Heuristic Continuous Base Flow Separation
Jozsef Szilagyi

Abstract: A digital filtering algorithm for continuous base flow separation is compared to physically based simulations of base flow. It is shown that the digital filter gives comparable results to model simulations in terms of the multiyear base flow index when a filter coefficient is used that replicates the watershed-specific time delay of model simulations. This way, the application of the heuristic digital filter for practical continuous base flow separation can be justified when auxiliary hydrometeorological data (such as precipitation and air temperature) typically required for physically based base flow separation techniques are not available or not representative of the watershed. The filter coefficient can then be optimized upon an empirical estimate of the watershed-specific time delay, requiring only the drainage area of the watershed.


CE Database subject headings: Base flow; Digital filters; Runoff; Streamflow; Algorithms.

Introduction

Detailed knowledge of groundwater contribution to streams, i.e., base flow, is important in many water management areas: water supply, wastewater dilution, navigation, hydropower generation (Dingman 1994) and aquifer characterization (Brutsaert and Nipher 1977; Troch et al. 1993; Szilagyi et al. 1998; Brutsaert and Lopez 1998). Also, base flow can directly be related to aquifer recharge (Birtles 1978; Wittenberg and Sivapalan 1999; Szilagyi et al. 2003), which is crucial in ascertaining safe yields of water development schemes, such as irrigation planning in the Great Plains (Sophocleous 2000).

The importance of having knowledge of base flow is reflected in the number of published works, as reviewed by Tallaksen (1995). With the widespread use of PCs, traditional, event-based methods that contain varying degrees of subjectivity, such as graphical base flow separation (Barnes 1939; Hewlett and Hibbert 1963; Szilagyi and Parlangie 1998), have been replaced by automated techniques that can result in continuous base flow modeling. Present-day automated techniques consist mainly of two types: digital filtering methods (Nathan and McMahon 1990; Arnold et al. 1995; Arnold and Allen 1999) and conceptual hydrologic models (e.g., Jakeman et al. 1990; Szilagyi and Parlangie 1999). The former have “no true physical basis” (Arnold and Allen 1999) but have the distinct advantage of requiring only streamflow measurements. The latter are physically based but require precipitation data as a minimum in addition to measured streamflow. Often, available precipitation data are insufficient because the precipitation station is either not located within the watershed, or it is within the watershed but not at a representative location. In larger catchments, more than one station is typically needed to obtain a good estimate on the amount of water available to runoff. Many times the precipitation record has discontinuities that can easily thwart efforts to perform continuous base flow separation using physically based techniques. Clearly, there is a practical need for a technique that uses the most basic information available: streamflow and the corresponding drainage area. The digital filtering technique of Nathan and McMahon (1990) is such a “minimalist” approach. Because their method is not based on any physical law, a question arises whether the ensuring base flow hydrograph is realistic at all, or, in other words, can the results be backed by a more complex, physically based approach? Unfortunately, there is no trivial way of validating the results of the filter algorithm by measurements. Isotope or chemical tracer techniques may one day prove useful in validation efforts in spite of the currently existing discrepancy in base flow interpretation between physical and tracer techniques (Rice and Hornberger 1998).

Baseflow recession can generally be described by the following equation (Brutsaert and Nipher 1977):

\[
\frac{dQ}{dt} = -a Q^b
\]

where \( a \) \([L^{3(1-\beta)} T^{-\beta-2}]\) and \( \beta \) (\( \beta >0 \)) are constants; \( Q_b \) \([L^3 T^{-1}]\) is the groundwater discharge to the stream. Under simplifying assumptions (Brutsaert and Lopez 1998), the theoretical value of \( \beta \) during recession may change from three to unity, when \( \beta \) becomes unity, the aquifer behaves as a linear reservoir, and \( a \) then equals \( k^{-1} \), the inverse of the storage coefficient \([T]\) in the linear storage equation \( S=k Q_b \), where \( S \) \([L^3] \) is water volume in storage. Naturally, not all aquifers behave as linear reservoirs, even after a sufficient period of streamflow recession (Brutsaert and Nipher 1977; Szilagyi and Parlangie 1998; Troch et al. 1993; Wittenberg and Sivapalan 1999), but many do, as reported by Vogel and Kroll (1992); Jakeman and Hornberger (1993), and Brutsaert and Lopez (1998). The analysis that follows is strictly valid for watersheds that exhibit this latter type of base flow recession property, although the results and conclusions can straightforwardly be generalized to a fully nonlinear aquifer case as well, where \( \beta \) is always larger than unity.

Jakeman and Hornberger (1993) pointed out that the information content of a rainfall–runoff model allows for only a handful of model parameters to be optimized. Perrin et al. (2001), in a
study involving 429 catchments worldwide, demonstrated that "very simple models can achieve a level of performance almost as high as models with more parameters." In fact, "inadequate complexity typically results in model over-parameterization and parameter uncertainty" (Perrin et al. 2001). In the light of these findings, the simplest possible physically based model for base flow simulation was sought. The model of Jakeman et al. (1990) and Jakeman and Hornberger (1993), from now on referred to as the Jakeman model, meets this criterion.

Methodology

Following Jakeman et al. (1990) and Jakeman and Hornberger (1993), any nonlinearity in the rainfall-runoff relationship can be dealt with by the transformation of the observed precipitation series into "excess" or "effective" rainfall $u$ [LT$^{-1}$] via an antecedent precipitation index $s(-)$

$$s_i = c(r_i + (1 - \tau^{-1})r_{i-1} + (1 - \tau^{-1})^2 r_{i-2} + \ldots)$$  

(2)

where $r$ [LT$^{-1}$] = observed rainfall; $\tau$ = the rate at which the catchment wetness declines in the absence of precipitation; $i$ = time index (incremented on a daily basis); and $c$ [TL$^{-1}$] = a normalizing parameter that ensures that the excess rainfall volume equals the volume of total runoff over the calibration period. Excess rainfall is obtained by

$$u_i = r_i s_i$$  

(3)

Seasonal changes in evapotranspiration are described by

$$\tau_f = \tau_0 e^{(30-t)}$$  

(4)

where $f$ [r$^{-1}$] = a temperature modulation factor; $t$ = temperature ($^\circ$C); and $\tau_0$ = the rate at which the catchment wetness declines at 30°C.

Effective rainfall is routed through two parallel linear reservoirs representing quick and slow (i.e., base flow) storm responses. The unit impulse response $[h(-)]$ of a linear reservoir in discrete time $i$ is (O’Connor 1976)

$$h_i = \frac{1}{1+k} \left( \frac{k}{1+k} \right)^i \quad i = 0, 1, 2, \ldots$$  

(5)

from which the impulse response of the two parallel discrete linear reservoirs follows as

$$h_{iq} + h_{ib} = \frac{v_q}{1+k_q} \left( \frac{k_q}{1+k_q} \right)^i + \frac{v_b}{1+k_b} \left( \frac{k_b}{1+k_b} \right)^i \quad i = 0, 1, 2, \ldots$$  

(6)

where the subscripts $q$ and $b$ represent quick and base flow storm responses, respectively. Note that in discrete time the storage coefficients $k_q$ and $k_b$ become unitless. The volumetric throughput coefficients $v_q$ and $v_b(-)$, add up to unity. The model response $Q_m$ [LT$^{-1}$] to effective rainfall is obtained via the convolution summation

$$Q_m = \sum_{i=0}^{m} h_i u_{m-i}, \quad m = 0, 1, 2, \ldots$$  

(7)
Altogether, the model has seven parameters \( f, \tau_0, k_q, k_b, v_q, v_b, \) and \( c \). A schematic of the system is shown in Fig. 1. As a demonstration, the model response to fictive precipitation is shown with arbitrarily assigned parameter values in Fig. 2. The base flow peaks occur almost simultaneously with the total runoff peaks. This is because effective rainfall is split into two parts and routed directly through the two linear reservoirs representing quick and base flow responses, without any time delay in the latter case. In reality, there generally is a time lag between the two peaks (Pilgrim and Cordery 1993; Szilagyi and Parlange 1998), depending on how long it takes the infiltrated water to reach the saturation zone.

The present modification of the original Jakeman model can account for this possible time lag by incorporating a third linear reservoir (with a storage coefficient \( k_s \)) representing soil storage (Besbes and de Marsily 1984; Wu et al. 1997; Wittenberg and Sivapalan 1999). A schematic of the model arrangement can be seen in Fig. 3. The unit impulse response of two serial discrete linear reservoirs is obtained via the Z-transform of the difference equation (Singh 1988)

\[
(1 + k_s \nabla)(1 + k_b \nabla)Q_i = u_i
\]  

(8)

where the difference operator \( \nabla \) is for time shifting, i.e., \( \nabla g_i = g_i - g_{i-1} \), where \( g \) is an arbitrary discrete function. Upon inverting the resulting transfer function \( H(z) \)

\[
H(z) = \frac{z^2}{(1 + k_s + k_b + k_s k_b)z^2 - (k_s + k_b + 2k_s k_b)z + k_s k_b}
\]  

(9)

the discrete unit impulse response results as

\[
h_i = \frac{-k_b \left( \frac{k_b}{1 + k_b} \right)^i k_s + k_s \left( \frac{k_s}{1 + k_s} \right)^i + k_s \left( \frac{k_s}{1 + k_s} \right)^i k_b - k_b \left( \frac{k_b}{1 + k_b} \right)^i}{k_s + k_b^2 + k_b k_s^2 - k_b - k_b^2 - k_b^2 k_s^2}
\]  

(10)

By adding a soil-storage component to the Jakeman model, the number of parameters has increased through \( k_s \), by one, from seven to eight. The soil-storage component delays the base flow peak as well as flattens it, thus making it look more realistic, as is seen in Fig. 4, where a \( k_s = 2 \) (day) was added to the previously prescribed model parameter set.

In the last modification of the model, the changing effect of the exponent in Eq. (1) is being investigated. Right after the start of the base flow recession, the exponent may reach a value of three, provided the aquifer became close to full saturation. Fig. 5 (from Szilagyi 1999) demonstrates this case, with the lower envelopes (that are thought to represent "pure" groundwater discharge) of the data points expressing a slope of three and unity. Numerical and analytical solutions of the Boussinesq equation that describe groundwater drainage also confirm (Brutsaert and Nieber 1977; Szilagyi 1999) this change of the exponent in Eq. (1). A time-varying exponent in Eq. (1) can only be modeled via a general nonlinear reservoir, \( S = k Q^n_b \), if \( n \) changes with time as well. Alternatively, rather than changing \( n \) through time, \( k \) may be changed with time in the linear reservoir representation, as was done by Aksoy et al. (2001). The critical base flow discharge

Fig. 3. Schematic representation of the modified Jakeman model

Fig. 4. Response of the modified Jakeman model to fictive precipitation with arbitrary parameters: \( k_q = 1 \) (day), \( k_s = 2 \) (day), \( k_b = 30 \) (day), \( f = 1 \) (°C⁻¹), \( \tau_0 = 1 \), \( v_q = 0.5 \). Solid line is modeled base flow; intermittent line is modeled total runoff.
\( Q_{b0} \), when this change starts (Fig. 5), is obtained by solving Eq. (1) simultaneously for the two lower envelope lines as \( Q_{b0} = (a_1/a_3)^{0.5} \), where \( a_1 \) and \( a_3 \) are with \( \beta = 1 \) and \( \beta = 3 \), respectively. For convenience, it is assumed here that \( k \) changes linearly from a maximum value of \( k_{b0} (= a_1^{-1}) \), when \( Q_b = Q_{b0} \), to a minimum value of 0.5 \( k_b \) when the aquifer becomes close to saturation. Under simplifying assumptions (Brutsaert and Lopez 1998), drainable water storage at full saturation, \( S_{\text{max}} \), can be estimated as \( S_{\text{max}} \approx 1.97 A^{1/2} \), where \( A \) = drainage area of the watershed. Since \( k \) is changing with time now, a simple convolution cannot be maintained; instead, base flow is simulated (Fig. 6) by numerically solving the linear storage equation with a time-varying storage coefficient. This means that through the calculation of \( S \) at each time step, the corresponding \( k(S) \) value is

Fig. 5. Measured daily discharge versus change in discharge between consecutive days, 6 days after rain.

Fig. 6. Response of the modified Jakeman model with time-varying storage coefficient to fictive precipitation with arbitrary parameters: \( k_q = 1 \) (day), \( k_s = 2 \) (day), \( k_b = 30 \) (day), \( f = 1 \) (°C \(^{-1}\)), \( \tau_0 = 1 \), \( v_q = 0.5 \). Solid line is modeled base flow; intermittent line is modeled total runoff. \( Q_0 \) and \( S_{\text{max}} \) are assumed to be 0.05 (mm·d \(^{-1}\)) and 10 (mm), respectively.
obtained with the help of the maximum and minimum values of \( k \), as \( k = c_1 S + c_2 \), where \( c_1 = k_b/2(S_b - S_{\text{max}}) \), and \( c_2 = k_b - S_b c_1 \). \( S_b \) is the drainable water storage at \( Q_b = Q_{b0} \).

This last modification of the Jakeman model \( \sim MJ \) will be used for the validation of the digital filter algorithm (Nathan and McMahon 1990), which estimates base flow \( (Q_b) \) as

\[
Q_{bi} = pQ_{b(i-1)} + \frac{1-p}{2}(Q_i + Q_{i-1})
\]

from measured or modeled streamflow \( (Q) \), where \( p \) is the filter parameter. The resulting base flow values are constrained by the concurrent streamflow values, so that whenever \( Q_{bi} > Q_i \), the \( Q_{bi} \) value is replaced by \( Q_i \). The validation is done by running the MJ model with Monte Carlo-simulated daily precipitation values in combination with deterministic daily temperature values, following Milly (1994) and Szilagyi (2001). The daily values of precipitation \( (P_d \ [\text{L}]) \) are assumed to follow an exponential distribution \( (\varphi) \)

\[
\varphi(P_d) = \lambda e^{-\lambda P_d}
\]

where \( \lambda^{-1} = [P_a/(365.25 \times SF)] \), with \( P_a \ [\text{L}] \) denoting the mean annual precipitation, and \( SF \ [\text{day}^{-1}] \) the mean storm frequency. \( SF \) is calculated as \( 2(P_d^2)/\text{var}(P_d) \), where the angular brackets denote temporal averaging, and \( \text{var} \) denotes the variance. The number of interstorm days \( (i_d) \) is assumed to follow a Poisson distribution.

---

**Fig. 7.** (a) First year of the simulated streamflow (intermittent line) and base flow values; (b) base flow hydrographs of the same period. Intermittent line is the filter result. Here \( k_q = 1 \ (\text{day}), \ k_s = 1 \ (\text{day}), \ k_b = 60 \ (\text{day}), \ v_q = 0.2 \).

**Fig. 8.** (a) First year of the simulated streamflow (intermittent line) and base flow values; (b) base flow hydrographs of the same period. Intermittent line is the filter result. Here \( k_q = 1 \ (\text{day}), \ k_s = 2 \ (\text{day}), \ k_b = 30 \ (\text{day}), \ v_q = 0.8 \).
\[ P(i_d = N) = \frac{\alpha}{N!} e^{-\alpha} \]  
\[ \text{where } \alpha = SF^{-1}. \]

**Results and Discussion**

Rather than fitting the MJ model to measured streamflow and comparing the filter results to the MJ-model-obtained base flow, a Monte Carlo-type simulation with the MJ model was preferred due to the much greater flexibility the latter approach offers. By making sure that the model-prescribed parameters are physically meaningful and driving the model with realistic precipitation and temperature inputs, realistic model simulations of base flow can be expected and compared to filter results. The MJ model, even in its original, simplest form, performed quite effectively in simulating daily streamflow of small catchments in the U.S., Europe, Asia, and in Australia (Jakeman et al. 1990; Jakeman and Hornberger 1993).

The modified MJ model was run in a Monte Carlo simulation mode with daily precipitation and daily mean temperature inputs, characteristic of a mild continental climate of central Europe, with a mean annual precipitation of 600 mm evenly distributed (i.e., no seasonal cycle) throughout the year, a mean annual temperature of 11°C, and a mean storm frequency of 0.2365/day. A choice of \[ \tau_0 = 1 \text{ and } f = 1°C^{-1}, \] in combination with a 5th order polynomial in Eq. (2), resulted in a 7% runoff ratio, which is typical of the lowland regions in central Europe. The daily mean temperatures (°C) followed the mean monthly temperatures in the model starting with January: -1.1, 1.5, 8.8, 11.8, 16.8, 20.2, 22.2, 21.4, 17.4, 11.3, 5.8, and 1.5. Each model simulation represented 10 years. The quick storm response parameter \( k_q \) and the soil storage coefficient \( k_s \) were each assigned two values: 1 and 2 day. The base flow storage coefficient \( k_b \) was allowed to have values of 30 and 60 day. The volumetric throughput parameter \( v_q \) was assigned the following values: 0.2, 0.4, 0.6, and 0.8. Note that \( v_b = 1 - v_q \), correspondingly. The values of the above parameters are representative of the catchments reported by Jakeman and Hornberger (1993). With decreasing \( v_q \) values, groundwater contribution to the streamflow increases, requiring increased subsurface storage capability in the watershed. This is accommodated for in the model by increasing the value of \( S_{\max} \) and \( Q_{b0} \) in the model accordingly, such as \( (5, 0.03), (10, 0.06), (15, 0.09), \text{ and } (20, 0.12) \), where the first value in each parenthesis is \( S_{\max} \) (mm), the second one is \( Q_{b0} \) (mm/day), and the first parenthesis corresponds to \( v_q = 0.8 \). The \( S_{\max} \) and \( Q_{b0} \) values are representative of the small catchments of the Washita Experimental Watershed complex in Oklahoma (Brutsaert and Lopez 1998).

The three storage coefficients and the \( v_q \) values amount to 32 different and unique combinations. With each combination of the model parameters, the MJ model was run for 10 years in daily time increments. From the resulting base flow hydrograph, the

**Table 1. Model Simulation and Optimized Filtering Results (\( N_d, p, BFI_{filt}/BFI \)).**

<table>
<thead>
<tr>
<th>( k_q ) (d)</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k_s ) (d)</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>( k_b ) (d)</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>( v_q = 0.2 )</td>
<td>3.34</td>
<td>3.15</td>
</tr>
<tr>
<td>( = 0.4 )</td>
<td>3.39</td>
<td>3.26</td>
</tr>
<tr>
<td>( = 0.6 )</td>
<td>3.67</td>
<td>3.51</td>
</tr>
<tr>
<td>( = 0.8 )</td>
<td>4.20</td>
<td>4.02</td>
</tr>
</tbody>
</table>

Fig. 9. (a) First year of the simulated streamflow (intermittent line) and base flow values; (b) base flow hydrographs of the same period. Intermittent line is the filter result. Here \( k_q = 2 \) (day), \( k_s = 2 \) (day), \( k_b = 30 \) (day), \( v_q = 0.8 \).
mean watershed-specific time delay (Linsley et al. 1958) \( N_d(d) \) could be calculated. \( N_d \) is the mean elapsed time between the peak of streamflow and the first instant when streamflow becomes dominated by base flow. The critical point when this latter happens was calculated as \( Q_{bi}/Q_1 \approx 1 - 0.25v_q \), which results in a \( Q_{bi}/Q_1 \) ratio of 80% when \( v_q = 0.8 \) and 95% when \( v_q = 0.2 \). A more stringent critical value is necessary when base flow dominates streamflow. Note that when \( v_q = 0.2, 80\% \) of the streamflow is made up by base flow on a long-term basis, which means that the base flow index, BFI (= \langle Q_{bi} \rangle / \langle Q'_1 \rangle , where the angle brackets denote temporal averaging) is 0.8 or 80\%, as well. Note also that the use of such a critical value is not necessary with the filter algorithm because of the constraint applied there, which makes streamflow become base flow fully “overnight.” In the MJ model, this can never happen due to the exponential decay in the quick flow component.

With the known \( N_d \) value from the MJ model, the filter parameter \( p \) was systematically changed until the filter model gave the closest possible matching value of \( N_d \) with the MJ model, which was generally within 1\%. Figs. 7, 8, and 9 display hydrographs for small (=3.15\( d \)), medium (=4.16\( d \)), and large \( N_d \) (=6.13\( d \)) cases, respectively. The resulting \( p \), \( N_d \), and BFI at/BFI values are listed in Table 1. As it can be seen, the MJ-simulated watershed-specific time delays ranged between 3.15 and 6.21 \( d \), the filter parameter value \( p \) ranged from 0.953 to 0.999, and the BFI at/BFI ratios changed between 0.48 and 1.22. Fig. 10 displays the distribution of the values.

Fig. 10 shows that the long-term base flow index, given by the filter algorithm, is within 20\% of the modeled BFI value in 80\% of the cases considered, with a mean value of only 6\% less than the modeled mean BFI value. This suggests that the filter algorithm of Nathan and McMahon (1990) is of practical value, provided one can estimate the watershed-specific time delay \( N_d \) for real watersheds. Fortunately, this is possible by the application of Linsley’s empirical equation (Linsley et al. 1958) \( N_d = A^{0.2} \), where \( N_d \) is in days and \( A \), the drainage area of the watershed, is in square miles. When applying the filter algorithm, the filter parameter must be adjusted until the resulting \( N_d \) value becomes sufficiently close to Linsley’s value. This has been done by Szilagyi et al. (2003) for 100-plus gauging stations in Nebraska where the spatial distribution of the long-term BFI index was of interest.

In conclusion, it can be stated that the filter algorithm, in spite of its lack of any physical basis, can have its place in practical applications when more complex and/or physically based base flow separation methods are hindered by data availability. The filter algorithm, with its suggested optimization, based on the watershed-specific time delay, requires only the most basic data: streamflow and the corresponding drainage area. Of course, at best, the practical value of the filter algorithm is only as good as the empirical equation of Linsley et al. (1958), which has been frequently used in a wide variety of applications in the past 4 decades. As illustrated previously with the help of model simulations, it gave comparable results to a more complex, physically based base flow separation technique under a variety of soil and aquifer properties characteristic of small watersheds in Oklahoma and North Carolina.

**Acknowledgment**

The writer is grateful to Charles Flowerday for his editorial help.

The views, conclusions, and opinions expressed in this paper are solely those of the writer and not the University of Nebraska, state of Nebraska or any political subdivision thereof.

**References**


