Water Resources Management **10:** 303–320, 1996.

Estimation of Baseflow Recession Constants

RICHARD M. VOGEL
Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155, U.S.A., e-mail: rvogel@tufts.edu.

and

CHARLES N. KROLL
Research Associate, Department of Civil and Environmental Engineering, Cornell University, Ithaca, N.Y. 14853, U.S.A.

(Received: 17 February 1995; in final form: 30 August 1995)

Abstract. Hydrograph recession constants are required in rainfall-runoff models, baseflow augmentation studies, geohydrologic investigations and in regional low-flow studies. The recession portion of a streamflow hydrograph is shown to be either an autoregressive process or an integrated moving average process, depending upon the structure of the assumed model errors. Six different estimators of the baseflow recession constant are derived and tested using thousands of hydrograph recessions available at twenty-three sites in Massachusetts, U.S. When hydrograph recessions are treated as an autoregressive process, unconditional least squares or maximum likelihood estimators of the baseflow recession constant are shown to exhibit significant downward bias due to the short lengths of hydrograph recessions. The precision of estimates of hydrograph recession constants is shown to depend heavily upon assumptions regarding the structure of the model errors. In general, regression procedures for estimating hydrograph recession parameters are generally preferred to the time-series alternatives. An evaluation of the physical significance of estimates of the baseflow recession constant is provided by comparing regional regression models which relate low-flow statistics to three independent variables: drainage area, basin slope and the baseflow recession constant. As anticipated, approximately unbiased estimators of the baseflow recession constant provide significant information regarding the geohydrologic response of watersheds.

Key words: baseflow recession constant, hydrograph, low flows, recession analysis, hydrogeology, streamflow.

1. Introduction

Tallaksen (1995) reviews the application of baseflow recession constants for forecasting low flows, hydrograph analysis, low-flow frequency analysis, and for describing aquifer characteristics. Estimates of hydrograph recession constants are required for the calibration of rainfall-runoff models and in some cases for fitting stochastic streamflow models (Kelman, 1980). Hydrograph separation procedures and associated baseflow recession constants are used routinely for modeling surface runoff (Bates and Davies, 1988) and for constructing unit hydrographs by separating the baseflow component of streamflow from the total streamflow to obtain direct runoff. With increasing attention focused on the quality and quantity of groundwater, understanding the contribution of groundwater to streamflow (baseflow) is often the focus of applied groundwater studies. Ponce and Lindquist (1990) review management strategies for eight general types of baseflow augmentation applications, all of which require characterization of baseflow recessions.

Riggs (1961), Bingham (1986), Vogel and Kroll (1992), Demuth and Hagemann (1994), this study, and others show that regional models which relate low-flow statistics to basin characteristics, can be significantly improved by using the baseflow recession constant as one of the independent basin parameters. Vogel and Kroll (1992) show that the baseflow recession constant is related to both the basin hydraulic conductivity and drainable soil porosity. Regional models for estimating low-flow statistics are used routinely at ungaged sites for the purposes of both water quality and water quantity management (see Vogel and Kroll, 1990, 1992, for discussions).

Since the introduction of graphical hydrograph separation procedures by Barnes (1939), a variety of approaches have been developed for separating baseflow from the total streamflow hydrograph. Barnes (1939) graphical hydrograph separation procedures are still included in most introductory hydrology textbooks in spite of the serious criticisms voiced by Kulandaiswamy and Seetharaman (1969) and Anderson and Burt (1980). Knisel (1963), Singh and Stall (1971) and Brutsaert and Nieber (1977) introduced alternative graphical procedures for hydrograph separation which contain fewer subjective judgements then the graphical procedures introduced by Barnes. A variety of analytic procedures have also been advanced to provide a more objective approach to hydrograph separation (James and Thompson (1970), Jones and McGilchrist (1978), Birtles (1978) and others). A review of studies relating to baseflow recession was performed by Hall (1968) and Tallaksen (1995).

Recession constants must be estimated from an individual or an ensemble of hydrographs; but in either case, derived estimates are random variables whose physical significance should only be determined after an evaluation of their statistical significance. To our knowledge, there are no studies which evaluate the statistical properties of estimated values of baseflow recession constants. One could perform bootstrap experiments with actual streamflow data, however still, one never knows the true values of the baseflow recession constant, hence such experiments cannot be definitive. Alternatively, one could perform Monte-Carlo experiments, making certain assumptions regarding the true underlying structure of hydrograph recessions. Again, such experiments are not completely definitive because one never knows the true structure of hydrograph recessions. Thus, it is very difficult to discern whether any of the available graphical or analytical estimation procedures provide, for example, minimum variance and/or unbiased estimates of the true baseflow recession constant. Our approach is to focus on both the physical and statistical significance of the resulting estimators. After presenting the theory of hydrograph recessions and six plausible estimators of the baseflow recession constant, experiments are performed to evaluate the alternative estimators in terms of their ability to explain the geohydrologic or groundwater outflow response of

23 watersheds in Massachusetts. These experiments are no more or less definitive than Monte-Carlo or Bootstrap experiments would be, yet they do allow us to discriminate among the alternate estimators.

2. Hydrograph Recessions

Barnes (1939) found that hydrograph recessions seem to follow the basic relations

$$B_{t+1} = K_b B_t, \tag{1a}$$

$$O_{t+1} = K_o O_t, \tag{1b}$$

where B_t and O_t are the baseflow and other (or remaining) component of streamflow, respectively, on day t and K_b and K_o are the baseflow and other recession constant, respectively. These two components of the recession streamflow sum to produce the total streamflow, thus

$$Q_t = B_t + O_t, \tag{2}$$

where Q_t is the average daily streamflow on day t. Interestingly, Equation (1a) is an approximate linear solution to the nonlinear differential equation which governs the unsteady flow from a large unconfined aquifer to a stream channel (Boussinesq, 1877; Singh, 1968; Singh and Stall, 1971; Brutsaert and Nieber, 1977; Vogel and Kroll, 1992, *etc.*). Boussinesq (1904) obtained an equation similar to (2) using the principle of superposition of linear solutions. Brutsaert and Nieber (1977), Vogel and Kroll (1992) and others show how (1a) or (1b) can be derived by treating the watershed as a linear reservoir.

Equations (1) and (2) may be expressed in a variety of different forms (Hall, 1968; Tallaksen, 1995), each of which leads to different estimation procedures. In the following sections we derive six different estimators of the baseflow recession constant, K_b, using the model described by (1) and (2). Using thousands of streamflow hydrographs at 23 basins in Massachusetts, we compare both the physical and statistical properties of the four baseflow recession constant estimation procedures.

3. A Hydrograph Recession is a Time Series

In the following sections, Equations (1) and (2) are rearranged to show that hydrograph recessions arise from an autoregressive process.

3.1. BASEFLOW RECESSION AS AN AR(1) PROCESS

Equations (1) and (2) describe a streamflow hydrograph without accounting for the inherent model error introduced by Boussinesq's linear approximation, measurement errors associated with all streamflow measurements or sampling error arising from the fact that all streamflow records are of finite and often small duration. Due to these naturally occurring errors Equation (1a) is more realistically

$$B_{t+1} = K_b B_t + \varepsilon_{t+1} \tag{3}$$

where the ε_t are independent normally distributed errors with zero mean and constant variance. The random errors are a result of both measurement and model errors. Equation (3) provides a reasonable approximation to the baseflow portion of a streamflow hydrograph recession because the other component(s) will have decayed to zero. Equation (3) is a first-order autoregressive process, AR(1), where K_b is the autoregressive parameter. Using the notation introduced by Box and Jenkins (1976) for an AR(1) process, $K_b = \phi_1$. James and Thompson (1970) derive the baseflow recession curve as an ARMA(1,1) process with $K_b = \phi_1 = \theta_1$. Of course James and Thompson did not call their model an ARMA(1,1) process because in 1970, the Box and Jenkins notation was unavailable. An ARMA(1,1) process with $\phi_1 = \theta_1$ contains parameter redundancy (see Box and Jenkins, 1976, pp. 248–250), thus the process reduces to white noise, making their analysis questionable.

3.2. BASEFLOW AS AN INTEGRATED MOVING AVERAGE PROCESS

Rewriting (3), assuming additive errors in log space, one obtains

$$B_t = K_b B_{t-1} e^{\varepsilon_t}. \tag{4}$$

Experiments performed later on indicate that the error structure imposed in (4) is more representative of actual streamflow records. Equation (4) can be rewritten using Box and Jenkins (1976) notation by taking logarithms and rearranging to obtain

$$y_t - y_{t-1} = \ln(K_b) + \varepsilon_t = \nabla y_t, \tag{5}$$

where $y_t = \ln(Bt)$ and $y_{t-1} = \ln(B_{t-1})$, again the residuals are independent, with zero mean and constant variance. Box and Jenkins (1976) term (5) an integrated moving average model IMA(0,1,0) with constant drift parameter $\ln(K_b)$. Box and Jenkins (1976) provide a discussion of the properties of nonstationary IMA models.

3.3. THE HYDROGRAPH RECESSION AS AN AR(2) PROCESS

If the baseflow and other component of a hydrograph are considered, then Equations (1) and (2) may be rewritten as

$$Q_{t+1} = K_b B_t + K_o O_t + \varepsilon_{t+1}, \tag{6}$$

where the residuals ε_t are independent and identically distributed random errors with zero mean and constant variance. Using (1a) and (1b), (6) may be rearranged as follows:

$$Q_{t+1} = K_b(B_t + O_t) - K_bO_t + K_o(O_t + B_t) - K_oB_t + \varepsilon_{t+1}$$
$$= (K_b + K_o)(B_t + O_t) - K_bK_oO_{t-1} - K_oK_bB_{t-1} + \varepsilon_{t+1}$$
$$= (K_b + K_o)Q_t - K_bK_oQ_{t-1} + \varepsilon_{t+1}. \tag{7}$$

Equation (7) may be rewritten as an AR(2) process using the notation introduced by Box and Jenkins (1976)

$$Q_{t+1} = \phi_1 Q_t + \phi_2 Q_{t-1} + \varepsilon_{t+1}. \tag{8}$$

If the model errors are assumed to be additive in log space, similar to the IMA(0,1,0) model, then (8) can be rewritten as

$$Q_{t+1} = (\phi_1 Q_t + \phi_2 Q_{t-1})e^{\varepsilon_{t+1}}. \tag{9}$$

Combining the relationships between ϕ_1, ϕ_2, K_o, and K_b implied by Equations (7) and (8) yields relationships between K_o and K_b and the parameters of the autoregressive process

$$K_b = \frac{1}{2}[\phi_1 + (\phi_1^2 + 4\phi_2)^{\frac{1}{2}}], \tag{10}$$

$$K_o = \frac{1}{2}[\phi_1 - (\phi_1^2 + 4\phi_2)^{\frac{1}{2}}]. \tag{11}$$

The discriminant $({\phi_1}^2 + 4\phi_2)$ must be nonnegative, otherwise K_o and K_b are complex numbers which have no physical meaning. A nonnegative discriminant implies that $(K_b - K_o)^2 \geq 0$, thus the discriminant is zero when $K_o = K_b = \frac{1}{2}\phi_1$. Box and Jenkins (1976) show that for (8) to be a stationary process, the parameters ϕ_1 and ϕ_2 must also satisfy

$$\phi_1 + \phi_2 < 1, \tag{12a}$$

$$\phi_2 - \phi_1 < 1, \tag{12b}$$

$$-1 < \phi_2 < 1. \tag{12c}$$

These stationarity conditions are always satisfied and the discriminant in (10) and (11) is always nonnegative when one considers the more restrictive conditions

$$0 \leq K_o \leq K_b \leq 1 \tag{13}$$

which are normally imposed to obtain physically reasonable recessions.

The use of ARMA models for the generation of the complete streamflow hydrograph is not new, for example Spolia and Chander (1974, 1979) showed that if a watershed is conceptualized as a cascade of n equal or unequal linear reservoirs, then the streamflow process is equivalent to an ARMA(n,n – 2) model. In this

TABLE I. Description of U.S. geological survey sites used in this study

Site No.	USGS gage No.	Record length yrs	Drainage area, A mi^2	Basin Slope S = 2Hd	$Q_{7,2}$ cfs	$Q_{7,10}$ cfs
1	01180500	73	52.70	1.24	1.412	4.38
2	01096000	34	63.69	0.95	4.489	10.16
3	01106000	37	8.01	0.17	0.047	0.18
4	01170100	16	41.39	0.99	4.488	7.54
5	01174000	34	3.39	0.40	0.021	0.11
6	01175670	23	8.68	0.17	0.227	0.50
7	01198000	19	51.00	0.66	3.275	5.42
8	01171800	11	5.46	0.33	0.468	0.90
9	01174900	22	2.85	0.46	0.093	0.18
10	01101000	38	21.30	0.15	0.207	0.80
11	01187400	31	7.35	0.54	0.227	0.50
12	01169000	44	89.00	1.23	7.836	13.85
13	01111300	20	16.02	0.25	0.207	0.84
14	01169900	17	24.09	1.08	3.449	5.44
15	01181000	48	94.00	1.11	5.580	10.81
16	01332000	52	40.90	1.07	5.014	7.83
17	01097300	20	12.31	0.13	0.152	0.58
18	01333000	34	42.60	1.67	4.330	8.26
19	01165500	65	12.10	0.55	0.602	1.23
20	01171500	45	54.00	1.11	6.135	10.08
21	01176000	71	150.00	0.53	15.517	31.34
22	01162500	47	19.30	0.41	0.438	1.40
23	01180000	28	1.73	0.30	0.057	0.11

Note: Basin average slope is estimated using $S = 2Hd$, (dimensionless) where H is basin relief and d is drainage density (see Vogel and Kroll, 1992; for a discussion of this estimator of S).

instance the hydrograph recession is considered to have two components ($n = 2$), which results in an ARMA(2,0) or more simply an AR(2) model. It is comforting to realize that Equation (8) can be derived from two rather different yet realistic physical interpretations of a watershed. For this two component recession, James and Thompson (1970) derived an ARMA(2,2) process with $\phi_1 = \theta_1$ and $\phi_2 = \theta_2$ which again contains parameter redundancy and the derived process reduces to white noise.

3.4. ESTIMATION OF BASEFLOW RECESSION CONSTANTS – AN EXPERIMENT

Previous studies which sought to develop estimation procedures for hydrograph recession constants used a relatively small number of individual hydrographs at each site, hence it is difficult to evaluate either the statistical or physical significance of the derived estimates. In this study, we test our estimators using 23 unregulated basins in Massachusetts with long-term U.S. Geological Survey streamflow records. The U.S. Geological Survey gage numbers are shown in Table I along with site numbers chosen for this and other studies (Vogel and Kroll, 1989, 1990 and 1992), the length of each gaged record, drainage area A, average basin slope S, and estimates of the 7-day 2-year and 7-day 10-year low-flow statistics $Q_{7,2}$, and $Q_{7,10}$ respectively. The simple estimator of basin slope $S = 2Hd$ used here, where H is basin relief and d is drainage density, compares favorably with more complex estimators (Zecharias and Brutsaert, 1985). See Vogel and Kroll (1992) for a more detailed discussion of the basin characteristics reported in Table I. These basin descriptors are employed later on to test the physical significance of the derived estimates of K_b.

3.5. EXPERIMENTAL DESIGN

An automatic hydrograph recession selection algorithm was developed so that the complete database of gaged streamflows described in Table I could be used to estimate values of K_b at each site. Table I describes 23 gaged basins with a total of 845 site-years of average daily streamflow, or a total of 308,425 daily flow values among all 23 sites. An automatic hydrograph recession algorithm is employed to search the daily flow record at each site to define a set of hydrograph recessions. A recession begins when a 3-day moving average begins to decrease and ends when a 3-day moving average begins to increase. Only recessions of 10 or more consecutive days are accepted as candidates. If the total length of a candidate recession is λ then some initial portion of that recession contains predominantly surfaceflow or stormflow. In this study the first 0.3λ days were removed from each hydrograph recession, however that choice is somewhat arbitrary because Vogel and Kroll (1992) found that almost any value of λ in the interval [0,0.8] is a reasonable choice to assure that the linear reservoir hypothesis (AR(1) model) provides an adequate approximation to the low-flow behavior of watersheds in this region.

4. Estimation Methods

4.1. TIME-SERIES ESTIMATORS

Estimation of K_b in (3), (5), (8) or (9) using time-series methods would result in m estimates of K_b at each site, corresponding to the m recessions obtained from the automatic hydrograph recession selection algorithm. If each estimate were unbiased, then the ensemble average of the m individual estimates of K_b would be a reasonably precise estimate of the true value of K_b since many sites have hundreds ($m > 100$) of individual recessions. However, each of the individual m estimates of K_b are based on very short recessions, leading to hundreds of biased estimates.

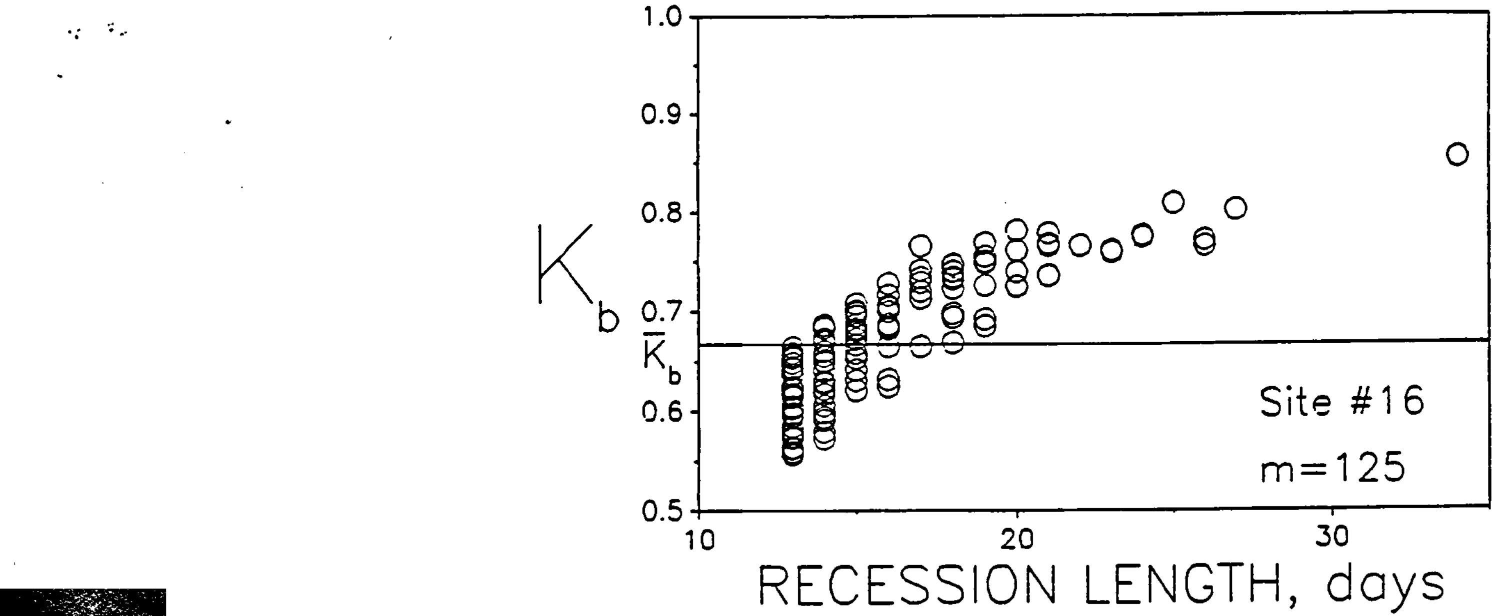

Fig. 1. Estimates of K_b for 125 hydrograph recessions at site 16 using unconditional least squares (ULS) estimators of ϕ_1 and ϕ_2 in Equation (8).

Box and Jenkins (1976, pp. 212–236) recommend the use of unconditional least squares (ULS) estimators for fitting ARMA models. For very large samples, perhaps 500 or more observations, the ULS estimates are approximately equal to the maximum likelihood estimates. While ULS estimates of the parameters of an ARMA model are asymptotically unbiased and have minimum variance, they contain substantial downward bias for the small samples encountered here. Typical hydrograph recessions encountered in this study ranged from about 6–35 days. For example, Figure 1 contains $m = 125$ independent estimates of K_b at site 16 obtained by application of ULS estimators for the AR(2) model parameters ϕ_1 and ϕ_2 in (8). Clearly the value of the estimates depend significantly upon the recession length, which is not the case for the estimators which follow. For example, two estimators of K_b described later on reveal that for this site, K_b is in the range [0.90–0.94]. Clearly each of the $m = 125$ independent ULS estimates of K_b illustrated in Figure 1 are downward biased, and the bias does not disappear completely, even for the longer hydrograph recessions.

For the AR(1) model in (3) a method-of-moments estimate of the parameter K_b is equal to an estimate of the first-order serial correlation coefficient ρ_1, which is also significantly downward biased for the samples encountered in this study (see Wallis and O'Connell, 1972). Furthermore, the unbiased estimators of ρ_1 recommended by Wallis and O'Connell (1972) are significantly downward biased when $\rho_1 > 0.9$ and $n < 30$ as is usually the case in this study. Since the small sample properties of even the most efficient estimators of ρ_1, ϕ_1, and ϕ_2 (and, hence, K_b)

contain substantial downward bias for the short and highly autocorrelated samples encountered here, we elected to treat Equations (3), (4), (8) and (9) as regression equations instead of as time-series.

4.2. REGRESSION ESTIMATORS FOR THE AR(1), IMA(0,1,0) AND AR(2) MODELS – K_{b1}, K_{b2}, K_{b3} AND K_{b4}

If Equations (3), (5), (8) and (9) are treated as regression equations, then one need not estimate K_b for each of the m individual recessions at a site. Instead, all the sequences of flows Q_{t+1}, Q_t and Q_{t-1} chosen by the automatic hydrograph recession selection algorithm may be combined and fit to (3), (5), (8) and (9) to produce a single estimate of K_b for a site. Since such an estimate is based on hundreds of observations, the resulting estimates are sure to be statistically significant as long as the model residuals have constant variance and are approximately normally distributed. For the AR(1) model in (3), an ordinary least squares regression estimate of K_b is

$$K_{b1} = \frac{\sum_{t=1}^{n-1} Q_{t+1}Q_t}{\sum_{t=1}^{n-1} Q_t^2}, \tag{14}$$

where n is the total number of consecutive observations obtained for a site using the automatic hydrograph recession selection algorithm. Equation (14) is related to Knisels (1963) procedure which estimates K_b by plotting consecutive values of Q_{t+1} versus Q_t and taking K_b as the maximum slope.

Treating the IMA(0,1,0) model in (5) as a regression problem one obtains the ordinary least squares estimator

$$K_{b2} = \exp\left[\frac{1}{m}\sum_{i=1}^{m}(y_t - y_{t-1})\right], \tag{15}$$

where again $y_t = \ln(B_t)$ and $y_{t-1} = \ln(B_{t-1})$. Equation (15) provides one of the simplest possible approaches to estimation of the baseflow recession constant.

For the AR(2) model multivariate ordinary least squares regression may be employed to estimate ϕ_1 and ϕ_2 in (8), however it was found that the resulting model residuals were heteroscedastic. We found that Var(ε_t) was proportional to the flows. To account for this heteroscedasticity, Equation (8) was rewritten as

$$\frac{Q_{t+1}}{Q_t} = \phi_1 + \phi_2\frac{Q_{t-1}}{Q_t} + \eta_t, \tag{16}$$

where now Var(η_t) = Var(ε_t)/Q^2 hence the residuals η_t are approximately homoscedastic, and simple ordinary least squares regression procedures may be employed to estimate ϕ_1 and ϕ_2 in (16). Substitution of those estimates of ϕ_1 and ϕ_2 into (8) and (10) leads to an AR(2) model estimate of K_b which we term K_{b3}. This procedure is essentially a weighted least squares approach to estimating the residuals in (8) where the weights are proportional to the streamflows.

When model errors for the AR(2) model are assumed to be additive in log space, as in (9), then an iterative least squares approach is required (see Kroll, 1989) which we term K_{b4} in this study.

In summary, the estimators K_{b1} and K_{b3} are the estimators for the AR(1) and AR(2) models, respectively, when the residuals are assumed to be additive in real space. The estimators K_{b2} and K_{b4} correspond to the IMA(0,1,0) and AR(2) models, when the residuals are additive in log space.

4.3. THE ESTIMATOR K_{b5} OBTAINED BY TREATING THE WATERSHED AS A LINEAR RESERVOIR

A steady-state solution to the difference equation in (1a) is

$$Q_t = Q_o(K_b)^t e^{\varepsilon_t}, \tag{17}$$

where Q_o is the initial baseflow and Q_t is the baseflow after t days, and ε are independent errors with zero mean and constant variance. Brutsaert and Nieber (1977) and Vogel and Kroll (1992) show that (3) is a solution to the continuity equation $dV/dt = I - Q$, when the outflow Q from the watershed is linearly related to the basin storage V, so that $Q = -V \ln(K_b)$ when inflow I to the watershed is zero (under low-flow conditions) resulting in the differential equation

$$dQ/dt = -\ln(K_b)\, Q\, e^{\varepsilon t}. \tag{18}$$

Equation (18) can be applied to observed streamflows by taking logarithms to obtain

$$\ln[-dQ/dt] = \ln[-\ln[K_b]] + \ln[Q] + \varepsilon_t, \tag{19}$$

where the ε_t are normally distributed errors with zero mean and constant variance. Vogel and Kroll (1992) use the numerical approximations $dQ/dt \approx Q_t - Q_{t-1}$ and $Q \approx (Q_t + Q_{t-1})/2$ to derive the least squares estimator

$$K_{b5} = \exp\left\{-\exp\left[\frac{1}{m}\sum_{t=1}^{m}\{\ln[Q_{t-1} - Q_t] - \ln[\tfrac{1}{2}(Q_t + Q_{t+1})]\}\right]\right\}, \tag{20}$$

where m is the total number of pairs of consecutive daily streamflows Q_t and Q_{t-1} at each site. Essentially K_{b5} is an ordinary least squares regression estimator of K_b in (19).

4.4. THE TRADITIONAL BASEFLOW RECESSION CONSTANT ESTIMATOR K_{b5}

Taking natural logarithms (17) becomes

$$\ln(Q_t) = \ln(Q_o) + \ln(K_b)\, t + \varepsilon_t \tag{21}$$

hence one would anticipate the natural logarithms of baseflow to be linearly related to time, with additive errors in log space. Most introductory textbooks on hydrology

TABLE II. Summary of estimators

Estimator	Explanation	Equations	Model error structure
K_{b1}	AR(1) Model	(3) (14)	Additive in real space
K_{b2}	IMA(0,1,0)	(5) (15)	Additive in log space
K_{b3}	AR(2) Model	(8) (16)	Additive in real space
K_{b4}	AR(2) Model	(9)	Additive in log space
K_{b5}	Linear reservoir	(19) (20)	Additive in log space
K_{b6}	Traditional	(21)	Additive in log space

advocate fitting (21) by plotting $\ln(Q_t)$ versus t and taking $\ln(K_b)$ as the minimum slope corresponding to the baseflow portion of the hydrograph. Such graphical procedures are often rather arbitrary, hence Singh and Stall (1971) introduced a more objective approach which selects K_b such that the other component of the recession (Equation (1b)) is also linear on semi-log graph paper.

In this study, the m candidate hydrograph recessions at each site, selected by the automatic hydrograph recession selection algorithm are fit to (21) using simple ordinary least squares regression. This leads to m estimates of K_b for a given site. Since ordinary least squares regression produces unbiased estimates of $\ln(K_b)$ in (21), an estimate of the mean of the m estimates of K_b should be a nearly unbiased estimate of the true value of K_b at a site as long as m is large. The average value of m regression estimates of K_b at a site is denoted by K_{b6}. The six estimators of K_b described in the above section are summarized in Table II.

5. Results

5.1. STATISTICAL COMPARISONS OF ESTIMATORS

Figure 2 contains the m individual estimates of K_b obtained at sites 4, 8, 13 and 15 using ordinary least squares estimates of $\ln(K_b)$ in (21). These four sites are representative of the range of drainage areas listed in Table I. In addition, the average of the m individual estimates which we term K_{b6} are depicted by a solid line at each site. Essentially, each circle in Figure 2 represents a single estimate of K_b obtained using the traditional estimator which is roughly equivalent to passing a straight line through the flow versus time relationship on semi-log graph paper. In contrast with Figure 1, the estimator of K_b appears to be an approximately unbiased estimator of the average value as evidenced by the fact that the estimates appear to form a cone with longer recession lengths giving rise to estimates of K_b which are closer to the mean value. Figure 2 dramatizes the variability of individual estimates of K_b derived from individual hydrograph recessions. Apparently, individual estimates of K_b derived from a single hydrograph are very poor estimates of the true value of K_b even for fairly long hydrograph recessions.

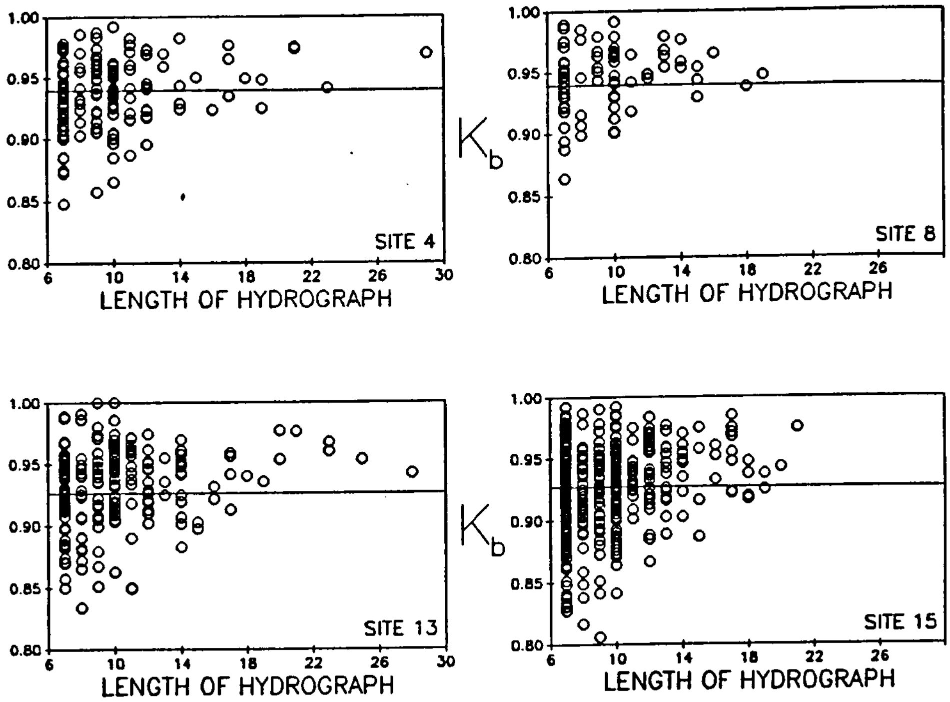

Fig. 2. Estimates of K_b for individual hydrograph recessions at sites 4, 8, 13 and 15 using the traditional approach described by Equation (24).

Table III and Figure 3 compare the six estimators K_{b1}, K_{b2}, K_{b3}, K_{b4}, K_{b5} and K_{b6} summarized in Table II for the 23 U.S. Geological Survey gaged basins summarized in Table I. These six estimators lead, uniformly, to rather different estimates of K_b at each site. It appears that the estimators exhibit some bias relative to one another, however it is difficult to evaluate that bias since the true values are unknown here. All four estimators have relatively low variance because of the large samples used.

The model residuals associated with the estimators K_{b2}, K_{b4}, K_{b5}, and K_{b6} based on (4), (9), (18) and (21), respectively, were homoscedastic and approximately normally distributed. These four estimators assume that the model errors are additive in log space. The model residuals were tested for normality using the probability plot correlation coefficient test (Vogel, 1986) with a 5% significance level. Interestingly, the model residuals associated with the estimators K_{b1} and K_{b3}, based on (3) and (8), respectively, were not homoscedastic and were not normally distributed. Recall that we attempted to correct for this heteroscedasticity associated with the estimator K_{b3} by using a weighted least squares estimator. We conclude that hydrograph recession model errors are additive in log space. Therefore hydrograph recessions are not really autoregressive processes as shown in (3) and (6), instead they more closely resemble an IMA(0,1,0) process as in (5) or the

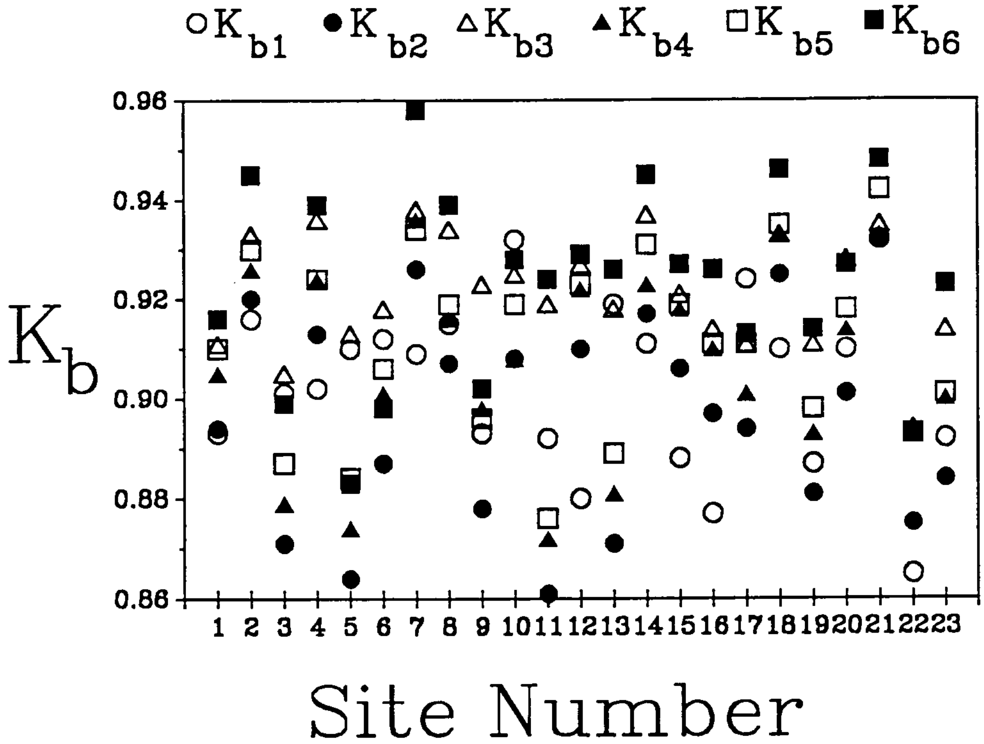

Fig. 3. Comparison of estimates of K_b at 23 unregulated U.S. Geological Survey sites in Massachusetts.

more complex higher order process given in (9) which is no longer a AR(2) process because the error structure is multiplicative rather than additive in real space.

5.2. PHYSICAL SIGNIFICANCE OF ESTIMATES OF K_b

The baseflow recession constant, K_b, is a nondimensional parameter which describes the rate at which streamflow decreases when the stream channel is recharged by groundwater. Kunkle (1962), Knisel (1963), Bingham (1986), Vogel and Kroll (1992), Demuth and Hagemann (1994), and others have shown that estimates of K_b are highly correlated with basin geohydrologic parameters. Riggs (1961), Bingham (1986), Vogel and Kroll (1992), Demuth and Hagemann (1994) and others, also found estimates of K_b for a basin to be highly correlated with low-flow statistics. This result should not be surprising, since low streamflows are groundwater outflow or baseflow and K_b is the only model parameter in the baseflow model described in (1a). Brutsaert and Nieber (1977) and Vogel and Kroll (1992) derive physically-based catchment models which relate groundwater outflow (baseflow) for a basin to the baseflow recession constant.

TABLE III. Comparison of baseflow recession constant estimators

Site No.	K_{b1}	K_{b2}	K_{b3}	K_{b4}	K_{b5}	K_{b6}
1	0.893	0.894	0.911	0.905	0.910	0.916
2	0.916	0.920	0.933	0.926	0.930	0.945
3	0.901	0.871	0.905	0.879	0.887	0.899
4	0.902	0.913	0.936	0.924	0.924	0.939
5	0.910	0.864	0.913	0.874	0.884	0.883
6	0.912	0.887	0.918	0.901	0.906	0.898
7	0.909	0.926	0.938	0.936	0.934	0.958
8	0.915	0.907	0.934	0.916	0.919	0.939
9	0.893	0.878	0.923	0.898	0.896	0.902
10	0.932	0.908	0.925	0.908	0.919	0.928
11	0.892	0.861	0.919	0.872	0.876	0.924
12	0.880	0.910	0.927	0.922	0.923	0.929
13	0.919	0.871	0.918	0.881	0.889	0.926
14	0.911	0.917	0.937	0.923	0.931	0.945
15	0.888	0.906	0.921	0.918	0.919	0.927
16	0.877	0.897	0.914	0.910	0.911	0.926
17	0.924	0.894	0.911	0.901	0.911	0.913
18	0.910	0.925	0.933	0.933	0.935	0.946
19	0.887	0.881	0.911	0.893	0.898	0.914
20	0.910	0.901	0.928	0.914	0.918	0.927
21	0.932	0.932	0.935	0.933	0.942	0.948
22	0.865	0.875	0.894	0.893	0.893	0.893
23	0.892	0.884	0.914	0.900	0.901	0.923

Most prior studies which have sought to develop regional regression relationships between low-flow statistics and drainage basin characteristics have met with limited success. In general, existing regression equations do not yield accurate low-flow predictions for ungaged sites due to the inability of those regional models to explain the geohydrologic response of catchments. Vogel and Kroll (1992) show that considerable improvements in regional regression relationships for predicting low-flow statistics may be obtained by including the baseflow recession constant. They show that the baseflow recession constant acts as a surrogate for basin scale hydraulic conductivity and drainable soil porosity.

In this section we compare the six estimators of K_b summarized in Table II for their ability to improve regional regression models which relate low-flow statistics to basin geohydrology and geomorphology. Following Vogel and Kroll (1992), we use ordinary least squares regression procedures to fit regional models of the form

$$Q_{7,T} = aA^b S^c K_b^d, \tag{22}$$

TABLE IV. Summary of estimated regional regression equations. $Q_{7,T} = aA^b S^c K_b^d$

T	Model	a	b	c	d	Standard Error SE%[e]	R^2 [f]
2	A	0.034(11.1)[g]	1.34 (14.1)			58.0	90.0
	A, S	0.090(6.8)	1.13 (12.2)	0.553(3.8)		44.0	93.9
	A, S, K_{b1}	*	0.949(13.2)	0.892(7.4)	16.0(6.4)	46.0	94.1
	A, S, K_{b2}	*	0.912(19.4)	0.535(4.8)	16.2(9.4)	34.0	96.1
	A, S, K_{b3}	*	1.03 (15.7)	0.487(3.8)	26.1(8.2)	37.9	95.4
	A, S, K_{b4}	*	0.930(19.5)	0.466(4.1)	19.0(9.6)	33.4	96.2
	A, S, K_{b5}	*	0.915(19.8)	0.534(4.9)	19.0(9.6)	33.4	96.2
	A, S, K_{b6}	*	0.885(17.4)	0.536(4.2)	21.0(8.1)	38.3	95.8
10	A	0.011(9.6)	1.43 (9.9)			98.8	81.4
	A, S	0.048(5.5)	1.12 (7.7)	0.819(3.6)		74.0	88.1
	A, S, K_{b1}	*	0.886(7.8)	1.25 (6.6)	19.9(5.1)	78.2	87.9
	A, S, K_{b2}	*	0.858(11.2)	0.774(4.2)	20.8(7.4)	58.2	91.6
	A, S, K_{b3}	*	1.02 (10.3)	0.693(3.5)	34.3(7.2)	59.7	91.3
	A, S, K_{b4}	*	0.888(11.8)	0.671(3.7)	24.9(7.9)	54.9	92.4
	A, S, K_{b5}	*	0.861(11.3)	0.773(4.3)	24.5(7.5)	57.8	91.7
	A, S, K_{b6}	*	0.843(11.9)	0.730(4.1)	28.3(7.9)	55.3	92.6

[e] Computed using SE% = $100[\exp(s_\varepsilon^2)-1]^{\frac{1}{2}}$, where s_ε^2 is an estimate of the variance of the residuals ε for each model.
[f] The values of R^2 are adjusted for the number of degrees of freedom which remain after parameter estimation.
[g] The values in parenthesis are the t-ratios of the estimated model parameters.
* The estimated values of $\ln(a)$ were not significantly different from zero using a 5% level t–test, hence the models were refit using ordinary least squares regression, constraining parameter a to be equal to unity.

where a, b, c, and d are model parameters to be estimated and A, S, and K_b are the drainage basin area, average basin slope, baseflow recession constant, respectively, and $Q_{7,T}$ is the 7-day T-year low-flow statistic. Vogel and Kroll (1992) show that (22) can be derived by extending a conceptual stream-aquifer model to a watershed scale.

The 7-day 10-year low flow $Q_{7,10}$, is the most widely used index of low flow in the United States (Vogel and Kroll, 1990). Estimates of $Q_{7,2}$ and $Q_{7,10}$ at each of the 23 sites described in Table I are obtained by fitting observed series of annual minimum 7-day low-flows to a two-parameter lognormal distribution using the method of maximum likelihood. Vogel and Kroll (1989) found that one could not reject the regional hypothesis that annual minimum 7-day low-flows at the 23 sites described in Table I arise from a two-parameter lognormal distribution assuming a type I error probability of 5%.

Table IV summarizes the adjusted R^2 values, standard errors of estimate, model parameters and their associated t-ratios (in parenthesis) corresponding to fitted regional regression models. Overall, except for the estimator K_{b1}, significant improvements in adjusted R^2, standard error of estimate and the t-ratios of model parameters were obtained by adding estimators of K_b as an explanatory variable. Overall, the estimators K_{b2}, K_{b4}, K_{b5}, and K_{b6} led to the most dramatic reductions in the standard error of prediction (SE%) and the most dramatic improvements in R^2. Recall that these four estimators assume additive model errors in log space. Interestingly, those four estimators also led to regression models with the most stable model parameter estimates represented by the model parameter estimates with the highest t-ratio's. In all cases, the values of R^2 were computed for regressions which contained a constant term $\ln(a)$, even though in most instances that constant term $\ln(a)$ was not significantly different from zero. This is because it is a well known fact that R^2 makes little sense unless it is computed for a regression model with a constant term.

The estimators K_{b1} and K_{b3} were the only estimators which assumed additive model errors in real space. Furthermore, it was found that model errors associated with both of those estimators were heteroscedastic and nonnormal. In the case of K_{b3}, we employed a weighted least squares estimator which removed the heteroscedasticity, hence K_{b3} performed slightly better than K_{b1} in both regression equations reported in Table IV.

For the regression equation for $Q_{7,2}$ the estimators K_{b4} and K_{b5} led to regression equations with the highest R^2, lowest standard error, and highest t-ratios of the model parameters b and d. For the regression equation for $Q_{7,10}$ the estimators K_{b4} and K_{b6} led overall to regression equations with the highest R^2, lowest standard error, and highest t-ratios of the model parameters b and d. Although the estimator K_{b2} did not perform as well as the others, it still performs credibly, especially when compared to K_{b1} and K_{b3}. Overall, our comparisons consistently reveal that the estimators K_{b2}, K_{b4}, K_{b5}, and K_{b6} provide significant improvements in the regional low-flow regression equations in terms of overall model precision (R^2 and SE%) and the precision of model parameter estimates (t-ratio's).

6. Conclusions

Six alternative estimators of the baseflow recession constant K_b, are derived and compared in terms of their ability to produce a physically plausible and approximately unbiased estimate of the average recession constant at a site. Recent research indicates that the baseflow recession constant can act as a surrogate for basin scale hydraulic conductivity and is useful in regional low-flow investigations (Vogel and Kroll, 1992). Physical plausibility of the recession constant estimators is determined from their ability to explain the variability of low-flow statistics for 23 catchments in Massachusetts. Two of the estimators K_{b1} and K_{b3} are based on a representation of a hydrograph recession as an AR(1) and an AR(2) process, respectively, with additive model errors in real space. Experiments document that these estimators yielded almost no information about the baseflow response of a watershed, when compared to the alternative estimators K_{b2}, K_{b4}, K_{b5}, and K_{b6}, all of which assume model errors be additive in log space. Although these four estimators were roughly comparable, the estimator K_{b4} consistently led to regional regression models of low-flow statistics with the highest R^2, lowest standard error SE%, and highest t-ratios of model parameter estimates. The estimator K_{b4} is a least squares regression estimate of an AR(2) model (Equation (9)) with additive residuals in log space. It is the most complex estimator considered here, requiring an iterative algorithm for its implementation. By comparison, the estimators K_{b5} and K_{b6} are much easier to implement and performed almost as well.

All of the estimators described here, except K_{b6}, are regression estimators which lump many individual hydrographs together, using a single estimator. The traditional estimator K_{b6} is the average of a large ensemble of individual regression estimates of K_b from (21) using the traditional approach of representing the baseflow portion of a hydrograph recession as a linear relationship between the natural logarithm of streamflow and time. Although the estimators K_{b4} and K_{b5} are comparable, the traditional estimator K_{b6} may be preferred in some instances because one can easily assess its variance, as was illustrated in Figure 2, which illustrates that a single estimate of K_b based on an individual hydrograph recession using traditional procedures (Equation (21)) is extremely variable, even for long hydrograph recessions. Nevertheless, the average of a large ensemble of individual estimates, K_{b6}, provides a reasonably stable estimate of K_b.

This study, in addition to Demuth and Hagemann (1994), Vogel and Kroll (1992), Bingham (1986) and Riggs (1961) indicate that incorporation of the baseflow recession constant in regional regression models for predicting low flow statistics holds substantial promise. Hopefully future research will lead to improvements in our ability to estimate hydrograph recession constants while concurrently leading to improvements in regional models which are useful for estimating low-flow statistics and groundwater outflow. Baseflow recession constants are also required in rainfall-runoff models, baseflow augmentation studies and geohydrologic investigations, so improvements in our ability to estimate K_b have a number of other applications in addition to their use in regional low-flow studies emphasized in this study.

References

Anderson, M. G. and Burt, T. P.: 1980, Interpretation of recession flow, *J. Hydrol.* **46**, 89–101.

Barnes, B. S.: 1939, The structure of discharge-recession curves, *Trans. Am. Geophys. Union* **20**(4), 721–725.

Bates, B. C. and Davies, P. K.: 1988, Effect of baseflow separation procedures on surface runoff models, *J. Hydrol.* **103**, 309–322.

Bingham, R. H.: 1986, Regionalization of low-flow characteristics of Tennessee streams, U.S. Geological Survey, Water-Resources Investigations Report 85–4191.

Birtles, A. B.: 1978, Identification and separation of major base flow components from a stream hydrograph, *Water Resour. Res.* **14**(5), 791–803.

Boussinesq, J.: 1877, Essai sur la theories des eaux courantes, *Memoires Presentes par Divers Savants a l'Academic des Sciences de l'Institut National de France*, Vol. 23, No. 1.

Boussinesq, J.: 1904, Recherches theoretique sur l'ecoulement des nappes d'eau infiltrees dans le sol et sur le dibit des sources, *J. Math. Pure Appl.* **10**, 5–78, pp. 363–394.

Box, G. E. P. and Jenkins, G. M.: 1976, *Time Series Analysis: Forecasting and Control*, Holden-Day, Oakland, California.

Brutsaert, W. and Nieber, J. L.: 1977, Regionalized drought flow hydrographs from a mature glaciated plateau, *Water Resour. Res.* **13**(3), 637–643.

Demuth, S. and Hagemann, I.: 1994, Estimation of flow parameters applying hydrogeological area information, in: *FRIEND: Flow Regimes from International Experimental and Network Data*, IAHS Publ. No. 221, pp. 151–157.

Hall, F. R.: 1968, Base flow recessions – A review, *Water Resour. Res.* **4**(5), 973–983.

James, L. D. and Thompson, W. O.: 1970, Least squares estimation of constants in a linear recession model, *Water Resour. Res.* **6**(4), 1062–1069.

Jones, P. N. and McGilchrist, C. A.: 1978, Analysis of hydrological recession curves, *J. Hydrol.* **36**, 365–374.

Kelman, J.: 1980, A stochastic model for daily streamflow, *J. Hydrol.* **47**, 235–249.

Knisel, W. G.: 1963, Baseflow recession analysis for comparison of drainage basins and geology, *J. Geophys. Res.* **68**(12), 3649–3653.

Kroll, C. N.: 1989, Estimation and usage of baseflow recession constants, M.S. Thesis, Tufts University, November, pp. 57–63.

Kulandaiswamy, V. C. and Seetharaman, S.: 1969, A note on Barnes' method of hydrograph separation, *J. Hydrol.* **9**, 222–229.

Kunkle, G. R.: 1962, The baseflow-duration curve, a technique for the study of groundwater discharge from a drainage basin, *J. Geophys. Res.* **67**(4), 1543–1554.

Ponce, V. M. and Lindquist, D. S.: 1990, Management of baseflow augmentation: A review, *Water Resour. Bull.* **26**(2), 259–268.

Riggs, H. C.: 1961, Regional low flow frequency analysis, *U.S. Geological Survey Professional Paper 424–B*, B–21–B–23, Washington, D.C.

Singh, K. P.: 1968, Some factors affecting baseflow, *Water Resour. Res.* **4**(5), 985–999.

Singh, K. P. and Stall, J. B.: 1971, Derivation of base flow recession curves and parameters, *Water Resour. Res.* **7**(2), 292–303.

Spolia S. K. and Chander, S.: 1974, Modelling of surface runoff systems by an ARMA model, *J. Hydrol.* **22**, 317–332.

Spolia S. K. and Chander, S.: 1979, Modelling of surface runoff systems by an ARMA model – a correction, *J. Hydrol.* **41**, 187.

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

Vogel, R. M.: 1986, The probability plot correlation coefficient test for the normal, lognormal, and Gumbel distributional hypotheses, *Water Resour. Res.* **22**(4), 587–590, (see correction in **23**(10), (1987), 2013).

Vogel, R. M. and Kroll, C. N.: 1989, Low-flow frequency analysis using probability plot correlation coefficients, *J. Water Resour. Plan. Manage.* **115**(3), 338–357.

Vogel, R. M. and Kroll, C. N.: 1990, Generalized low-flow frequency relationships for ungaged sites in Massachusetts, *Water Resour. Bull.* **26**(2), 241–253.

Vogel, R. M. and Kroll, C. N.: 1992, Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics, *Water Resour. Res.* **28**(9), 2451–2458.

Wallis, J. R. and O'Connell, P. E.: 1972, Small sample estimation of ρ_1, *Water Resour. Res.* **8**(3), 707–712.

Zecharias, Y. B. and Brutsaert, W.: 1985, Ground surface slope as a basin scale parameter, *Water Resour. Res.* **21**(12), 1895–1902.