https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

Assessing the Effect of Data Resolution on Optimized Hydrograph Separation

Edgar Emilio Villasano[1], Benjamin Hagedorn[1], Nicholas Ryan Januario[1]

[1]Department of Earth Science, California State University, Long Beach, CA 90840, USA.

Correspondence to: Edgar Emilio Villasano (Edgar.Villasano01@student.csulb.edu)

Abstract: Hydrograph separation is centered on quantifying groundwater contributions to streamflow, yet the influence of monitoring data resolution remains insufficiently characterized. In this study, we assess how temporal resolution influences chemically constrained hydrograph separation. We apply *HydrOHS*, a multi-objective optimization framework that calibrates the physically-grounded Eckhardt Recursive Digital Filter using SC-centered mass balance, to four modeling scenarios capturing two spatial settings and two temporal resolutions. *HydrOHS* reconstructs SC end-members and simultaneously minimizes SC mismatch, enforces chemically consistent flow component ordering, and constrains peak baseflow index (BFI) behavior. The four modeling scenarios produce similar optimized BFI_{max} parameter values (0.80-0.89) and consistent downstream increases in baseflow contribution, indicating that temporal resolution (daily- vs 15 min monitoring frequency) exerts a secondary influence relative to hydrologic variability and missing flow periods. Higher resolution datasets provide additional detail but are more sensitive to zero-flow gaps, which degrades chemical reconstruction. Importantly, all optimized BFI_{max} estimates substantially exceed the suggested "default" BFI_{max} value for hard-rock aquifer systems, demonstrating that uncalibrated parameters underestimate baseflow. These findings highlight the importance of site-specific, chemically constrained calibration and show that daily resolution datasets may be sufficient for regional groundwater-recharge assessments.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

1. Introduction

Hydrograph separation is the process of quantifying baseflow and quickflow-derived streamflow components. Baseflow represents continuous groundwater input sustained between precipitation events, whereas quickflow represents rapid non-baseflow input such as surface water runoff. Accurate quantification of both quickflow and baseflow is critical to the calibration of regional-scale groundwater recharge and surface runoff models (Pelletier & Andréassian, 2020; Rumsey et al., 2015). In these applications, models are calibrated by iteratively adjusting parameters controlling infiltration, soil storage, and percolation until simulated recharge and runoff matched watershed-scale observations of baseflow and quickflow (e.g., Nielsen & Westenbroek, 2019).

However, baseflow and quickflow cannot be quantified directly and are instead measured indirectly via various methodologies broadly classified as tracer-based and non-tracer-based separation methods (Gonzales et al., 2009, Hagedorn, 2020). Graphical methods estimate baseflow via the manual (and subjective) "drawing" of a separation line along hydrograph minima on assumed timings of baseflow dominance, a process that is inherently subjective and difficult to reproduce (Wahl & Wahl, 1995; Sloto & Crouse, 1996). Recession curve approaches center on fitting exponential or power-law decay functions to the falling limb of a hydrograph to identify baseflow, typically assuming that baseflow follows a characteristic recession pattern, which may not hold true under variable hydrologic conditions (Brutsaert & Nieber, 1977; Tallaksen, 1995; Chapman, 1999; Eckhardt, 2008; Thomas et al., 2015). Because both graphical and recession curve methods depend on subjective choices or restrictive structural assumptions, they offer limited reproducibility and physical defensibility (Nathan & McMahon, 1990; Sloto & Crouse, 1996; Rimmer & Hartmann, 2014; Pelletier & Andréassian, 2020).

Recursive digital filter (RDF) methods center on applying mathematical algorithms to measured streamflow data to isolate slow baseflow contributions from rapid runoff contributions, typically requiring the calibration of one or more input parameters (Eckhardt, 2005; Nathan & McMahon, 1990). Among available RDFs, the Eckhardt (ECK) RDF is widely regarded as the most physically defensible, as it explicitly represents baseflow recession as a linear reservoir process (Eckhardt 2023). The ECK RDF is specifically distinguished by its dependency on the site-specific recession constant and the BFI_{max} parameter, the latter constraining long-term groundwater contributions and preventing the filter from assigning more baseflow than what is determined to be physically plausible. However, the difficulty of measuring or estimating appropriate BFI_{max} values for specific watersheds motivates the development of an optimized hydrograph separation (OHS) framework that utilizes chemical mass balance (CMB) constraints to derive appropriate RDF parameters (Liu et al., 2004; Zhang et al., 2013; Hagedorn, 2020).

Among available hydrograph separation methods, CMB is typically regarded as being the most physically grounded, given its reliance on site-specific end-member chemistry rather than estimated input parameters, with uncertainty arising primarily from appropriate tracer selection and representative end-member sample representation (Hagedorn & Whittier, 2015; Kronholm & Capel, 2015). Numerous studies have shown that groundwater and quickflow chemistry shift with hydrologic conditions such as transitions from saline regional groundwater to diluted bank-storage return flows, or seasonal changes in snowmelt, rainfall, and surface runoff, making it impractical to capture these dynamics through opportunistic sampling alone (Cartwright et al., 2014; Hagedorn & Whittier, 2015; Yang et al., 2021). Discrepancies between RDF- and CMB-derived flow components reflect the misclassification stemming from temporally unrepresentative CMB end-members rather than different water sources contributing to baseflow and quickflow. When groundwater or snowmelt chemistry vary through seasons or individual events, CMB based on single, static SC end-member values will misattribute hydrograph components, affecting the magnitude of baseflow estimations. This limitation underscores the need to derive temporally constrained baseflow and quickflow end-members from a multi objective optimization workflow, where the ECK RDF informs the CMB, and vice versa, using continuous and concurrently measured streamflow-, stream SC-, and precipitation SC data.

Previous attempts to combine RDF- and CMB- approaches have relied on subjective determinations of baseflow and quickflow SC end-members, limiting reproducibility and physical defensibility (Miller et al., 2014; Saraiva Okello et al., 2018; Raffensperger et al., 2017; Foks et al., 2019). Hagedorn (2020) avoided this limitation by linking baseflow SC to instances in the SC monitoring record where the baseflow index (BFI =

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

ratio of baseflow to total flow) equals 1, so that the baseflow SC on instances with BFI <1 could be derived via interpolation. However, in this study, quickflow SC was derived inversely through calibration rather than basing it on precipitation chemistry. Hagedorn (2020) furthermore applied the Lyne and Hollick multi-pass filter, whose structure is less physically grounded than the Eckhardt (ECK) RDF (Eckhardt, 2005).

In this study, we present a new hydrograph separation algorithm, *HydrOHS*, that tunes the more physically-grounded ECK RDF via CMB in a refined multi objective optimization (MOO) framework. *HydrOHS* optimizes four objective functions simultaneously, through the Non-Dominated Sorting Genetic Algorithm II (NSGA-II), a widely adopted solver recognized for reliable performance in multi-objective optimization problems (Ishibuchi et al. 2016; Ma et al. 2023). The first objective function minimizes the mismatch between observed and modeled streamflow specific conductance (SC), while the second and third objective functions enforce chemically consistent ordering between baseflow, quickflow, and total streamflow. The fourth objective function constrains peak SC and BFI to ensure hydrologically plausible partitioning. Baseflow and quickflow SC values are respectively estimated by interpolating stream SC on baseflow-dominated days identified by the RDF and by processing concurrent precipitation SC data from nearby precipitation chemistry monitoring gages during optimization.

Through the calibration of *HydrOHS*, this study aims to address two primary goals. First, we assess the value of utilizing higher resolution (15-minute) data compared to daily resolution data for hydrograph separation. Second, we compare the BFI_{max} guidance values suggested by Eckhardt (2005) with the site-specific values derived through our MOO calibration. Although the 2005 values (Table 1) are widely adopted defaults, subsequent work - most notably Eckhardt (2012) - demonstrates that BFI_{max} should be calibrated whenever additional catchment information is available. This comparison, therefore, highlights the importance of physically constrained, site-specific parameterization.

Table 1. Stream classification and corresponding recommended BFI_{max} values for the Eckhardt RDF (after Eckhardt, 2005).

Stream Classification	BFI_{max} Value
Perennial streams, porous aquifers	0.80
Ephemeral streams, porous aquifers	0.50
Perennial streams, hard rock aquifers	0.25

2. Methods

2.1 Theoretical Overview

2.1.1 Eckhardt Recursive Digital Filter

Among available RDFs, the Eckhardt (ECK) filter has seen widespread application due to its partial physical basis and representation of groundwater system behavior. In contrast to purely empirical low pass RDFs, the ECK RDF incorporates a partial physical basis by assuming a linearity reservoir assumption in which groundwater discharge is proportional to aquifer storage (Eckhardt, 2005; Eckhardt, 2023). Although this simplifying assumption provides a practical representation of groundwater recession, it is not universally valid, as many aquifer systems exhibit nonlinear recession behavior (Brutsaert & Nieber, 1977; Kirchner, 2009), particularly in karst and highly fractured bedrock aquifer settings where storage-discharge relationships can be nonlinear.

Other commonly used RDFs differ from the ECK RDF in that they rely solely on signal-processing assumptions rather than hydrologic constraints. For example, the Lyne and Hollick RDF (Lyne & Hollick, 1979) functions as a low-pass filter that partitions quickflow based solely on signal smoothing with no physical basis. As such, it is often applied over multiple passes (forward and backward) to produce a smooth baseflow hydrograph. The Chapman and Maxwell RDF (Chapman & Maxwell, 1996) improves on this by imposing a fixed post-peak recession form, but unlike the ECK RDF, it does not explicitly limit long-term groundwater contributions.

The ECK RDF is reliant on two primary input parameters: the groundwater recession constant and the maximum value of the baseflow index (BFI_{max}). The mathematical expression for the ECK RDF is:

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

$$b_i = \frac{(1-\beta)ab_{i-1}+(1-a)\beta y_i}{1-a\beta} \tag{1}$$

where b_i corresponds to baseflow at time interval i, y_i corresponds to total streamflow at time interval i, a is the recession constant, and β is BFI_{max}.

The recession constant reflects the effective drainage efficiency of the groundwater reservoir under the assumption of the linear aquifer storage-discharge relationship (Eckhardt, 2005). It is bound between 0 and 1, with values approaching 1 indicating slower groundwater release and more sustained baseflow contributions. The recession constant can be estimated via streamflow recession analysis by isolating baseflow dominated recession limbs in a hydrograph (Nathan & McMahon, 1990; Eckhardt, 2005; Eckhardt, 2008, Kirchner, 2009; Cheng et al., 2016).

The BFI_{max} parameter represents the maximum fraction of streamflow attributable to baseflow and is bound between 0 and 1. BFI_{max} acts as an upper limit on long-term groundwater contributions, ensuing that the filter cannot attribute more baseflow than is physically plausible. While the recession constant can be constrained via hydrograph analysis, BFI_{max} cannot be measured and defined objectively. To address this limitation, Eckhardt (2005) recommended three representative BFI_{max} values by comparing filter-derived BFI values to independently estimated BFI values reported in earlier studies that applied conventional hydrograph separation methods. For perennial porous aquifer systems, the reference BFI_{max} values were sourced from Arnold and Allen (1999), who used graphical separation, recession-curve analysis, and minimum-flow methods for baseflow estimation. For ephemeral porous aquifer systems, the reference BFI_{max} values were sourced from Arnold and Allen (1996), who applied the same group of hydrograph separation techniques. For perennial hard-rock aquifer systems, the reference BFI_{max} values were taken from Kaviany (1978), who applied the graphical method from Natermann (1951) and the monthly minimum runoff method from Kille (1970).

However, long-term baseflow contributions are influenced by numerous additional catchment characteristics beyond these three classifications (Gonzales et al., 2009; Collischonn & Fan, 2013; Rimmer & Hartmann, 2014; Saraiva Okello et al., 2018; Helfer et al., 2024). Consequently, in multiple studies researchers have attempted to calibrate BFI_{max} using physically grounded constraints rather than relying on the generalized Eckhardt (2005) default values. Collischonn and Fan (2013) estimated BFI_{max} by applying a backwards filter prior to the application of the ECK RDF to derive a discharge-based upper bound to baseflow. Rimmer and Hartmann (2014) calibrated filter parameters using chemical tracers in an OHS framework, demonstrating that BFI_{max} varies substantially across applied tracers. Saraiva Okello et al. (2018) calibrated BFI_{max} using chemical hydrograph separation based on electrical conductivity, with annual upper and lower percentiles as proxies for baseflow and quickflow SC, respectively, showing that the applied filter parameters vary drastically by monitoring site. Most recently, Helfer et al. (2024) demonstrated that BFI_{max} should be constrained via multiple hydrological signatures such as recession slope, flow variability, and seasonal groundwater response. While these studies introduce physically informed constraints, most are centered on static and/or subjectively derived quickflow and baseflow end-members, single-objective calibration or still rely on site-specific empirical values, limiting their objectivity and reliability.

2.1.2 Chemical Mass Balance

ECK RDF baseflow estimations can be optimized by SC-centered CMB structured as:

$$SC_{yi} = \frac{SC_{bi}b_i + SC_{qi}q_i}{y_i} \tag{2}$$

where SC_{bi} and SC_{qi} are SC values of baseflow and quickflow components at time interval i respectively, SC_{yi} is the modeled SC value of total streamflow, and q_i corresponds to estimated quickflow; this CMB implementation assumes complete mixing and conservative SC behavior.

Traditional SC-based CMB hydrograph separation approaches require a priori estimation of baseflow SC and quickflow SC from groundwater sampling, precipitation chemistry, and/or assumed end-member values (Sanford et al., 2012; Miller et al., 2015; Saraiva Okello et al., 2018). These approaches are often limited by sparse sampling data, seasonal dilution effects, and end-member uncertainty. Consequently, discrepancies

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

between RDF- and CMB- derived baseflow estimates reported in previous studies largely reflect the use of temporally unrepresentative and subjective end-members rather than structural limitations of either method (Miller et al., 2015; Saraiva Okello et al., 2018; Helfer et al., 2024).

Given the challenge of collecting time series surface runoff (quickflow) SC (SC_y) data, we believe that constructing it from time series precipitation chemistry data through the adjustments of various smoothing, offsetting, and lagging parameters in a MOO framework is the most representative approach. Likewise, we believe interpolating stream SC signals on peaks in the baseflow index (BFI), when quickflow contributions are minimal, will yield a representative time series baseflow SC (SC_b) end-member. Importantly, all decision variables that control SC_y and SC_b (e.g., amplification, smoothing, offsetting, lagging, peak identification and interpolation method) should be determined through the calibration of a MOO model, where both chemical accuracy and realistic flow partitioning aspects are equally weighted. This approach ensures that both SC_y and SC_b estimations are dynamic and remain anchored to the observed SC record rather than relying exclusively on subjectively derived end-member values.

2.2 The *HydrOHS* Algorithm

HydrOHS applies MOO - specifically the Non-Dominated Sorting Genetic Algorithm II, NSGA-II (Deb et al., 2002) - to minimize four objective functions, each capturing a distinct aspect of RDF and CMB-derived hydrograph separation. NSGA-II is a widely adopted evolutionary algorithm that evolves candidate solutions through selection, recombination, and mutation while ranking individuals according to non-domination levels and preserving solution diversity via a crowding-distance criterion. Each individual solution represents a candidate hydrograph separation parameterization, defined by a unique combination of 11 decision variable values controlling baseflow recession behavior, baseflow index constraints, chemical end-member tuning, temporal lags, and peak detection thresholds (Table 2). These variables collectively determine how accurate and realistically the ECK RDF informs CMB – and vice versa - to partition baseflow and quickflow portions of a stream hydrograph.

Table 2. Decision variables (x_1 - x_{11}) and associated parameter bounds used in the *HydrOHS* algorithm framework.

ID	Parameter	Bounds
x_1	Missing Precipitation SC Interpolation	(1, 2)[a]
x_2	Recession Constant	(0.96, 0.99)[b]
x_3	BFI_{max}	(0, 1)
x_4	SC_y Smoothing Parameter	(0, 1)[c]
x_5	SC_y Multiplier	(0, 10)
x_6	SC_y Offset	(-100, 100)
x_7	SC_y Timestep Lag	(-200, 200)[d]
x_8	BFI Peak Minimum Height	(0, 1)[e]
x_9	BFI Peak Minimum Spacing	(1, 100)[d, e]
x_{10}	SC_b Interpolation	(1, 2)
x_{11}	Maximum Peak BFI Value	(0, 1)

Note. Superscripts indicate parameter-specific bound constraints:

a. Logical index bounds for method of interpolation (pchip vs. linear)
b. Continuous bounds determined for each site via manual log linear recession analysis (Kirchner, 2009). Bounds shown are an example and should vary by site.
c. Smoothing parameter for smoothing splines.
d. Bounds shown are full integers for daily-resolution applications.
e. Parameter for peak identification in the baseflow index (BFI = baseflow/streamflow)

2.2.1. Algorithm Structure

Figure 1 presents a schematic overview of the *HydrOHS* workflow, outlining the sequence of major pre-processing, optimization, and post-processing steps. Optimization bounds for decision variables are listed in

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

Table 2. The full R source code is publicly accessible in the algorithm's GitHub repository, archived on Zenodo (Villasano, 2026).

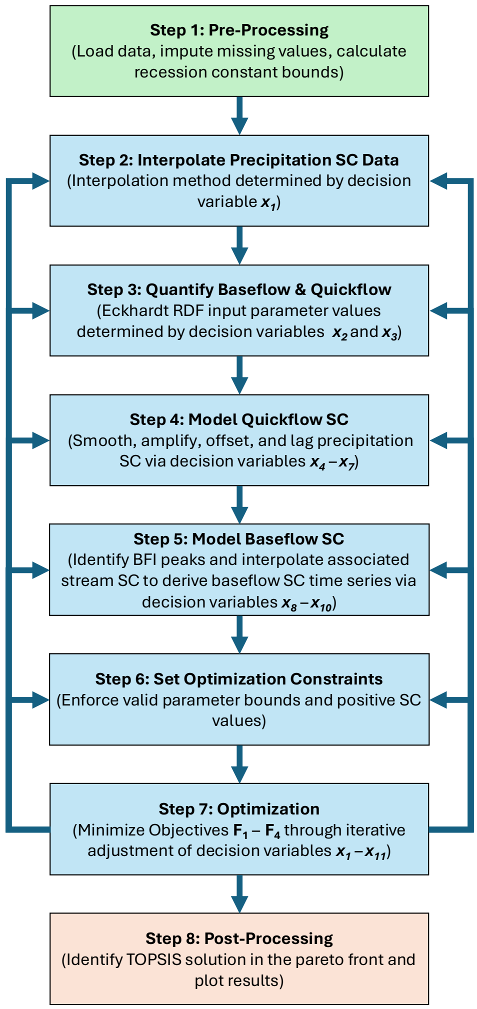

Figure 1: Workflow of the *HydrOHS* algorithm illustrating the pre-processing, iterative processing (i.e., optimization) routine, and post-processing stages. Looped arrows denote the steps included in the iterative multi objective optimization process. Decision variables (x_1 – x_{11}) are defined in Table 2. Objectives F_1 – F_4 correspond to equations (6) – (9).

2.2.2. Preprocessing

The preprocessing stage loads the streamflow and precipitation chemistry datasets required for algorithm application and establishes bounds for the recession constant parameter of the ECK RDF. These bounds are derived through log-linear recession limb analysis. Specifically, five years in the streamflow record that contain well-defined recession periods are selected. For each year, a peak streamflow period is manually delineated, after which the corresponding recession limb is isolated by selecting the monotonic falling limb. In practice, this is aided largely by locating the minimum streamflow value following a peak and using it to define the end of the recession segment.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

For each recession limb, discharge was converted to its natural logarithm and regressed against time step using ordinary least squares, where each index increment corresponds to one dataset-specific time interval. Under the assumption of a linear groundwater reservoir, the slope of the log-linear recession curve represents *-kt*, where *k* is the continuous recession rate and *t* is the dataset-specific time interval (Brutsaert & Neiber, 1977). Kirchner (2009) details that such rain-free, uninterrupted recession segments provide a robust means of inferring storage-discharge behavior, ensuring that the fitted slope reflects groundwater drainage rather than event-driven quickflow. However, he also emphasized that linear recession behavior is unlikely except in the limiting case of very low flows, and that some watersheds exhibit some degree of nonlinearity.

Consequently, the recession constant (α) for each of the 5 recession limbs is computed as:

$$\alpha = e^{-kt} \tag{3}$$

It should be noted that recession constants derived from higher resolution data are systematically higher compared to those derived from coarser resolution data for the same underlying recession behavior. This occurs because the recession constants estimated from log-linear recession analysis reflect the amount of discharge decline occurring over a single timestep interval, and shorter timestep intervals capture proportionally smaller increments of recession per observation. Because of this time resolution dependence, recession-limb analysis was performed independently for each of the temporal dataset resolutions of this study (daily vs 15-minute frequency). For each scenario, the minimum and maximum values of the five recession constant estimations were used as the upper and lower bounds for the recession constant parameter during optimization. This ensures that each modeling scenario explores recession behavior consistent within the α uncertainty range defined by the 5 selected recession limbs.

The a-priori determination of α optimization bounds prevents *HydrOHS* from compensating for extreme BFI_{max} estimates with unrealistic recession constant estimates. While this approach stabilizes the optimization procedure, it restricts the exploration of certain trade-off solutions that would otherwise be represented in Pareto-optimal solution spaces and considered in further processing.

The next step of preprocessing is missing data imputation. *HydrOHS* applies Piecewise Cubic Hermite Interpolating Polynomial (pchip) interpolation to fill any missing streamflow and stream SC values. We note that there is some degree of subjectivity in the missing data interpolation as numerous methods exist for this task (Fritsch & Carlson, 1980; De Boor, 1978; Press, 2007). We thus recommend using *HydrOHS* only on datasets containing limited and well-distributed missing streamflow and stream SC data, as modest levels of missing data are generally acceptable (<10%) (Teegavarapu, 2024), but long continuous gaps of missing data introduce substantially greater uncertainty than short, dispersed occurrences (See et al., 2020).

Lastly, a 30-day warm-up buffer of measured streamflow was applied at the start of modeling-period during the preprocessing stage. The buffer is primarily required to ensure minimal influence of the arbitrary assignment of b_{i-1} (Eq. 1) at the start of the monitoring period.

2.2.3. Processing

The processing stage encompasses the algorithm's core optimization procedure. Candidate solutions corresponding to unique parameter combinations are iteratively adjusted to minimize four objective functions and satisfy five constraints for joint flow and SC calibration.

The first step is the processing of precipitation SC data. Precipitation chemistry data can be obtained from networks such as the National Atmospheric Deposition Program (NADP) National Trends Network (NTN) (NADP, 2024), which provides weekly precipitation SC measurements at monitoring stations across the United States. Because precipitation SC measurements are typically sparse compared to streamflow and streamflow SC observations, precipitation SC is interpolated across the full timeseries via either pchip or linear interpolation, with the interpolation method selection being treated as an adjustable parameter (x_1, bounded between 1 and 2; Table 1)

The second step of the processing stage is the application of the ECK RDF to the streamflow data (Eq. 1). Whereas the ECK RDF's recession constant input parameter (x_2, bounded estimated recession constants defined from five recession limbs of the streamflow record is constrained prior to optimization in the preprocessing stage, the BFI$_{max}$ parameter (x_3, non-measurable and bounded between 0 and 1) is not, enabling the algorithm to explore all values permitted by its processing constraints. Because the preprocessing stage ensures that each modeling scenario dataset begins with uninterrupted, fully interpolated streamflow, the recursive filter can be applied directly without additional adjustments. The resulting baseflow and quickflow estimations for the timeseries serve as the basis for BFI values and as input for subsequent CMB calculations.

The third step involves the estimation of the quickflow SC from the precipitation SC record. Since baseflow and quickflow cannot be quantified directly and instead require indirect estimation, quickflow SC cannot be measured directly either. Thus, precipitation SC serves as the most applicable proxy for the SC concentration of quickflow contributions. The transformation applied to the precipitation SC data is governed by three parameters: a smoothing parameter (x_4, bounded between 0 and 1), a multiplier (x_5, bounded between 0 and 10), and an additive offset (x_6, bounded between -100 and 100 day-equivalents; Table 1). This conversion is expressed as:

$$SC_{yi} = spline(x_5 \times SC_{rain}, x_4) + x_6 \tag{4}$$

where SC_{yi} is estimated SC_y at time interval *i* and SC_{Rain} is precipitation SC. The algorithm first multiplies the precipitation SC dataset by an "amplifier", x_5, to account for the difficult-to-measure effects of evaporation and water rock interaction on increasing precipitation SC to quickflow SC. Next, the amplified SC is smoothed via smoothing spline, set by the x_4 parameter, to attenuate signal peaks. As the third step, the smoothed signal is vertically offset by the x_6 parameter to, again, give the algorithm more versatility in converting a dilute precipitation signal to a more concentrated quickflow signal or vice versa. Lastly, a temporal lag (x_7, bounded between -200 and 200 day-equivalents; Table 1) is applied to account for the delayed hydrologic response between precipitation inputs and their chemical expression as quickflow in the system. Chemical expression in this context specifically refers to the appearance of the precipitation-derived SC signal in surface runoff, consistent with the conservative-tracer assumption commonly applied in studies of event-water dynamics (Hagedorn, 2020; Hagedorn & Meadows, 2021). Although event water may travel through a combination of overland flow, shallow subsurface interflow, or small tributary pathways, SC is generally conservative along these routes, and changes in its concentration primarily reflect hydrologic mixing. This lag is formally defined as:

$$SC*_{yi} = SC_{yi+x_7} \tag{5}$$

Where $SC*_{yi}$ is lagged SC_y. By shifting the precipitation-derived SC_y signal forward or backward in time, the algorithm aligns the reconstructed quickflow chemistry with the timing of estimated quickflow responses. Examples of optimized SC_y are shown in section 4.1.

After establishing the SC_y, the algorithm estimates SC_b. Baseflow-dominated datetime intervals are identified by applying peak detection to the BFI timeseries using the *findpeaks* function from the *pracma* R package (Borchers, 2023). The peak detector is controlled by two decision variables: peak height (x_8, bounded between 0 and 1) and peak distance (x_9, bounded between 0 and 10 day-equivalent), which determine the minimum amplitude and temporal separation between detected BFI peaks (Table 1). The datetime intervals identified as peaks are subsequently allocated to a sub-dataframe. These identified peak values serve as anchors for estimating baseflow SC values during peak baseflow events via the chemical mass balance formulation. (Eq. 2). The resulting peak SC_b values are then used as a basis for subsequent baseflow SC interpolation across the remainder of the timeseries (Fig. 7). The interpolation method (pchip vs linear) is, again, handled as an adjustable parameter (x_{10}).

With all necessary components established, the algorithm reconstructs streamflow SC (SC_q) on non-BFI peak instances via the CMB approach (Eq. 2). Each candidate solution is evaluated via the algorithm's four objective functions, with the pareto-optimal set of solutions being generated via NSGA-II. The first objective (F_1) minimizes discrepancies between measured and modeled streamflow SC using a mean absolute error (MAE) metric augmented by a heteroscedasticity penalty. It is defined as:

$$F_1 = (1 + pen)\frac{1}{n}\sum_{i=1}^{n}|SC_i - SC_{qi}| \tag{6}$$

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

where n represents the number of calibration observations and *pen* represents the heteroscedasticity penalty, *pen*, is computed as the Gini coefficient of the absolute residuals using the *ineq* R package (Zeileis et al., 2009). For each optimization iteration, residuals are calculated as the difference between modeled and measured streamflow SC. The Gini coefficient is calculated by ranking the absolute residuals from smallest to largest and evaluating the relative differences between all pairs of residuals, producing a single value that quantifies the total inequality among residuals. The inclusion of this penalty mitigates the influence of periods with low or high variable flow, ensuring that calibration is not biased towards overestimation or underestimation during specific flow regimes. The MAE was selected over root mean square error (RMSE) and Nash-Sutcliffe efficiency (NSE) for this application due to weighing all residuals linearly, reducing disproportionate influence of outliers, and preserving the units of the observed streamflow (Willmott & Matsuura, 2005). By combining the MAE with *pen*, F_1 achieves robust calibration that balances overall fit with sensitivity to variable flow conditions.

The second and third objective functions (F_2 & F_3), defined as:

$$\mathrm{F}_2 = \sum_{\mathrm{i}=1}^{\mathrm{n}} 1(SC_{qi} > SC_{bi}) \; - \sum_{\mathrm{i}=1}^{\mathrm{n}} 1(SC_{qi} < SC_{bi}) \tag{7}$$

$$\mathrm{F}_3 = \sum_{\mathrm{i}=1}^{\mathrm{n}} 1(SC_{yi} > SC_{qi}) \; - \sum_{\mathrm{i}=1}^{\mathrm{n}} 1(SC_{yi} < SC_{qi}) \tag{8}$$

both evaluate the correctness of modeled flow component ordering using sign agreement. F_2 minimizes the number of occurrences in which observed streamflow SC exceeds baseflow SC estimations. This objective is computed as the difference between the number of unrealistic SC exceedances ($SC_{qi} > SC_{bi}$) and realistic non-exceedances ($SC_{qi} < SC_{bi}$) across all calibration observations. Through this objective, calibration discourages unrealistic baseflow estimations and promotes chemical mass balance consistency during baseflow-dominated periods. F_3 similarly minimizes instances where modeled quickflow SC exceeds observed streamflow SC.

The fourth objective (F_4) introduces the decision variable x_{11}, which represents the BFI corresponding to the maximum peak SC value. Collectively, F_4 and x_{11} reward parameter combinations where the highest-quality baseflow SC signal (max SC_b) coincides with a high BFI value. This allows the optimization to explore different upper limits on the BFI timeseries while maintaining consistency with the hydrologic structure inferred from the SC records. The objective is described as:

$$\mathrm{F}_4 = 1 - x_{11} \tag{9}$$

Because all objectives are minimized, this formulation maximizes x_{11}.

Constraints are imposed in *HydrOHS* to ensure physically meaningful parameter values and chemically consistent SC component estimates throughout the optimization process. The first constraint, C_1, limits the permissible range of the BFI_{max} parameter x_3, and is expressed as:

$$0 < x_3 < 1 \tag{10}$$

These two extreme values respectively correspond to the complete absence or domination of baseflow input, which would prevent meaningful hydrograph partitioning across a multi-year system application.

The three subsequent constraints (C_2-C_4), collectively ensure that all interpolated SC components remain physically meaningful. Specifically, precipitation, quickflow, and baseflow SC must remain positive throughout the time series. Candidate solutions violating any of the algorithm's constraints are rejected during optimization. C_2-C_4 are expressed as:

$$\min(SC_{Rain}) > 0 \tag{11}$$

$$\min(SC_y) > 0 \tag{12}$$

$$\min(SC_b) > 0 \tag{13}$$

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

The last constraint (C_5) restricts decision variable x_{11} so that it cannot exceed the BFI value associated with the maximum estimated baseflow conductivity SC_b. If BFI_{peak*} denotes the BFI peak where SC_b is at a maximum, the constraint is:

$$x_{11} - BFI_{peak*} \leq 0 \tag{14}$$

The final step of the processing stage applies the NSGA-II optimization to generate a Pareto-optimal set of solutions, each representing alternative trade-offs among objectives F_1- F_4. Each candidate solution, represented as a set of decision variable values, evolves over multiple generations through NSGA-II's selection, recombination, and mutation procedure. A parent population (μ) and offspring population (λ) are maintained to generate a new set of candidate solutions for each generation. Individual solutions are ranked by non-domination levels and the tradeoff satisfaction of the four objectives, with diversity among the individuals being preserved with a crowding-distance criterion that discourages solution clustering in specific regions of the objective space. Via this iterative process, NSGA-II produces a Pareto-optimal set of solutions that represent alternative trade-off between objectives F_1- F_4. This Pareto-optimal set of solutions is expressed as a Pareto front, which illustrates how the Pareto-optimal set of solutions performed across all objectives.

2.2.4. Postprocessing.

In this step, a single solution is identified from the Pareto-optimal set of solutions. In contrast to bi-objective frameworks where geometric knee-point identification is traceable (e.g., Hagedorn, 2020 and references therein), knee-point identification in higher-dimensional Pareto spaces is difficult to employ because the 'bend' becomes less clearly defined beyond a two-dimensional environment (Deb & Jain, 2014). As such, we used the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) (Hwang & Yoon, 1981). TOPSIS ranks Pareto-optimal solutions according to their relative distance to both ideal best and worst solutions in a multi-objective space. F_1, which quantifies the fit between measured and modeled stream SC, is assigned a TOPSIS weight of 1 to emphasize accurate reconstruction of the hydrograph across all flow conditions. F_2 and F_3 are each assigned a weight of 0.5 because they both center on flow component ordering. Assigning them full weight values would result in the algorithm slightly prioritizing physical realism over numerical accuracy. Like F_1, F_4 receives a full weight of 1 given its independent role of enforcing hydrologic plausibility by aligning modeled BFI and measured SC peaks. Given that these weights are only applied post-optimization during the TOPSIS evaluation, they are not influencing the search for Pareto-optimal solutions.

3. Algorithm Application

3.1 Study Area

The South Fork Tule River is a mountainous stream system in Tulare County, California representing one of the three major forks of the larger Tule River System (Fig. 2). The headwaters of the South Fork Tule River begin in the Sequoia National Forest on the western flanks of the Sierra Nevada Mountains. The river flows directly through the Tule River Reservation (225 km^2) populated by the Tule River Indian Tribe before draining into Lake Success, a constructed reservoir east of the Reservation.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

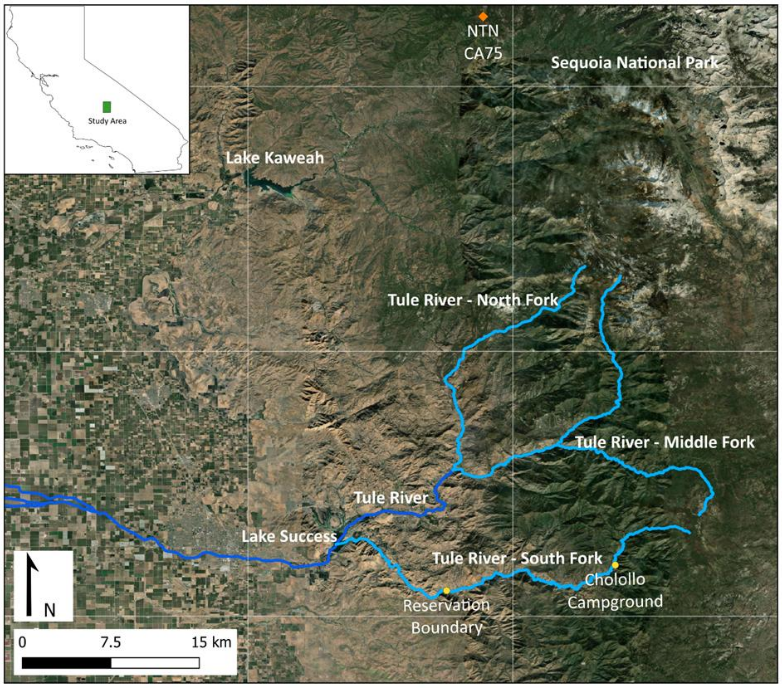

Figure 2: Site map of the South Fork Tule River showing USGS stream gages Cholollo Campground (Site 11203580) and Reservation Boundary (Site 11204100), along with the National Atmospheric Deposition Program (NADP) National Trends Network (NTN) precipitation station CA75 (Sequoia National Park-Giant Forest) (NADP, 2024). River lines are from the National Hydrography Dataset (USGS, 2025). California locator map boundaries are from Natural Earth (2025). Basemap imagery is from Esri World Imagery (2025). Map created in QGIS 3.40.12 (QGIS Development Team, 2025).

Two stream gage stations along the South Fork Tule River, Cholollo Campground (Site 11203580) and Reservation Boundary (Site 11204100), were selected from available U.S. Geological Survey (USGS) streamflow network gages in California (USGS, 2025) due to fitting three specific criteria. Firstly, both stations are free of influence from upstream urbanization, agricultural land use, dams, reservoirs and major diversions. Secondly, both stations have multi-year discharge and SC data available at both daily and 15-minute resolutions, allowing for the assessment of the merits of utilizing high resolution data for hydrograph separation. Finally, the positions of the two stations relative to one another – Cholollo Campground upstream at 1120 m above sea level (asl) and Reservation Boundary downstream at 288 m asl – allow for an analysis of the downstream evolution of concurrent baseflow and quickflow model estimates.

The South Fork Tule River is underlain primarily by shallow crystalline bedrock of the Sierra Nevada batholith, with thin soils and fractured-rock aquifer characteristics across most of the domain (Meadows & Hagedorn, 2022). Alluvial deposits are largely restricted to the western lowlands of the watershed and only occur locally along the channel. Given this limited extent, these deposits provide relatively little groundwater storage compared to large, slow draining systems typically associated with porous aquifer settings.

Daily and 15-minute resolution discharge data extending from March 2008 to July 2024 indicate highly variable flow and periods of extremes correlating to droughts and extreme precipitation events over the past 16 years (Huang et al., 2018; Maurer, 2007; Miller & Urban, 1999; Null et al., 2010). Streamflow values range from 0.01 m^3/s to 108 m^3/s upstream at Cholollo Campground and 0.00 m^3/s to 379 m^3/s downstream at Reservation Boundary. Streamflow peaks typically occur between the months of January through April. High

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

streamflow values likely correlated to major flood events and occurred in 2010, 2017, 2019, and 2023. Minimum values in streamflow correlating to major droughts occurred from 2011 to 2016 and in 2021. Data from the Reservation Boundary site indicates two periods of recorded zero flow from July to September 2014 and from June to October 2015 at both 15 minute and daily monitoring resolutions. As we will show and discuss later, the prolonged ~2 month no flow period at Reservation Boundary in the fall of 2021 has important implications on the application of *HydrOHS*.

Daily and 15-minute resolution SC data ranging from March 2008 to July 2024 also indicate extreme variability for both sites across all four datasets, ranging from 26.0 µS/cm to 459 µS/cm upstream at Cholollo Campground and from 21.0 µS/cm to 684 µS/cm downstream at Reservation Boundary. Corresponding with a multi-year drought event, annual SC peak values showcase a notable increase from 2011 to 2016 before beginning to decrease in 2017. Similar increases in SC peak values from 2019 to 2021 correspond to a later drought event.

Precipitation chemistry data from the National Atmospheric Deposition Program (NADP) National Trends Network (NTN) monitoring station nearest to the South Fork Tule River, Site CA75 (Sequoia National Park-Giant Forest), was used for SC_y approximation. Importantly, the precipitation chemistry data is at a weekly resolution and concurrent to the USGS stream monitoring data. Because this station is the closest precipitation station to both stream gages, its data was used for *HydrOHS* algorithm applications on both Cholollo Campground and Reservation Boundary data. Precipitation SC values from this station range from 1.4 µS/cm on January 1, 2011 to 44.9 µS/cm, on April 25, 2009.

3.2 Algorithm Application Considerations

We applied *HydrOHS* to four datasets (Cholollo Campground daily and 15-minute resolution; Reservation Boundary daily and 15-minute resolution) using the parameter bounds detailed in Table 2. Algorithm performance is evaluated at a parent population size of $\mu = 1000$ and an offspring population size of $\lambda = 1000$ over 100 generations per model run. This configuration is based on preliminary convergence testing with daily resolution datasets that indicated Pareto front stabilization between 50 and 100 generations. We select $\lambda = 1000$ to allow TOPSIS to select the more evenly spaced Pareto optimum solution because $\lambda = 100$ produced slightly more fragmented and clustered fronts.

Notably, the daily resolution Cholollo Campground dataset contains no missing streamflow values, and the 15-minute resolution Cholollo Campground dataset contained only minor non-continuous gaps in streamflow, with only 2.16% of the total streamflow record missing. In most cases, these gaps are related to time zone conversions or equipment malfunction. In contrast, both Reservation Boundary datasets contain an extensive period of missing streamflow data occurring from April – August 2021 corresponding to drought and partially zero flow conditions. Despite the occurrence of these gaps, the daily resolution Reservation Boundary dataset contains only 1.88% total missing streamflow, whereas the 15-minute resolution Reservation Boundary dataset contains 3.63% total missing streamflow. These gaps required the application of interpolation during the preprocessing stage and were subsequently flagged during algorithm evaluation to evaluate the importance of missing data imputation on algorithm performance.

4. Results and Discussion

4.1 Algorithm Performance

Recession limb analysis was conducted during the preprocessing stage of each application. The resulting recession limbs and fitted regressions are presented to verify that the watershed satisfies the linear-reservoir assumption required by the ECK RDF. For each modeling scenario, the recession limb with the lowest estimated recession constant was plotted in log(Q)- time space, and a linear regression was fitted to the selected limb (Fig. 3). All four modeling scenarios exhibit strong linearity ($R^2 = 0.955$-0.978), indicating that the steepest recession segments follow an approximate exponential decline in discharge. Because the lowest recession limb represents the most conservative portion of the recession behavior, linearity in these limbs implies that the remaining recession periods also conform to near-linear reservoir dynamics. Minor flat segments in the 15-minute scenario limbs reflect finite measurement precision rather than deviations from exponential decay.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

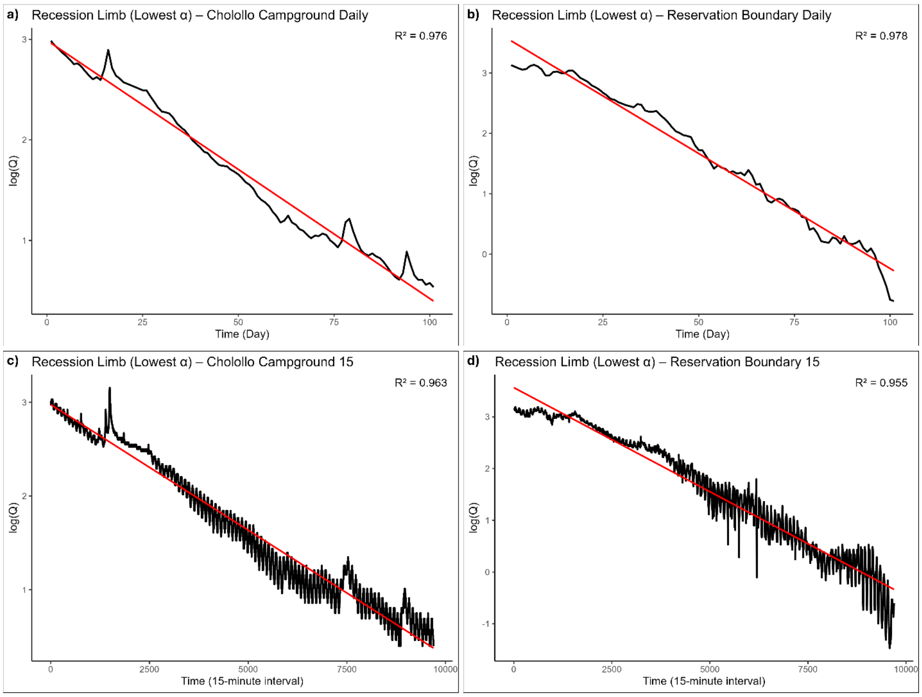

Figure 3: Recession limbs with the lowest recession constant estimate plotted in log(Q)-time space for the Chololло Campground Daily (a), Reservation Boundary Daily (b), Chololло Campground 15-minute (c), and Reservation Boundary 15-minute (d) modeling scenarios. Black represents observed recession behavior whereas red represents the fitted linear regression of each modeling scenario.

To further evaluate recession behavior, log (|dQ/dt|) was plotted against log (Q) following the recession-slope framework of Kirchner (2009) (Fig. 4). The daily modeling scenarios produced slopes of 0.92 for Chololло Campground and 0.72 for Reservation Boundary, consistent with the expectation that a linear reservoir exhibits a slope near unity. The 15-minute modeling scenarios yielded slopes of 0.55 and 0.75 for Chololло Campground and Reservation Boundary respectively, with increased scatter reflecting sensitivity to short-term fluctuations in high-resolution discharge data. Despite this variability, the slope magnitudes remain broadly consistent across temporal resolutions. Together, these diagnostics confirm that the South Fork Tule River behaves near-linearly and therefore is suitable for the ECK RDF application in *HydrOHS*.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

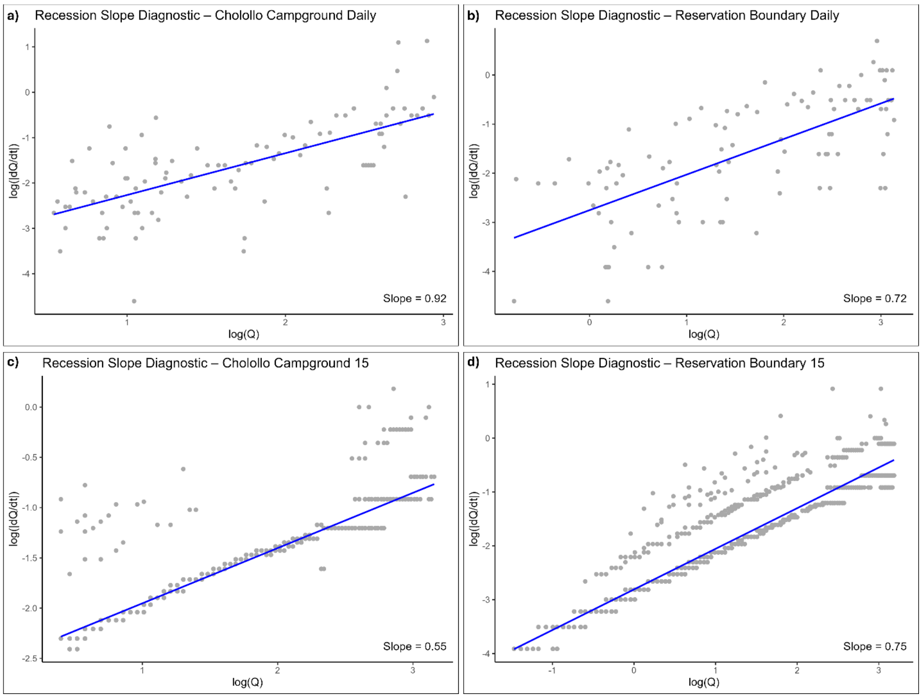

Figure 4: Recession-slope diagnostics for the Chollolo Campground Daily (a), Reservation Boundary Daily (b), Chollolo Campground 15-minute (c), and Reservation Boundary 15-minute (d) modeling scenarios following the framework of Kirchner (2009). Grey points represent raw recession data whereas blue represents fitted linear regressions.

The Pareto-optimal solution spaces for all four modeling scenarios with highlighted TOPSIS-selected solutions are shown in Figure 5. To facilitate three-dimensional visualization, only objectives F_1, F_2, and F_4 are displayed. The full set of objective function values (F_1-F_4) for each modeling scenario is provided in Table 3.

Table 3. Objective function values (F_1-F_4) for the TOPSIS-selected Pareto-optimal solutions from each modeling scenario.

Model Scenario	F_1	F_2	F_3	F_4
Chollolo Campground Daily	6.81	-3.90 x 10^3	-4.53 x 10^3	0.070
Chollolo Campground 15-Minute	5.98	-3.86 x 10^5	-5.08 x 10^5	1.10 x 10^{-5}
Reservation Boundary Daily	23.6	-2.96 x 10^3	-4.59 x 10^3	0.0024
Reservation Boundary 15-Minute	30.3	-1.80 x 10^5	-4.01 x 10^5	0.118

Note. Lower values indicate better agreement with the constraints defined in the four model objectives.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

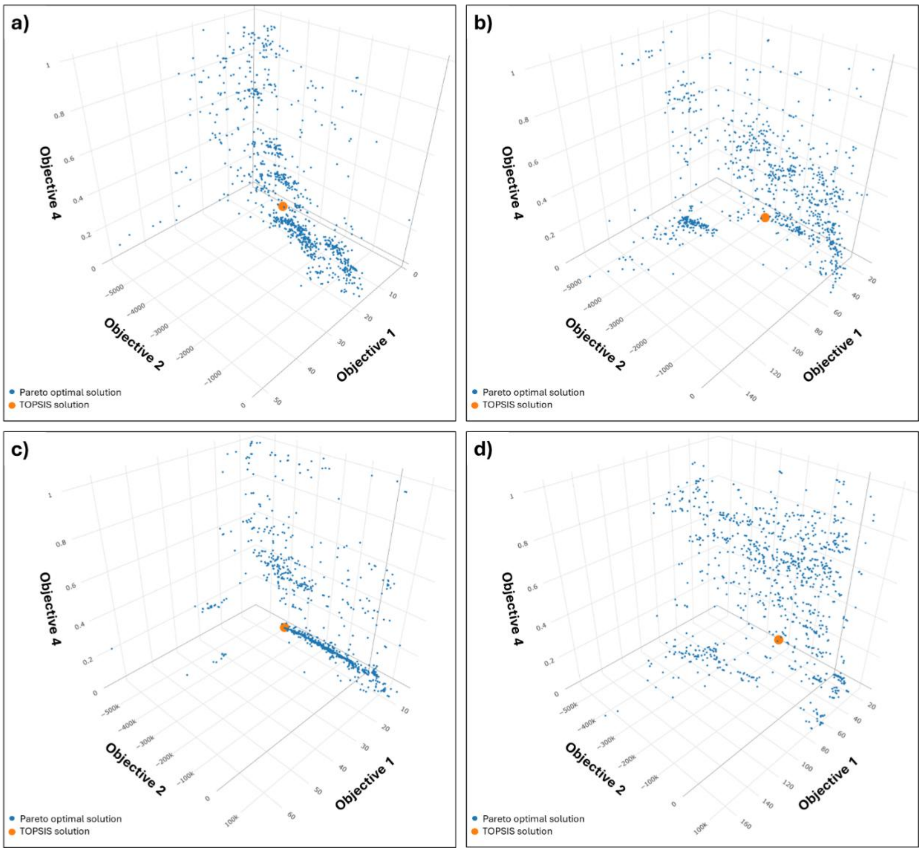

Figure 5: Pareto-optimal solution spaces after 100 generations for an offspring population of 1000 for the Chollollo Campground Daily (a), Reservation Boundary Daily (b), Chollollo Campground 15-minute (c), and Reservation Boundary 15-minute (d) modeling scenarios. Selected TOPISIS solutions are highlighted in orange.

All four Pareto-optimal solution spaces exhibit non-convex and disconnected fronts, reflecting associations in the conflicting objectives in the optimization (Deb, 2001). Notably, both Chollollo Campground modeling scenarios produce relatively tighter clustering along the F_1 axis, whereas the downstream Reservation Boundary modeling scenarios exhibit a substantially broader spread of Pareto-optimal solutions along the F_1 axis. This wider dispersion likely reflects greater hydrologic variability and complexity in the lower basin extent drained by the Reservation Boundary station. The downstream station exhibits broader ranges of both baseflow and quickflow components (Fig. 7) along with periods of zero flow conditions absent in the Chollollo Campground data records. These attributes expand the range of feasible tradeoffs between F_1 and the other three objectives. In contrast, the upstream Chollollo Campground station drains a much smaller catchment with no recorded periods of zero flow conditions, likely resulting in more stable recession behavior and thereby a narrower set of optimal solutions along F_1. These hydrologic contrasts are consistent with the watershed characteristics summarized in Table 4, which highlight the substantially lower recharge (baseflow) to precipitation ratio at the downstream Reservation Boundary station.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

Table 4. Mean recharge to precipitation ratio (mean precipitation calculated according to the general procedure in Thornton et al, 2021) for each modeling scenario.

Model Scenario	Mean Recharge (Baseflow) to Precipitation Ratio (%)
Cholollo Campground Daily	38.4
Cholollo Campground 15-Minute	39.0
Reservation Boundary Daily	18.5
Reservation Boundary 15-Minute	17.3

The differences in F_1 behavior between the modeling scenarios are reflected in the regression of measured and modeled streamflow SC (Fig. 6). RMSE values range from 7.18 to 48.09 and percent errors between 3.21% and 7.17%. Both Reservation Boundary modeling scenarios exhibit higher RMSE and percent error values than the Cholollo Campground counterparts, consistent with the broader dispersion of Pareto-optimal solutions along the F_1 axis. A striking underprediction of high modeled SC_q for both Cholollo Campground and Reservation Boundary 15-minute resolution modeling scenarios highlights prolonged missing data during low or zero flow periods and apparently unrealistic imputation (highlighted in red in Figs. 6c and d).

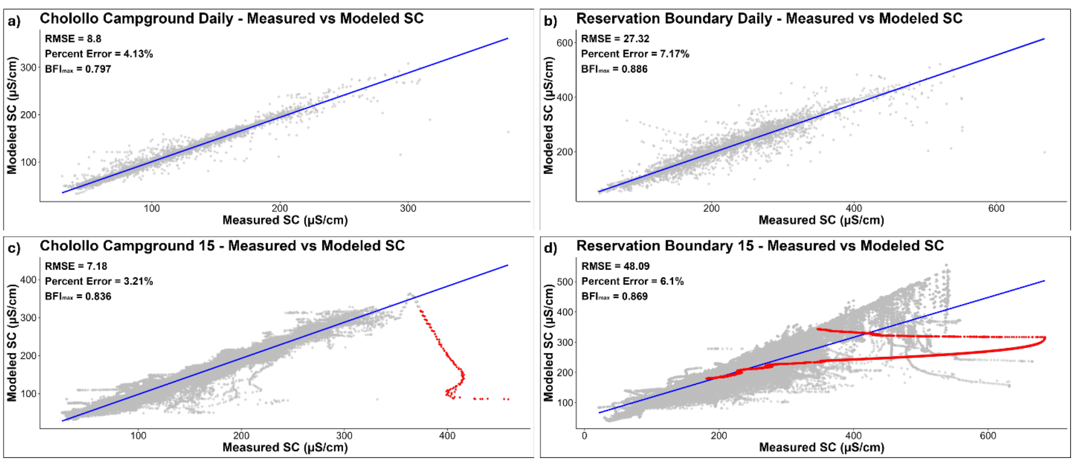

Figure 6: Comparison of measured and modeled streamflow SC using linear regression for the Cholollo Campground Daily (a), Reservation Boundary Daily (b), Cholollo Campground 15-minute (c), and Reservation Boundary 15-minute (d) modeling scenarios. Model performances summarized via root mean square error (RMSE) and percent error. Red intervals represent drought/low flow periods when the algorithm drastically underpredicts imputed streamflow SC.

F_2 and F_3 scale directly with the number of chemical-ordering comparisons in each calibration dataset and thus differ primarily between daily and 15-minute temporal resolutions rather than between stations. In contrast, F_1 and F_4 reflect aggregated hydrologic performance and structural constraints and are consequently less sensitive to dataset size and resolution. Across all four modeling scenarios, there are few instances in which the expected chemical ordering of a stream system is broken (Fig. 7). While violations of F_3 can be attributed to the limits of converting precipitation SC to a proxy for quickflow SC, violations of F_2 represent instances where the algorithm's peak optimized peak BFI values (which are dependent on the optimized ECK

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

RDF parameters and peak detection parameters) and subsequent SC_b interpolation (which is dependent on the optimized interpolation method) are unable to account for elevated streamflow SC values.

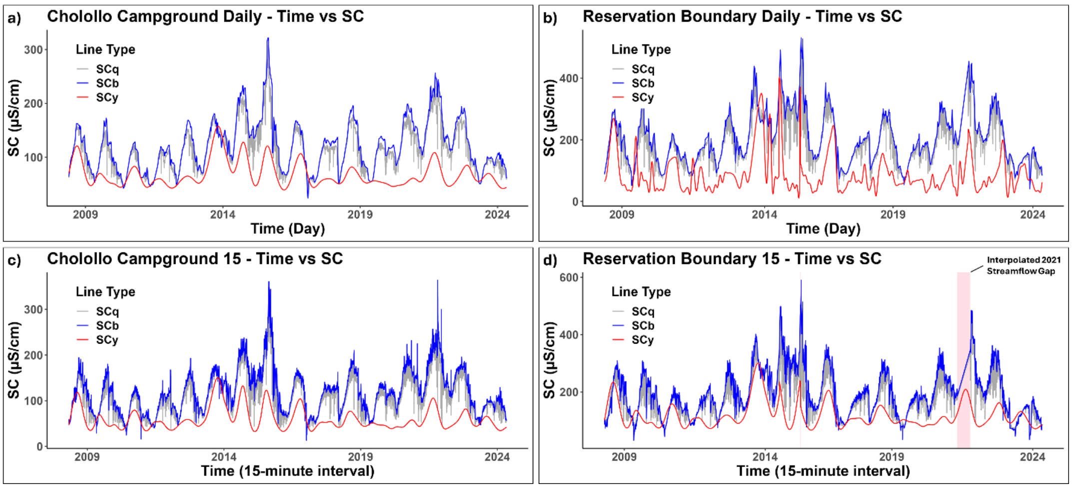

Figure 7: Modeled baseflow SC (SC_b), modeled quickflow SC (SC_y), and modeled streamflow SC (SC_q) for the Cholollo Campground Daily (a), Reservation Boundary Daily (b), Cholollo Campground 15-minute (c), and Reservation Boundary 15-minute (d) modeling scenarios. The period of prolonged (~2 month) zero flow condition at Reservation Boundary in 2021 is highlighted in d.

F_4 shows no consistent trend across the four modeling scenarios, and while all values are generally low (Fig. 5), the TOPSIS-selected values vary without a clear pattern by station or temporal resolution. This performance reflects the role of F_4 in keeping the optimized BFI time series to remain below the optimized instantaneous BFI peak at the maximum SC_b value. The low F_4 values suggest that this prevents the algorithm from artificially deflating BFI_{max} to improve SC_b-centered chemical-ordering performance (F_2). Because F_4 is minimized rather than set as a hard constraint and fixed to zero, the algorithm retains flexibility to accommodate for analytical accuracy/precision issues in the SC record and uncertainty in peak BFI estimation (see Step 5 in Fig. 1). As a result, F_4 values remain small across all four modeling scenarios, with minor difference being primarily attributable to variations in optimized BFI_{max} and peak SC_b values rather than any trends between modeling scenarios.

While the Reservation Boundary modeling scenarios display broader Pareto-optimal distributions than the upstream Cholollo Campground scenarios, the TOPSIS solutions do not indicate uniformly poorer performances across all four objectives. The downstream station scenarios perform worse primarily with respect to F_1, reflecting greater hydrologic variability and more frequent reduced flow conditions in the lower basin that ultimately affects the algorithm's ability to recreate streamflow SC in select periods of the timeseries. In contrast, F_2 and F_3 do not exhibit meaningful station-level differences. These hydrologic and structural differences expand the feasible trade-off space of the Reservation Boundary modeling scenarios without implying systematic underperformance across all four objectives.

4.2. Result Implications

A comparison between daily and 15-minute resolution modeling scenarios reveals no consistent directional trend in BFI_{max} estimates. At Reservation Boundary the daily modeling scenario yields a slightly higher BFI_{max} estimate (0.886) than the corresponding 15-minute modeling scenario (0.869), whereas the opposite pattern occurs at Cholollo Campground, where the daily scenario produces a lower BFI_{max} estimate (0.797) than the 15-minute scenario (0.836) (Fig. 6). These opposing patterns indicate that temporal resolution does not systematically influence BFI_{max} estimation.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

Although the Reservation Boundary scenarios yield slightly higher BFI_{max} estimates than the upstream Chalollo Campground scenarios (Fig. 6), the magnitude of these differences is minor and consistent with the shared hydrogeology setting of the two stations. The short spatial separation between Reservation Boundary and Cholollo Campground (Fig. 2) limits the potential for notable systematic downstream increases in baseflow contribution. Localized alluvial occurrences in the lower extent of the Reservation Boundary watershed may slightly enhance baseflow, but not to the degree of producing a more meaningful downstream trend in BFI_{max}. Thus, the observed BFI_{max} differences are more reflective of minor station-specific variation rather than a notable hydrologic gradient along the river.

All four estimated BFI_{max} values substantially exceed both the BFI_{max} values recommended by Eckhardt (2005) for perennial systems with hard-rock aquifers (0.25) and ephemeral systems with porous aquifers (0.50). Instead, all four estimates most closely aligned with the recommended value for perennial systems with porous aquifers (0.80). However, the vagueness of this definition, without any quantitative criteria (e.g., aquifer thickness, proportion of zero flow occurrences over the monitoring period, etc.), makes it difficult to assign the study setting into any of these 3 groups. In this regard, porous aquifer systems should exhibit large groundwater storage and slow drainage whereas hard-rock systems are characterized by limited storage and rapid recession. Although alluvial deposits are present along portions of the South Fork Tule River channel, they probably do not provide the degree of groundwater storage typically associated with porous aquifer systems. Recent studies of the area emphasize that the system's watershed is dominated by shallow crystalline bedrock and Mesozoic plutonic rock of the Sierra Nevada batholith, with alluvium restricted primarily to the lower western valley, as illustrated in the geologic map of Meadows and Hagedorn (2022). Additionally, the historic flashy hydrologic behavior described for the system (Hagedorn, 2020) is consistent with low-storage hard-rock aquifer systems. Comparable studies applying the ECK RDF classify their systems as perennial hard-rock aquifer systems based on similar low storage crystalline bedrock conditions, reinforcing that the South Fork Tule River fits the same classification (Mohammed & Scholz, 2018; Koffi et al., 2026).

Although the Reservation Boundary station exhibits periods of zero flow during extreme drought periods at both daily and 15-minute resolutions, this behavior is typical of low-storage crystalline bedrock systems and does not exclude the South Fork Tule River from classification as a perennial hard-rock aquifer system as described in Eckhardt (2005). Similar to aquifer classification, the Eckhardt (2005) perennial versus ephemeral classification is based on vague hydrologic functioning rather than strict year-round flow. Because Eckhardt (2005) did not include ephemeral hard-rock systems, no recommended BFI_{max} value exists for said classification. As such, the discrepancy between the modeling scenario BFI_{max} estimates and the recommended value for perennial hard-rock aquifer systems (0.25) cannot be resolved by changing the system's classification to ephemeral. Instead, the discrepancy highlights the need to calibrate BFI_{max} using physically constrained, site-specific information over general empirical values.

Comparing the baseflow outputs from the optimized BFI_{max} to those values derived from the Eckhardt's (2005) recommended BFI_{max} values, it becomes clear the Eckhardt's guidance values produce drastic baseflow underestimates (Table 5). This contrast is most pronounced when compared to output for a perennial hard-rock aquifer system (0.25), which uniformly yielded a mean baseflow value less than half of their corresponding optimized mean baseflow value (Table 5). This discrepancy underscores the need for chemically constrained optimization for accurate representation of a system's hydrologic behavior. Any large scale groundwater models calibrated against ECK RDF baseflow values derived upon default parameters should thus be treated with caution.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

Table 5. Comparison of each modeling scenario's optimized mean baseflow values against mean baseflow values derived using the Eckhardt (2005) recommended BFI_{max} values.

Model Scenario	Mean Baseflow at Optimized BFI_{max} (m^3/s)	Mean Baseflow $BFI_{max} = 0.25$ (m^3/s)	Mean Baseflow $BFI_{max} = 0.50$ (m^3/s)	Mean Baseflow $BFI_{max} = 0.80$ (m^3/s)
Chololло Campground Daily	0.392	0.127	0.256	0.393
Chololло Campground 15-Minute	0.400	0.129	0.259	0.386
Reservation Boundary Daily	1.07	0.257	0.638	0.975
Reservation Boundary 15-Minute	1.01	0.253	0.634	0.937

4.3 Model Limitations

Missing data in a modeling scenario induces structural complications that can substantially affect baseflow estimation and the estimation of modeled SC. Consistent with established limitations of the ECK RDF, *HydrOHS* requires a complete streamflow timeseries and therefore fills missing values via pchip interpolation in the preprocessing stage. Dependent on missing data period and type (continuous vs random), the imputed missing data may not adequately capture actual conditions.

A prime example of this is apparent in the Reservation Boundary 15-minute modeling scenario, which exhibits the largest interpolated missing streamflow gap from late April to mid-October 2021 (Fig. 8). Notably, this phenomenon is more apparent in the 15-minute dataset than in the daily resolution dataset (compare Fig. 8c with Fig. 8g). In the daily dataset, the missing period is bounded by zero flow measurements so missing data is realistically interpolated as zero. *HydrOHS* handles zero flow periods by removing all zero flow occurrences prior to BFI estimation after the application of the ECK RDF in the processing stage, creating an artificial, "truncated" continuous time series with an arbitrary time scale. This is not problematic when zero flow occurrences are rare in the monitoring record. Complications arise, however, when the zero flow missing data periods are not demarked with actual zero readings, but bounded by actual low flow measurements, as is the case for the Reservation Boundary 15-minute modeling scenario. Here, the missing data interpolation overwrites the true zero flow period, producing non representative flow and SC periods. These structurally inconsistent inputs degrade the BFI peak detection and subsequent SC_b interpolation far more than either zero flow periods or missing non-zero flow streamflow gaps alone.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

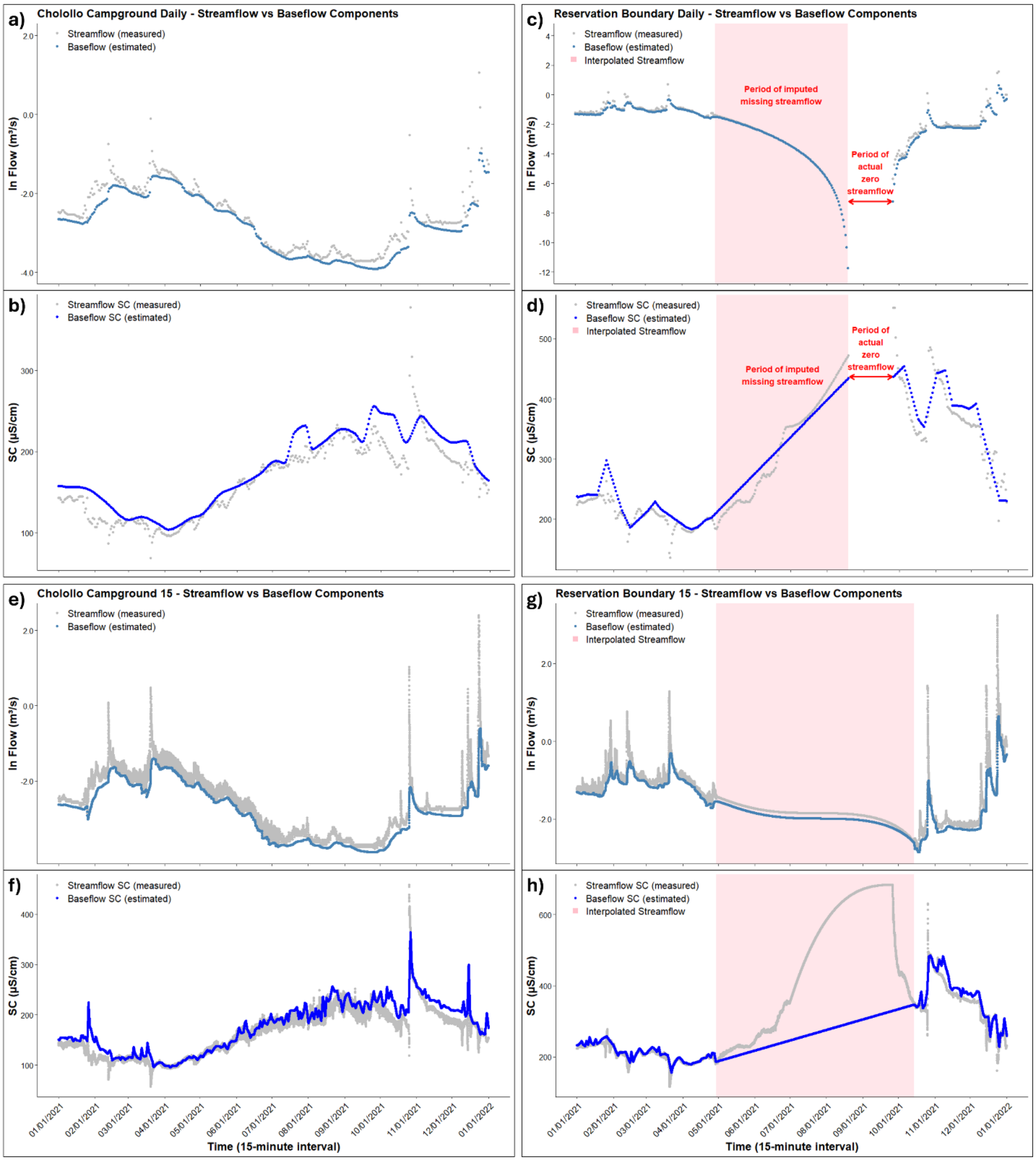

Figure 8: Streamflow/baseflow (a, c, e, g) and corresponding SC component (b, d, f, h) time series for all four modeling scenarios, highlighting periods of imputed missing streamflow in the 15-minute resolution dataset (pink) and periods of actual zero streamflow in the daily resolution dataset (red arrow) at Reservation Boundary. Note that the former period is bound by low flow measurements, while the latter is bound by actual zero flow values. Also note the shorter period of missing SC data in the 15-minute dataset

Compounding this issue, the gaps of missing streamflow and SC are not always concurrent. In the Reservation Boundary 2021 data, for example, the SC record contains a shorter interpolated gap, meaning that several intervals of high SC remain in the dataset that correspond to interpolated non-representative streamflow values. These mismatched low streamflow and high SC pairings can result in unrealistic relationships that the

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

algorithm cannot reconcile (see red data in Fig. 6). Such periods will disproportionately deflate MAE statistics of F_1 due to their sensitivity to outliers.

In addition to interpolated missing data, extended zero flow periods introduce a secondary structural complication for *HydrOHS* application. As mentioned before, *HydrOHS* filters out zero flow periods, but these effects propagate into the modeled streamflow SC signal. As illustrated by the continuous zero flow period from July to September 2014 and the near-continuous zero flow period from June to October 2015 (Fig. 9), SC_b interpolation can struggle to encompass the sudden shifts to high SC values that occur immediately after streamflow resumes. Typically, streamflow that increases from zero to low values should be accompanied by SC that "plummets" from high to lower values due to solute concentration from evaporated stagnant water at the gage or potentially from equipment error (Cartwright, 2022). The low, unstable streamflow values corresponding to high SC values will be missed by *HydrOHS* if these first high SC and resuming measured flow values are not captured by the BFI peak detection. The result is interpolated SC_b underestimating the magnitude of the post-zero-flow period SC behavior and a discrepancy between subsequent modeled streamflow SC and measured streamflow SC values.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

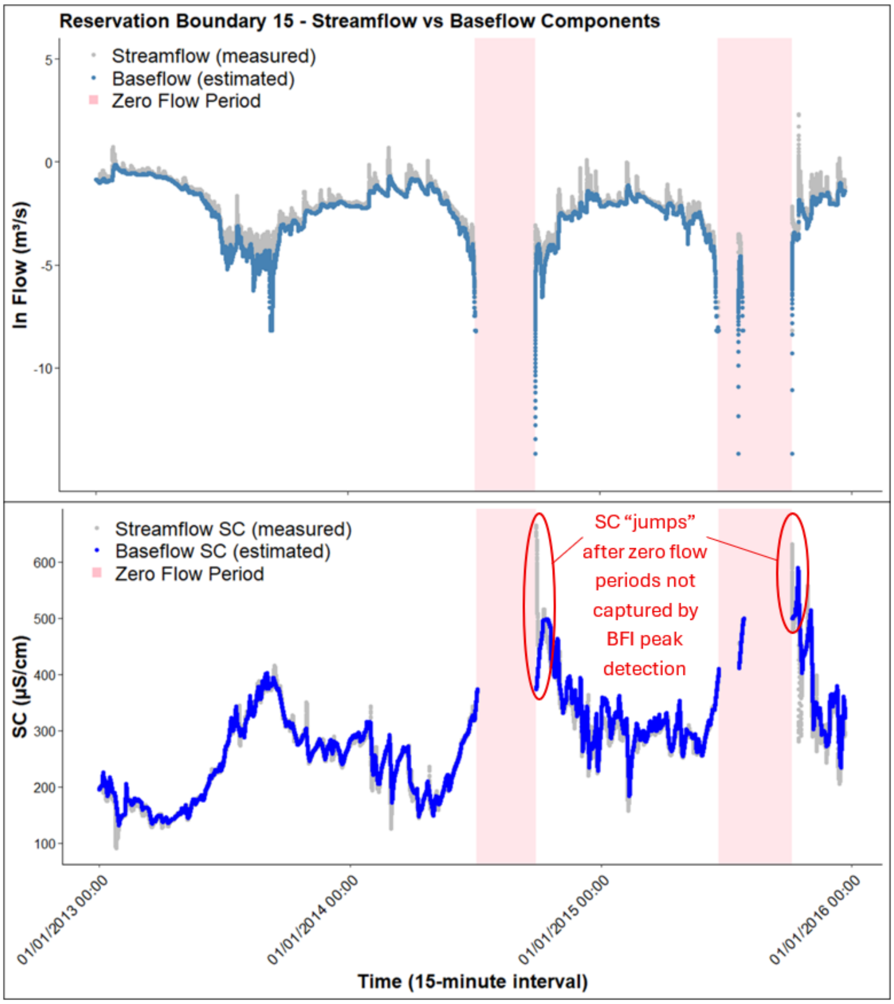

Figure 9: Streamflow/baseflow (top) and corresponding specific conductance (bottom) time series highlighting the 2014 and 2015 periods of zero streamflow in the Reservation Boundary 15-modeling scenario which are proceeded by SC maxima values.

Algorithm performance is also influenced by the occurrence of spurious measurements, particularly in the SC record. As an example, during October 2021 in the Cholollo Campground 15-minute modeling scenario, non-interpolated streamflow SC values almost triple (increase from 150 to 427 µS/cm) in just two 15-minute intervals. This SC "spike" corresponds with only a minor decrease in streamflow and should thus be considered noise rather than signal. Notably, in the Cholollo Campground daily modeling scenario, the SC spike and its surrounding transitional behavior are aggregated into a single anomalous daily value minimizing the effect of that spurious spike.

Another limitation of *HydrOHS* relevant to both temporal resolutions is the availability of precipitation chemistry data. For all four modeling scenarios in this study, precipitation SC data was sourced from the nearest NADP monitoring station, located over 45 km away from either stream gage station along the South Fork Tule River. *HydrOHS* does not explicitly account for spatial separation between data sources when estimating quickflow SC. Instead, the algorithm converts measured precipitation SC into a likely quickflow SC signal, reflecting the assumption that quickflow represents precipitation input with relatively short residence time. This assumption is necessary given that quickflow cannot be directly measured. However, the absence of a clear threshold for acceptable distances between precipitation chemistry and stream chemistry data sources remains a limitation for future *HydrOHS* applications.

5. Conclusions

Here we present a physically constrained OHS algorithm, *HydrOHS,* that informs ECK RDF outputs with CMB (and vice versa) in a MOO framework. The results of the algorithm calibration process can be summarized as three primary findings. First, baseflow contributions within the study system showed a minor change between the upstream and downstream gage stations, suggesting some degree of spatial variability in hydrogeological conditions in the system. Second, while higher resolution datasets can improve calibration stability and capture hydrologic behavior nuances otherwise aggregated at lower resolutions, daily resolution datasets remain sufficient for robust hydrograph separation applications. Finally, algorithm-derived BFI_{max} estimates diverge greatly from the BFI_{max} values recommended in Eckhardt (2005). This divergence underscores the key implication of the study that relying on literature defined BFI_{max} recommendations can introduce structural bias into hydrograph separation and groundwater model calibration. *HydrOHS* demonstrates that BFI_{max} is highly catchment specific and should be calibrated whenever supporting hydrologic and chemical information is processed in a non-subjective multi objective optimization framework.

Moving forward, several attributes would strengthen future applications of *HydrOHS*. Longer time series data would improve algorithm performance by reducing sensitivity to short-term anomalies. As illustrated by the disparities between Cholollo Campground and Reservation Boundary performances, applications with minimal degrees of missing streamflow data during droughts can greatly influence algorithm performance. While no definitive threshold for acceptable distance has yet been established, closer and more representative precipitation chemistry stations would reduce uncertainty in utilizing precipitation SC data for quickflow SC estimation, thereby improving algorithm performance in relation to chemical component ordering. Additionally, independent baseflow tracers such as radon (^{222}Rn) could provide a valuable external constraint on baseflow contributions, offering a chemically independent check on BFI_{max} estimation.

Beyond its application to the study watershed, the *HydrOHS* algorithm offers potential use towards groundwater resource assessments at any scale or setting. Ongoing application of the algorithm in future research will further evaluate its performance across diverse hydrogeologic environments and assess its potential as a tool for reliable water resource assessments.

6. Code and Data Availability

Finalized code for HydrOHS, Cholollo Campground daily resolution example data (Chololo Campground streamflow data, Chololo Campground SC data, and NADP NTN precipitation chemistry data), and outlines for code application are publicly available via the GitHub repository (https://github.com/E-Villasano/HydrOHS) and archived through Zenodo (Villasano 2026).

7. Author contribution

EEV: Conceptualization, Methodology, Data Curation, Software, Validation, Analysis, Investigation, Writing, Visualization, Writing – Original Draft, Writing – Review & Editing. **BH:** Conceptualization, Methodology, Data Curation, Software, Investigation, Supervision, Funding Acquisition, Visualization, Writing – Review & Editing. **NRJ:** Methodology, Software, Investigation, Visualization.

8. Competing Interests

The authors declare that they have no conflict of interest.

9. Acknowledgements

We would like to recognize the Collaboration of Native Nations for Climate Transformation and Stewardship for their support of this research. We also thank the Tule River Indian Tribe of the Tule River Reservation for enabling granting access to the portion of the South Fork Tule River located within the Tule River Reservation.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

10. References

Arnold, J. G. and Allen, P. M.: Estimating hydrologic budgets for three Illinois watersheds, J. Hydrol., 176, 57-77, https://doi.org/10.1016/0022-1694(95)02782-3, 1996.

Arnold, J. G. and Allen, P. M.: Automated methods for estimating baseflow and ground water recharge from streamflow records, J. Am. Water Resour. Assoc., 35, 411–424, https://doi.org/10.1111/j.1752-1688.1999.tb03599.x, 1999.

Borchers, H. W.: pracma: Practical Numerical Math Functions, https://doi.org/10.32614/CRAN.package.pracma, 2023.

Brutsaert, W. and Nieber, J. L.: Regionalized drought flow hydrographs from a mature glaciated plateau, Water Resour. Res., 13, 637-643, https://doi.org/10.1029/WR013i003p00637, 1977.

Cartwright, I.: Implications of variations in stream specific conductivity for estimating baseflow using chemical mass balance and calibrated hydrograph techniques, Hydrol. Earth Syst. Sci., 26, 183-195, https://doi.org/10.5194/hess-26-183-2022, 2022.

Cartwright, I., Gilfedder, B., and Hofmann, H.: Contrasts between estimates of baseflow help discern multiple sources of water contributing to rivers, Hydrol. Earth Syst. Sci., 18, 15–30, https://doi.org/10.5194/hess-18-15-2014, 2014.

Chapman, T.: A comparison of algorithms for stream flow recession and baseflow separation, Hydrol. Process., 13, 701–714, https://doi.org/10.1002/(SICI)1099-1085(19990415)13:5<701::AID-HYP774>3.0.CO;2-2, 1999.

Chapman, T. and Maxwell, A. I.: Baseflow Separation - Comparison of Numerical Methods with Tracer Experiments, Hydrology and Water Resources Symposium 1996: Water and the Environment; Institution of Engineers, Australia, 1996.

Cheng, L., Zhang, L., and Brutsaert, W.: Automated Selection of Pure Base Flows from Regular Daily Streamflow Data: Objective Algorithm, J. Hydrol. Eng., 21, 06016008, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001427, 2016.

Collischonn, W. and Fan, F. M.: Defining parameters for Eckhardt's digital baseflow filter, Hydrol. Process., 27, 2614–2622, https://doi.org/10.1002/hyp.9391, 2013.

De Boor, C.: A practical guide to splines / Carl de Boor., Springer-Verlag, New York, 1978.

Deb, K.: Multi-Objective Optimization using Evolutionary Algorithms, 1st ed., John Wiley & Sons, New York, 2001.

Deb, K. and Jain, H.: An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints, IEEE Trans. Evol. Computat., 18, 577–601, https://doi.org/10.1109/TEVC.2013.2281535, 2014.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Computat., 6, 182-197. https://doi.org/10.1109/4235.996017, 2002.

Eckhardt, K.: How to construct recursive digital filters for baseflow separation, Hydrol. Process., 19, 507-515, https://doi.org/10.1002/hyp.5675, 2005.

Eckhardt, K.: A comparison of baseflow indices, which were calculated with seven different baseflow separation methods, J. Hydrol., 352, 168–173, https://doi.org/10.1016/j.jhydrol.2008.01.005, 2008.

Eckhardt, K.: Technical Note: Analytical sensitivity analysis of a two parameter recursive digital baseflow separation filter, Hydrol. Earth Syst. Sci., 16, 451-455. https://doi.org/10.5194/hess-16-451-2012, 2012.

Eckhardt, K.: Technical note: How physically based is hydrograph separation by recursive digital filtering?, Hydrol. Earth Syst. Sci., 27, 495-499. https://doi.org/10.5194/hess-27-495-2023, 2023.

Esri.: World Imagery, ArcGIS Online, https://www.arcgis.com, last access: March 2025.

Foks, S. S., Raffensperger, J. P., Penn, C. A., and Driscoll, J. M.: Estimation of Base Flow by Optimal Hydrograph Separation for the Conterminous United States and Implications for National-Extent Hydrologic Models, Water, 11, 1629, https://doi.org/10.3390/w11081629, 2019.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

Fritsch, F. N. and Carlson, R. E.: Monotone Piecewise Cubic Interpolation, SIAM J. Numer. Anal., 17, 238-246, https://doi.org/10.1137/0717021, 1980.

Gonzales, A. L., Nonner, J., Heijkers, J., and Uhlenbrook, S.: Comparison of different base flow separation methods in a lowland catchment, Hydrol. Earth Syst. Sci., 13, 2055-2068, 2009.

Hagedorn, B.: Hydrograph separation through multi objective optimization: Revealing the importance of a temporally and spatially constrained baseflow solute source, J. Hydrol., 590, 125349, https://doi.org/10.1016/j.jhydrol.2020.125349, 2020.

Hagedorn, B. and Meadows, C.: Trend Analyses of Baseflow and BFI for Undisturbed Watersheds in Michigan—Constraints from Multi-Objective Optimization, Water, 13, 564, https://doi.org/10.3390/w13040564, 2021.

Hagedorn, B. and Whittier, R. B.: Solute sources and water mixing in a flashy mountainous stream (Pahsimeroi River, U.S. Rocky Mountains): Implications on chemical weathering rate and groundwater–surface water interaction, Chem. Geol., 391, 123-137, https://doi.org/10.1016/j.chemgeo.2014.10.031, 2015.

Helfer, F., Bernardi, F. K., De Barros, C. A. P., Piccilli, D. G. A., Minella, J. P. G., Tassi, R., and Schlesner, A. A.: Calibrated Eckhardt's filter versus alternative baseflow separation methods: A silica-based approach in a Brazilian catchment, J. Hydrol., 644, 132073, https://doi.org/10.1016/j.jhydrol.2024.132073, 2024.

Huang, X., Hall, A. D., and Berg, N.: Anthropogenic Warming Impacts on Today's Sierra Nevada Snowpack and Flood Risk, Geophys. Res. Lett., 45, 6215–6222, https://doi.org/10.1029/2018GL077432, 2018.

Hwang, C. L. and Yoon, K.: Multiple attribute decision making: methods and applications: a state-of-the-art survey / Ching-Lai Hwang, Kwangsun Yoon., Springer-Verlag, Berlin, 1981.

Ishibuchi, H., Imada, R., Setoguchi, Y., and Nojima, Y.: Performance comparison of NSGA-II and NSGA-III on various many-objective test problems, in: 2016 IEEE Congress on Evolutionary Computation (CEC), 3045-3052, https://doi.org/10.1109/CEC.2016.7744174, 2016.

Kaviany E.: Zur Hydrogeologie im Niederschlagsgebiet der Dill (Hessen), Gießener Geologische Schriften, 19, 248, 1978.

Kille K.: Das Verfahren MoMNQ—ein Beitrag zur Berechnung der mittleren langj¨ahrigen Grundwasserneubildung mit Hilfe der monatlichen Niedrigwasserabfl¨usse, Zeitschrift der Deutschen Geologischen Geseltschaft, Sonderheft Hydrogeologie/Hydrogeochemic, 89–95, 1970.

Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resour. Res., 45, https://doi.org/10.1029/2008WR006912, 2009.

Koffi, K. V., Obuobie, E., Soro, D. D., and Wohnlich, S.: Automated Baseflow Separation and Estimation of Groundwater Recharge Using Master Recession Curve Technique in the Sissili Basin, Open J. Mod. Hydrol., 16, 158–176, https://doi.org/10.4236/ojmh.2026.162010, 2026.

Kronholm, S. C. and Capel, P. D.: A comparison of high-resolution specific conductance-based end-member mixing analysis and a graphical method for baseflow separation of four streams in hydrologically challenging agricultural watersheds, Hydrol. Process., 29, 2521-2533. https://doi.org/10.1002/hyp.10378, 2015.

Liu, F., Williams, M. W., and Caine, N.: Source waters and flow paths in an alpine catchment, Colorado Front Range, United States, Water Resour. Res, 40, 2004WR003076, https://doi.org/10.1029/2004WR003076, 2004.

Lyne, V. and Hollick, M.: Stochastic time-variable rainfall-runoff modelling, in: Hydrology and Water Resources Symposium, Perth, Australia, Institution of Engineers, Australia, 1979.

Ma, H., Zhang, Y., Sun, S., Liu, T., and Shan, Y.: A comprehensive survey on NSGA-II for multi-objective optimization and applications, Artif. Intell. Rev., 56, 15217-15270, https://doi.org/10.1007/s10462-023-10526-z, 2023.

Maurer, E. P.: Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California, under two emissions scenarios, Climatic Change, 82, 309–325, https://doi.org/10.1007/s10584-006-9180-9, 2007.

Meadows, C. and Hagedorn, B.: Temporal and Spatial Patterns of Groundwater Recharge Across a Small Watershed in the California Sierra Nevada Mountains, Front. Water, 4, 815228, https://doi.org/10.3389/frwa.2022.815228, 2022.

Miller, C. and Urban, D. L.: Forest Pattern, Fire, and Climatic Change in the Sierra Nevada, Ecosystems, 2, 76–87, https://doi.org/10.1007/s100219900060, 1999.

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

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

Mohammed, R. and Scholz, M.: Flow–duration curve integration into digital filtering algorithms for simulating climate variability based on river baseflow, Hydrol. Sci. J., 63, 1558-1573, https://doi.org/10.1080/02626667.2018.1519318, 2018.

Natermann E.: Die Linie des langfristigen Grundwassers (AuL) und die Trockenwetterabflußlinie (TWL), Die Wasserwirtschaft, Sonderh, 1950, 12–20, 1951.

Nathan, R. J. and McMahon, T. A.: Evaluation of automated techniques for base flow and recession analyses, Water Resour. Res., 26, 1465–1473, https://doi.org/10.1029/WR026i007p01465, 1990.

National Atmospheric Deposition Program (NRSP-3): National Trends Network (NTN) site CA75, Sequoia National Park–Giant Forest, California, Wisconsin State Laboratory of Hygiene, University of Wisconsin–Madison [data set], https://nadp.slh.wisc.edu/sites/ntn-CA75/, 2024.

Natural Earth.: Admin 1- States and Provinces, https://www.naturalearthdata.com, last access: March 2025.

Nielsen, M. G. and Westenbroek, S. M.: Groundwater recharge estimates for Maine using a Soil-Water-Balance model—25-year average, range, and uncertainty, 1991 to 2015, US Geol. Surv., Reston, VA, https://doi.org/10.3133/sir20195125, 2019.

Null, S. E., Viers, J. H., and Mount, J. F.: Hydrologic Response and Watershed Sensitivity to Climate Warming in California's Sierra Nevada, PLoS ONE, 5, e9932, https://doi.org/10.1371/journal.pone.0009932, 2010.

Pelletier, A. and Andréassian, V.: Hydrograph separation: an impartial parametrisation for an imperfect method, Hydrol. Earth Syst. Sci., 24, 1171–1187. https://doi.org/10.5194/hess-24-1171-2020, 2020.

Press, W. H.: Numerical recipes 3rd edition: The art of scientific computing, Cambridge University Press, Cambridge, UK, 2007.

QGIS Development Team: QGIS Geographic Information System, QGIS [code], https://www.qgis.org, 2025.

Raffensperger, J. P., Baker, A. C., Blomquist, J. D., and Hopple, J. A.: Optimal hydrograph separation using a recursive digital filter constrained by chemical mass balance, with application to selected Chesapeake Bay watersheds, U.S. Geol. Surv., 2017-5034, 2017.

Rimmer, A. and Hartmann, A.: Optimal hydrograph separation filter to evaluate transport routines of hydrological models, J. Hydrol., 514, 249-257, https://doi.org/10.1016/j.jhydrol.2014.04.033, 2014.

Rumsey, C. A., Miller, M. P., Susong, D. D., Tillman, F. D., and Anning, D. W.: Regional scale estimates of baseflow and factors influencing baseflow in the Upper Colorado River Basin, J. Hydrol.: Reg. Stud., 4, 91–107, https://doi.org/10.1016/j.ejrh.2015.04.008, 2015.

Sanford, W. E., Nelms, D. L., Pope, J. P., and Selnick, D. L.: Quantifying components of the hydrologic cycle in Virginia using chemical hydrograph separation and multiple regression analysis, U.S. Geol. Surv., 2011-5198, 2012.

Saraiva Okello, A. M. L., Uhlenbrook, S., Jewitt, G. P. W., Masih, I., Riddell, E. S., and Van Der Zaag, P.: Hydrograph separation using tracers and digital filters to quantify runoff components in a semi-arid mesoscale catchment, Hydrol. Process., 32, 1334–1350, https://doi.org/10.1002/hyp.11491, 2018.

See, C. R., Green, M. B., Yanai, R. D., Bailey, A. S., Campbell, J. L., and Hayward, J.: Quantifying uncertainty in annual runoff due to missing data, PeerJ, 8, e9531, https://doi.org/10.7717/peerj.9531, 2020.

https://doi.org/10.5194/egusphere-2026-3227
Preprint. Discussion started: 20 July 2026

Sloto, R. A. and Crouse, M. Y.: HYSEP: A computer program for streamflow hydrograph separation and analysis, U.S. Geol. Surv., 96-4040, 1996.

Tallaksen, L. M.: A review of baseflow recession analysis, J. Hydrol., 165, 349-370, 1995.

Teegavarapu, R. S. V.: Imputation methods for missing hydrometeorological data estimation, Springer International Publishing AG, 2024.

Thomas, B. F., Vogel, R. M., and Famiglietti, J. S.: Objective hydrograph baseflow recession analysis, J. Hydrol., 525, 102-112. https://doi.org/10.1016/j.jhydrol.2015.03.028, 2015.

Thornton, P. E., Shrestha, R., Thornton, M., Kao, S.-C., Wei, Y., and Wilson, B. E.: Gridded daily weather data for North America with comprehensive uncertainty quantification, Sci. Data, 8, 190, https://doi.org/10.1038/s41597-021-00973-0, 2021.

U.S. Geological Survey: USGS water data for the Nation: U.S. Geological Survey National Water Information System database, https://doi.org/10.5066/F7P55KJN, last access: August 2025.

Villasano, E. E.: HydrOHS, https://doi.org/10.5281/zenodo.20534164, 2026

Wahl, K. L. and Wahl, T. L.: Determining the Flow of Comal Springs at New Braunfels, Texas, Texas Water '95; American Society of Civil Engineers: San Antonio, TX, USA, 1995.

Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res., 30, 79–82. https://doi.org/10.3354/cr030079, 2005.

Yang, W., Xiao, C., Zhang, Z., and Liang, X.: Can the two-parameter recursive digital filter baseflow separation method really be calibrated by the conductivity mass balance method?, Hydrol. Earth Syst. Sci., 25, 1747–1760, https://doi.org/10.5194/hess-25-1747-2021, 2021.

Zeileis, A., Kleiber, C., and Zeileis, M. A.: Package 'ineq'. Tech. Rep., 2009.

Zhang, R., Li, Q., Chow, T. L., Li, S., and Danielescu, S.: Baseflow separation in a small watershed in New Brunswick, Canada, using a recursive digital filter calibrated with the conductivity mass balance method, Hydrol. Process., 27, 2659–2665. https://doi.org/10.1002/hyp.9417, 2013.