sHydrology

sHydrology is a *sh*iny implementation of *hydrologi*cal analysis tools. All of the analysis included as part of the sHydrology suite is computed automatically given a streamflow timeseries and its known catchment area. This document outlines all theory involved in the sHydrology toolset. The code is open source, written in R, and can be accessed through the author's GitHub account.

The sHydrology toolset has been categorized into four main functionalities:

1. Hydrographs: This provides the user the ability to review the hydrograph including any flags associated with their dataset. From within the viewer window range, some basic statistics can be viewed on the fly and compared with the full record.

   (a) Separation: Here, a number of published hydrograph separation algorithms are applied to the data on the fly. Both the range and median "baseflow" estimates can be view with respect to the total flow hydrograph.

   (b) Disaggregation: These tools allow the user to separate the hydrograph into it's constituent parts. Signals from the hydrograph that can be identified include baseflow hydrograph, baseflow recession, rising & falling limbs, and event volumes (i.e., hydrograph discretization).

2. Longterm tend analysis: This set of plots give insight into the long term trends of the hydrograph. Seasonal influences and stationarity of the flow regime can be investigated here both on total flow and separated baseflow.

3. Statistics: A suite of common statistical analyses are included here to allow the user to identify lone term characteristics from peak flow, drought conditions to flow regimes.

4. Data download: The data computed on the timeseries from the variety of tools can be exported by the user such that they can perform further analysis or re-plot the data using their preferred graphical interface.

Note: Many of static plots can be downloaded simply by right-clicking the plot and save as image to disk.

Hydrograph Timeseries

The first page shown when entering the sHydrology app is the raw timeseries. From this (and only this) dataset do the remaining pages perform their analysis. The user may select a subset range within this timeseries and the main hydrograph page will automatically update to provide the user with statistics from within the selected range that can be quickly compared with the full timeseries. This gives the user quick access to revealing change to the flow regime and thus (at least, qualitatively) assess gauge stationarity.

2.1 Automatic hydrograph separation

> "Division of a hydrograph into direct and groundwater runoff as a basis for subsequent analysis is known as *hydrograph separation* or *hydrograph analysis*. Since there is no real basis for distinguishing between direct and groundwater flow in a stream at any instant, and since definitions of these two components are relatively arbitrary, the method of separation is usually equally arbitrary." — Linsley et. al., 1975

For hydrograph separation, it is generally assumed that total flow (q) at any particular time (t) of a streamflow hydrograph can be partitioned into two primary components:

1. The *slow flow* component (b), which is itself composed of the gradual release of water from watershed stores in addition to groundwater discharging into streams (i.e., "groundwater runoff" in Linsley et. al., 1975). The slow flow component is commonly referred to as "baseflow." and,

2. The *quick flow* component (f), which originates from rainfall and/or snow melt events (i.e., "direct runoff" in Linsley et. al., 1975).

Together, the slow and quick flow components sum to total flow: $q = b + f$. Conceptually, after a period of time following a precipitation event, streamflow tends to decrease steadily as it is composed entirely of baseflow ($f = 0$). Upon the onset of a heavy rain event, the hydrograph quickly rises, as quick flow is added to the baseflow signature. One could imagine that should this rain event never occur, the underlying baseflow would have continued uninterrupted (such as in Reed et. al., 1975). The difference between this "underlying" baseflow and actual flow is perceived as total runoff in the form of quick flow.

Hydrologists found the need to separate the hydrograph into its constitutive components as it was found that runoff created from a precipitation event (i.e., rainfall and/or snow melt) tended to correlate best with the quick flow component only, as opposed to the total flow hydrograph (Beven, 2012). Consequently, a number of automatic hydrograph separation routines were proposed, all (as in Linsley et. al., 1975) being "equally arbitrary."

In general, two metrics can be determined from hydrograph separation that need definition. The first is the baseflow index (BFI), which is the ratio of long term baseflow discharge to total discharge:

$$\text{BFI} = \frac{\sum b}{\sum q} \tag{2.1}$$

The second is the baseflow recession coefficient (k), which describes the withdrawal of water from storage within the watershed (Linsley et. al., 1975). The recession coefficient is a means of determining the amount baseflow recedes after a given period of time:

$$b_t = k b_{t-1}, \tag{2.2}$$

where b_{t-1} represents the slow flow calculated at one timestep prior to b_t. (Note, this assumes that total flow measurements are reported at equal time intervals, when unequal intervals are used, $k^{\Delta t}$ must be used, where Δt is the time interval between successive b calculations—for now, constant intervals are assumed.)

Linsley et. al.(1975) also offered an approximate means of determining the time (in days) after peak flow discharge to when quick flow ceases ($f = 0$) and thus total flow is entirely composed of the slow flow component and thus can be predicted using Equation 2.2. As a "rule of thumb" (Linsley et. al., 1975) the number of days (N) when quick flow terminates is approximated by:

$$N = 0.827A^{0.2}, \tag{2.3}$$

where A is the watershed area [km^2]. The above empirical relation is included here as a number of automatic hydrograph separation algorithms discussed below utilizes this approximation. Hydrograph components and "quick flow cessation" (N) is implicitly conceptualized when performing automatic hydrograph separation routines (Figure 2.1).
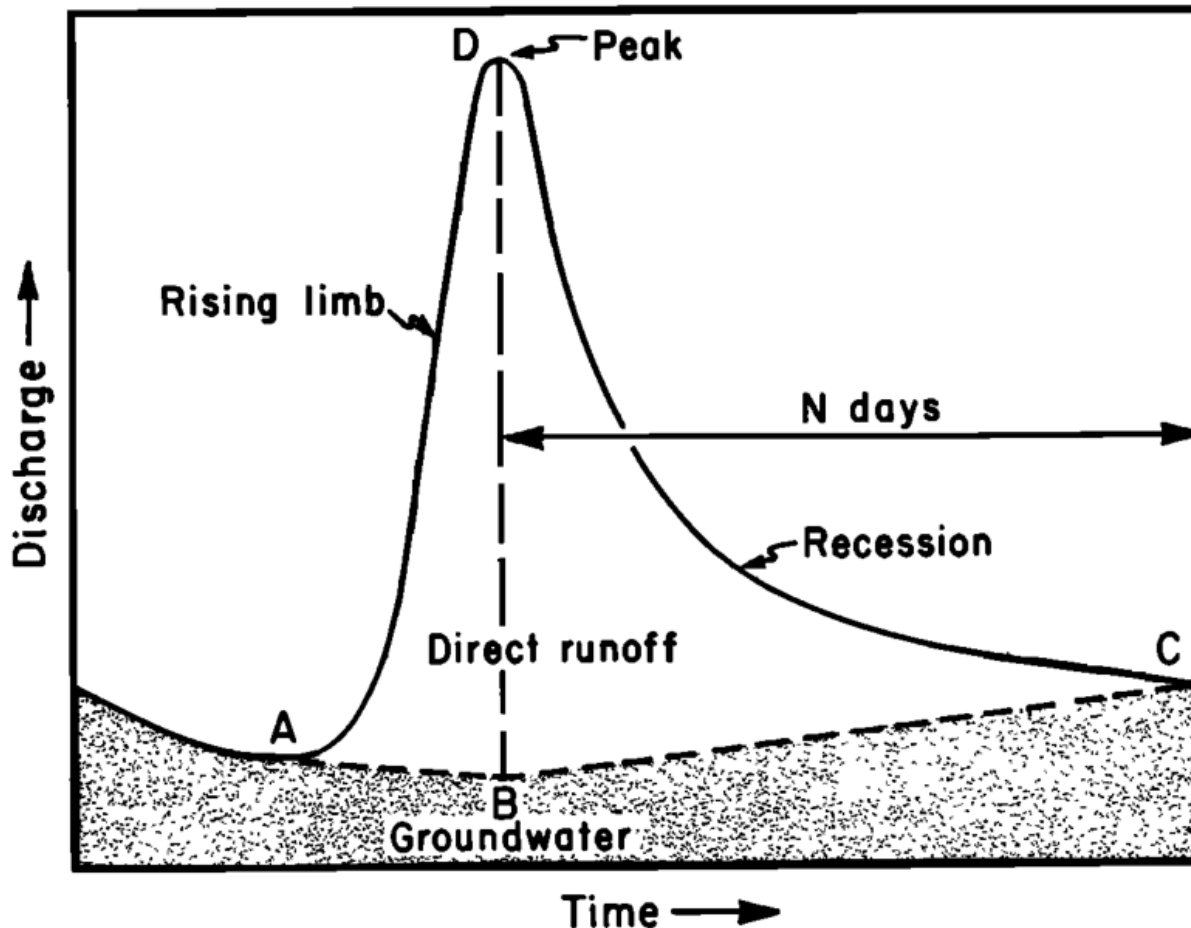

Figure 2.1: The conceptual hydrograph separated into its quick (direct) and slow (groundwater) components (after Linsley and Franzini, 1964).

2.1.1   Digital filters

Digital filters represent a set of automatic hydrograph separation algorithms that require no input other than the measured stream flow signal (q). Considering the streamflow hydrograph as a signal is quite apt when dealing with digital filters, as they themselves were inspired from signal processing of Lyne and Hollick, 1979 (Nathan and McMahon, 1990). With respect to the quick and slow hydrograph components (Figure 2.1), hydrograph separation is nothing more than the application of a low-pass filter to the total streamflow signal.

Another point to note is that many authors have applied these digital filters in multiple passes, either in two-passes (forward–backward) or three-passes (forward–backward–forward) to increase the smoothing of the resulting slow flow signal (Chapman, 1991).

With digital filters, there is no physical interpretation to the algorithm, it only produces a baseflow signal that resembles what one would expect. The general form of all digital filters used for hydrograph separation follows:

$$b_t = \alpha b_{t-1} + \beta \left(q_t + \gamma q_{t-1} \right), \tag{2.4}$$

where q_{t-1} represents the total flow measured at one timestep prior to q_t, and α, β and γ are parameters. The above equation is a three-parameter equation, however most implementations do not require every parameter be specified or, in other cases, two or more parameters can be specified as a function of another.

For example, the Lyne and Hollick (1979) equation (the earliest of digital filters used for hydrograph separation), is a one-parameter equation found by a single smoothing parameter a suggested to be set between the values of 0.9–0.95 (Nathan and McMahon, 1990), where:

$$\alpha = a \qquad \beta = \frac{1-a}{2} \qquad \gamma = 1.0 \tag{2.5}$$

Chapman (1991), after noting some conceptual discrepancies with the Lyne and Hollick (1979) equation, modified the equation into a parameter-less form as a function of the recession coefficient k, which can be determined from the q time series (equation 2.2). The Chapman (1991) algorithm takes the form:

$$\alpha = \frac{3k-1}{3-k} \qquad \beta = \frac{1-k}{3-k} \qquad \gamma = 1.0 \tag{2.6}$$

Chapman and Maxwell (1996) later simplified the above equation by assuming that slow flow is the weighted average of quick flow and the slow flow from the previous timestep (Chapman, 1999), that is $b_t = kb_{t-1} + (1-k)f_t$, leading to:

$$\alpha = \frac{k}{2-k} \qquad \beta = \frac{1-k}{2-k} \qquad \gamma = 0.0 \tag{2.7}$$

Boughton (1993) used a similar approach to Chapman and Maxwell (1996), except added an adjustment parameter C, such that $b_t = kb_{t-1} + Cf_t$. The Boughton (1993) form of the digital filter thus requires:

$$\alpha = \frac{k}{1+C} \qquad \beta = \frac{C}{1+C} \qquad \gamma = 0.0 \tag{2.8}$$

While also investigating the generalized digital filter, Eckhardt (2005) discovered an interpretation of the Boughton (1993) algorithm that eliminated the C parameter and introduced the concept of $\text{BFI}_{\max}$: the maximum value of the baseflow index that can be achieved using the digital filter. The Eckhardt (2005) digital filter is found by:

$$\alpha = \frac{(1-\text{BFI}_{\max})k}{1-k\text{BFI}_{\max}} \qquad \beta = \frac{(1-k)\text{BFI}_{\max}}{1-k\text{BFI}_{\max}} \qquad \gamma = 0.0 \tag{2.9}$$

or made equivalent to Boughton (1993) by setting:

$$C = \frac{(1-k)\text{BFI}_{\max}}{1-\text{BFI}_{\max}} \tag{2.10}$$

The commonly cited Jakeman and Hornberger (1993) algorithm closely follows that of Boughton (1993) and Chapman and Maxwell (1996), except it was formulated from a component of the IHACRES data-based model rather than being intended strictly for hydrograph separation (Chapman, 1999). Nonetheless, the IHACRES model can be shown to fit the general digital filter of equation 2.4 above, using 3 parameters, where:

$$\alpha = \frac{a}{1+C} \qquad \beta = \frac{C}{1+C} \qquad \gamma = \alpha_s \tag{2.11}$$

Note that setting $\alpha_s < 0$ is conceptually correct, as it implies that the rate of change of slow flow is positively correlated the rate of change of total flow (Chapman, 1999).

Lastly, Tularam and Ilahee (2008) most recently presented an digital filter that also resembled that of Chapman and Maxwell (1996), with the slight difference of assuming that slow flow is the weighted average of the slow flow of the previous timestep and total flow, not quick flow (i.e., $b_t = ab_{t-1} + (1-a)q_t$). This formulation is essentially the same as Lyne and Hollick (1979) with the exception that Tularam and Ilahee (2008) does not average total flow of the current and previous timestep. The one-parameter Tularam and Ilahee (2008) form yields:

$$\alpha = a \qquad \beta = 1-a \qquad \gamma = 0.0 \tag{2.12}$$

Digital filter equations in their published form:

Lyne and Hollick (1979):

$$b_t = ab_{t-1} + \frac{1-a}{2}(q_t + q_{t-1}) \tag{2.13}$$

Chapman (1991):

$$b_t = \frac{3k-1}{3-k}b_{t-1} + \frac{1-k}{3-k}(q_t + q_{t-1}) \tag{2.14}$$

Chapman and Maxwell (1996):

$$b_t = \frac{k}{2-k}b_{t-1} + \frac{1-k}{2-k}q_t \tag{2.15}$$

Boughton (1993):

$$b_t = \frac{k}{1+C}b_{t-1} + \frac{C}{1+C}q_t \tag{2.16}$$

Eckhardt (2005):

$$b_t = \frac{(1-\text{BFI}_{\max})kb_{t-1} + (1-k)\text{BFI}_{\max}q_t}{1 - k\text{BFI}_{\max}} \tag{2.17}$$

Jakeman and Hornberger (1993):

$$b_t = \frac{a}{1+C}b_{t-1} + \frac{C}{1+C}(q_t + \alpha_s q_{t-1}) \tag{2.18}$$

Tularam and Ilahee (2008):

$$b_t = ab_{t-1} + (1-a)q_t \tag{2.19}$$

2.1.2 Moving-window methods

A second class of hydrograph separation schemes are here considered "moving window methods" also known as "manual separation techniques" in Arnold and Allen (1999). These methods do not follow an equation, per se, rather a methodology based on the explicit/manual selection of discharge values assumed representative of baseflow discharge within a window of a set number of days.

In total, 10 estimates of baseflow discharge are computed using variants of 4 methods. Many of these methods are included in stand-alone software packages and have been re-coded here. The methods include:

1. The UKIH/Wallingford technique (Institute of Hydrology, 1980). This method operates by locating minimum discharges in a (user specified) N-day window. This set of minimum discharge is then further screened, automatically, for discharges that are considered representative of "baseflow," which are deemed "turning points." Linear interpolation is then conducted between subsequent turning points yielding the final (baseflow) discharge. In a similar fashion to the digital filters, this method extracts a filtered/smoothed hydrograph of total flow minima, and is therefore often also referred to as the "smoothed minima technique."

   Piggott et. al.(2005) discussed how the UKIH technique can yield alternate baseflow estimates depending on the origin of the N-day window. They proposed staggering N-sets of UHIK baseflow estimates to create an overall aggregate baseflow hydrograph. Three versions of this modification are included here:

   (a) Sweeping minimum: returns the daily minimum of the staggered hydrographs;

   (b) Sweeping maximum: returns the daily maximum of the staggered hydrographs; and,

   (c) Sweeping median: returns the median of the N-staggered hydrographs.

2. The HYSEP technique (Sloto and Crouse, 1996). This method depends on the computed days of quick flow termination N (equation 2.3). Like the UHIK method, the HYSEP techniques then proceed to determine minimum discharges within the N-day window. Three methods of producing baseflow estimates are computed in HYSEP and are reproduced here, they include:

(a) Fixed interval: where baseflow is assumed to be the minimum discharge reported within sequential, non-overlapping N-day windows. Like the UKIH method, results from the fixed interval method is dependent on the ("fixed") window origin;

(b) Sliding interval: where baseflow is assumed to be the minimum discharge found within a moving N-day window. In contrast, this method tends to yield a higher BFI; and,

(c) Local minimum: linearly-interpolates total flow minima within a moving N-day window.

3. The PART technique (Rutledge, 1998). This method aims to reproduce the conceptual hydrograph represented in Figure 2.1. Using a combination of quick flow termination estimates (equation 2.3), recession coefficients (equation 2.2), and a parameter termed the "antecedent requirement," a combination of forward and backward filtering techniques are used in producing the final hydrograph separation estimates. Three estimates using the PART method are produced here, based on the suggested antecedent requirement choices offered by Rutledge (1998):

(a) antecedent requirement $= 1$;

(b) antecedent requirement $= 2$; and,

(c) antecedent requirement $= 3$;

4. The Clarifica Inc., (2002) technique. This method separates the total flow hydrograph by performing two sweeps on the hydrograph. The first is a 6-day moving minimum, followed by a 5-day moving average (3-days previous, 1-day ahead). This method was designed for use within southern Ontario watersheds and tends to produce higher estimates of baseflow during peak flow events.

2.2   Hydrograph disaggregation

The hydrograph disaggregation algorithm is a means of categorizing the hydrograph into three primary constituent parts, namely: the rising limb, the falling limb, and the baseflow recession component.

The rising limb identifies the portion of the hydrograph that has been influenced by some forcing, typically a rainfall event, a snowmelt episode, and perhaps, some planned release of a significant volume of water from storage.

The falling limb is subsequent to the rising limb and is likely indicative of the cessation of the forcing, and thus can be though of the excess precipitation that has yet to drain from the basin.

The baseflow recession component should be though of that portion of the hydrograph that represents the slow release of water stored in the watershed. It is not uncommon to assume that this portion of the hydrograph is entirely composed of groundwater discharge, however there may be exceptions. The baseflow recession component is identified by locating portions of the hydrograph that closely follow the computer baseflow recession coefficient (Equation 2.2) within a predefined degree of error.

Hydrograph disaggregation can be useful to hydrologic modellers needing to identify model parameters that directly relate to specific portions of the hydrograph. The disaggregation also enables the ability to isolate streamflow volumes associated with specific events and thus can be used to convert a continuous hydrograph into a discrete form (see below §2.3).

2.3   Hydrograph discretization

A common challenge with the interpretation of hydrology is the comparison of streamflow response to a rainfall/snowmelt event. Complications arise since streamflow response can have a greater dependence on antecedent conditions over the amount of precipitation. In addition, snowmelt events can add further complication as they are rarely measured, yet has a significant impact on streamflow in Canadian watersheds. By assuming that initial flow prior to an event is a surrogate for antecedent state, one could isolate the portion of the hydrograph caused by that particular event. This method is an extension of the work of Reed et. al., (1975) who used the method to derive parameters for their rainfall-runoff model (Figure 2.2). Arnold and Allen (1999) also refer this to the "recession curve displacement method."

The algorithm locates the onset of a rising limb and projects streamflow recession from this point as if the event had never occurred (e.g., at points T' and T in Figure 2.2). This projected streamflow, termed "underlying flow" by Reed et. al.(1975), is subtracted from the total observed flow to determine the volume associated with the event. Comparing these event volumes with rainfall/snowmelt volumes can be used to calculate runoff coefficients but will also highlight non-linearities of watershed hydrology (see Beven et. al., 2011).

In particular, events may be observed in the streamflow hydrograph that may have no corresponding rainfall event on record (or vice versa). This will be problematic during rainfall-runoff model calibration as this particular event will do nothing but hinder the fitness of the model as there will be no information (i.e., data) to drive the model. Identifying and eliminating these "red herrings" (Bevin and Westerberg, 2011) from a calibration dataset is a crucial step in preventing model overfitting.
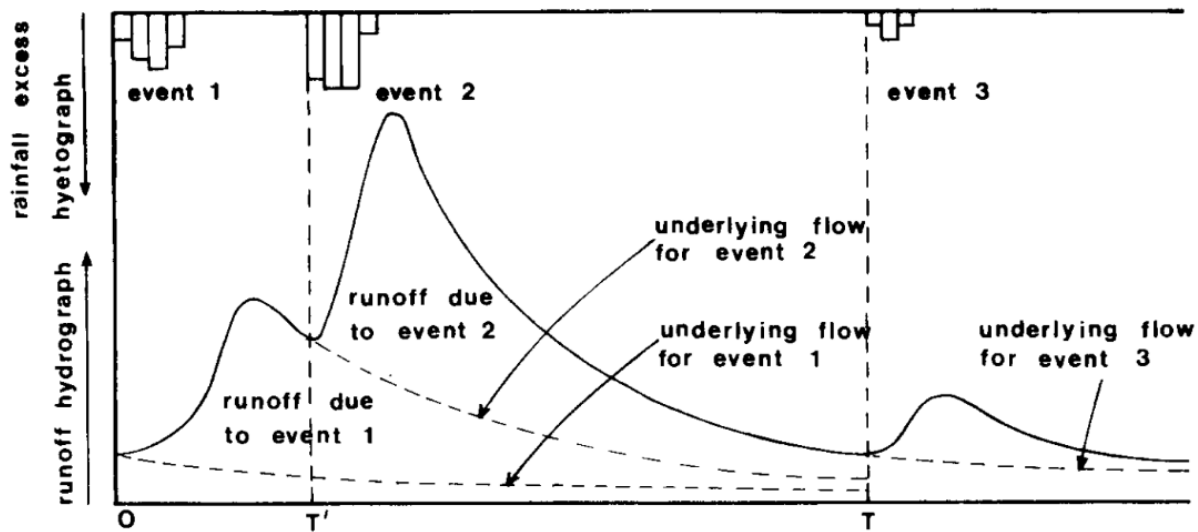
Figure 2.2: Isolation of hydrograph influence after a series of events (After Reed et. al., 1975).

References

Arnold, J.G. and P.M. Allen, 1999. Automated Methods for Estimating Baseflow and Groundwater Recharge from Stream Flow Records Journal of American Water Resources Association 35(2): 411–424.

Beven, K.J., P.J. Smith, A. Wood, 2011. On the colour and spin of epistemic error (and what we might do about it). Hydrology and Earth System Sciences 15: 3123-–3133.

Bevin, K.J. and I. Westerberg, 2011. On red herrings and real herrings: disinformation and information in hydrological inference. Hydrological Processes 25: 1676–1680.

Beven, K.J., 2012. Rainfall-runoff modelling: the primer, 2nd ed. John Wiley & Sons, Ltd. 457pp.

Boughton, W.C., 1993. A hydrograph-based model for estimating the water yield of ungauged catchments. Hydrology and Water Resources Symposium, Institution of Engineers Australia, Newcastle: 317–324.

Chapman, T.G. and A.I. Maxwell, 1996. Baseflow separation—comparison of numerical methods with tracer experiments.Institute Engineers Australia National Conference. Publ. 96/05: 539–545.

Chapman, T.G., 1991. Comment on the evaluation of automated techniques for base flow and recession analyses, by R.J. Nathan and T.A. McMahon. Water Resource Research, 27(7): 1783–1784

Chapman T.G., 1999. A comparison of algorithms for stream flow recession and baseflow separation. Hydrological Processes 13: 710–714.

Clarifica Inc., 2002. Water Budget in Urbanizing Watersheds: Duffins Creek Watershed. Report prepared for the Toronto and Region Conservation Authority.

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes 19: 507–515.

Institute of Hydrology, 1980. Low Flow Studies report. Wallingford, UK.

Jakeman, A.J. and Hornberger G.M., 1993. How much complexity is warranted in a rainfall-runoff model? Water Resources Research 29: 2637–2649.

Linsley, R.K., M.A. Kohler, J.L.H. Paulhus, 1975. Hydrology for Engineers 2nd ed. McGraw-Hill. 482pp.

Linsley, R.K., J.B. Franzini, 1964. Water-resources engineering. McGraw-Hill, New York. 654pp.

Lyne, V. and M. Hollick, 1979. Stochastic time-variable rainfall-runoff modelling. Hydrology and Water Resources Symposium, Institution of Engineers Australia, Perth: 89–92.

Nathan, R.J. and T.A. McMahon 1990. Evaluation of Automated Techniques for Baseflow and Recession Analyses. Water Resource Research 26(7): 1456–1473.

Piggott, A.R., S. Moin, C. Southam, 2005. A revised approach to the UKIH method for the calculation of baseflow. Hydrological Sciences Journal 50(5): 911–920.

Reed, D.W., P. Johnson, J.M. Firth, 1975. A Non-Linear Rainfall-Runoff Model, Providing for Variable Lag Time. Journal of Hydrology 25: 295–305.

Rutledge, A.T., 1998. Computer Programs for Describing the Recession of Ground-Water Discharge and for Estimating Mean Ground-Water Recharge and Discharge from Streamflow Records—Update, Water-Resources Investigation Report 98–4148.

Sloto, R.A. and M.Y. Crouse, 1996. HYSEP: A Computer Program for Streamflow Hydrograph Separation and Analysis U.S. Geological Survey Water-Resources Investigations Report 96-4040.

Tularam, A.G., Ilahee, M., 2008. Exponential Smoothing Method of Base Flow Separation and its Impact on Continuous Loss Estimates. American Journal of Environmental Sciences 4(2): 136–144.