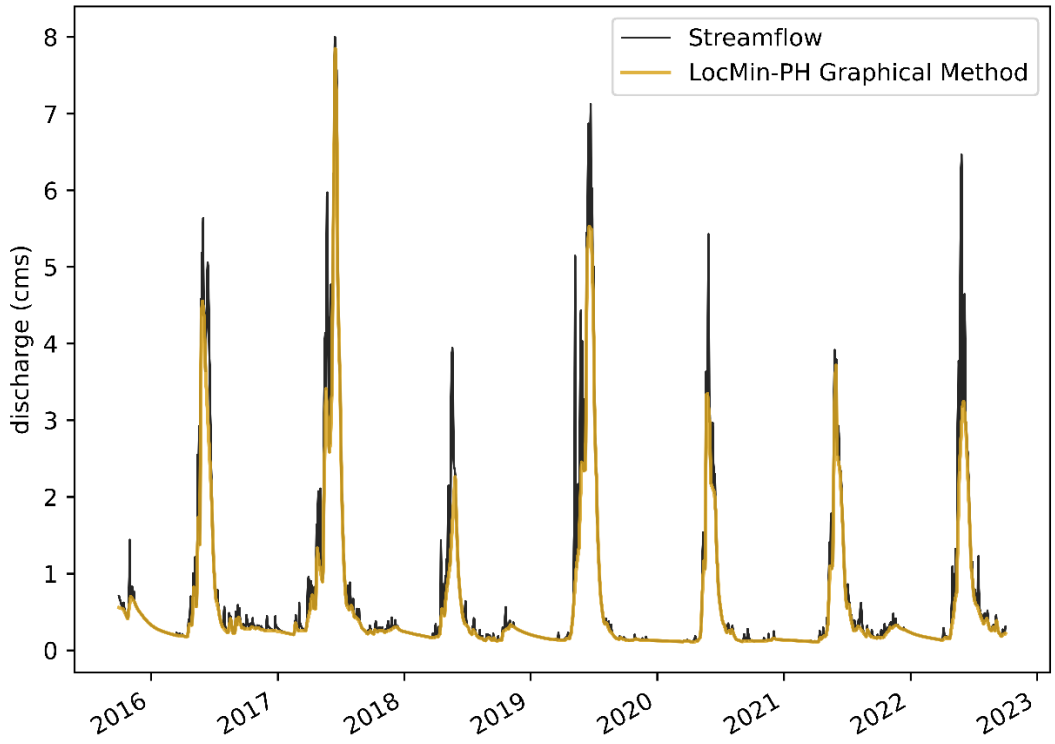
4.2 PH-Graphical Method

In our case, catchment area A is 20.1 square miles (Equation 21). The N value is 1.8, and the 2N is 3.6. The closest odd integer to 3.6 is 3, so the 2N* that will be used now is 3, not 3.6. It is worth noting that 3 is also the lower limit of the 2N interval, and results might be affected (see Equation 21).

The results of baseflow separation using the three different approaches—fixed, slide, and LocMin—are presented in Figure 16. Only the simulated streamflow by the ATS numerical model is used to separate the baseflow in this method. Statistically, the average baseflow of all the approaches is very high (above 0.55 m^3/s) compared to that of the streamflow average (0.68 m^3/s). LocMin approach has the lowest average of 0.56. Moreover, the lowest maximum of all the approaches is 7.76 while the streamflow maximum is 8. This means that some major peaks were considered to be mostly baseflow.

For further evaluation, as all three methods showed similar results, the LocMin approach will be used in the final comparison. Figure 17 provides more details for the LocMin baseflow curve, specifically for the period between May and October of 2016 which has both types of peaks, short-term and long-term peaks. Before May 20 and after July 01, the streamflow peaks were short and the baseflow was separated well, likely generated by rainfall or quick snowmelt events. However, between the two dates, the model struggles with long-term peaks generated by snowmelt in snow-dominated catchments.


Streamflow
Slide-PH Graphical Method
discharge (cms)
0
1
2
3
4
5
6
7
8
2016
2017
2018
2019
2020
2021
2022
2023


Streamflow
LocMin-PH Graphical Method
discharge (cms)
0
1
2
3
4
5
6
7
8
2016
2017
2018
2019
2020
2021
2022
2023

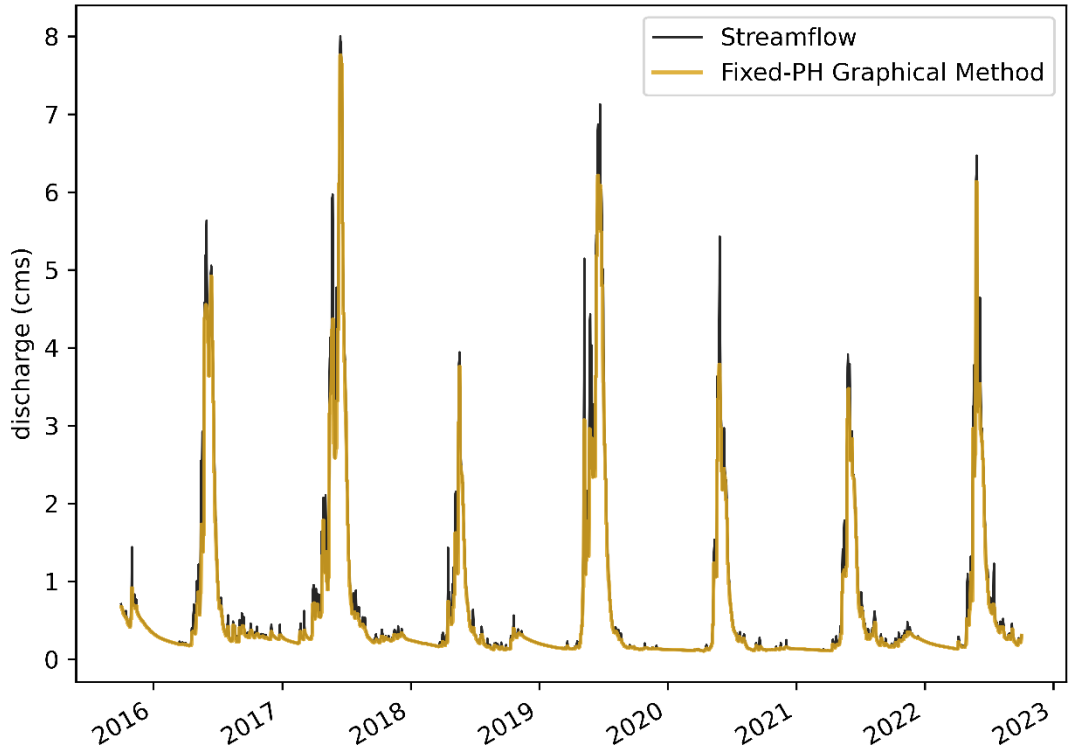


Figure 16: The three approaches of PH graphical baseflow separation method. The top plot shows the baseflow obtained using the slide interval approach. The middle plot shows the results of the LocMin approach. The bottom plot shows the results from the fixed interval approach.

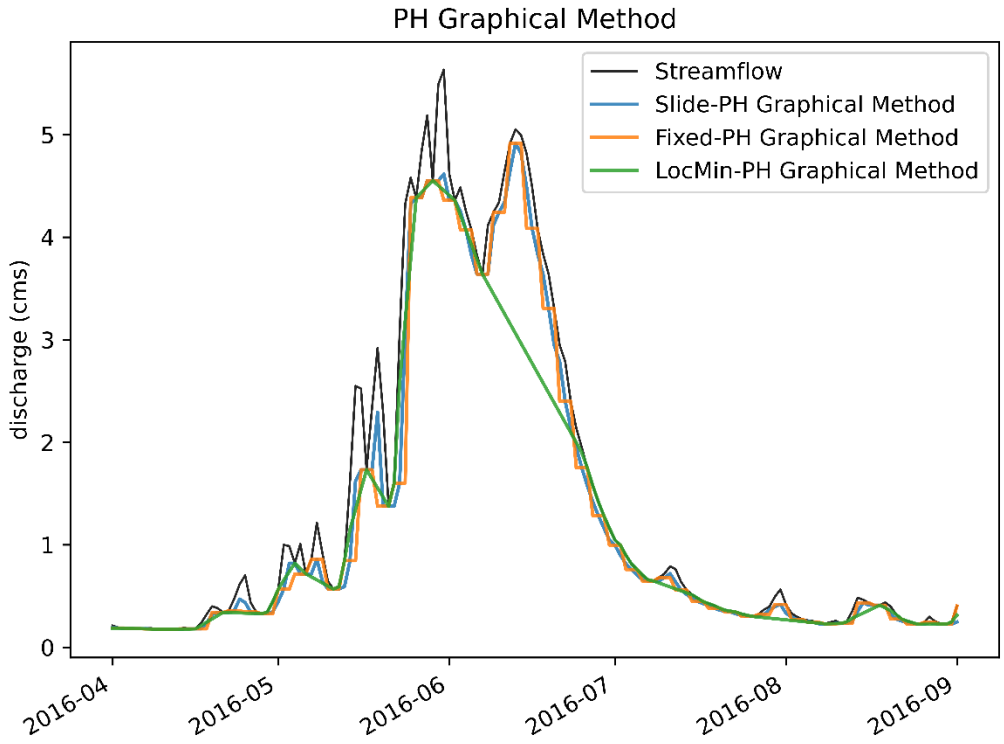


Figure 17: Baseflow curve obtained from the three approaches of PH graphical method for the period between April 2016 and October 2016.

4.3 UKIH Graphical Method

The UKIH method separated the baseflow from the total simulated streamflow for the entire period (Figure 18). In 2016, streamflow peaks were short before May 15 and after July 01 and long between them. The UKIH method effectively separates the baseflow of most of these short and long peaks. However, the baseflow is characterized by connected segments of straight lines.

Focusing on the long peaks created by snowmelt, the timing and magnitude variations of the baseflow are discussed in this section. The baseflow recorded the highest value of approximately 3.64 m^3/s on June 07, 2016, even though this was an average year. In contrast, 2017 and 2019 were wet years, yet the baseflow reached only 2.7 m^3/s on June 02 and 2.3 m^3/s on June 07, respectively. In dry years, the peak reached around 1.1 m^3/s on May 13, 2018, and around 2.6 m^3/s on June 03, 2021. Thus, the peak record of some dry years is higher than the peak of wet years, and the highest baseflow was observed in an average year rather than in a wet year. This variation also appears between years of the same category, such as two dry years. Regarding the timing, the variation also occurred between dry years.

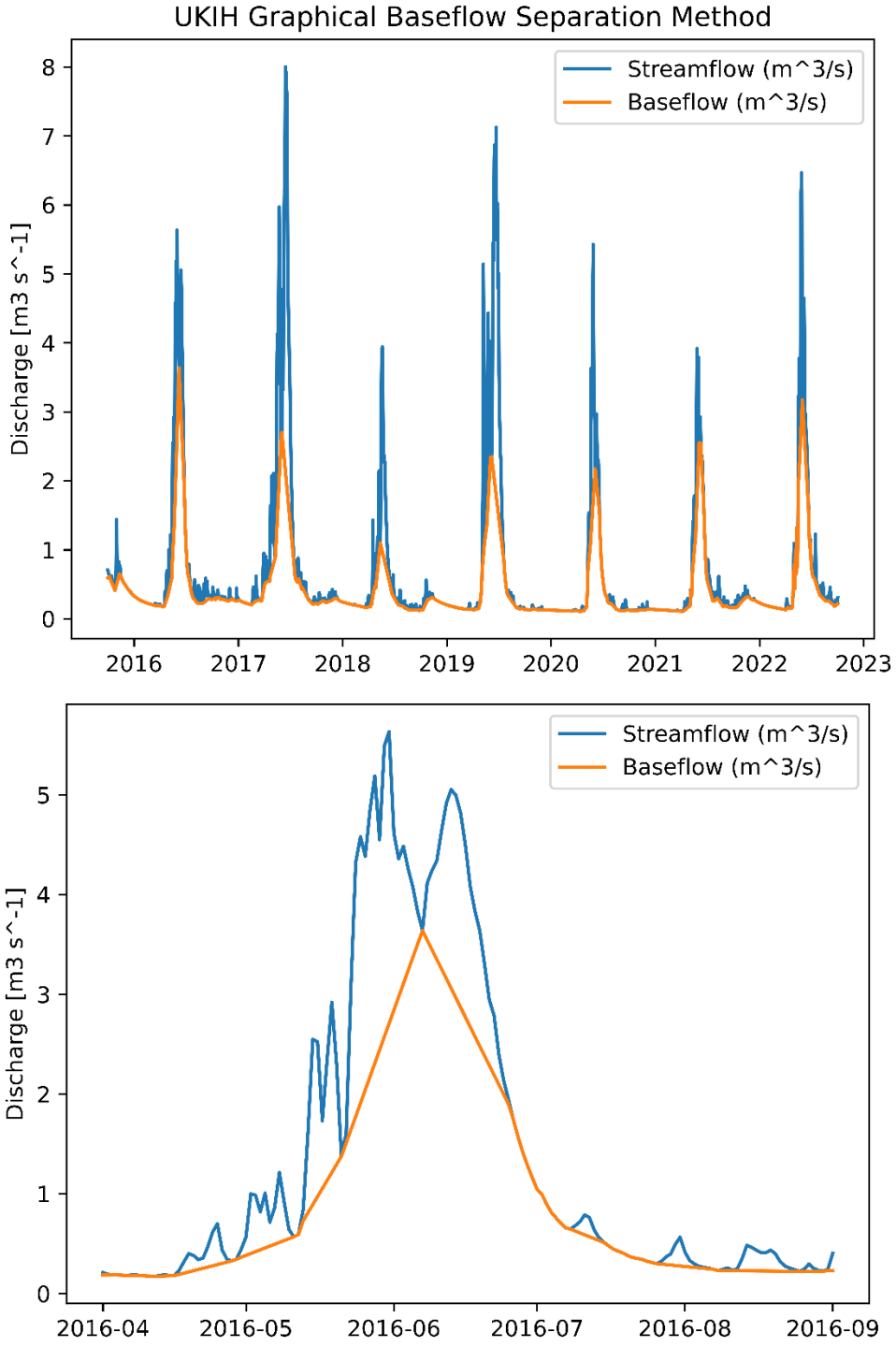
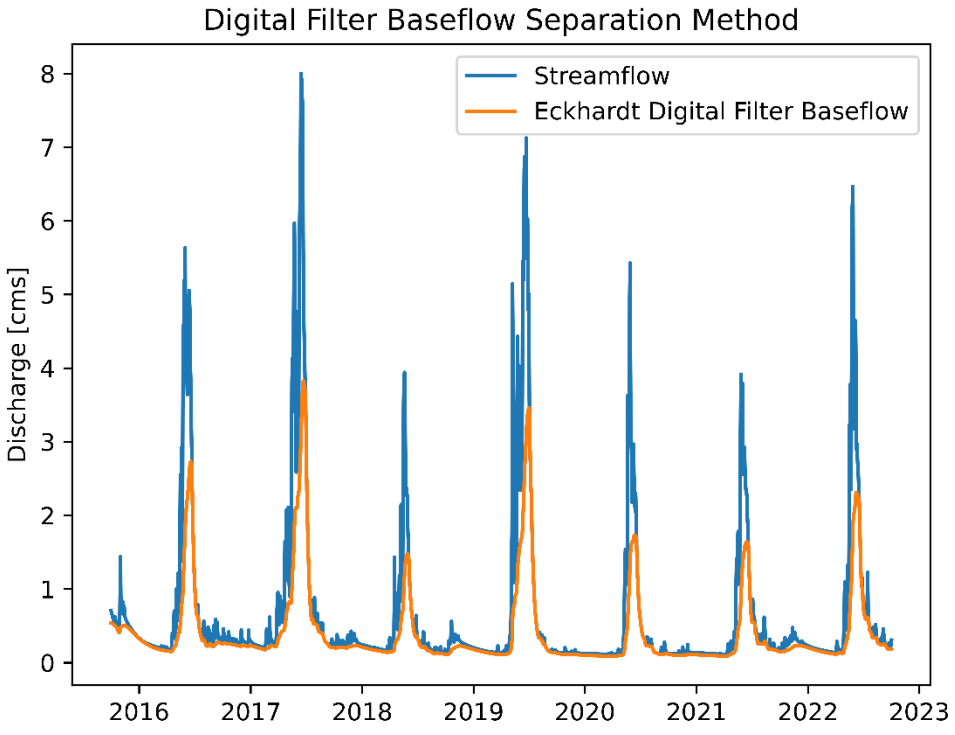


Figure 18: UKIH baseflow separation method. The upper figure shows the entire period and the lower one zooms in to show the separated baseflow in the average year 2016.

4.4 Eckhardt Digital Filter Method

This section describes the resultant baseflow using the digital filter method, specifically the Eckhardt approach (Figure 19). The BFI_{max} value of 0.763 is obtained using the backward filtering technique (Equation 25). The lower plot zooms in 2016 to demonstrate the performance on both short-term and long-term peaks. In 2016, the baseflow was separated for both types of peaks, with the baseflow representing a high ratio of the streamflow during high flow periods and a lower ratio during low flow periods.

More quantitative analysis shows that the maximum estimated baseflow recorded is 3.82 m^3/s, which is around 48% of the maximum streamflow (8 m^3/s). However, the mean baseflow is 0.4 m^3/s, constituting 59% of the average streamflow (0.68 m^3/s). This indicates that during very high streamflow events, the baseflow does not represent a high ratio. And because the average baseflow-to-average streamflow ratio is high, this means that lower peak streamflow events include a higher baseflow ratio.



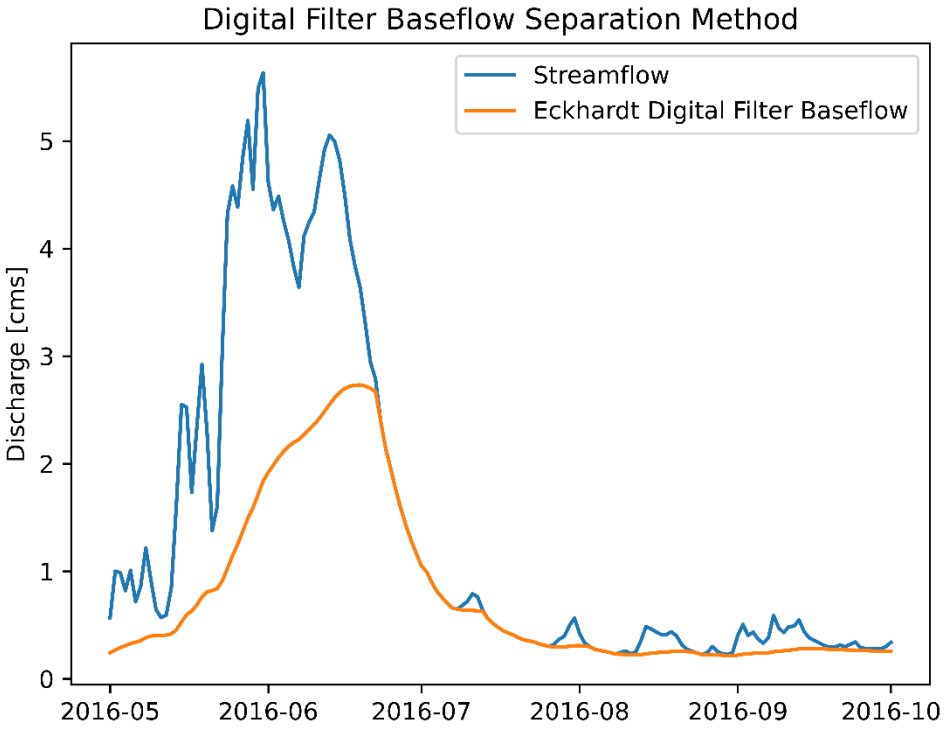


Figure 19: Baseflow Separation using Digital Filter Methods. The BFI_{max} and recession constant used are 0.763 and 0.99 respectively. The upper figure includes the entire period of study, while the lower one is a zoom-in to parts of 2016.

4.5 Conductance Mass Balance (CMB) Method

The streamflow is simulated and obtained from the numerical model, and the specific conductance (SC) is obtained from the 09111250-gage station on USGS. However, the conductivity observations are discrete, with only around 26 measurements available for Coal Creek from 2015 to 2022 (Figure 20). A power relationship between SC and Q is obtained based on those 26 measurements:

$$SC = 155.59 \times Q^{-0.31} \qquad \text{Equation 26}$$

Where SC is the specific conductance in µS/cm and Q is the daily mean discharge in m^3/s. The R^2 value of this power relationship (Figure 21) is around 0.38, indicating that only around 40% of the SC data can be explained by the discharge (Figure 21). Particularly, data with SC lower than 100 µS/cm and higher than 300 µS/cm are the least likely to follow the trendline.

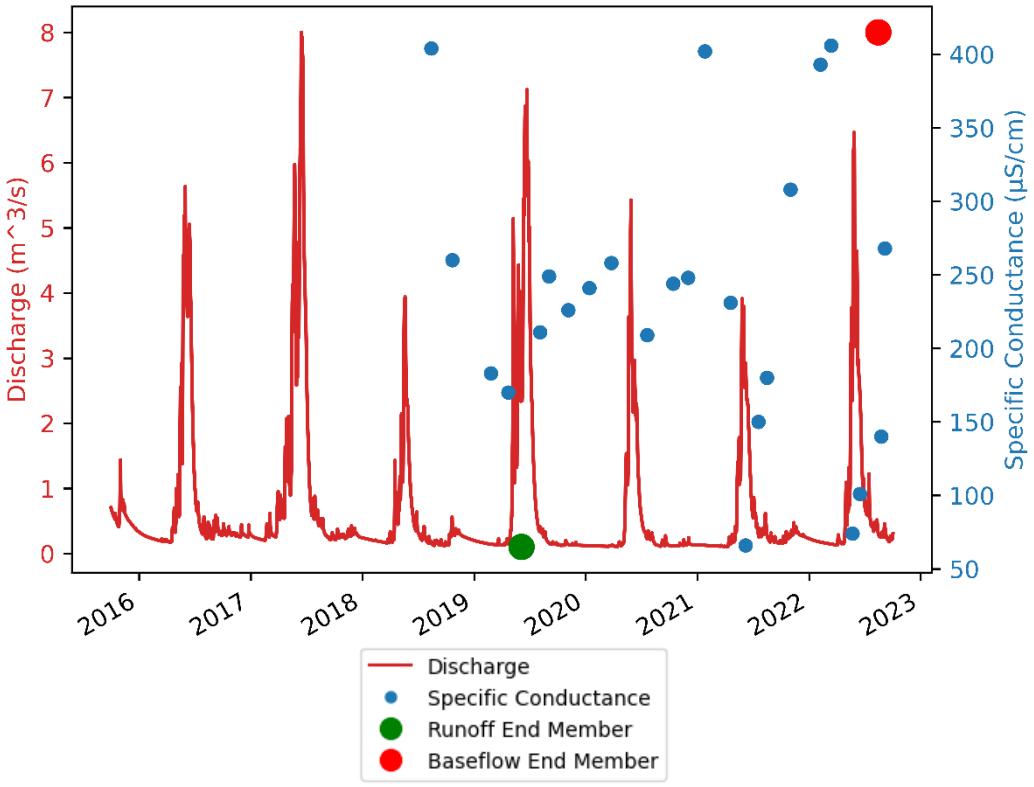


Figure 20: Specific conductance measured at 09111250-gage station extracted from USGS and simulated discharge by the ATS numerical model. The SC is available after 2019. The end members are also shown in this figure

The baseflow end member used in this study (415 µS/cm) is the highest conductance recorded in the summer in Coal Creek on August 15, 2022, when streamflow reached 0.12 m^3/s. The runoff end member is the lowest SC measured for Coal Creek which, is 65 µS/cm on June 05, 2019, when streamflow reached 3.3 m^3/s (Figure 20).

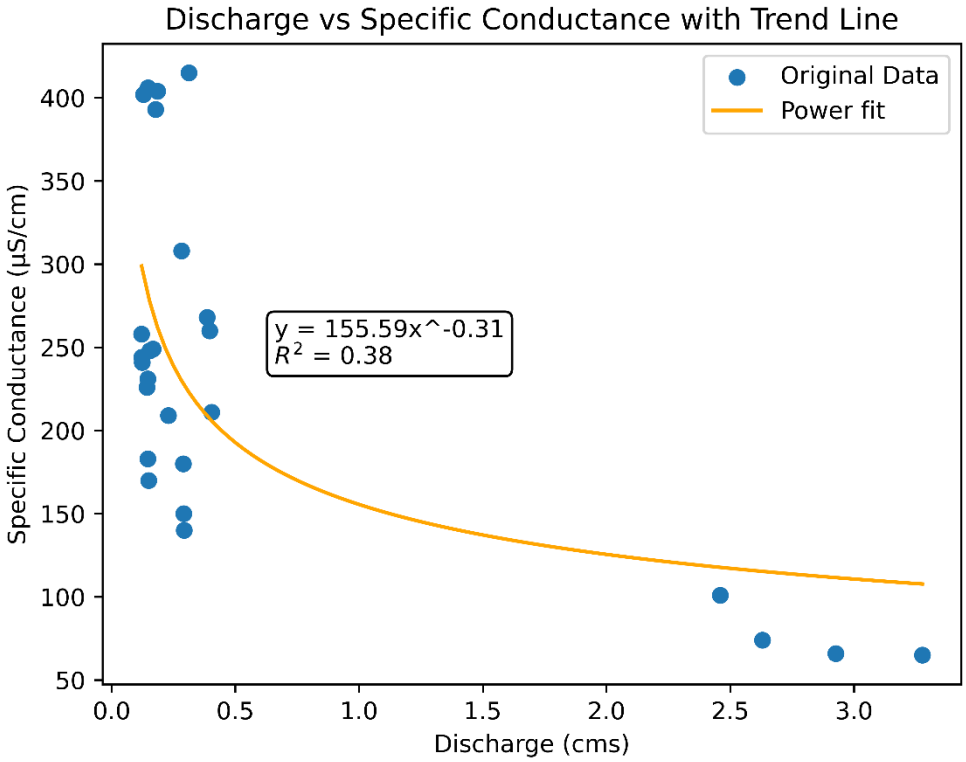


Figure 21: Specific Conductance-Discharge power relationship and the R^2

The developed power law relationship (Equation 26) was used to estimate the SC when no observations were available, e.g., converting the discrete data to continuous data. The estimated SC, as shown in Figure 22, does not capture the extreme values observations, very high and very low specific conductivity during dry and wet years, but only captures the average variation in the SC during an average year. These challenges generate a margin of error when using the estimated SC to estimate the baseflow. Therefore, the end members were extracted from the observed data but not from the estimated SC.

The observed very high SC of around 400 µS/cm in the dry year 2018 is very different from the estimated SC of around 250 µS/cm, although it is one of the highest estimated SC. Conversely, during the highest flow in 2019, the SC was measured at around 60 µS/cm but estimated to be around 100 µS/cm. However, all other observations between 100 and 250 µS/cm are closely estimated by the power functions, validating the estimation method. The

main reason for this discrepancy is that data with SC between 250 and 350 µS/cm are better interpolated.

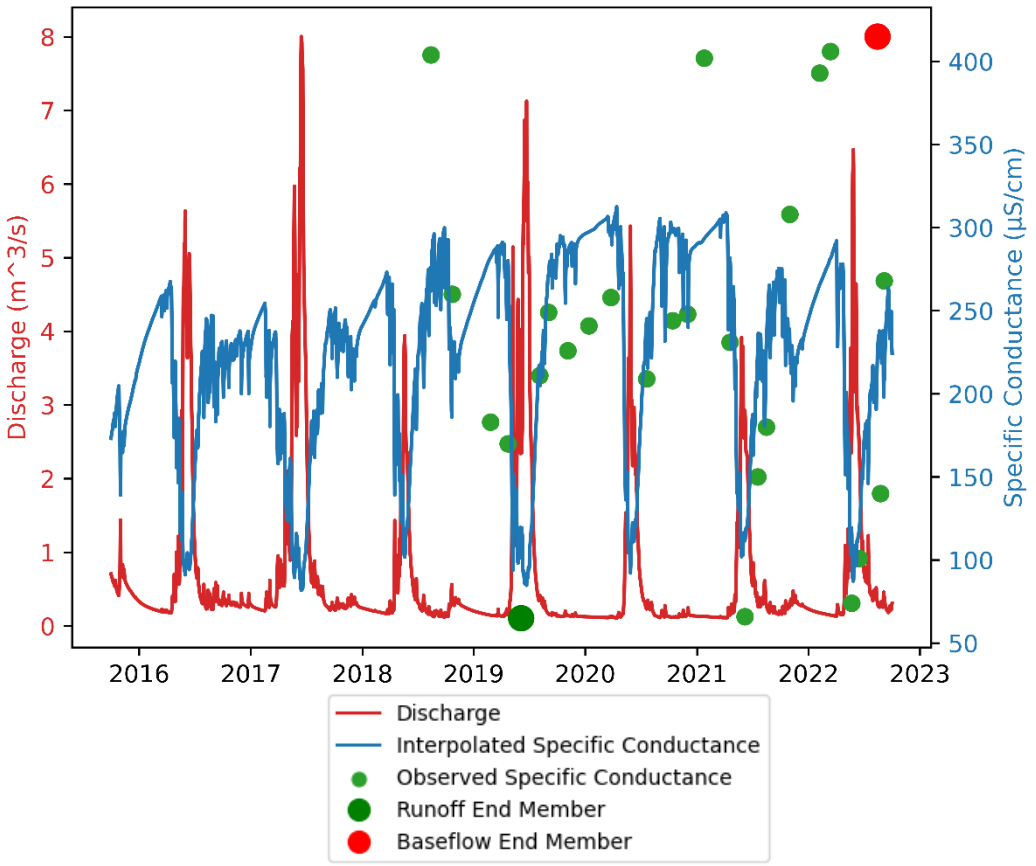
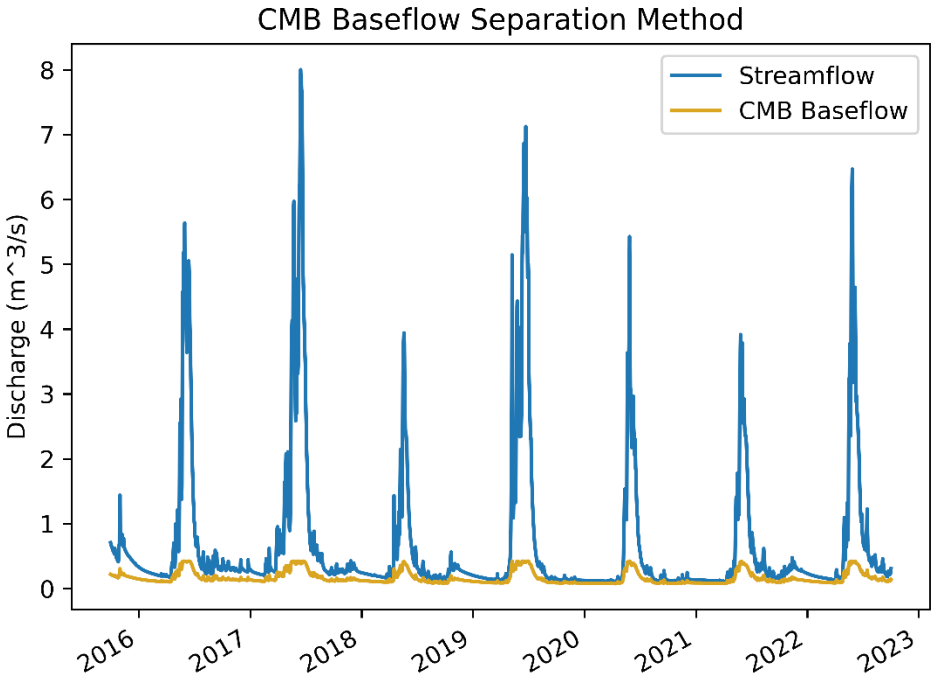


Figure 22: The CMB Method. The blue curve is the estimated daily specific conductivity using the power function with discharge provided in Figure 21. The green dots are the observed SC. All the observed SC between 100 and 250 µS/cm significantly matching the estimated SC. The values below and above this range are not well estimated.

Using Equation 25, the daily baseflow is calculated and then plotted as shown in Figure 23. From 2016 to 2019, the baseflow is lower than the low flow but shows some variation with changes in the streamflow, most likely due to precipitation. Although the baseflow during high flow follows some fluctuations that could be linked to infiltration and snowmelt processes impacting streamflow, the baseflow is significantly low. Furthermore, the difference in streamflow from one year to another is not reflected in a significant difference in the baseflow. For example, the baseflow ratio of 2017 over 2021 is 0.98, not indicating a substantial difference between dry and wet years. In contrast, the streamflow

ratio of the same years, 0.5, indicates a significant difference between the two years (Table 3).

This lack of a significant difference in baseflow between dry and wet years could be linked to the accumulation of errors starting from the interpolation of the continuous SC. To determine if the interpolation also affected the baseflow estimation not considered as the yearly highest and lowest streamflow, a statistical analysis and comparison is provided in the next section for all the methods.



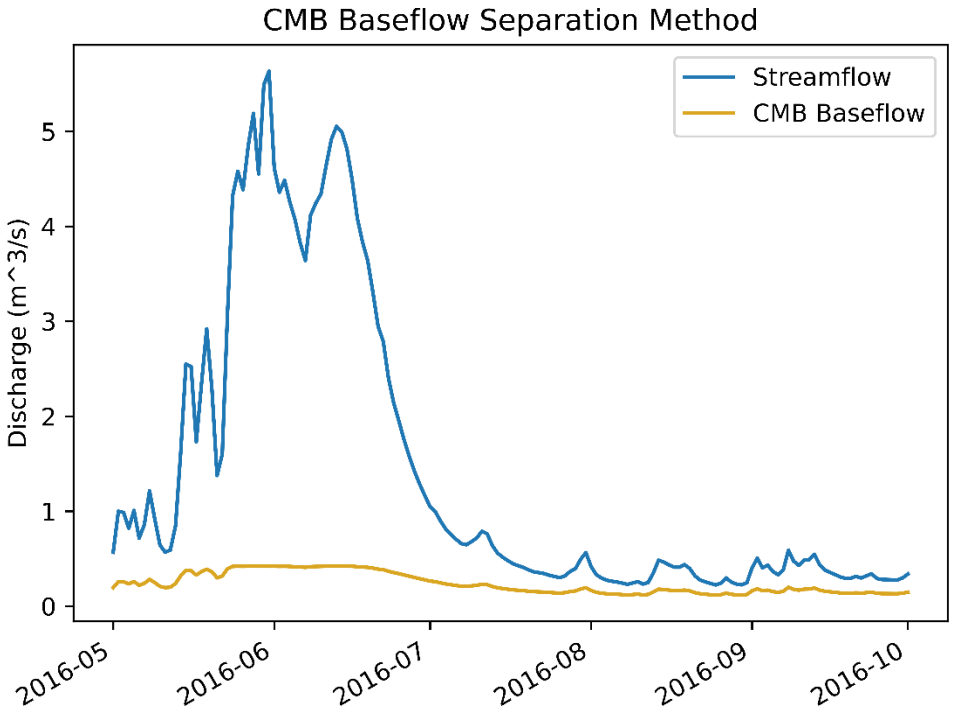


Figure 23: Discharge Hydrograph with Baseflow Separation using CMB Methods.

Table 3: Highest streamflow and baseflow of wet (2017) and dry (2022) years.

	Highest baseflow (m3/s)	Date of highest baseflow	Highest streamflow (m3/s)	Date of highest streamflow
2017	0.42	09/06/2017	8.00	15/06/2017
2021	0.41	27/05/2021	3.92	27/05/2021
2021/2017 ratio	0.98		0.49	

4.6 Baseflow Separation Methods Evaluation

Using the numerical baseflow as the control, the performance of the graphical, digital filter, and CMB baseflow separation methods are evaluated using common metrics. The quantitative statistics are provided in Table 4 andTable *5* including the RMSE, NSE, and KGE for each method baseflow with the numerical method baseflow estimation. In addition, it includes the minimum, maximum, mean, and SD of each method baseflow. The fraction of

baseflow (BFI) was also added for each method. The BFI is the percentage of total baseflow volume to total streamflow volume.

During the overall period (Figure 24), different methods of baseflow separation show varied performance compared to the numerical model baseflow. both UKIH graphical and Eckhardt digital filter methods performed relatively well with high mKGE (0.72 and 0.68, respectively) and NSE (0.58 and 0.7, respectively) values. On the other hand, both the PH graphical and CMB methods did not perform satisfactorily in estimating baseflow with both mKGE and NSE values less than 0.3. Among them, the PH graphical method has consistently overestimated baseflow with an average BFI of 85%, whereas the CMB method has consistently underestimated baseflow even during low-flow periods with an average BFI of 24% as compared to an average BFI of 53% for the numerical method. This is further evidenced by the bias term β from the decomposed mKGE, where $\beta<1$ indicates underestimation (i.e., mean baseflow is smaller than the mean of the reference) and $\beta>1$ indicates overestimation (i.e., mean baseflow is larger than the mean of the reference). CMB baseflow shows the smallest β (=0.41) and the PH graphical baseflow shows the largest β (=1.55).

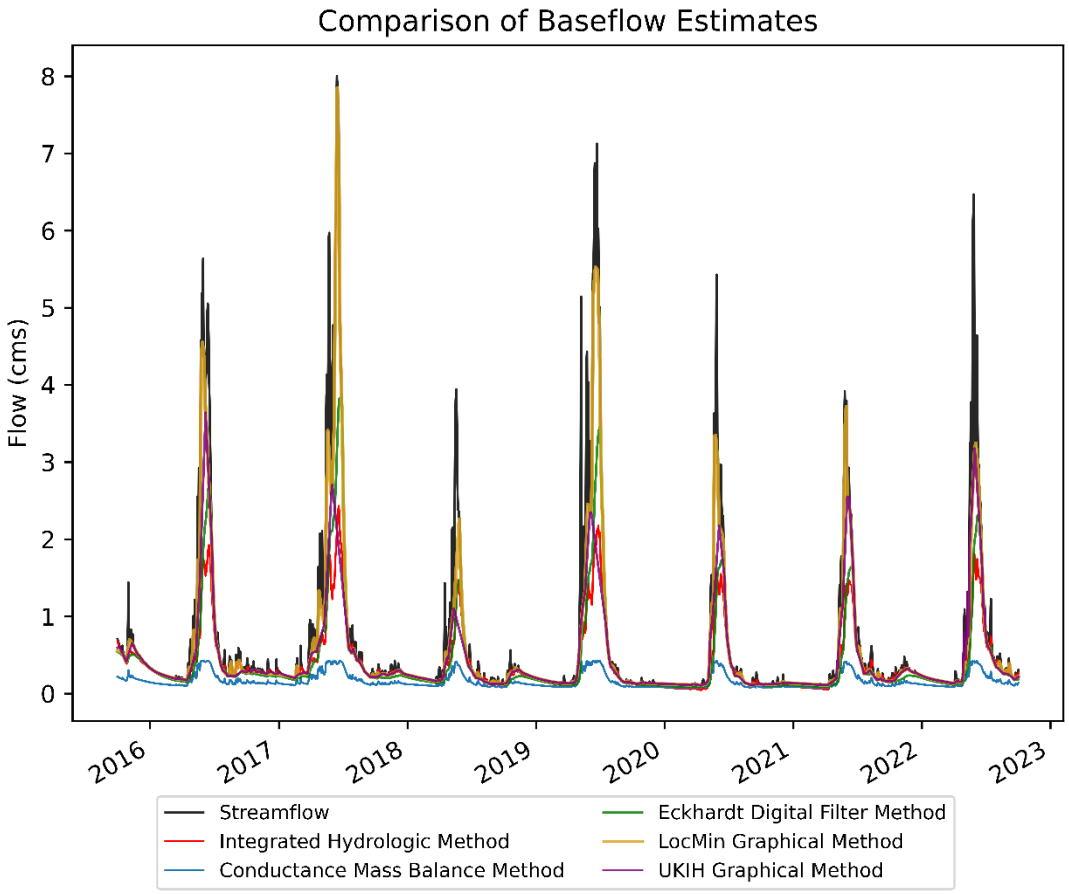


Figure 24: Baseflow curves from all the used methods in this study.

Table 4: Statistics of all the used method in this study. The RMSE, NSE, and KGE of baseflow from each method with the numerical hydrological model baseflow are calculated. The fraction of baseflow (ratio of baseflow volume to streamflow volume).

Method	Min (m3/s)	Max (m3/s)	Mean (m3/s)	SD	volume (*10^8 m3)	BFI (%)
Numerical	0.05	2.43	0.4	0.44	0.79	53
PH Graphical	0.1	7.85	0.58	0.96	1.28	85
UKIH Graphical	0.1	3.64	0.46	0.59	1.02	68
Eckhardt Digital	0.09	3.82	0.42	0.6	0.93	62
CMB	0.07	0.42	0.16	0.1	0.36	24

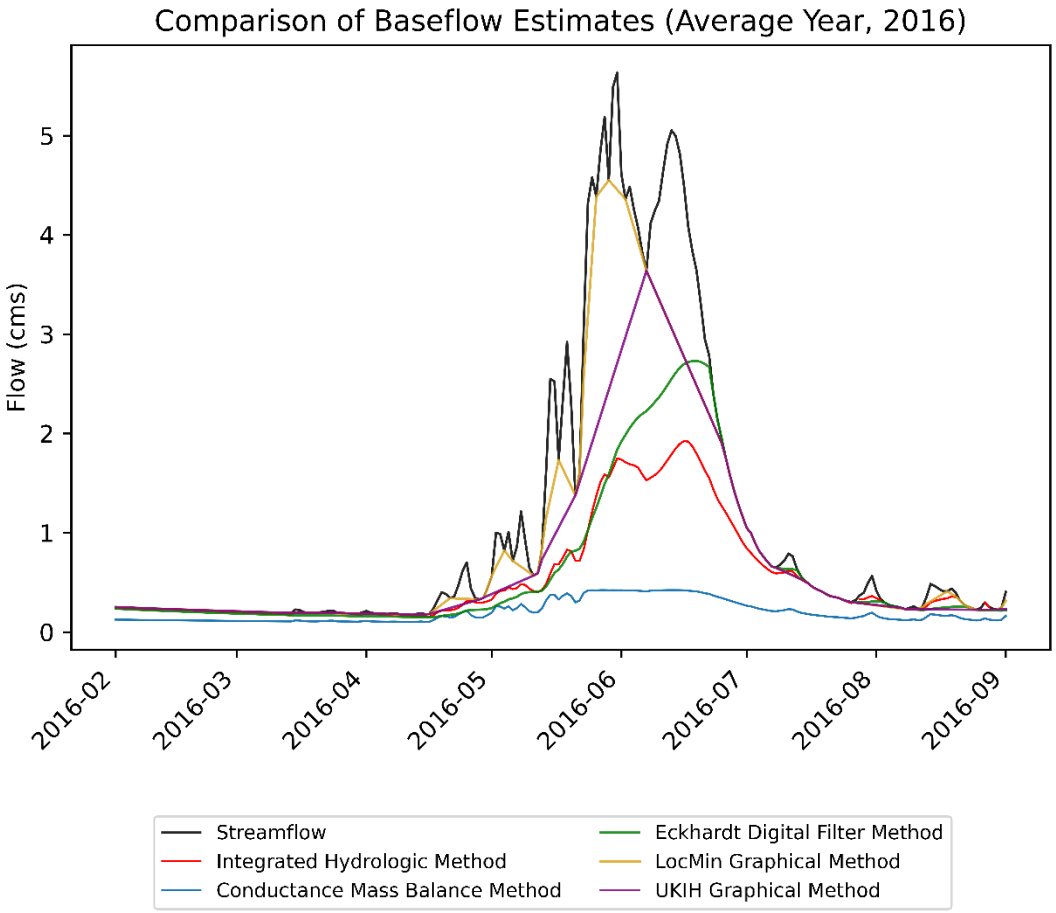
Table 5: mKGE components, NSE, and RMSE of each method compared to the numerical baseflow

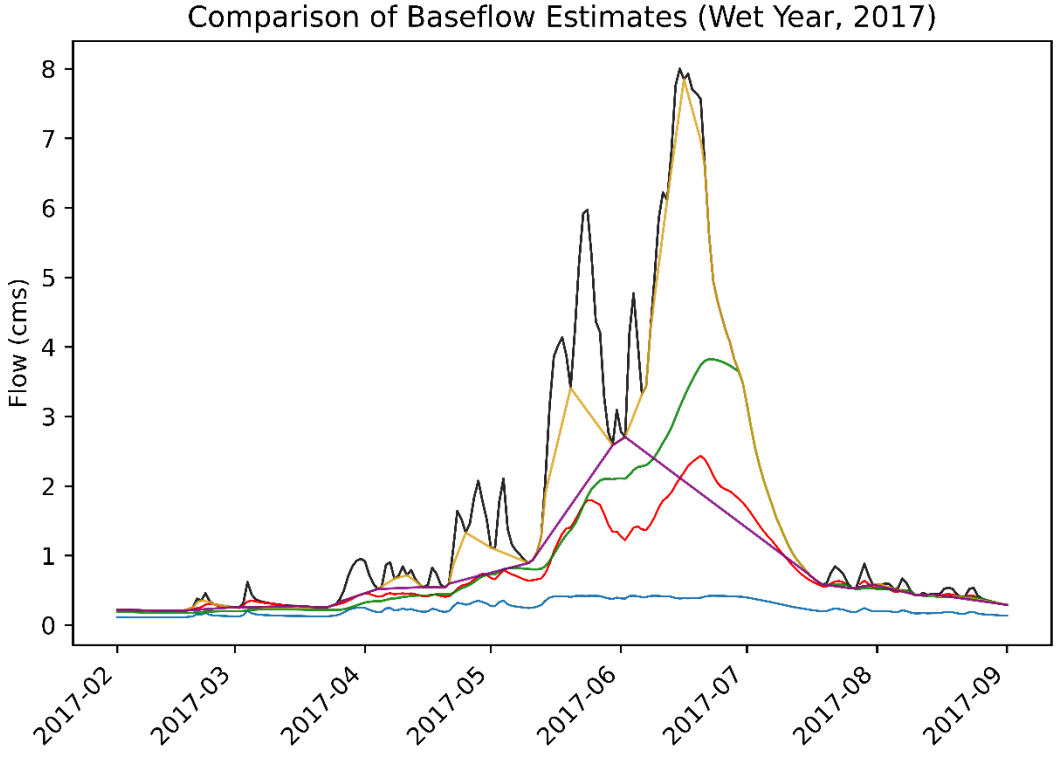
Method	KGE r	KGE γ	KGE β	mKGE	NSE	RMSE
PH Graphical	0.94	1.49	1.55	0.26	-1.24	0.66
UKIH Graphical	0.92	1.16	1.21	0.72	0.58	0.29
Eckhardt Digital	0.96	1.3	1.1	0.68	0.7	0.27
CMB	0.94	0.54	0.41	0.25	0.09	0.42

In the year-by-year analysis (Figure 25 and Table 6), typical dry, average, and wet years are selected to show a detailed comparison of the two best performing methods: the Eckhardt digital filter and the UKIH graphical baseflow separation methods. Generally, the Eckhardt baseflow closely aligns with the numerical baseflow across different years, demonstrating better performance in average (e.g., mKGE = 0.85 in 2020) and dry years (e.g., mKGE = 0.90 in 2021) but worse performance in wet years (e.g., mKGE = 0.51 in 2017). The performance of the UKIH method is highly variable across different years. It usually performs well in wet years (e.g., mKGE = 0.83 and 0.84 for 2019 and 2017, respectively) but not as well in other years. This is likely due to the inherent limitation of the method, which uses a short 5-day window to find the local minima of the hydrograph. Consequently, the UKIH method does not perform well when the streamflow hydrograph exhibits a unimodal flow peak. For example, during the dry year of 2021, the UKIH method does not clearly separate baseflow, which follows the general unimodal shape of the streamflow. Consequently, the UKIH method significantly overestimates baseflow with a BFI of 82%, which is 29.3% higher than the numerical baseflow. Additionally, the shape of the UKIH baseflow is often arbitrary and sometimes unnatural compared to the Eckhardt baseflow. For example, during the wet year

of 2017, the peak UKIH baseflow occurred in early June, around 17 days ahead of the peak streamflow. Furthermore, the UKIH baseflow shows a single peak instead of the double peaks seen in the streamflow. In comparison, the Eckhardt baseflow shows a smoother curve with peak baseflow following peak streamflow, although it may also struggle to capture multiple baseflow peaks. For example, in the average year of 2016, the Eckhardt baseflow shows a single peak in late June with a monotonic increase, missing another peak in early June. This is further evidenced by the correlation term r (i.e., timing) of the decomposed mKGE. The UKIH baseflow shows the lowest correlation (r=0.92) whereas the Eckhardt baseflow shows the highest correlation (r=0.96).

To compare baseflow interannual variability, yearly BFI values were calculated for the Eckhardt digital filter and UKIH methods (Table 6). The yearly UKIH BFI shows a higher variability similar to the numerical BFI, whereas the Eckhardt BFI displays a more stable pattern that does not show a significant reduction in BFI in dry years, which may underestimate the important contribution of baseflow to streamflow during drought.


Comparison of Baseflow Estimates (Average Year, 2016)
Flow (cms)
0
1
2
3
4
5
2016-02
2016-03
2016-04
2016-05
2016-06
2016-07
2016-08
2016-09
Streamflow
Integrated Hydrologic Method
Conductance Mass Balance Method
Eckhardt Digital Filter Method
LocMin Graphical Method
UKIH Graphical Method


Comparison of Baseflow Estimates (Wet Year, 2017)
Flow (cms)
0
1
2
3
4
5
6
7
8
2017-02
2017-03
2017-04
2017-05
2017-06
2017-07
2017-08
2017-09

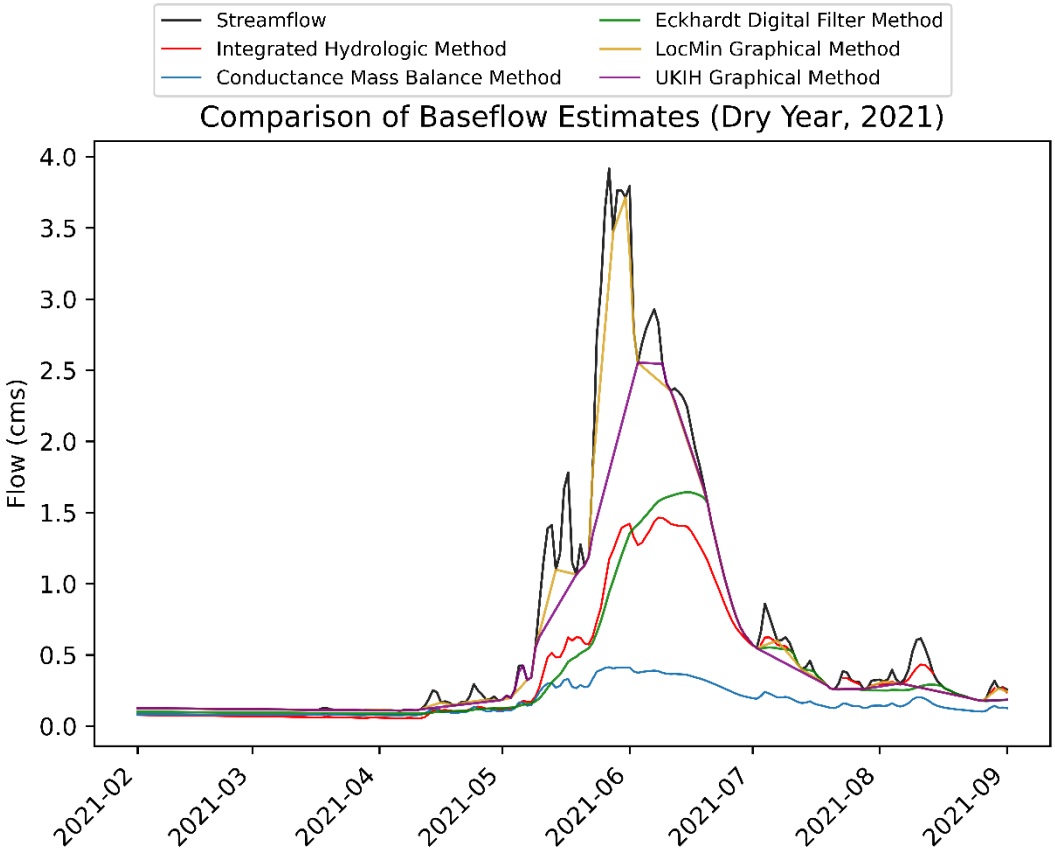


Figure 25: All separation methods used in this study in addition to the numerical method. The upper, middle, and lower plots show the results of the periods from 02-2016 to 09-2016, 02-2017 to 09-2017, and 02-2021 to 09-2021 respectively.

Table 6: Yearly BFI percentage, mKGE, and NSE of the digital filter and UKIH methods in addition to the numerical method

season	Water Year	BFI % (numerical model ATS)	BFI % (Eckhardt)	BFI % (UKIH)
dry	2018	61.0	66.7	66.3
	2021	52.7	63.6	82.0
wet	2017	49.6	64.3	56.5
	2019	46.8	61.9	56.2
average	2016	56.6	66.1	74.7

	2020	51.5	63.2	74.2

		mKGE		NSE	
season	Water Year	Eckhardt	UKIH	Eckhardt	UKIH
average years	2016	0.61	0.49	0.62	0.10
	2020	0.85	0.66	0.90	0.78
wet years	2017	0.51	0.84	0.22	0.74
	2019	0.61	0.83	0.50	0.77
dry years	2018	0.82	0.79	0.94	0.77
	2021	0.90	0.55	0.94	0.42

4.7 Uncertainties, Limitations, and Future Work

Starting from the quote "All models are wrong, but some are useful" (Box and Draper, 1987), the application of numerical models comes with inherent uncertainties in terms of parameters and assumptions. Even after calibration, the numerical model shows the discrepancy between the simulated and observed streamflow. For example, the simulated low flow is overestimated during dry periods (Figure 5). To ensure a fair comparison and minimize the uncertainties, all estimated baseflow are based on the simulated streamflow in their methods. Additionally, due to the limitation of the model, simulated baseflow is not a direct output but rather based on the mass balance in the subsurface (i.e., baseflow = net

groundwater discharge = total groundwater discharge – re-exfiltration). In the analysis, we developed a linear relationship between re-exfiltration and total groundwater discharge based on results during dry periods. The re-exfiltration during wet periods was then estimated by the linear regression, which may introduce extrapolation errors. Future work could take a stream-centric mass balance approach for estimating the baseflow by delineating the streambed region in the model (Shuai et al 2023).

Both the PH and UKIH graphical methods can be sensitive to the choice of window size (i.e., N) for determining the local minima or turning points for the baseflow hydrograph. As Miller et al (2015) showed in their analysis of the optimal N value in the UKIH method, the BFI value decreased with increasing N value and a larger N value (> 10 days) was suggested for use in snow-dominated areas. In our study, a 5-day window size was used in separating baseflow using the UKIH method, causing the estimated baseflow to be high. Unfortunately, the N value used in the PH graphical method is typically determined based on the catchment area. As a result, a small N is often used in small catchments such as the Coal Creek Watershed. Future work should be focused on optimizing the N value before applying the graphical methods.

The digital filter method relies on two parameters: the recession constant and BFI_{max}. In this study, the recession constant was calculated first and then used to calculate BFI_{max}. Thus, uncertainties arose from the recession analysis test and the assumption that streamflow on days without rainfall and snowmelt represents the recession streamflow, leading to uncertainties in BFI_{max} calculations and, consequently, the baseflow estimation. To test the sensitivity of the estimated baseflow to BFI_{max}, BFI_{max} was varied by +/- 0.1, ranging from 0.663 to 0.863. The estimated baseflow increases with increasing BFI_{max} as evident in

Figure 26. For a small change of BFI_{max} from 0.763 to 0.663, the baseflow decreased significantly (Table 7).

The CMB method relies on continuous SC observation to estimate baseflow. In our site, only 37 discrete SC samples are available, which may cause large discrepancy in estimated baseflow (Miller et al., 2015). In our study, we developed a regression model to estimate the continuous SC based on the discrete SC following the Miller et al. (2015) approach. A low R^2 was obtained from the power inverse relationship, which cause large uncertainties in estimating the continuous SC data. Furthermore, the limited SC data may impact the end member selection since the discrete sample did not cover the time series of the streamflow, especially during high flows. To test the sensitivity of baseflow and runoff end member, we varied the end member values by +/- 10%. The results showed that the 58.5 runoff end member has the highest peak and the 71.5 runoff end member has the lowest peak while the baseflow end member curves are located between them. Thus, baseflow is more sensitive to runoff end members than baseflow end members (Figure 27).

Lastly, our findings are based on a single catchment with a relatively small area, which may not represent other snow-dominated systems with different characteristics. However, due to the high computational cost associated with calibrating integrated hydrologic models, the numerical method could not be easily applied across various catchments. Future work could leverage the advancement in machine learning to develop surrogate models for estimating baseflow in snow-dominated systems.

A summary of all used methods is provided in Table 8.

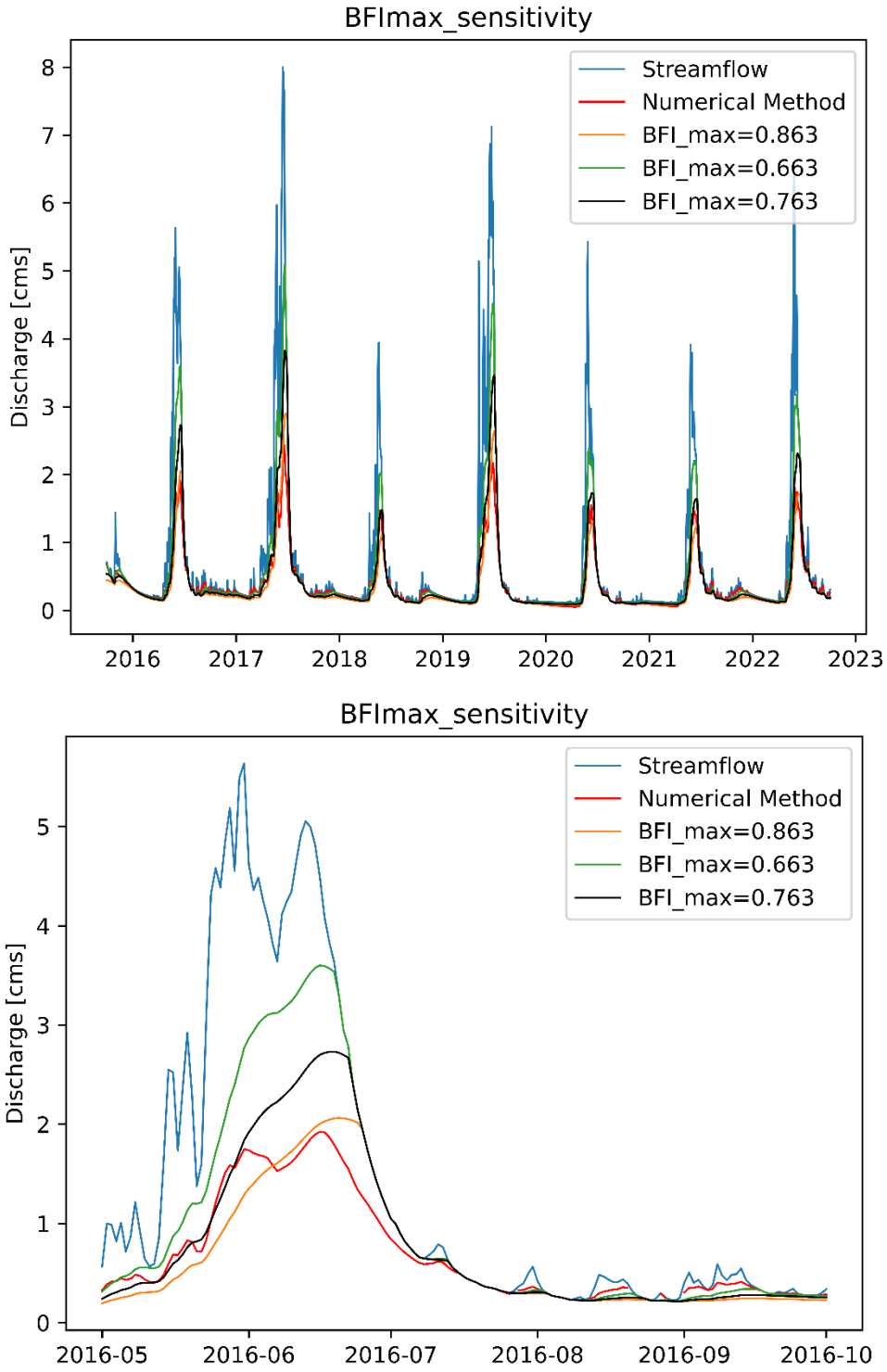


Figure 26: Sensitivity test for Eckhardt method by changing the BFImax values.

Table 7: The BFI percentage obtained from the Eckhardt method using different BFImax.

BFI_max	0.663	0.763	0.863	Numerical Model

BFI (%)	74	62	53	53

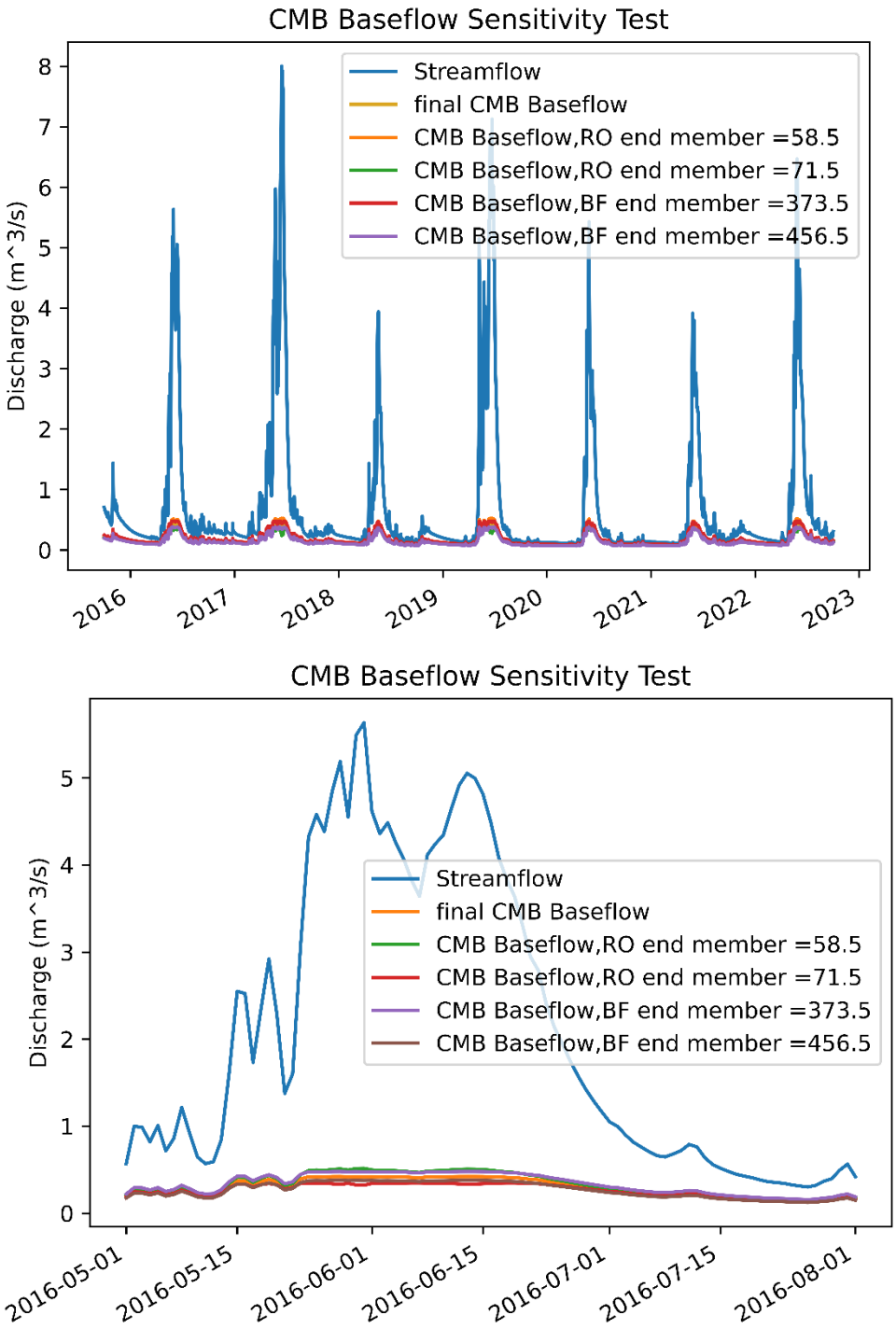


Figure 27: Sensitivity test for CMB by changing the runoff or baseflow end members. Only one of them is changed at a time.

Table 8: Brief description of all the used methods in this study

Method	Assumptions	Advantages	Limitations	Recommendations for baseflow separation
Numerical (e.g., ATS)	- Re-infiltration equals re-exfiltration. - A linear relationship exists between the total exfiltration and the re-exfiltration	- Physically based - All storage and water fluxes are quantifiable	- Difficult to calibrate - Computationally expensive - Data intensive	- Could be used for baseflow separation regardless of catchment characteristics
PH Graphical	- Event length (i.e., 2N) can be calculated from the catchment area	- Easy to apply and automate - Minimal data requirement	- Event length has no physical connection to catchment characteristics - Could not separate baseflow when peak flow	- Not suitable for baseflow separation in small, snow-dominated catchments - Baseflow tends to be overestimated
UKIH Graphical	- Event length is typically fixed at five days	- Easy to apply and automate - Minimal data requirement	- Event length has low physical connection to catchment characteristics - Not effective in separating hydrographs with a unimodal shape	- Not recommended for snow-dominated catchments
Eckhardt Digital Filter	- Baseflow is the low-frequency component of the hydrograph - Linear reservoir assumption - The ratio of groundwater recharge to quick flow is constant over time	- Easy to apply and automate - Minimal data requirement	- Fitting parameters have low physical connection to catchment characteristics	- Shows promise for the qualitative estimation of baseflow in snow-dominated catchments, though the relative importance of baseflow during dry periods may be underestimated
CMB	- Baseflow and runoff end members are distinctively different and often constant - Contributions from other end members are negligible - An inverse power function exists	- SC data is relatively easy to obtain in the field	- Uncertainties in end members - End members may be spatially and temporally variable - Not all sites have SC measurements	- The estimated baseflow is highly sensitive to the choice of end members, posing a challenge for applying this method without detailed knowledge about the site - Baseflow during dry periods tend to be underestimated

	between the SC and streamflow.			

5 CONCLUSION

An integrated hydrologic model is developed to quantify the baseflow component of the streamflow in a snow-dominated watershed from 2015 to 2022. The calculated baseflow was then used as a control to evaluate four commonly used conceptual baseflow separation methods: the PH graphical, the UKIH graphical, the Eckhardt digital filter, and the CMB methods. Overall, the Eckhardt digital filter method is promising for the qualitative estimation of baseflow in snow-dominated catchments, although it may underestimate the relative importance of baseflow contribution to streams during dry years. The UKIH graphical method tends to overestimate baseflow in snow-dominated catchments, particularly when the hydrograph exhibits a unimodal peak, which is typical in snow-dominated watersheds. The PH graphical method significantly overestimates baseflow regardless of the hydrograph's shape due to the short event window (e.g., typically less than five days) used in determining local minima (i.e., the turning point). Conversely, the CMB method significantly underestimates baseflow due to limited SC observation and uncertainties in end members, and further testing is warrantied to determine its suitability for baseflow separation in snow-dominated catchments.

Integrated hydrologic models, when calibrated, provide a physical and quantitative way to estimate baseflow. However, their applications may be limited due to their high

computational costs associated with model calibration. Future studies should focus on improving existing automated baseflow separation methods by optimizing the fitting parameters and relating them to watershed hydroclimatic characteristics. Additionally, future work should focus on the evaluation of baseflow separation methods in catchments that cover a range of diverse hydroclimate conditions.

6 DATA AVAILABILITY STATEMENT

All data and codes used in the analysis are available under the following link:

https://github.com/hydroaggie/CoalCreek

7 REFERENCES

Barlow, P. M., McHugh, A. R., Kiang, J. E., Zhai, T., Hummel, P., Duda, P., & Hinz, S. (2022). *U.S. Geological Survey Hydrologic Toolbox — A graphical and mapping interface for analysis of hydrologic data: U.S. Geological Survey Techniques and Methods*. https://doi.org/10.3133/tm4D3

Bhardwaj, S. S., Jha, M. K., & Uniyal, B. (2024). Assessing Efficacy of Baseflow Separation Techniques in a Himalayan River Basin, Northern India. *Environmental Processes*, *11*(1). https://doi.org/10.1007/s40710-024-00680-z

Bosch, D. D., Arnold, J. G., Allen, P. G., Lim, K. J., & Park, Y. S. (2017). Temporal variations in baseflow for the Little River experimental watershed in South Georgia, USA. *Journal of Hydrology: Regional Studies*, *10*, 110–121. https://doi.org/10.1016/j.ejrh.2017.02.002

Box, G. E., & Draper, N. R. (1987). Empirical model-building and response surfaces. John Wiley & Sons.

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

Caine, N. (1989). Hydrograph separation in a small alpine basin based on inorganic solute concentrations. *Journal of Hydrology*, *112*(1–2), 89–101. https://doi.org/10.1016/0022-1694(89)90182-0

Caissie, D., Pollock, T. L., & Cunjak, R. A. (1996). Variation in stream water chemistry and hydrograph separation in a small drainage basin. Journal of Hydrology, 178(1–4), 137–157. https://doi.org/10.1016/0022-1694(95)02806-4

Chapman, T. G. (1991). Comment on "Evaluation of automated techniques for base flow and recession analyses" by R. J. Nathan and T. A. McMahon. In *Water Resources Research* (Vol. 27, Issue 7, pp. 1783–1784). https://doi.org/10.1029/91WR01007

Collischonn, W., & Fan, F. M. (2013). Defining parameters for Eckhardt's digital baseflow filter. Hydrological Processes, 27(18), 2614–2622. https://doi.org/10.1002/hyp.9391

Cook, P. G., Wood, C., White, T., Simmons, C. T., Fass, T., & Brunner, P. (2008). Groundwater inflow to a shallow, poorly-mixed wetland estimated from a mass balance of radon. Journal of Hydrology, 354(1–4), 213–226. https://doi.org/10.1016/J.JHYDROL.2008.03.016

Delottier, H., Therrien, R., Young, N. L., & Paradis, D. (2022). A hybrid approach for integrated surface and subsurface hydrologic simulation of baseflow with Iterative Ensemble Smoother. Journal of Hydrology, 606, 127406. https://doi.org/10.1016/J.JHYDROL.2021.127406

Dinçer, T., Payne, B. R., Florkowski, T., Martinec, J., & Tongiorgi, E. (1970). Snowmelt runoff from measurements of tritium and oxygen-18. *Water Resources Research*, *6*(1), 110–124. https://doi.org/10.1029/WR006i001p00110

Eckhardt, K. (2005). How to construct recursive digital filters for baseflow separation. *Hydrological Processes*, *19*(2), 507–515. https://doi.org/10.1002/hyp.5675

Eckhardt, K. (2008). A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. *Journal of Hydrology*, *352*(1–2), 168–173. https://doi.org/10.1016/J.JHYDROL.2008.01.005

Godsey, S. E., Kirchner, J. W., & Tague, C. L. (2014). Effects of changes in winter snowpacks on summer low flows: Case studies in the Sierra Nevada, California, USA. *Hydrological Processes*, *28*(19), 5048–5064. https://doi.org/10.1002/hyp.9943

Gonzales, A. L., Nonner, J., Heijkers, J., & Uhlenbrook, S. (2009a). Comparison of different base flow separation methods in a lowland catchment. In *Hydrol. Earth Syst. Sci* (Vol. 13). www.hydrol-earth-syst-sci.net/13/2055/2009/

Gonzales, A. L., Nonner, J., Heijkers, J., & Uhlenbrook, S. (2009b). Comparison of different base flow separation methods in a lowland catchment. In *Hydrol. Earth Syst. Sci* (Vol. 13). www.hydrol-earth-syst-sci.net/13/2055/2009/

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving

hydrological modelling. Journal of Hydrology, 377(1–2), 80–91. https://doi.org/10.1016/J.JHYDROL.2009.08.003

Gustard, A., Bullock, A., & Dixon, J. M. (1992). Low flow estimation in the United Kingdom. Institute of Hydrology.

Hall, F. R. (1968). Base-Flow Recessions—A Review. Water Resources Research, 4(5), 973–983. https://doi.org/10.1029/WR004i005p00973

Hayashi, M., Quinton, W. L., Pietroniro, A., & Gibson, J. J. (2004). Hydrologic functions of wetlands in a discontinuous permafrost basin indicated by isotopic and chemical signatures. *Journal of Hydrology*, *296*(1–4), 81–97. https://doi.org/10.1016/J.JHYDROL.2004.03.020

Huyck, A. A. O., Pauwels, V. R. N., & Verhoest, N. E. C. (2005). A base flow separation algorithm based on the linearized Boussinesq equation for complex hillslopes. Water Resources Research, 41(8), 1–18. https://doi.org/10.1029/2004WR003789

Institute of Hydrology, 1980a, Low flow studies: Wallingford, Oxon, United Kingdom, Report No. I, 41 p.

Institute of Hydrology, 1980b, Low flow studies: Wallingford, Oxon, United Kingdom, Report Wallingford, UK: Institute of Hydrology. http://nora.nerc.ac.uk/id/eprint/9093/.

Jiang, P., Shuai, P., Sun, A., Mudunuru, M. K., & Chen, X. (2023). Knowledge-informed deep learning for hydrological model calibration: an application to Coal Creek Watershed in Colorado. *Hydrology and Earth System Sciences*, *27*(14), 2621–2644. https://doi.org/10.5194/hess-27-2621-2023

Julander, R. P., & Clayton, J. A. (2018). Determining the proportion of streamflow that is generated by cold season processes versus summer rainfall in Utah, USA. Journal of Hydrology: Regional Studies, 17, 36-46.

Kendall, C., & Caldwell, E. A. (1998). Fundamentals of Isotope Geochemistry. Isotope Tracers in Catchment Hydrology, 51–86. https://doi.org/10.1016/B978-0-444-81546-0.50009-4

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *424–425*, 264–277. https://doi.org/10.1016/J.JHYDROL.2012.01.011

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

Konrad, C. P. (2022). *BFS-A Non-Linear, State-Space Model for Baseflow Separation and Prediction*. https://doi.org/10.3133/sir20225114

Koskelo, A. I., Fisher, T. R., Utz, R. M., & Jordan, T. E. (2012). A new precipitation-based method of baseflow separation and event identification for small watersheds (<50 km2). *Journal of Hydrology*, *450–451*, 267–278. https://doi.org/10.1016/J.JHYDROL.2012.04.055

Ledoux, E., Girard, G., & Villeneuve, J. P. (1984). PROPOSITION D'UN MODELE COUPLE POUR LA SIMULATION CONJOINTE DES ECOULEMENTS DE SURFACE ET DES ECOULEMENTS SOUTERRAINS SUR UN BASSIN HYDROLOGIQUE. *HOUILLE BLANCHE*, *39*(1–2). https://doi.org/10.1051/lhb/1984005

Lee, G., Shin, Y., & Jung, Y. (2014). Development of web-based RECESS model for estimating baseflow using SWAT. *Sustainability (Switzerland)*, *6*(4), 2357–2378. https://doi.org/10.3390/su6042357

Lott, D. A., & Stewart, M. T. (2013). A Power Function Method for Estimating Base Flow. *GroundWater*, *51*(3), 442–451. https://doi.org/10.1111/j.1745-6584.2012.00980.x

Matsubayashi, U., Velasquez, G. T., & Takagi, F. (1993). Hydrograph separation and flow analysis by specific electrical conductance of water. *Journal of Hydrology*, *152*(1–4), 179–199. https://doi.org/10.1016/0022-1694(93)90145-Y

Mau, D. P., & Winter, T. C. (1997). Estimating ground-water recharge from streamflow hydrographs for a small mountain watershed in a temperate humid climate, New Hampshire, USA. *Ground Water*, *35*(2), 297–304.

Miller, M. P., Johnson, H. M., Susong, D. D., & Wolock, D. M. (2015). A new approach for continuous estimation of baseflow using discrete water quality data: Method description and comparison with baseflow estimates from two existing approaches. *Journal of Hydrology*, *522*, 203–210. https://doi.org/10.1016/j.jhydrol.2014.12.039

Miller, M. P., Susong, D. D., Shope, C. L., Heilweil, V. M., & Stolp, B. J. (2014). Continuous estimation of baseflow in snowmelt-dominated streams and rivers in the upper Colorado River Basin: A chemical hydrograph separation approach. *Water Resources Research*, *50*(8), 6986–6999. https://doi.org/10.1002/2013WR014939

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

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

Partington, D., Brunner, P., Simmons, C. T., Therrien, R., Werner, A. D., Dandy, G. C., & Maier, H. R. (2011). A hydraulic mixing-cell method to quantify the

groundwater component of streamflow within spatially distributed fully integrated surface water–groundwater flow models. *Environmental Modelling & Software*, *26*(7), 886–898. https://doi.org/10.1016/J.ENVSOFT.2011.02.007

Partington, D., Brunner, P., Simmons, C. T., Werner, A. D., Therrien, R., Maier, H. R., & Dandy, G. C. (2012). Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water-groundwater flow model. *Journal of Hydrology*, *458–459*, 28–39. https://doi.org/10.1016/J.JHYDROL.2012.06.029

Pettyjohn, W. A., & Henning, R. (1979). *Preliminary Estimate of Ground-Water Recharge Rates, Related Streamflow and Water Quality in Ohio*.

Pinder, G. F., & Jones, J. F. (1969). Determination of the ground-water component of peak discharge from the chemistry of total runoff. *Water Resources Research*, *5*(2), 438–445. https://doi.org/10.1029/WR005i002p00438

Piggott, A. R., Moin, S., & Southam, C. (2005). A revised approach to the UKIH method for the calculation of baseflow/Une approche améliorée de la méthode de l'UKIH pour le calcul de l'écoulement de base. Hydrological Sciences Journal, 50(5).

Rossman, N. R., & Zlotnik, V. A. (2013). Review: Regional groundwater flow modeling in heavily irrigated basins of selected states in the western United States. In Hydrogeology Journal (Vol. 21, Issue 6, pp. 1173–1192). https://doi.org/10.1007/s10040-013-1010-3

Rumsey, C. A., Miller, M. P., Susong, D. D., Tillman, F. D., & Anning, D. W. (2015). Regional scale estimates of baseflow and factors influencing baseflow in the Upper Colorado River Basin. *Journal of Hydrology: Regional Studies*, *4*(PB), 91–107. https://doi.org/10.1016/j.ejrh.2015.04.008

Sanford, W. E., Nelms, D. L., Pope, J. P., & Selnick, D. L. (2011). *Quantifying components of the hydrologic cycle in Virginia using chemical hydrograph separation and multiple regression analysis*. https://doi.org/10.3133/sir20115198

Santhi, C., Allen, P. M., Muttiah, R. S., Arnold, J. G., & Tuppad, P. (2008). Regional estimation of base flow for the conterminous United States by hydrologic landscape regions. *Journal of Hydrology*, *351*(1–2), 139–153. https://doi.org/10.1016/j.jhydrol.2007.12.018

Shuai, P., Chen, X., Mital, U., Coon, E. T., & Dwivedi, D. (2022). The effects of spatial and temporal resolution of gridded meteorological forcing on watershed hydrological responses. *Hydrology and Earth System Sciences*, *26*(8), 2245–2276. https://doi.org/10.5194/hess-26-2245-2022

Shuai, P., Jiang, P., Coon, E. T., & Chen, X. (2023). The importance of explicitly representing the streambed in watershed models. Hydrological Processes, 37(12), e15043.

Singh, S. K., Pahlow, M., Booker, D. J., Shankar, U., & Chamorro, A. (2019). Towards baseflow index characterisation at national scale in New Zealand. *Journal of Hydrology*, *568*, 646–657. https://doi.org/10.1016/j.jhydrol.2018.11.025

Singh, V. P. (2018). Hydrologic modeling: progress and future directions. In *Geoscience Letters* (Vol. 5, Issue 1). SpringerOpen. https://doi.org/10.1186/s40562-018-0113-z

Sloto, R. A., & Crouse, M. Y. (1996). *HYSEP: A COMPUTER PROGRAM FOR STREAMFLOW HYDROGRAPH SEPARATION AND ANALYSIS*. https://doi.org/10.3133/wri964040

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

Stewart, M., Cimino, J., & Ross, M. (2007). Calibration of base flow separation methods with streamflow conductivity. *Ground Water*, *45*(1), 17–27. https://doi.org/10.1111/j.1745-6584.2006.00263.x

Tan, X., Liu, B., & Tan, X. (2020). Global Changes in Baseflow Under the Impacts of Changing Climate and Vegetation. *Water Resources Research*, *56*(9). https://doi.org/10.1029/2020WR027349

Therrien, R., McLaren, R., Sudicky, E., & Panday, S. (2006). HydroGeoSphere: A three-dimensional numerical model describing fully integrated subsurface and surface flow and solute transport. *HydroGeoLogic Inc., Herndon, VA*.

Thornton, P. E., Shrestha, R., Thornton, M., Kao, S. C., Wei, Y., & Wilson, B. E. (2021). Gridded daily weather data for North America with comprehensive uncertainty quantification. *Scientific Data*, *8*(1). https://doi.org/10.1038/s41597-021-00973-0

Xie, J., Liu, X., Wang, K., Yang, T., Liang, K., & Liu, C. (2020). Evaluation of typical methods for baseflow separation in the contiguous United States. *Journal of Hydrology*, *583*. https://doi.org/10.1016/j.jhydrol.2020.124628

Yang, W., Xiao, C., Zhang, Z., & Liang, X. (2021). Can the two-parameter recursive digital filter baseflow separation method really be calibrated by the conductivity mass balance method? *Hydrology and Earth System Sciences*, *25*(4), 1747–1760. https://doi.org/10.5194/hess-25-1747-2021

Zhang, J., Zhang, Y., Song, J., & Cheng, L. (2017). Evaluating relative merits of four baseflow separation methods in Eastern Australia. Journal of hydrology, 549, 252-263.