WATER RESOURCES RESEARCH, VOL. 28, NO. 9, PAGES 2451–2458, SEPTEMBER 1992

Regional Geohydrologic-Geomorphic Relationships for the Estimation of Low-Flow Statistics

RICHARD M. VOGEL

Department of Civil Engineering, Tufts University, Medford, Massachusetts

CHARLES N. KROLL

Department of Civil and Environmental Engineering, Cornell University, Ithaca, New York

Many investigators have sought to develop regional multivariate regression models which relate low-flow statistics to watershed characteristics. Normally, a multiplicative model structure is imposed and multivariate statistical procedures are employed to select suitable watershed characteristics and to estimate model parameters. Since such procedures have met with only limited success, we take a different approach. A simple conceptual stream-aquifer model is extended to a watershed scale and evaluated for its ability to approximate the low-flow behavior of 23 unregulated catchments in Massachusetts. The conceptual watershed model is then adapted to estimate low-flow statistics using multivariate regional regression procedures. Our results indicate that in central western Massachusetts, low-flow statistics are highly correlated with the product of watershed area, average basin slope and base flow recession constant, with the base flow recession constant acting as a surrogate for both basin hydraulic conductivity and drainable soil porosity.

INTRODUCTION

Estimates of low-flow statistics are needed in water quality management and water supply planning and for the determination of minimum downstream release requirements from hydropower, irrigation, water supply, cooling plant and other facilities. Water quality management applications of low-flow statistics include the determination of wasteload allocations, discharge permits, and the siting of treatment plants and sanitary landfills.

Many investigations have attempted to develop regional hydrologic models for the purpose of estimating low-flow statistics at ungaged sites from readily available geomorphic, geologic, climatic and topographic parameters. For example, *Thomas and Cervione* [1970], *Tasker* [1972], *Parker* [1977], *Dingman* [1978], *Male and Ogawa* [1982], *Cervione et al.* [1982], *Downer* [1983], *Fennessey and Vogel* [1990], and *Vogel and Kroll* [1990] have developed regional low-flow models in the New England region.

Usually such models take the form

$$Q_{d,T} = b_0 X_1^{b_1} X_2^{b_2} X_3^{b_3} \cdots \qquad (1)$$

where $Q_{d,T}$ is the d-day, T-year low-flow statistic obtained from gaged flow records, the X_i are measurable drainage basin characteristics and the b_i are parameter estimates obtained from multivariate regression procedures. Such models are generally developed using long-term streamflow data and associated basin characteristics from many sites. Regional statistical models of this type, frequently referred to as "state equations," are used widely in the United States for estimating flood flow statistics at ungaged sites. *Newton and Herrin* [1983] recommend such statistically based regional regression equations over the use of deterministic watershed models for estimating flood flows at ungaged sites. Their recommendations are based upon a large nationwide comparison of alternative methods for estimating flood flows at ungaged sites developed by several federal agencies. Unfortunately, most studies which have attempted the same approach to estimate low-flow statistics at ungaged sites have met with only limited success. *Thomas and Benson* [1970] found that average prediction errors for low-flow regional regression models in the Potomac River basin were at least twice as large as for analogous flood flow regional regression models in the same basin.

Since the application of multivariate regression procedures to select and fit models of the form given in (1) has met with limited success for estimating low-flow statistics at ungaged sites, an alternate approach is proposed. Instead, we hypothesize that the development of a physically based model which links the low-flow response of a basin to appropriate geohydrologic and geomorphic parameters will lead to improvements in the assumed structure and variables employed in regression models to be used for estimating low-flow statistics. Such physically based models need not take the form of (1), and they may contain basin parameters which are difficult to estimate in practice; nevertheless, the development of such a model may improve our understanding of the low-flow process sufficiently so that improved "state equations" may be developed. As *Wallis* [1965] so clearly showed, it is difficult to uncover basic physical relationships using multivariate statistical procedures, without prior knowledge of the physical relationships.

MODEL DEVELOPMENT

A Stream-Aquifer Model of Low Flow

This section develops a simple stream-aquifer low-flow model. The next section extends the model to an entire watershed and finally to a region or system of watersheds. Similar to the work of *Brutsaert and Nieber* [1977], the stream-aquifer system is conceptualized as outflow into a

Paper number 92WR01007.
0043-1397/92/92WR-01007$05.00

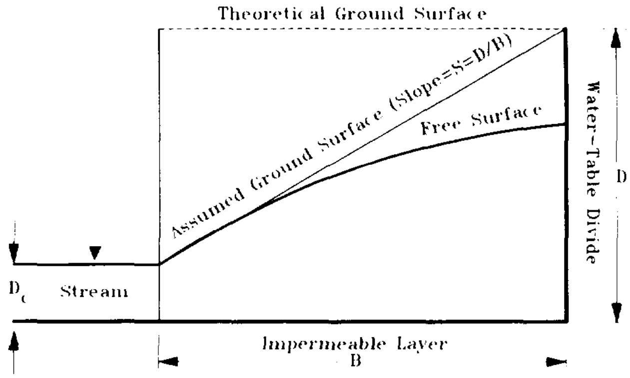


Fig. 1. Conceptualization of Dupuit-Boussinesq stream-aquifer model.

fully penetrating stream channel from an unconfined rectangular aquifer placed on a horizontal impermeable layer (see Figure 1). If one ignores both the impact of capillarity above the free water surface and evapotranspiration, and if one employs the Dupuit assumption, then the stream-aquifer system may be represented by the Boussinesq equation [*Boussinesq*, 1877]. Brutsaert and Nieber refer to this system as the Dupuit-Boussinesq aquifer. Using *Brutsaert and Nieber*'s [1977] notation, the groundwater outflow q, per unit length of stream channel, corresponding to the solution to the linearized Boussinesq equation after the initial transient decay is

$$q = \left[\frac{2kD(D - D_c)}{B}\right] \exp\left[\frac{-\pi^2 kDt}{4fB^2}\right] \tag{2}$$

where k is the hydraulic conductivity in Darcy's law, D the aquifer thickness, B the aquifer breadth, D_c the water depth in the channel, f the drainable porosity of the soil and t time after a initial saturation of the aquifer.

A Watershed Model of Low Flow

A watershed is conceptualized as being composed of a large number of stream-aquifer elements where each stream-aquifer element is described by the linearized Dupuit-Boussinesq aquifer model in (2) and Figure 1. Equation (2) may be extended to an entire watershed by employing a few hydrologic similarity conditions. For example, the total watershed runoff Q arises from the lateral inflow contributions to all stream channels; hence

$$Q = 2Lq \tag{3}$$

where L is the total length of all stream channels in the watershed. If the entire watershed is underlain by aquifers which contribute to streamflow, then the unit watershed area required to sustain each unit of stream length is the inverse of the drainage density [*Schumm*, 1956]. If only a fraction of the watershed, α, is underlain by aquifers which contribute to streamflow, then the average breadth of an individual aquifer, B, is

$$B = \alpha A/2L = \alpha/2d \tag{4}$$

where d is the drainage density and A is the watershed area. Although the linearized Dupuit-Boussinesq aquifer model assumes a rectangular aquifer cross section, actual aquifers are more closely approximated in cross section by a triangle, as depicted in Figure 1, in which case the aquifer thickness at the basin divide, D, may be approximated by

$$D \cong SB \tag{5}$$

where S is the average slope of the actual land surface for each stream-aquifer unit depicted in Figure 1. Under low-flow conditions $D_c \ll D$, therefore $D - D_c \cong D$. Combining equations (2), (3), (4) and (5) converts the stream-aquifer model into a watershed model for low-flow and yields

$$Q = 2\alpha kAS^2 \exp\left[\frac{-\pi^2 kSdt}{2f\alpha}\right] \tag{6}$$

Equation (6) represents a simple approximation of the time response of watershed runoff under low-flow conditions and is used to develop a regional statistical model for estimating low-flow statistics in Massachusetts.

The Watershed as a Linear Reservoir

Equation (6) can be rewritten in the form of the well-known base flow equation,

$$Q = Q_0 K_b^t \tag{7}$$

where K_b is the base flow recession constant and Q_0 is the initial streamflow at time $t = 0$. Here K_b is a nondimensional time constant for the system. Typically, values of K_b are in the range [0.8–0.95]. *Hall* [1968] provides a review of base flow equations of the form given in (7). From (6) and (7) one obtains a physical model for the base flow recession constant

$$K_b = \exp\left[-\pi^2 kSd/2f\alpha\right] \tag{8}$$

Dooge [1973], *Brutsaert and Nieber* [1977] and others have shown that (7) is a solution of the continuity equation

$$dV/dt = I - Q \tag{9}$$

when the outflow Q from the watershed is linearly related ($n = 1$) to the basin storage V:

$$Q = aV^n \tag{10}$$

Here a and n are constants and the inflow I to the watershed is assumed to be zero under low-flow conditions. The general storage model described by (9) and (10) may be rewritten in the form

$$dQ/dt = -na^{1/n}Q^{(2n-1)/n} \tag{11a}$$

$$dQ/dt = -na^{1/n}Q^b \tag{11b}$$

where a, b and n are constants with $b = (2n - 1)/n$. Under the linear reservoir hypothesis, $n = 1$ in (10) and (11), in which case (7) and (11b) yield

$$a = -\ln(K_b) \tag{12a}$$

$$b = 1 \tag{12b}$$

Therefore one would expect the constant b in (11b) to be equal to 1 if the watershed acts like a linear reservoir under low-flow conditions. Furthermore, since the linear reservoir model in (7) is equivalent to the conceptual watershed model in (6), one would hope that if $b = 1$ for a set of watersheds,

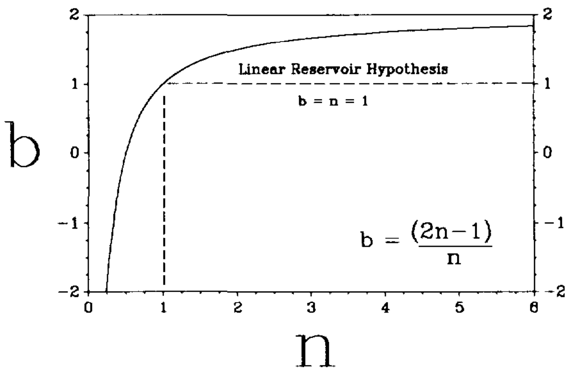


Fig. 2. Plot of the relation $b = (2n - 1)/n$ in (11).

then both the linear reservoir hypothesis (equation (7)) and the simple conceptual watershed model (equation (6)) should provide an adequate approximation to the relationship between watershed characteristics and the low-flow response of watersheds in that region.

The general storage model described by (11) is physically valid for any value of n in the range $(0, \infty)$, leading to an infinite number of possible solutions. Figure 2 describes the relationship $b = (2n - 1)/n$. For values of n near unity, b can take on a wide range of values and the linear reservoir hypothesis ($n = b = 1$) is only one.

EXPERIMENTS

In the following sections, experiments are performed to evaluate the adequacy of applying the linear reservoir hypothesis and the linearized Dupuit-Boussinesq aquifer model to approximate the low-flow response of watersheds in Massachusetts. In addition, we extend the watershed model derived in the previous section to a regional hydrologic model for estimating low-flow statistics in the Massachusetts region. The following experiments employ records from 23 unregulated basins in or near Massachusetts with long-term U.S. Geological Survey streamflow records. Stations with the following attributes were selected: (1) All of the rivers are perennial and all streamflows are greater than zero. (2) No significant withdrawals, diversions, or artificial recharge areas are contained in the basins; hence we consider the streamflows to be essentially unregulated.

Table 1 reports the U.S. Geological Survey gage numbers, record lengths, streamflow statistics, and basin characteristics along with site numbers used in this and four other low-flow studies [*Vogel and Kroll*, 1989, 1990, 1991; *Fennessey and Vogel*, 1990]. Further information regarding these sites, including a location map, is provided in those studies.

Experiment 1: Do Watersheds in Massachusetts Act Like Linear Reservoirs?

In the previous section we showed that if the Dupuit-Boussinesq representation of a watershed (equation (6)) provides an adequate approximation to the low-flow re-

TABLE 1. Geomorphic and Hydrologic Watershed Characteristics

USGS Gage No.	Site No.	Record Length, yr	A, square miles	H, feet	d, miles^{-1}	K_b	$Q_{7,2}$, feet^3/s	$Q_{7,10}$, feet^3/s
01180500	1	73	52.70	1765	1.86	0.906	4.38	1.41
01096000	2	34	63.69	1161	2.15	0.928	10.16	4.48
01106000	3	37	8.01	227	2.00	0.884	0.18	0.05
01170100	4	16	41.39	1873	1.40	0.924	7.54	4.49
01174000	5	34	3.39	531	1.97	0.863	0.11	0.02
01175670	6	23	8.68	417	1.09	0.901	0.50	0.23
01198000	7	19	51.00	1317	1.33	0.932	5.42	3.27
01171800	8	11	5.46	530	1.65	0.920	0.90	0.47
01174900	9	22	2.85	585	2.09	0.893	0.18	0.09
01101000	10	38	21.30	277	1.43	0.918	0.80	0.21
01187400	11	31	7.35	877	1.64	0.867	0.50	0.23
01169000	12	44	89.00	1667	1.95	0.920	13.85	7.83
01111300	13	20	16.02	393	1.69	0.888	0.84	0.34
01169900	14	17	24.09	1298	2.19	0.931	5.44	3.45
01181000	15	48	94.00	1739	1.70	0.919	10.81	5.58
01332000	16	52	40.90	2068	1.37	0.909	7.83	5.01
01097300	17	20	12.31	248	1.40	0.907	0.58	0.15
01333000	18	34	42.60	2658	1.66	0.933	8.26	4.32
01165500	19	65	12.10	797	1.83	0.895	1.23	0.60
01171500	20	45	54.00	1476	1.98	0.916	10.08	6.13
01176000	21	71	150.00	801	1.74	0.941	31.34	15.49
01162500	22	63	19.30	718	1.52	0.885	1.40	0.44
01180000	23	28	1.73	643	1.25	0.900	0.11	0.06

Average basin slope S, is approximated in this study using (17). Further information regarding the location and characteristics of these 23 sites is contained in the works by *Vogel and Kroll* [1989, 1990, 1991] and *Fennessey and Vogel* [1990]. Variables are defined as A, watershed area; H, watershed relief; d, drainage density; K_b, base flow recession constant estimated from (15); $Q_{7,2}$, 7-day 2-year discharge estimated from streamflow record; and $Q_{7,10}$, 7-day 10-year discharge estimated from streamflow record. 1 square mile equals 2.59 km^2; 1 foot equals 0.3048 m; 1 mile^{-1} equals 0.6215 km^{-1}; 1 foot^3/s equals 0.0283 m^3/s.

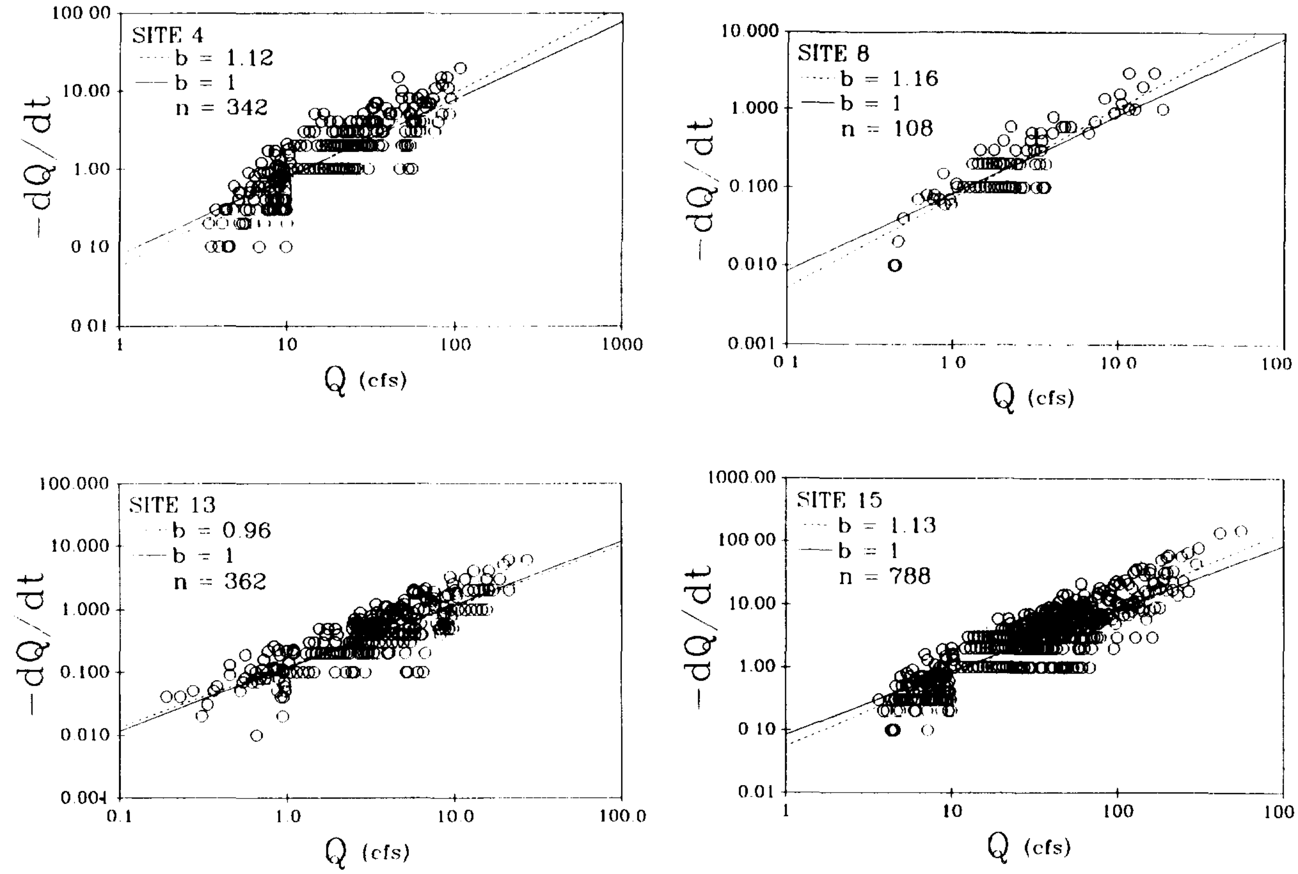


Fig. 3. Comparison of the theoretical relationships $dQ/dt = -aQ^b$ when b is constrained to unity (solid lines) with the relationships estimated when b is unconstrained (dashed lines) using ordinary least squares estimators in (14) at four of the 23 sites.

sponse of stream-aquifer systems, then the watershed should behave like a linear reservoir under low-flow conditions with $b = n \cong 1$ in (11). In this section a value for b in (11) is estimated for each of the watersheds in Table 1.

The 23 watersheds contain a total of 845 site-years of average daily streamflow data, or over 300,000 average daily flow values. 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 candidate hydrograph recessions. If the total length of a candidate recession is l then some initial portion of that recession contains predominantly surface or storm runoff. In this study, the first λl days were removed from each hydrograph recession, where λ is a fraction which was allowed to vary over the interval $0.0 \leq \lambda \leq 0.8$. To avoid spurious observations, we only accepted pairs of streamflow (Q_t, Q_{t-1}) when $Q_t \geq 0.7Q_{t-1}$.

Since the focus of this study is on the development of a regional hydrologic model for estimating $Q_{d,T}$, the d-day T-year low flow, only candidate recessions during the summer are considered. Annual minimum 7-day low flows occur almost exclusively during the period July–October at these 23 sites; hence we define the summer by that period.

Equation (11b) may be applied to actual streamflows by taking logarithms and adding an error term to obtain

$$\ln[-dQ/dt] = \ln[a] + b\ln[Q] + \varepsilon \qquad (13)$$

where the ε are normally distributed errors with zero mean and constant variance. If $n = b = 1$, $\ln[a] = \ln[-\ln[K_b]]$. Note that if $b \neq 1$, the linear reservoir hypothesis no longer holds. Similar to *Brutsaert and Nieber* [1977], (13) is fit to the candidate recessions at each site using the numerical approximation

$$\ln[Q_{t-1} - Q_t] = \ln[a] + b\ln[\tfrac{1}{2}(Q_t + Q_{t+1})] + \varepsilon \qquad (14)$$

where Q_{t-1} and Q_t are the daily streamflows on two successive days. Figure 3 depicts the relationship between dQ/dt and Q at four of the 23 sites using $\lambda = 0.3$. The four sites are chosen to span the range of watershed areas described in Table 1. The dashed lines on each plot in Figure 3 depict (14) fit using ordinary least squares estimators of $\ln[a]$ and b. In each case, the estimated value of b is close to 1. For comparison, the solid line in each plot depicts (14) when the parameter b is constrained to unity. When $b = 1$, one can easily show that a (constrained) least squares estimator of K_b in (14) is

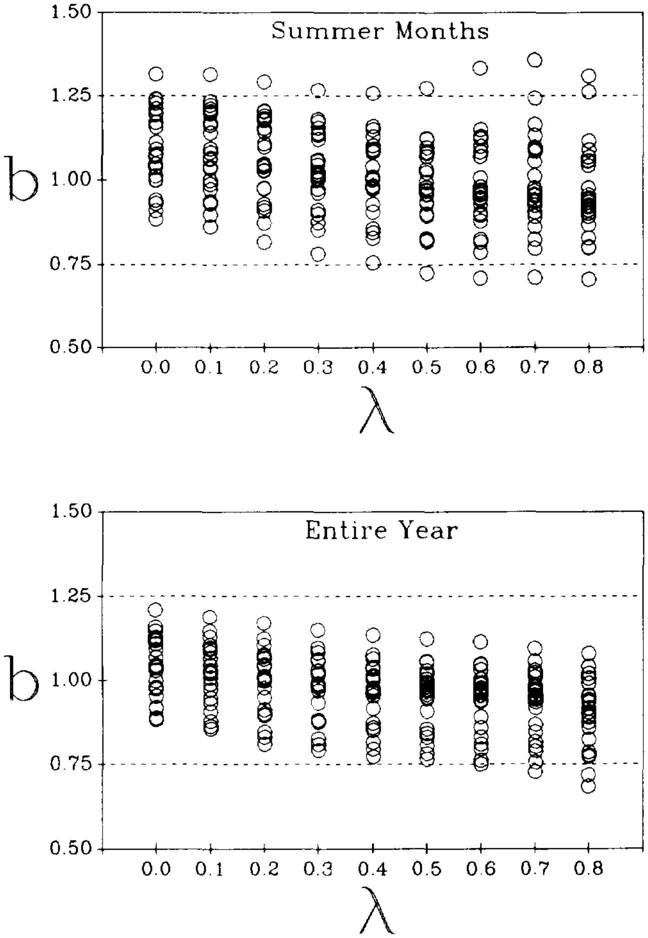


Fig. 4. Plot of ordinary least squares estimates of b in $dQ/dt = -aQ^b$ at 23 sites (open circles) as a function of λ, the fraction of each hydrograph recession which is removed. The dotted lines represent the average value of b across all sites as a function of λ. 1 cfs (foot^3/s) equals 0.0283 m^3/s.

$$K_b = \exp\left\{-\exp\left[\frac{1}{m}\sum_{t=1}^{m}\{\ln(Q_{t-1} - Q_t) - \ln[\tfrac{1}{2}(Q_t + Q_{t+1})]\}\right]\right\} \quad (15)$$

where m is the total number of pairs of consecutive daily streamflows Q_{t-1} and Q_t at each site. The estimates of K_b derived from (15) are given in Table 1.

Figure 4 displays the estimated values of b in (14) for all 23 sites, as a function of λ, the fraction of each candidate recession which was removed during the summer and for the entire year. The values of b, denoted by the open circles in Figure 4, appear to be scattered about $b = 1$. Estimated b values appear to be more variable during the summer than for the entire year, especially for large λ. The increased variability is probably due to sampling error. As λ increases, the candidate recessions are shorter, and, in addition, the samples of recessions during the summer contain about 25% as much data as for the entire year; hence the estimated values of b during the summer contain much more sampling noise than corresponding estimates of b for the entire year. The estimates of b based upon the entire year, at all 23 sites, remain within the interval $0.75 \le b \le 1.25$ for $\lambda \le 0.5$, indicating that $b = 1$ is a reasonable approximation for this region. For the summer months $\lambda = 0.3$ appears to be a reasonable choice to assure that the linear reservoir hypothesis provides an adequate approximation to the low-flow behavior of watersheds in this region.

Zecharias and Brutsaert [1988] also employed the assumption $n = b = 1$ in the Allegheny Mountain section of the Appalachian Plateau. *Brutsaert and Nieber* [1977] concluded that $b = 1.5$ ($n = 2$) for six basins in the Finger Lakes region of New York. However, they examined relatively few streamflows at each site and they did not fit (14) using ordinary least squares regression as was done here.

The linear reservoir hypothesis and the Dupuit-Boussinesq aquifer model which both lead to a fixed watershed time constant, K_b, appear to provide a reasonable approximation to the low-flow behavior in the 23 watersheds examined here.

Experiment 2: Estimation of Physically Based Regional Low-Flow Models

In this section, the conceptual watershed model in (6) is modified to obtain regional regression equations for estimating the low-flow statistics $Q_{7,2}$ and $Q_{7,10}$. The statistic $Q_{7,10}$, the most widely used index of low flow in the United States [*Riggs et al.*, 1980], is defined as the annual minimum 7-day average daily streamflow which will be less than $Q_{7,10}$, on average, once every 10 years. The statistic $Q_{7,2}$ is the median of the annual minimum 7-day low-flow series. Estimates of $Q_{7,2}$ and $Q_{7,10}$ were obtained by fitting a two-parameter lognormal distribution to sequences of annual minimum 7-day streamflows [see *Vogel and Kroll*, 1989]; the resulting estimates are provided in Table 1.

Equations (6) and (8) can be combined to yield

$$Q = 2\alpha k A S^2 K_b^t \quad (16)$$

Here the flow at time t depends on the five parameters α, k, A, S, and K_b. Interestingly, (16) is similar in structure to (1), the most widely used model form employed in both low-flow and flood flow regional regression models. *Zecharias and Brutsaert* [1985] discussed simplified procedures for the estimation of the average basin slope S. They showed that average basin slope estimates derived with *Strahler*'s [1950*a*, *b*] method were highly correlated with estimates derived from more complex methods in the Allegheny Mountain section of the Appalachian Plateau. We employ *Strahler*'s [1950*a*, *b*] slope-determining method which is

$$S \cong 2Hd \quad (17)$$

where H is equal to the basin relief, defined as the difference in elevation between the basin summit and the basin outlet and d is drainage density. We follow *Zecharias and Brutsaert* [1985] by defining the basin summit elevation as the average of the highest peak and the two adjacent peaks on either side of it. The values of relief H, drainage density d, and watershed area A, are given in Table 1. Here H and A were measured from 1:24,000 scale topographic maps and the total length of streams, L, was obtained from a geographic information system based upon digitized 1:24,000 scale topographic maps.

TABLE 2. Summary of Estimated Regional Regression Equations: $Q_{7,T} = b_0 A^{b_1} S^{b_2} K_b^{b_3}$

T	Model	b_0	b_1	b_2	b_3	Standard Error (SE%)*	R^2†
2	A	0.033 (11)‡	1.34 (14)			54.9	90.0
	A, K_b	···§	0.99 (17)		24.1 (12.8)	47.1	92.9
	A, S	0.089 (6.8)	1.13 (12)	0.556 (3.8)		40.4	93.9
	A, S, K_b	···	0.89 (21)	0.557 (5.4)	17.5 (10)	26.9	96.9
10	A	0.011 (9.6)	1.44 (9.9)			93.2	81.4
	A, K_b	···	0.98 (11)		32.2 (12)	74.4	85.9
	A, S	0.048 (5.5)	1.12 (7.8)	0.83 (3.7)		67.3	88.3
	A, S, K_b	···	0.84 (13)	0.79 (4.9)	22.9 (8.5)	44.2	93.4

*Computed using SE% = $100[\exp(s_\epsilon^2) - 1]^{1/2}$, where s_ϵ^2 is *Hardison's* [1971, equation (3)] unbiased estimator of the variance of the residuals ϵ, in (18).

†The values of R^2 are adjusted for the number of degrees of freedom which remain after parameter estimation.

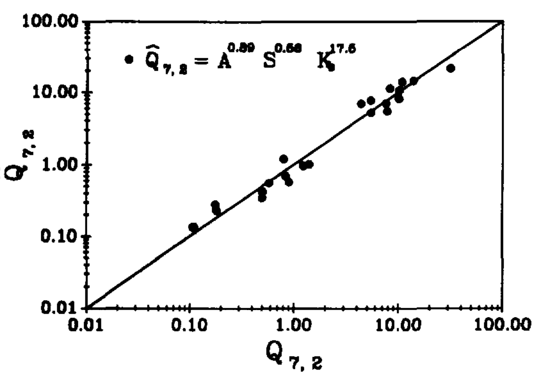
‡The values in parentheses are the t ratios of the estimated model parameters.

§Three dots denote that the estimated values of ln (b_0) were not significantly different from zero using a 5% level t test; hence the models were refit using ordinary least squares regression, constraining b_0 to be equal to unity.

Multivariate ordinary least squares regression procedures were used to fit models of the form

$$Q_{7,T} = b_0 A^{b_1} S^{b_2} K_b^{b_3} e^{\epsilon} \quad (18)$$

where T is the recurrence interval and ϵ are normally distributed residuals with zero mean and constant variance σ_ϵ^2. The estimated models are summarized in Table 2 and the model predictions are displayed in Figure 5. As is typical, watershed area, A, explains most of the variability of these low-flow statistics. However, substantial improvements in the regression models are obtained by including all three independent variables: A, S and K_b. In all cases, estimated model parameters are significantly different from zero and the t ratios (shown in parentheses) are always greater than 4.9 for the full model described by (18).



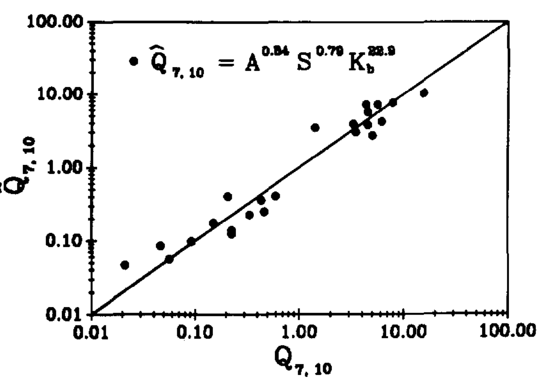


Fig. 5. Plot of predicted values of $Q_{7,2}$ and $Q_{7,10}$ based upon regional regression equations (equation (18)) versus at-site sample estimates of these low-flow statistics given in Table 1.

Stedinger and Tasker [1985] showed that ordinary least squares (OLS) regression procedures can lead to large upward bias in estimates of the model error variance σ_ϵ^2. To correct this upward bias we employ the simple unbiased estimator suggested by *Hardison* [1971, equation (3)] which we term s_ϵ^2. *Stedinger and Tasker* [1986, equation (19)] also derive Hardison's unbiased estimator of model error variance and verify that it leads to reasonably unbiased estimates of σ_ϵ^2. The estimator s_ϵ^2 is a function of the standard OLS model error variance estimator, the average cross correlation among the 23 sequences of annual minimum 7-day low flows ρ, the record lengths (given in Table 1) and, in this case, the standard deviation of the logarithms of each sequence of annual minimum 7-day low flows. Here $\rho = 0.35$ which is identical to the value used by *Vogel and Kroll* [1990] when they compared the use of generalized least squares (GLS) and OLS regression procedures using the same data base employed here.

Equation (18) cannot be used to estimate low-flow statistics at an ungaged site because estimates of K_b would not be available. However, models are in use which incorporate K_b, where corresponding estimates are obtained from maps [*Bingham*, 1986]. Models which require estimates of independent variables from maps introduce an additional source of error, map error. Alternatively, K_b could be estimated from a modest gaging program by examining a few hydrograph recessions at the ungaged site. This idea is analogous to *Potter and Faulkner*'s [1987] suggestion to estimate the time to peak T for an ungaged watershed from a modest gaging program. They found the ratio A/T to be highly correlated with flood flow statistics, yet T is unavailable at a gaged site analogous to the situation here for K_b.

The relationships described in Table 2 and Figure 5 could possibly be improved by incorporating an estimate of the basin hydraulic conductivity, k, as shown in (16). However, the relationships in Table 2 and Figure 5 could never be perfect, because the statistics $Q_{d,T}$ and K_b will always

TABLE 3. Comparison of Regional Regression Models for Estimation of $Q_{7,10}$ Developed Here With Previous Studies in New England

State	Regional Regression Model*	Source
Connecticut	$Q_{7,10} = 0.457A^{0.90}S_m^{-0.66}E^{0.75}A_s^{0.38}$	*Thomas and Cervione* [1970]
Connecticut	$Q_{7,10} = 0.67A_s + 0.01A_t$	*Cervione et al.* [1982]
Maine	$Q_{7,10} = 0.000201A^{1.09}P^{4.41}$	*Parker* [1977]
Vermont	$Q_{7,10} = (0.068A^{0.827}E_m^{0.224}) - 1$	*Downer* [1983]
Massachusetts	$Q_{7,10} = (0.0457A^{0.83}10^{0.0028G}) - 0.1$	*Tasker* [1972]
Massachusetts	$Q_{7,10} = 10^{-28.7}A^{1.36}P^{21.9}I^{-5.78}G^{0.568}F^{-4.36}U^{-0.344}$	*Male and Ogawa* [1982]
Massachusetts	$Q_{7,10} = A^{0.84}S^{0.79}K_b^{22.9}$	this study

*Definition of model parameters: A denotes watershed area, in square miles; A_s, area of basin underlain by stratified drift, in square miles; A_t, area of basin underlain by till-mantled bedrock, in square miles; E, mean watershed elevation in, feet above mean sea level (msl); E_m, minimum watershed elevation in feet above msl; F, forest cover, as percent of A; G, groundwater factor; I, 24-hour, 2-year rainfall, in inches; K_b, base flow recession constant, dimensionless; P, mean annual precipitation, in inches; S, average basin slope, dimensionless, equal to $2Hd$ (see Table 1); S_m, main channel slope, in feet per mile; and U, urban area, as percentage of A. 1 square mile equals 2.59 km^2; 1 foot equals 0.3048 m; 1 inch equals 2.54 cm; 1 foot per mile equals 0.1894 m/km.

contain sampling variability due to the limited record lengths upon which those statistics are based. Measurement errors associated with the streamflows and drainage basin characteristics further confound the fitted relationships. Model error is inevitable since real watersheds always behave differently than their mathematical representations.

COMPARISON WITH PREVIOUS STUDIES IN NEW ENGLAND

Table 3 summarizes regression equations developed for estimating $Q_{7,10}$ at ungaged sites in the states of Connecticut, Massachusetts, Vermont and Maine. The equations derived in this study are included for comparison. Since each regression model is based upon different sets of gaging stations, with different record lengths, cross correlations, and parameter estimation procedures, we cannot compare the precision of these models. Interestingly, watershed area A is the only independent variable which appears in every model; otherwise the models are conspicuously different considering that all models purport to model the same process in the New England region. This result is exactly what *Wallis* [1965] predicted would happen when the physical model is unspecified.

Interestingly, the exponents for the watershed area term in the models developed by *Tasker* [1972], *Downer* [1983], and us are all very close in magnitude, 0.83, 0.827 and 0.84, respectively. This result is revealing considering the fact that these three studies were developed for three different regions, southeastern Massachusetts [*Tasker*, 1972], Vermont [*Downer*, 1983] and central western Massachusetts (this study).

CONCLUSIONS

The primary objective of this study was to develop improved regional regression equations for estimating low-flow statistics from watershed characteristics by first formulating a conceptual watershed model for low flows. Three of the independent variables derived from that conceptual model, drainage area A, average basin slope S, and the base flow recession constant K_b, explained 97% and 93% of the variability associated with the low-flow statistics $Q_{7,2}$ and $Q_{7,10}$, respectively, in the central western Massachusetts region.

Unfortunately, K_b is usually unavailable at ungaged sites; hence recent studies by *Vogel and Kroll* [1990] and *Fennessey and Vogel* [1990] have developed regional regression models for estimating low-flow statistics without resorting to the use of K_b. However, it is possible to estimate K_b from a modest gaging program by examining the recession characteristics of a few hydrographs or by constructing a map of K_b similar to *Bingham* [1986].

Perhaps the most important lesson to be learned here is that a simple physically based watershed model of groundwater outflow can suggest variables and the functional form for regional regression equations which estimate low-flow statistics at ungaged sites. Further research is required to examine alternate conceptual watershed models so that more general conclusions may be reached regarding which drainage basin characteristics are useful to include in low-flow regional regression models.

Acknowledgments. This research was supported by a cooperative agreement between Tufts University and the U.S. Geological Survey with matching funds from the Massachusetts Division of Water Pollution Control in the Department of Environmental Protection. We are grateful to Thomas Farmer and S. William Wandle, Jr., for providing the geographic information system estimates of drainage densities and to Katherine D. Martel for providing the estimates of relief in Table 1. We are grateful to Gary Tasker and the four anonymous reviewers whose comments led to important revisions.

REFERENCES

Bingham, R. H., Regionalization of low-flow characteristics of Tennessee streams, U.S. Geol. Surv. Water Resour. Invest., *85-4191*, 1986.

Boussinesq, J., Essai sur la theorie des eaux courantes: Du mouvement non permanent des eaux souterraines, *Acad. Sci. Inst. Fr.*, *23*, 252–260, 1877.

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

Cervione, M. A., R. L. Melvin, and K. A. Cyr, A method for estimating the 7-day 10-year low flow of streams in Connecticut, *Conn. Water Resour. Bull.*, *34*, 1982.

Dingman, S. L., Synthesis of flow-duration curves for unregulated streams in New Hampshire, *Water Resour. Bull.*, *14*(6), 1481–1502, 1978.

Dooge, J. C. I., Linear theory of hydrologic systems, *Tech. Bull. U.S. Dep. Agric.*, *1468*, 327 pp., 1973.

Downer, R. N., 7-day, 10-year low-flow model for ungaged streams in Vermont, report prepared for the Agency of Environmental Conservation, Montpelier, Vt., 1983.

Fennessey, N., and R. M. Vogel, Regional flow-duration curves for ungauged sites in Massachusetts, *J. Water Resour. Plann. Manage.*, *116*(4), 530–549, 1990.

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

Hardison, C. H., Prediction error of regression estimates of streamflow characteristics at ungaged sites, *U.S. Geol. Surv. Prof. Pap.*, *750-C*, C228–C236, 1971.

Male, J. W., and H. Ogawa, Low flows of Massachusetts streams, *Publ. 125*, Water Resour. Res. Cent., Univ. of Mass., Amherst, 1982.

Newton, D. W., and J. C. Herrin, Estimating flood frequencies at ungaged locations, in *Proceedings of the Conference in Hydraulic Engineering*, pp. 528–533, American Society of Civil Engineers, New York, 1983.

Parker, G. W., Methods for determining selected flow characteristics for streams in Maine, *U.S. Geol. Surv. Open File Rep.*, *78-871*, 1977.

Potter, K. W., and E. B. Faulkner, Catchment response as a predictor of flood quantiles, *Water Resour. Bull.*, *23*(5), 857–861, 1987.

Riggs, H. C., et al., Characteristics of low flows, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, *106*(HY5), 717–731, 1980.

Schumm, S. A., Evolution of drainage systems and slopes in Badlands at Perth Amboy, New Jersey, *Bull. Geol. Soc. Am.*, *67*, 607, 1956.

Stedinger, J. R., and G. D. Tasker, Regional hydrologic analysis, 1, Ordinary, weighted, and generalized least squares compared, *Water Resour. Res.*, *21*(9), 1421–1432, 1985.

Stedinger, J. R., and G. D. Tasker, Regional hydrologic analysis, 2, Model-error estimators, estimation of sigma and log-Pearson type 3 distributions, *Water Resour. Res.*, *22*(10), 1487–1499, 1986.

Strahler, A. N., Equilibrium theory of erosional slopes approached by frequency distribution analysis, I, *Am. J. Sci.*, *248*, 673–696, 1950*a*.

Strahler, A. N., Equilibrium theory of erosional slopes approached by frequency distribution analysis, II, *Am. J. Sci.*, *248*, 800–814, 1950*b*.

Tasker, G. D., Estimating low-flow characteristics of streams in southeastern Massachusetts from maps of groundwater availability, *U.S. Geol. Survey Prof. Pap.*, *800-D*, D217–D220, 1972.

Thomas, D. M., and M. A. Benson, Generalization of streamflow characteristics from drainage-basin characteristics, *U.S. Geol. Surv. Water Supply Pap.*, *1975*, 1970.

Thomas, M. P., and M. A. Cervione, A proposed streamflow data program for Connecticut, *Conn. Water Resour. Bull.*, *23*, 1970.

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

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

Vogel, R. M., and C. N. Kroll, The value of streamflow record augmentation procedures in lowflow and floodflow frequency analysis, *J. Hydrol.*, *125*, 259–276, 1991.

Wallis, J. R., Multivariate statistical methods in hydrology: A comparison using data of known functional relationship, *Water Resour. Res.*, *1*(4), 447–461, 1965.

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

Zecharias, Y. B., and W. Brutsaert, Recession characteristics of groundwater outflow and base flow from mountainous watersheds, *Water Resour. Res.*, *24*(10), 1651–1658, 1988.

C. N. Kroll, Department of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853.

R. M. Vogel, Department of Civil Engineering, Tufts University, Medford, MA 02155.

(Received June 18, 1991;
revised April 22, 1992;
accepted April 29, 1992.)