Tsallis entropy–Based flow duration curve

3 AUTHORS:

Vijay P. Singh
Texas A&M University

731 PUBLICATIONS **11,831** CITATIONS

SEE PROFILE

Huijuan Cui
Chinese Academy of Sciences

15 PUBLICATIONS **47** CITATIONS

SEE PROFILE

Aaron Byrd
Engineer Research and Development Center …

22 PUBLICATIONS **24** CITATIONS

SEE PROFILE

Tsallis Entropy-Based Flow Duration Curve

V. P. Singh, H. Cui, A. R. Byrd

ABSTRACT. *The flow duration curve (FDC) is employed for addressing a multitude of problems in water resources engineering, such as prediction of the distribution of future flows, forecasting of future recurrence frequencies, comparison of watersheds, construction of load duration curves, and determination of low flow thresholds. Usually, the FDC is constructed empirically for a given set of flow data, and the FDC so constructed is found to vary from one year to another and from one gauging station to another within the same watershed. This article attempts to analytically derive the FDC by maximizing the Tsallis entropy based on the knowledge that the mean discharge is known, thus obviating the need for any fitting. The mean discharge is found to be strongly related to the drainage area. The Tsallis entropy-based FDC is tested using field data and is found to be in agreement with the observed curve. The entropy method permits a probabilistic characterization of the FDC and hence a quantitative assessment of its uncertainty. With this method, the flow duration curve can also be forecasted for different recurrence intervals. The entropy is found to monotonically increase with the increase in time interval, indicating that the flow system becomes more complex but the degree of complexity decreases with increasing time interval after a certain time, eventually reaching a constant value, reflecting a reduced influence of land use change and other human influences on the flow regime.*

Keywords. Duration curve, Entropy, Flow duration curve, Lagrange multipliers, Principle of maximum entropy, Tsallis entropy.

For a gauging station, the flow duration curve (FDC) is a plot of streamflow values against the percent of time that the streamflow either equals or exceeds a specific value. The plot considers the full range of flows and is constructed over a specified period of time scaled between 0% and 100%. The time interval for constructing an FDC depends on the need, but daily average discharge values are usually used; sometimes weekly, monthly, or seasonally averaged values are also used. However, averaging over longer time intervals obscures details of the variations in flow, and the effect of varying time interval is not the same for all streams. The differences between an FDC based on daily discharge values and that based on monthly values can be as high as 35%, as noted by Foster (1934). For large streams where the flow from day to day is almost uniform, weekly FDC may be almost the same as daily FDC, whereas for flashy streams with sudden floods lasting for a few hours in a day, the daily and weekly FDCs will be greatly different.

FDCs are constructed using the entire range of flow conditions for any given stream. If the FDC of a stream is based on long-term flow, then it can be employed for predicting the distribution of future flows for water supply (Mitchell, 1957), hydropower (Hickox and Wessenauer, 1933), sediment load (Miller, 1951), and pollutant load (Searcy, 1959). It can also be utilized to compare watersheds and hence their clustering, to construct load duration curves for total maximum daily loads (TMDLs) (U.S. EPA, 2007), to forecast future recurrence frequencies, to determine the low flow threshold for defining droughts, and to construct power duration curves.

Flow duration curves are generally constructed empirically. A typical semi-log FDC exhibits a sigmoidal shape, curving upward near the flow duration of 0% and downward at a frequency near 100%, with a nearly constant slope in between. The overall slope of a flow duration curve is an indication of streamflow variability at the gauge, reflecting, in turn, the integrated effect of watershed characteristics. For practical applications, the U.S. EPA (2007) classified the flow region into five different classes: 0% to 10% interval for high flows, 10% to 40% for moist conditions, 40% to 60% for mid-range conditions, 60% to 90% for dry conditions, and 90% to 100% for low flows.

This study employs the Tsallis entropy to analytically derive the flow duration curve (FDC). Since the FDC requires information on mean flow as well as the lower and upper limits of flow, these flows are related to the drainage area. The derived FDC is then expressed in terms of one parameter that is uniquely related to the ratio of mean flow to the maximum flow. This study also investigates the transferability of the FDC from one location to another on the same river or within the same basin and if it can be used

The authors are **Vijay P. Singh, ASABE Member,** Distinguished Professor and Caroline and William N. Lehrer Distinguished Chair in Water Engineering, Department of Biological and Agricultural Engineering and Zachry Department of Civil Engineering, Texas A&M University, College Station, Texas; **Huijuan Cui,** Graduate Student, Water Management and Hydrologic Science Program, Texas A&M University, College Station, Texas; **Aaron R. Byrd,** Research Hydraulic Engineer and Chief of Hydrologic Systems Branch, Coastal and Hydraulics Laboratory, U.S. Army Corps of Engineers Engineer Research and Development Center, Vicksburg, Mississippi. **Corresponding author:** V. P. Singh, 321 Scoates Hall, Texas A&M University, College Station, TX 77843-2117; phone: 979-845-7028; email: vsingh@tamu.edu.

Transactions of the ASABE

Vol. 57(3): 837-849   2014 American Society of Agricultural and Biological Engineers   ISSN 2151-0032   DOI 10.13031/trans.57.10483   837

in forecasts corresponding to a given probability. It also investigates the variation of entropy with time interval for which the FDC is constructed.

TSALLIS ENTROPY-BASED DERIVATION OF FLOW DURATION CURVE

Entropy is a measure of uncertainty of a random variable or its probability distribution. Because the system under consideration is represented by the random variable, entropy can be considered as a measure of the uncertainty associated with the system. Since information and uncertainty are inversely related, that is, information reduces uncertainty and more information means less uncertainty, entropy can also be considered as a measure of entropy. Many forms of entropy have been proposed in the statistical mechanics literature (Singh, 2013), but the most popular form is the Boltzmann-Gibbs-Shannon entropy (Shannon, 1948). In recent years, another form of entropy, called the Tsallis entropy (Tsallis, 1988), has been proposed. The Tsallis entropy is a generalization of the Boltzmann-Gibbs-Shannon entropy and is receiving considerable attention these days. The greatest advantage of the Tsallis entropy is that it satisfies three fundamental properties: universality, ubiquity, and eternity. This is what motivated the use of the Tsallis entropy in this study. Because uncertainty is quantified by a probability distribution, entropy allows us to determine the probability distribution, and with the use of physical reasoning we can determine a physical relationship, such as an FDC.

For deriving an FDC, it is assumed that temporally averaged discharge (Q) is a random variable, varying from a minimum value (Q_{min}) to a maximum value (Q_{max}), with a probability density function (PDF) denoted as $f(Q)$. The time interval for which the discharge is averaged depends on the purpose for constructing an FDC, but frequently it is taken as one day. The procedure for deriving the FDC entails essentially two main steps: maximizing the Tsallis entropy and hypothesizing the cumulative distribution of discharge in terms of time. The first step involves the specification of constraints, use of the method of Lagrange multipliers for optimization of the Tsallis entropy, derivation of the probability distribution of discharge, and determination of the Lagrange multipliers. Each of these steps is now discussed.

The Tsallis entropy (Tsallis, 1988), denoted as H, for discharge Q can be defined as:

$$H(Q) = H\big[f(Q)\big]$$
$$= \frac{1}{m-1} \int_{Q_{min}}^{Q_{max}} f(Q) \left\{ 1 - \big[f(Q)\big]^{m-1} \right\} dQ \quad (1a)$$

where m is the entropy index. If discharge is available as a discrete series, the Tsallis entropy takes on the form:

$$H(Q) = H[P]$$
$$= \frac{1}{m-1} \sum_{i=1}^{N} p(Q_i) \left\{ 1 - [p(Q_i)]^{m-1} \right\} \quad (1b)$$

where $p(Q_i) = p_i$ is the probability that $Q = Q_i$, $P = \{p_i, i = 1, 2,..., N\}$ is the probability distribution of Q, and N is the number of values that Q takes on between its maximum and minimum. Equation 1a expresses a measure of uncertainty about $f(Q)$ measured by $\{(1-[f(Q)]^{m-1})/(m-1)\}$ or the average information content of sampled Q. Therefore, $f(Q)$ must be derived first, which can be accomplished by maximizing $H(Q)$, subject to specified constraints. In order to determine the $f(Q)$ that is least biased toward what is not known and most biased toward what is known (with regard to discharge), the principle of maximum entropy (POME), developed by Jaynes (1957, 1982), is invoked. POME requires the specification of certain information on discharge, expressed in terms of what are called constraints, and leads to the most appropriate probability distribution that has the maximum entropy or uncertainty (Singh, 1998).

For purposes of simplicity from a practical standpoint, it is assumed that all that is known is the mean discharge, which can be expressed as:

$$C_1 = \int_{Q_{min}}^{Q_{max}} Q f(Q) dQ = \overline{Q} = Q_m = Q_{mean} \quad (2)$$

The mean discharge is a relatively stable quantity, and its value can be obtained from measurements. Since $f(Q)$ is a PDF, it must satisfy:

$$C_1 = \int_{Q_{min}}^{Q_{max}} f(Q) dQ = 1 \quad (3)$$

For maximizing the Tsallis entropy defined by equation 1, the method of Lagrange multipliers can be employed. To that end, the Lagrange function can be constructed using equations 1 to 3 as:

$$L = \frac{1}{m-1} \int_{Q_{min}}^{Q_{max}} f(Q) \left\{ 1 - \big[f(Q)\big]^{m-1} \right\} dQ$$
$$- \lambda_0 \left[\int_{Q_{min}}^{Q_{max}} f(Q) dQ - 1 \right] - \lambda_1 \left[\int_{Q_{min}}^{Q_{max}} Q f(Q) dQ - \overline{Q} \right] \quad (4)$$

Differentiating equation 4 with respect to $f(Q)$ and equating the derivative to 0, we obtain the entropy-based PDF of Q as:

$$f(Q) = \left[\frac{m-1}{m} \left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q \right) \right]^{\frac{1}{m-1}} \quad (5)$$

It is interesting to note that at $Q = 0$, $f(Q)$ becomes:

$$\{[(m-1)/m][(1/(m-1)) - \lambda_0]\}^{1/(m-1)}$$

The PDF given by equation 5 has two unknown Lagrange multipliers, but they can be determined using equations 2 and 3.

Substituting equation 5 into equation 3, we obtain:

$$\int_{Q_{\min}}^{Q_{\max}} \left[\frac{m-1}{m} \left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q \right) \right]^{\frac{1}{m-1}} dQ = 1 \qquad (6)$$

Solution of equation 6 can be written as:

$$\left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q_{\max} \right)^{\frac{m}{m-1}}$$

$$-\left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q_{\min} \right)^{\frac{m}{m-1}} \qquad (7)$$

$$= -\left(\frac{m}{m-1} \right)^{\frac{m}{m-1}} \lambda_1$$

Likewise, substitution of equation 5 into equation 2 yields:

$$\int_{Q_{\min}}^{Q_{\max}} Q \left[\frac{m-1}{m} \left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q \right) \right]^{\frac{1}{m-1}} dQ = \overline{Q} \qquad (8)$$

Integration of equation 8 results in:

$$-\lambda_1 \overline{Q} \left(\frac{m}{m-1} \right)^{\frac{m}{m-1}} = Q_{\max} \left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q_{\max} \right)^{\frac{m}{m-1}}$$

$$-Q_{\min} \left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q_{\min} \right)^{\frac{m}{m-1}} + \frac{m-1}{2m-1} \frac{1}{\lambda_1} \qquad (9)$$

$$\times \left[\left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q_{\max} \right)^{\frac{2m-1}{m-1}} - \left(\frac{1}{m-1} - \lambda_0 - \lambda_1 Q_{\min} \right)^{\frac{2m-1}{m-1}} \right]$$

For purposes of simplification, let:

$$\lambda_* = \frac{1}{m-1} - \lambda_0$$

Then equations 7 and 9 can be cast, respectively, as:

$$(\lambda_* - \lambda_1 Q_{\max})^{\frac{m}{m-1}} - (\lambda_* - \lambda_1 Q_{\min})^{\frac{m}{m-1}}$$

$$= -\left(\frac{m}{m-1} \right)^{\frac{m}{m-1}} \lambda_1 \qquad (10)$$

$$-\lambda_1 \overline{Q} \left(\frac{m}{m-1} \right)^{\frac{m}{m-1}} = Q_{\max} (\lambda_* - \lambda_1 Q_{\max})^{\frac{m}{m-1}}$$

and $-Q_{\min} (\lambda_* - \lambda_1 Q_{\min})^{\frac{m}{m-1}} + \frac{m-1}{2m-1} \frac{1}{\lambda_1} \qquad (11)$

$$\times \left[(\lambda_* - \lambda_1 Q_{\max})^{\frac{2m-1}{m-1}} - (\lambda_* - \lambda_1 Q_{\min})^{\frac{2m-1}{m-1}} \right]$$

Equations 10 and 11 are implicit in the Lagrange multipliers λ_0 (or λ_*) and λ_1 but can be solved numerically without any difficulty.

The cumulative probability distribution function (CDF) of Q can be obtained by integrating equation 5 from $Q_{\min}$ to Q as:

$$F(Q) = \left(\frac{m-1}{m} \right)^{m/(m-1)}$$

$$\times \frac{1}{\lambda_1} \left[(\lambda_* - \lambda_1 Q_{\min})^{m/(m-1)} - (\lambda_* - \lambda_1 Q)^{m/(m-1)} \right] \qquad (12)$$

If $Q_{\min} = 0$, equation 12 reduces to:

$$F(Q) = \left(\frac{m-1}{m} \right)^{m/(m-1)}$$

$$\times \frac{1}{\lambda_1} \left[\lambda_*^{m/(m-1)} - (\lambda_* - \lambda_1 Q)^{m/(m-1)} \right] \qquad (13)$$

Equation 12 can also be written for Q explicitly as:

$$Q = \frac{\lambda_*}{\lambda_1} - \frac{1}{\lambda_1} \frac{m}{m-1}$$

$$\times \left\{ -\lambda_1 \frac{m-1}{m} F(Q) + \left[\frac{m-1}{m} (\lambda_* - \lambda_1 Q_{min}) \right]^{\frac{m}{m-1}} \right\}^{\frac{m}{m-1}} \qquad (14a)$$

and equation 13 as:

$$Q = \frac{\lambda_*}{\lambda_1} - \frac{1}{\lambda_1} \frac{m}{m-1}$$

$$\times \left\{ -\lambda_1 \frac{m-1}{m} F(Q) + \left(\frac{m-1}{m} \lambda_* \right)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}} \qquad (14b)$$

Equations 14a and 14b are quantile-probability relationships.

The maximum Tsallis entropy or uncertainty of discharge can be obtained by substituting equation 5 into equation 1:

$$H(Q) = \frac{1}{m-1} \left\{ (Q_{\max} - Q_{\min}) + \left(\frac{m-1}{m} \right)^{m/(m-1)} \right.$$

$$\times \frac{1}{(2m-1)\lambda_1} \left[(\lambda_* - \lambda_1 Q_{\max})^{(2m-1)/(m-1)} \right.$$

$$\left. - (\lambda_* - \lambda_1 Q_{\min})^{(2m-1)/(m-1)} \right\} \qquad (15)$$

In order to derive the FDC, it is assumed that all temporally averaged values of discharge Q measured at the gauging station under consideration between $Q_{\min}$ and $Q_{\max}$ are equally likely. In reality, this is not highly unlikely because different values of discharge occur at different times, and each value is equally likely. The cumulative probability distribution of discharge can then be expressed as one mi-

nus the percent time (or the ratio of time to the period of time under consideration, say 365 days for daily discharge). The probability of discharge being equal to or less than a given value of Q, or the cumulative distribution function of discharge (CDF), $F(Q) = P$(discharge $\leq$ a given value of Q), P = probability, can be expressed as:

$$F(Q) = 1 - \left(\frac{t}{T}\right) = 1 - \tau, \ \tau = t/T \tag{16}$$

where t is time (say in days), τ is dimensionless time, and T is the duration under consideration (say 365 days). It should be noted that on the left side the argument of function F in equation 16 is variable Q, whereas on the right side the variable is τ. The CDF of Q is not linear in terms of Q, unless Q and τ are linearly related. It may also be noted that a similar hypothesis was employed when using the entropy theory for deriving infiltration equations by Singh (2010a, 2010b), soil moisture profiles by Al-Hamdan and Cruise (2010) and Singh (2010c), and velocity distributions by Chiu (1987), Barbe et al. (1991), Singh and Hao (2011), Hao and Singh (2011), and Cui and Singh (2012, 2013a, 2013b). Of course, it is plausible that $F(Q)$ might have a different form that needs to be verified using empirical data.

DETERMINATION OF ENTROPY INDEX

The value of entropy index m is greater than 0, but the question arises as to what the actual value of m is or should be. Figure 1 shows the flow duration curve (FDC) at station 02131000 estimated using m = 4/3, 2, 5/2, 3, 13/4, and 4, where the Lagrange multipliers were computed using equations 10 and 11. It is seen from the figure that the high discharge part of the FDC is closer to the observations for m = 5/2, 3, and 13/4, while the low discharge part of the FDC is closer to the observations for m = 4. The estimated sum of squared errors for the FDC corresponding to m = 4/3, 2, 5/2, 3, 13/4, and 4 was, respectively, 65.57, 48.39, 31.71, 12.52, 17.33, and 21.24 m^3 s^{-1}. Thus, m = 3 was selected for use in this study.
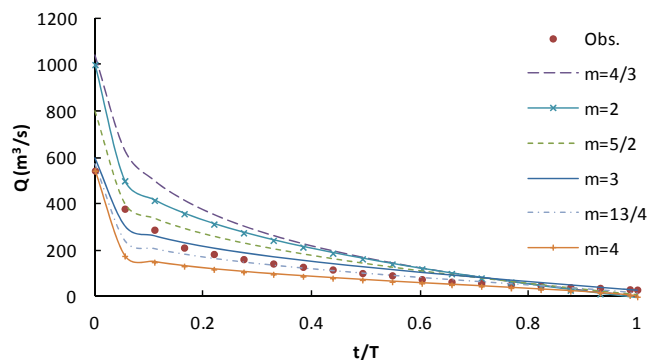

Figure 1. Flow duration curve of year 2006 for different m values.

REPARAMETERIZATION

Analysis becomes simpler and useful relations are obtained if reparameterization is done as follows. Let:

$$M = \frac{\lambda_1 Q_{max}}{\lambda_1 Q_{max} - \lambda_*} \tag{17}$$

or

$$Q_{max} = \frac{\lambda_* M}{\lambda_1 (M-1)} \tag{18}$$

Considering $Q_{min} = 0$, the ratio of $f(0)$ to $f(Q_{max})$ can be expressed in terms of M as:

$$\frac{f(0)}{f(Q_{max})} = \frac{\left[\frac{m-1}{m}(\lambda_*)\right]^{\frac{1}{m-1}}}{\left[\frac{m-1}{m}(\lambda_* - \lambda_1 Q_{max})\right]^{\frac{1}{m-1}}} = (1-M)^{\frac{1}{m-1}}$$

or

$$M = 1 - \left[\frac{f(0)}{f(Q_{max})}\right]^{m-1} \tag{19}$$

Equation 19 defines M and shows that if $M = 0$, $f(0) = f(Q_{max})$, and the distribution of discharge would tend to be uniform. On the contrary, if $M = 1$, $f(0) = 0$, and $f(Q_{max})$ would tend to infinity, which means that the probability distribution of discharge would be highly non-uniform. Thus, M can be used as an index of the uniformity of the probability distribution of discharge.

When discharge tends to reach Q_{max}, $F(Q_{max}) = 1$, and equation 14b with Q_{min} considered as 0 yields:

$$Q_{max} =$$

$$\frac{\lambda_*}{\lambda_1} - \frac{1}{\lambda_1}\frac{m}{m-1}\left[-\lambda_1\frac{m-1}{m} + \left(\frac{m-1}{m}\lambda_*\right)^{\frac{m}{m-1}}\right]^{\frac{m-1}{m}} \tag{20}$$

Dividing equation 14a by Q_{max}, we obtain:

$$\frac{Q}{Q_{max}} = \frac{\lambda_*}{\lambda_1 Q_{max}} - \frac{1}{\lambda_1 Q_{max}}\frac{m}{m-1}$$

$$\times \left\{-\lambda_1\frac{m-1}{m}F(Q) + \left[\frac{m-1}{m}\lambda_*\right]^{\frac{m}{m-1}}\right\}^{\frac{m-1}{m}} \tag{21}$$

$$= \frac{\lambda_*}{\lambda_1 Q_{max}}\left(1 - \frac{1}{\lambda_*}\left\{-\lambda_1\left(\frac{m-1}{m}\right)^{\frac{1}{m-1}}F(Q) + \lambda_*^{\frac{m}{m-1}}\right\}^{\frac{m-1}{m}}\right)$$

Noting from equation 18 that:

$$\frac{\lambda_*}{\lambda_1 Q_{max}} = 1 - \frac{1}{M}$$

which when substituted in equation 21, the result is:

$$\frac{Q}{Q_{\max}} = 1 - \frac{1}{M} - \left(1 - \frac{1}{M}\right)$$

$$\times \left\{ -\left(\frac{1}{\lambda_*}\right)^{\frac{m}{m-1}} \lambda_1 \left(\frac{m}{m-1}\right)^{\frac{1}{m-1}} F(Q) + 1 \right\}^{\frac{m-1}{m}} \tag{22}$$

$$= 1 - \frac{1}{M} \left(1 - \left\{ \left(\frac{m}{m-1}\right)^{\frac{m}{m-1}} \frac{M}{Q_{\max}} [F(Q)-1] + 1 \right\}^{\frac{m-1}{m}} \right)$$

Substituting $Q_{\min} = 0$, at $F(Q) = 0$ equation 22 reduces to:

$$0 = 1 - \frac{1}{M} \left\{ 1 - \left[1 - \left(\frac{m}{m-1}\right)^{\frac{m}{m-1}} \frac{M}{Q_{\max}} \right]^{\frac{m-1}{m}} \right\} \tag{23}$$

Rearranging equation 23, we obtain:

$$\left(\frac{m}{m-1}\right)^{\frac{m}{m-1}} \frac{M}{Q_{\max}} = 1 - (1-M)^{\frac{m}{m-1}} \tag{24}$$

Equation 22 can then be simplified by inserting equation 24 as:

$$\frac{Q}{Q_{\max}} = 1 - \frac{1}{M}$$

$$\times \left(1 - \left\{ (1-M)^{\frac{m}{m-1}} + \left[1 - (1-M)^{\frac{m}{m-1}} \right] F(Q) \right\}^{\frac{m-1}{m}} \right) \tag{25}$$

In equation 25, the Lagrange multipliers are replaced with M, and hence the flow duration curve can be determined with only one parameter, M.

The mean flow can be determined by taking the first moment of equation 5 as:

$$\overline{Q} = \int_{Q_{\min}}^{Q_{\max}} Q f(Q) dQ = \int_{Q_{\min}}^{Q_{\max}} \left[\frac{m-1}{m} (\lambda_* - \lambda_1 Q) \right]^{1/(m-1)} Q dQ$$

$$= -\frac{1}{\lambda_1} \left(\frac{m-1}{m}\right)^{\frac{m}{m-1}} \left\{ Q_{\max} (\lambda_* - \lambda_1 Q_{\max})^{\frac{m}{m-1}} \right. \tag{26}$$

$$\left. + \frac{m-1}{2m-1} \frac{1}{\lambda_1} \left[(\lambda_* - \lambda_1 Q_{\max})^{\frac{2m-1}{m-1}} - \lambda_*^{\frac{2m-1}{m-1}} \right] \right\}$$

Now the ratio of the mean flow to the maximum flow can be expressed as a function of parameter M as:

$$\frac{\overline{Q}}{Q_{\max}} = -\left(1 - \frac{1}{M}\right) \left(\frac{m-1}{m}\right)^{\frac{m}{m-1}} \left\{ \lambda_*^{\frac{1}{m-1}} Q_{\max} (1-M)^{\frac{m}{m-1}} \right.$$

$$\left. + \frac{m-1}{2m-1} \frac{\lambda_*^{\frac{m}{m-1}}}{\lambda_1} \left[(1-M)^{\frac{2m-1}{m-1}} - \lambda_*^{\frac{2m-1}{m-1}} \right] \right\} \tag{27}$$

Equation 27 can be cast as:

$$\frac{\overline{Q}}{Q_{\max}} = \Psi(M) \tag{28a}$$

where

$$\Psi(M) = -\left(1 - \frac{1}{M}\right) \left(\frac{m-1}{m}\right)^{\frac{m}{m-1}} \left\{ \lambda_*^{\frac{1}{m-1}} Q_{\max} (1-M)^{\frac{m}{m-1}} \right.$$

$$\left. + \frac{m-1}{2m-1} \frac{\lambda_*^{\frac{m}{m-1}}}{\lambda_1} \left[(1-M)^{\frac{2m-1}{m-1}} - \lambda_*^{\frac{2m-1}{m-1}} \right] \right\} \tag{28b}$$

To establish the relationship in equation 28a, the M value was computed from equation 18 by solving for the Lagrange multipliers with the use of equations 10 and 11, and it is plotted against the ratio of mean to maximum discharge in figure 2. Annual mean and maximum discharge values for five recent years collected from 13 stations on the Pee Dee River were used for verification. More details about the data are discussed later. It can be seen from figure 2 that M is linearly related to the ratio between the mean flow and maximum flow, which using regression can be written as:

$$M = 2.246 - 4.891 \frac{\overline{Q}}{Q_{\max}} \tag{29}$$

with a squared correlation coefficient of 0.9972.

VALIDATION

DATA

In order to evaluate the entropy-based flow duration curve, 13 stream gauging stations located on the Pee Dee River in South Carolina were selected. The Pee Dee River is large and wild, having a length of about 373 km. Most of the land bordering the river is floodplain forest, and the lower part of the river floodplain has been developed for rice. The climate of the Pee Dee River basin is typical of the southeastern U.S., where the average temperature is about 20.5°C and the average annual precipitation is 1195 mm. The gauging stations are operated by the U.S. Geological Survey. Table 1 gives relevant information on these stations and their lengths of records. The drainage areas for these stations vary from 11 km^2 to 36,519 km^2 and represent a broad range of flow conditions.
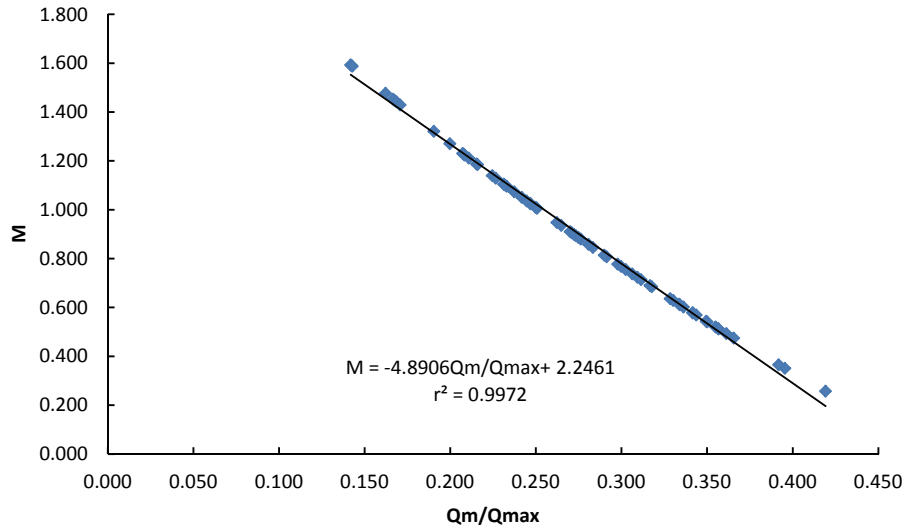
Figure 2. Relationship between *M* and the ratio of mean to maximum discharge.

The chart shows: M = -4.8906Qm/Qmax + 2.2461, r² = 0.9972

Table 1. Characteristics of 13 gauging stations.

Gauging Station	Drainage Area (km^2)	Years of Record	Length of Record (days)
02130561	19684	1990-2013	8202
02130840	134	2005-2012	2566
02130900	280	1959-2012	19362
02130910	448	1960-2012	18996
02130980	1114	2001-2013	4161
02131000	22870	1938-2011	26903
02131010	22922	1996-2011	5909
02131500	1748	1942-2013	14632
02132000	2668	1929-2013	30640
02135000	7226	1942-2013	26165
02135200	36519	2001-2011	9474
02136000	3243	1929-2013	30640
02136361	11	1993-2013	7126

GOODNESS-OF-FIT TESTS

Four goodness-of-fit tests were employed to evaluate the goodness of predicted FDCs:

Coefficient of correlation (r^2):

$$r^2 = \left\{ \frac{\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(Q_o(i)-\overline{Q}_o)(Q_c(i)-\overline{Q}_c)}}{\left[\frac{1}{N-1}\sum_{i=1}^{N}(Q_o(i)-\overline{Q}_o)\right]^{0.5}\left[\frac{1}{N-1}\sum_{i=1}^{N}(Q_c(i)-\overline{Q}_c)\right]^{0.5}} \right\}^2 \quad (30)$$

Average bias (Bias):

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}\left|[Q_o(i)-Q_c(i)]\right| \quad (31)$$

Root mean square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum[Q_o(i)-Q_c(i)]^2}{N-1}} \quad (32)$$

Nash-Sutcliffe efficiency (NSE):

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}[Q_o(i)-Q_c(i)]^2}{\sum_{i=1}^{N}[Q_o(i)-\overline{Q}_o]^2} \quad (33)$$

where $Q_o(i)$ is the ith observed discharge, $Q_c(i)$ is the ith computed discharge, $\overline{Q}_o$ and $\overline{Q}_c$ are the average values of observed and computed discharges, respectively, and N is the number of observations.

BEHAVIOR OF MEAN, MINIMUM, AND MAXIMUM FLOWS

For all 13 stations, the values of mean (Q_{mean}), minimum (Q_{min}), and maximum (Q_{max}) flows, and the ratio of mean flow to maximum flow were obtained for each year of record. For a sample station (02131000), figure 3 plots histograms of Q_{min}, Q_{mean}, and Q_{max}. The average values of minimum, mean, and peak flows for each station were computed and are tabulated in table 2. It can be seen that the discharge values for different stations show significant differences; for example, the mean value varies from 0.17 m^3 s^{-1} to greater than 283 m^3 s^{-1}. Ogden and Dawdy (2003) and Gupta et al. (2010) showed a power law relating peak discharge to drainage area. Thus, average values of minimum, mean, and peak flows were plotted against drainage area, as shown in figure 4. The log-log plots show a power relationship of Q_{min}, Q_{mean}, and Q_{max} values with drainage area, and the power law fitted well with a coefficient of determination around 0.9. Furthermore, the ratio of mean flow to maximum flow was also plotted versus the drainage area, but the relationship was found to be weak.

DETERMINATION OF LAGRANGE MULTIPLIERS

The Lagrange multipliers were computed by solving equations 10 and 11, which involve Q_{mean} and Q_{max}, and
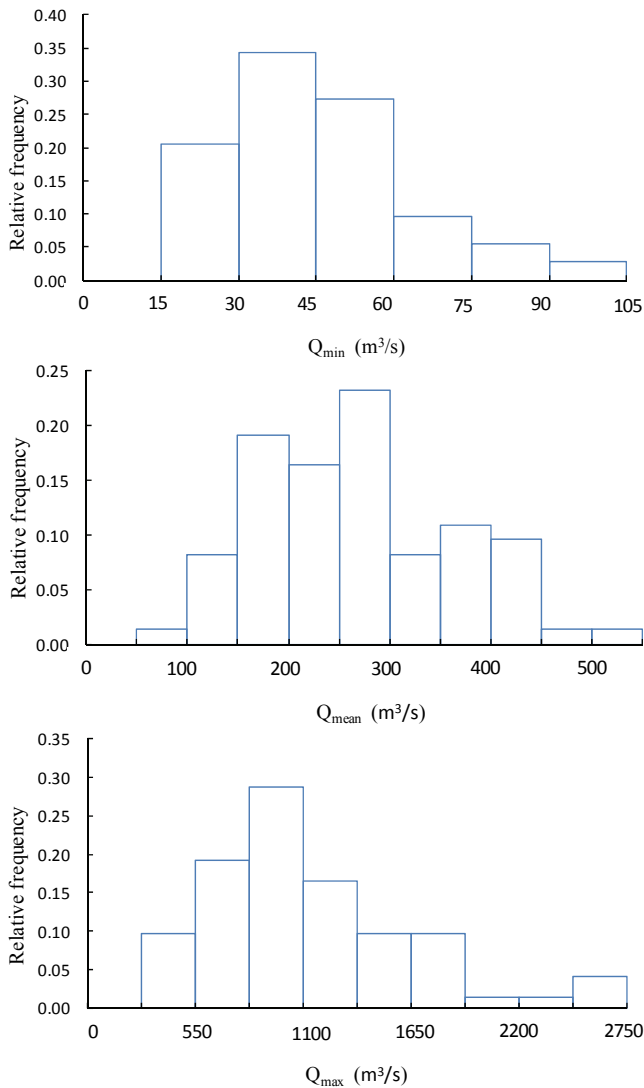
Figure 3. Frequency histograms of the Q_{min}, Q_{mean}, and Q_{max} values of station 02131000.

Table 2. Yearly flow statistics.

Gauging Station	$\overline{Q}_{mean}$ (m³ s⁻¹)	$\overline{Q}_{max}$ (m³ s⁻¹)	$\overline{Q}_{min}$ (m³ s⁻¹)	$\overline{Q}_{mean} / \overline{Q}_{max}$
2130561	187.41	498.08	77.83	0.376
02130840	1.13	2.97	0.40	0.381
02130900	3.76	14.26	0.96	0.264
02130910	5.35	17.01	1.67	0.314
02130980	9.45	22.50	4.08	0.420
02131000	301.17	1172.95	43.07	0.257
02131010	32.57	101.31	7.08	0.322
02131500	18.28	91.13	4.47	0.201
02132000	25.98	140.09	5.15	0.185
02135000	76.81	357.46	12.82	0.215
02135200	219.61	551.85	189.30	0.398
02136000	24.06	194.14	1.84	0.124
02136361	0.17	1.36	0.03	0.125

were plotted against Q_{mean} and Q_{max}, as shown in figure 5. Figure 5a shows the variation of λ_* with mean discharge for different values of λ_1, and figure 5b shows the variation of
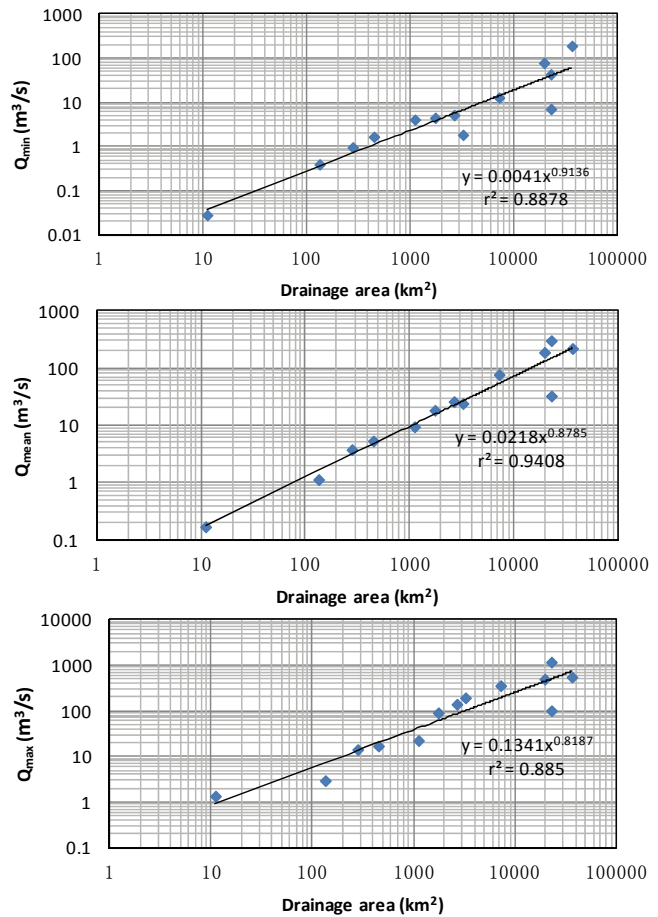
Figure 4. Power relationships between Q_{min}, Q_{mean}, and Q_{max} versus drainage area.

λ_1 with maximum discharge. It can be seen from the figures that the Lagrange multipliers retain the same sign. For positive values, λ_* decreases with increasing Q_{mean} and λ_1 decreases with increasing Q_{max}, while for negative values their behavior is opposite. Comparing the two figures, it is seen that λ_* has a wider distribution than λ_1; furthermore, λ_1 drops quickly to below the value of 0.1. The relations between the two Lagrange multipliers for different values of maximum discharge and mean discharge are shown in figures 5c and 5d. In both figures, λ_* increases with λ_1 but with different slopes. The slope is milder for $Q_{mean} = 10$ and 100 m³ s⁻¹ but much faster for $Q_{mean} = 500$ and 1000 m³ s⁻¹. The Lagrange multiplier λ_* also increases with Q_{mean} or Q_{max}. For constant λ_1, λ_* is larger for higher Q_{mean} or Q_{max}.

Histograms of Lagrange multipliers were also computed, as plotted in figure 6, which shows that the value of λ_1 is highly concentrated within the values between 0 and 0.025, whereas λ_* is distributed widely. The Lagrange multipliers computed for each station are listed in table 3. The mean values for all basins obtained for λ_1 and λ_* were respectively, 0.012 and 0.175, with standard deviations of 0.051 and 0.167, and skewness coefficients of 2.957 and 1.712.
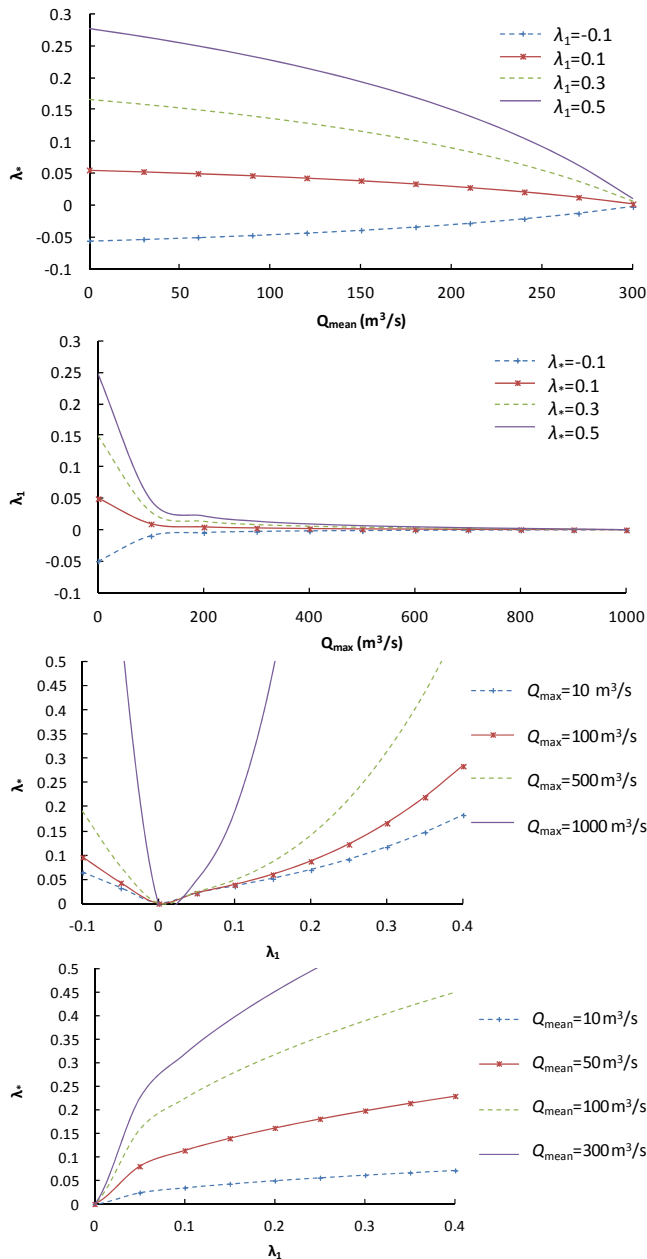
Figure 5. Behavior of Lagrange multipliers for different values of Q_{mean} and Q_{max}.
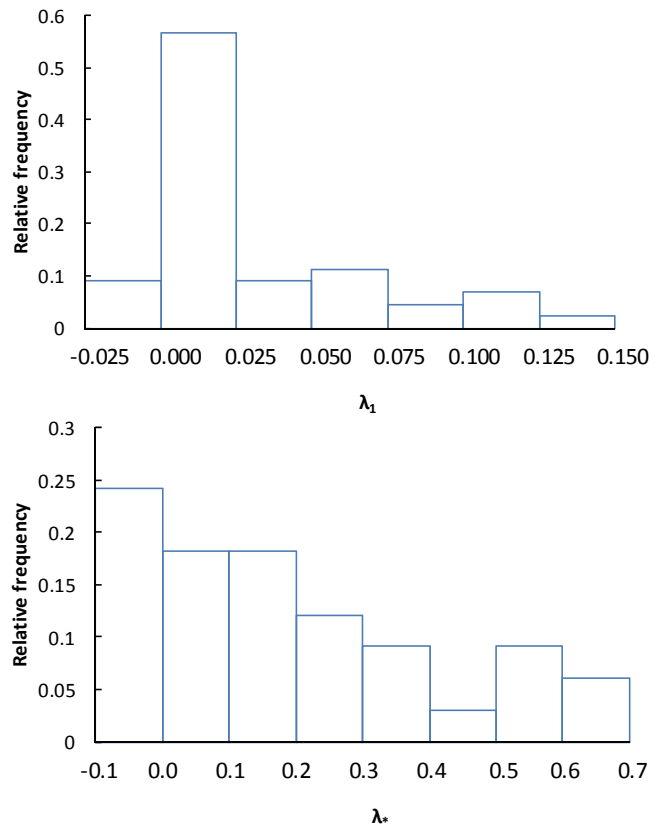

Figure 6. Histograms of Lagrange multipliers.

Table 3. Lagrange multipliers and parameter M for each station.

| Gauging Station	λ_1	λ_*	M Mean	SD	CV
02130561	0.001	0.248	0.406	0.439	1.081
02130840	0.009	-0.042	0.383	0.446	1.166
02130900	0.012	-0.008	0.955	0.270	0.283
02130910	0.001	-0.007	0.708	0.342	0.483
02130980	0.001	-0.005	0.191	0.513	2.683
02131000	-0.050	0.585	0.990	0.261	0.263
02131010	-0.002	0.097	0.674	0.363	0.524
02131500	0.006	0.118	1.265	0.190	0.150
02132000	0.004	0.14	1.339	0.172	0.128
02135000	0.003	0.177	1.195	0.207	0.173
02135200	0.001	0.274	0.774	0.452	1.584
02136000	0.001	0.095	1.640	0.105	0.064
02136361	0.173	0.091	1.635	0.106	0.065
Average	0.012	0.175	0.798	0.293	0.665

COMPUTATION OF M

To estimate the flow duration curve for a given year, the first step was to compute the M value from the given values of mean and maximum discharges. It is noted that equation 23 was derived by assuming $Q_{min} = 0$; however, Q_{min} is not small enough to be neglected, as seen from table 2. Thus, the modified discharge using $Q' = Q - Q_{min}$ was preferred to compute M. The M value was computed using equation 23 for each year of record, and then the mean, standard deviation, and coefficient of variance of the computed values were calculated, as shown in table 3. The M values computed from equation 23 had different ranges for each station, and a histogram of M was plotted for each station, as shown in figure 7 for three sample stations. The histograms varied from one station to another, but seemed to fit

the normal distribution in all cases. In general, M varied from 0.2 to 1.6, and its standard deviation was around 0.2 to 0.4. Combining values of five stations, a histogram of the M values was plotted, as shown in figure 8. Again, M seemed to follow a bell-shaped distribution, with a mean value of 0.798 and a standard deviation of 0.493. However, when computing the flow duration curve for a given station, the M value listed in table 3 was used. The average values of M were also plotted against the drainage area, but the relationship was weak.

COMPUTATION OF FLOW DURATION CURVE FOR GAUGED STATIONS

After obtaining M, the next step was to calculate the flow discharge using equation 21. Figure 9 shows the com-
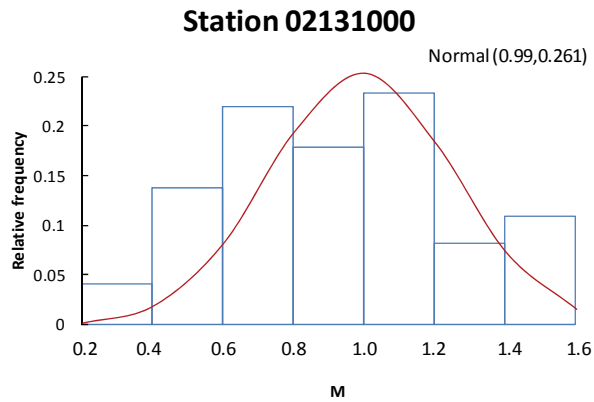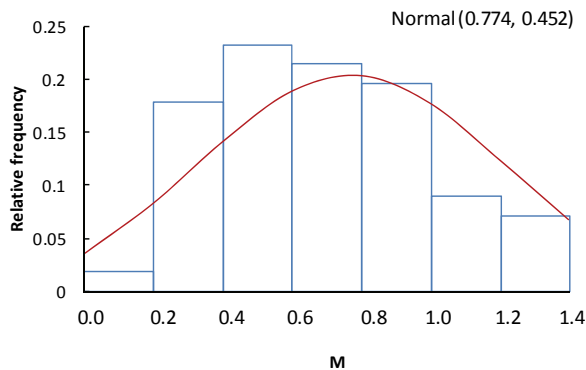
Station 02131000
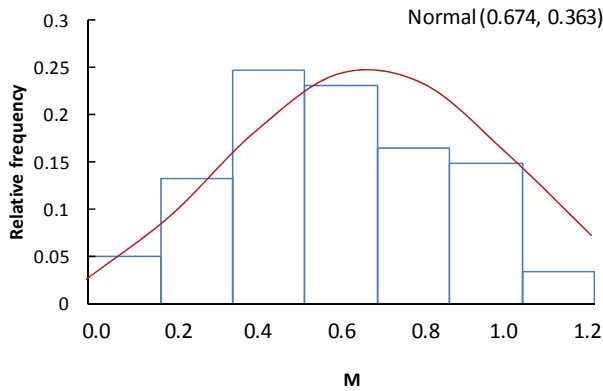
Station 02131010

Station 02135200

Figure 7. Histograms of *M* values for three stations.

Figure 8. Histogram of *M* values for combined stations.

Figure 9. Flow duration curves computed for three sample stations: (a) station 02131000, (b) station 02135210, and (c) station 02135200.

puted flow duration curves for three selected stations for the year 2006. For computation, the *M* values listed in table 3 were used for each station. The Tsallis entropy-based flow duration fitted the observations closely but had about 9% error in estimation as compared with maximum observed values. For station 02135210, the estimated FDC during the first $0.3t/T$ deviated from the observed curve but was still within the 95% confidence interval, as seen in figure 9b. All the computed values fell inside the 95% confidence interval for all stations. The RMSE values for computed FDCs for the three sample stations in figure 9 were 12.51, 22.17, and 10.43 m^3 s^{-1}, respectively.
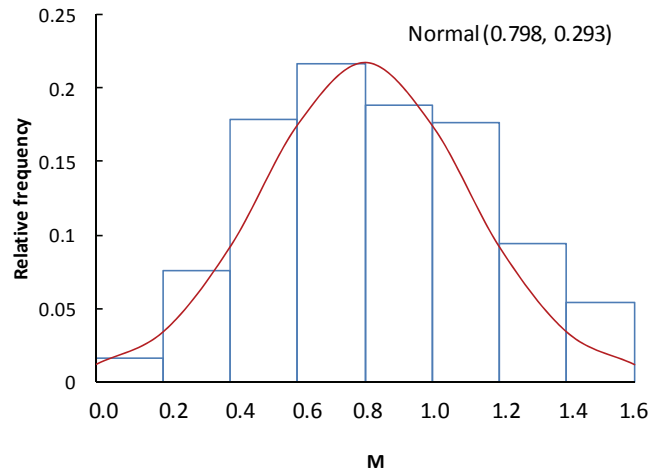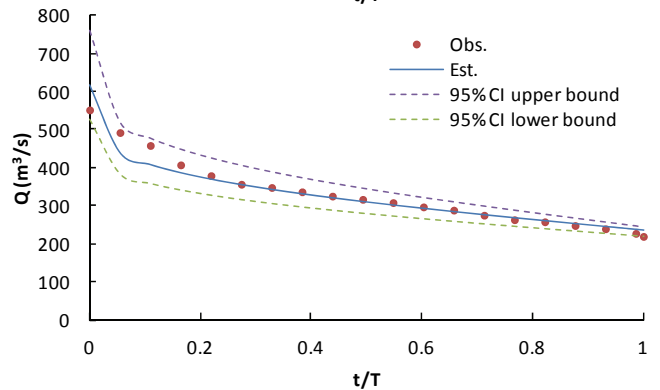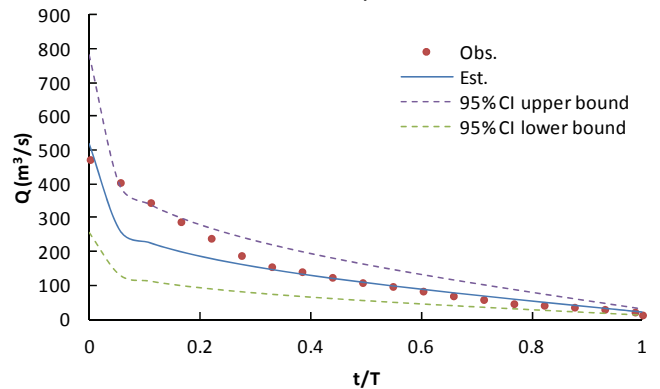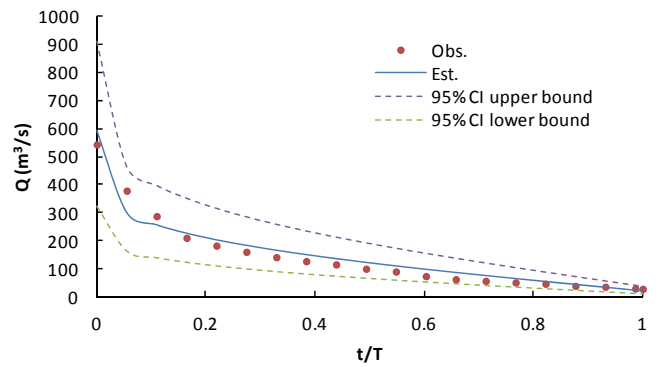
Table 4. Performance measures for FDCs for different gauging stations for year 2006.

Gauging Station	r^2	Bias ($m^3 s^{-1}$)	RMSE ($m^3 s^{-1}$)	NSE	Entropy
02130561	0.922	10.471	12.855	0.986	0.29
02130840	0.969	3.135	0.567	0.964	0.34
02130900	0.913	0.328	0.213	0.941	0.08
02130910	0.947	0.317	0.128	0.967	0.10
02130980	0.958	0.607	0.050	0.988	0.13
02131000	0.972	41.310	28.675	0.914	0.29
02131010	0.964	6.631	2.500	0.974	0.31
02131500	0.913	3.815	4.329	0.953	0.34
02132000	0.970	5.274	6.854	0.941	0.23
02135000	0.915	8.873	8.884	0.952	0.39
02135200	0.951	1.069	1.438	0.924	0.28
02136000	0.911	7.575	18.047	0.930	0.35
02136361	0.957	3.340	7.862	0.958	0.31

COMPUTATION OF GOODNESS-OF-FIT VALUES

For all 13 stations, the entropy-based flow duration curves compared well with the observed flow duration curves. The coefficient of correlation, average bias, root mean square error, and Nash-Sutcliffe efficiency were computed for the flow duration curves for all 13 stations and are given in table 4. For the flow duration curves of the five gauging stations using the Tsallis entropy, the r^2 varied from 0.986 to 0.994, Bias varied from 0.32 to 41.31 $m^3 s^{-1}$, RMSE varied from 0.05 to 28.67 $m^3 s^{-1}$, and NSE varied from 0.932 to 0.971. In general, the Tsallis entropy-based flow duration curves were in agreement with the observed curves, as reflected by the statistical measures given in table 4. In particular, the coefficient of correlation and Nash-Sutcliffe efficiency were very close to 1, which showed the goodness of fit.

The entropy values given in table 4 reflect the relative goodness of fit. The entropy values were quite comparable, and agreements between the observed and computed flow duration curves were also comparable. Since the differences between entropy values for different curves were small, it was difficult to judge the fit based on the entropy values alone. The lower values reflect less uncertainty, meaning a better fit, and this was observed somewhat.

COMPARISON WITH FDCs COMPUTED USING SHANNON ENTROPY

The flow duration curves based on the Tsallis entropy were compared with the flow duration curves computed using the Shannon entropy. The Shannon entropy can be considered as a simplification of the Tsallis entropy, where $m = 1$. The flow duration curve estimated by the Shannon entropy was shown to represent the observation well (Singh et al., 2013). Thus, one may be interested to determine whether the Tsallis entropy will be comparable to the Shannon entropy. Figure 10 compares the two entropy-based flow duration curves and the observed FDCs. It can be seen from the figure that the Tsallis entropy-based curves fit the observations a little better. The Shannon entropy-based flow duration curves had a larger estimation error. For the computation in figure 10a, the RMSE decreased from 41.25 $m^3 s^{-1}$ for the Shannon entropy to 12.51 $m^3 s^{-1}$ for the Tsallis entropy.
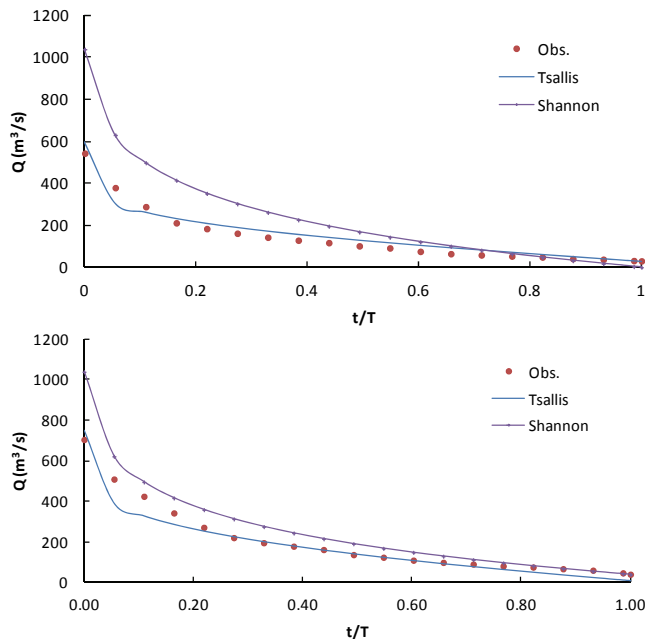

Figure 10. Comparison of flow duration curves computed using Shannon entropy and Tsallis entropy of (a) year 2006, and (b) year 2008 for station 02131000.

PREDICTION OF FLOW DURATION CURVE FOR UNGAUGED STATIONS

The FDC can be forecast ahead of time for a given station, once the entropy parameter has been determined. For a sample gauging station (02131000), the observed data for the years 1938 to 2006 were used as past information, from which the distributions of Q_{min}, Q_{mean}, and Q_{max} as well as M were obtained. The data from 2007 to 2011 were used for forecasting. To predict the flow duration curve, Q_{min}, Q_{mean}, and Q_{max} need to be forecasted. Since the future values of peak, minimum, and mean discharge are subject to uncertainty, they can only be predicted for given probability values. However, instead of using time series forecasting, Q_{min}, Q_{mean}, and Q_{max} of 1, 2, 10, and 50 year recurrence intervals were computed from the information given by the series 1938 to 2006. For example, the Q_{max} values of the 1, 2, 10, and 50 year recurrence intervals for station 02131000 were 441.7, 1039.2, 1915.9, and 2732.2 $m^3 s^{-1}$, respectively. The M values were then computed, and flow duration curves were constructed for the different recurrence intervals. The predicted flow duration curves of 1, 2, and 10 year recurrence intervals are shown in figure 11 with observations for the years 2009 and 2011. It can be seen from figure 11b that the observed FDC for 2011 is in close agreement with the predicted one-year recurrence interval FDC, whereas the observed FDC for 2009 (fig. 11a) is in close agreement with the predicted two-year recurrence interval. The result were even closer to the actual recurrence intervals, since it was found that the FDC for 2009 was a 2.2-year value, and the FDC for 2011 was a 1.3-year value.

Based on the power relationship between discharge and drainage area, the proposed method can also used to predict the FDC of a nearby basin without any information on historical discharge. For example, assume that there are no discharge records for stations 02135000 and 02130840,
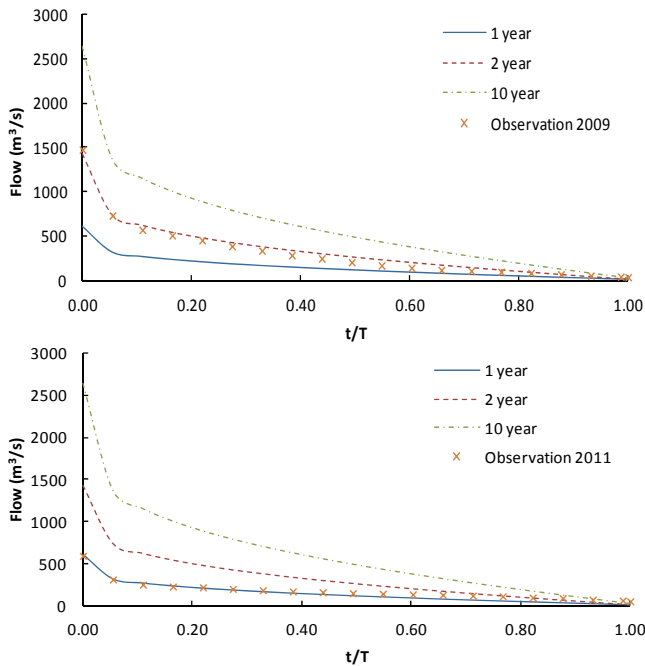
Figure 11. Prediction of FDC for gauging station 02131000 for (a) year 2009 and (b) year 2011.
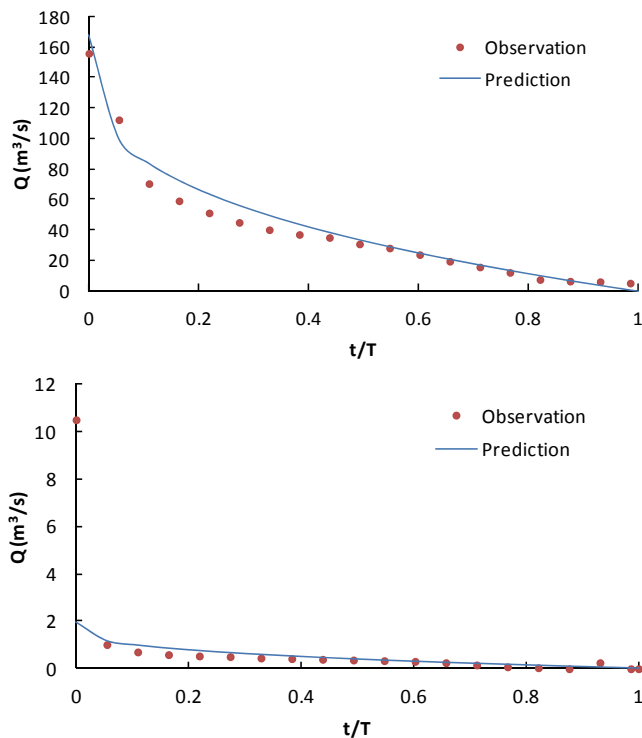

Figure 12. Prediction of FDC for (a) station 02135000 and (b) station 02130840 without using observations.

although we know that their drainage areas are 7726 km^2 and 134 km^2. Therefore, Q_{min}, Q_{mean}, and Q_{max} were computed using regression, as shown in figure 4. Once Q_{mean} and Q_{max} were known, M was computed as before, using equation 23, which produced M values of 0.243 and 0.407 for these two stations. Thus, the FDC was predicted using the above values and compared with the observations for year 2011, as shown in figure 12. It can be seen from the
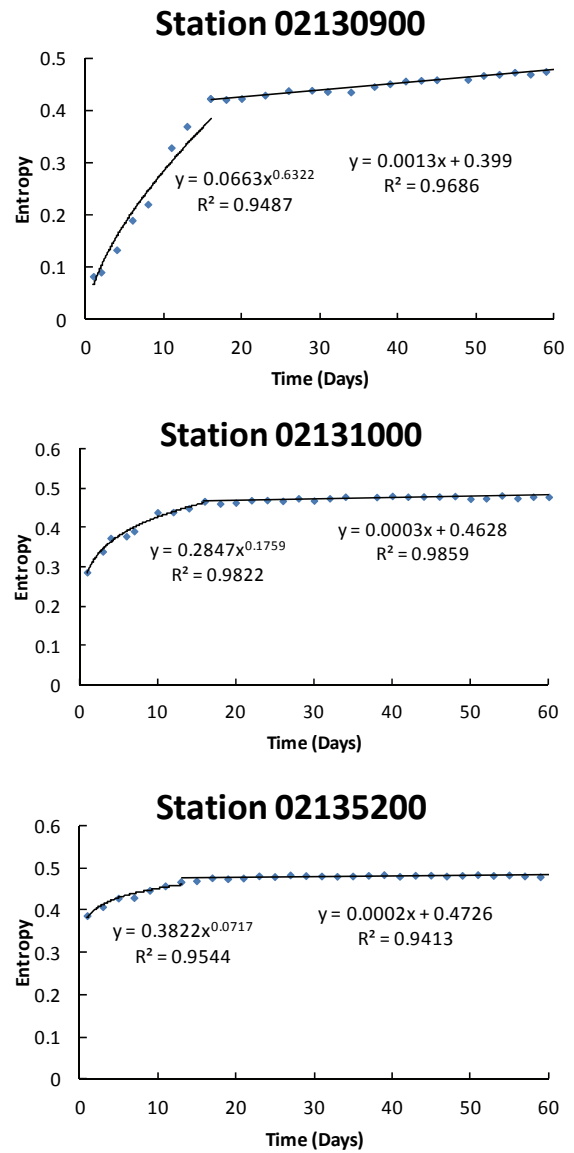

Figure 13. Variation of entropy with time scale.

figure that the predicted FDC is in close agreement with the observed FDC, and the RMSE values were only 3.39 and 0.86 m^3 s^{-1}. However, this method did not accurately predict the maximum value in figure 12b, where the observed value was unusually large.

SENSITIVITY OF ENTROPY

The flow duration curves discussed in the previous section were computed by using daily streamflow values. Flow duration of more than one day of streamflow can also be estimated using the same processes. The sensitivity of entropy to the computation interval was evaluated, as shown for three sample stations in figure 13. The entropy was computed by equation 1b from the whole given series for all 13 stations. It can be seen from the sample figure that the entropy increases with increasing time interval. However, the rate of increase is high during the first phase for about 15 days, but as the interval increases the rate significantly decreases. After about 15 days, the rate of entropy

Table 5. Exponent and interval for each station.

Gauging Station	Exponent	Interval (days)
02130561	0.10	8
02130840	0.25	17
02130900	0.63	16
02130910	0.31	11
02130980	0.28	14
02131000	0.18	16
02131010	0.20	15
02131500	0.12	10
02132000	0.43	13
02135000	0.25	15
02135200	0.07	13
02136000	0.14	12
02136361	0.21	15

increase declines significantly, and then entropy almost reaches a constant value. This suggests that the flow regime becomes more complex at higher intervals, reflecting a reduced influence of anthropogenic changes on the flow regime. The opposite is true for smaller intervals.

The entropy of the first phase was fitted by a power law. The exponents for the first part were 0.63, 0.18, and 0.07 for stations 02130900, 02131000, and 02135200, respectively. For all 13 stations, the exponent averaged 0.24, as shown in table 5. The entropy of the second phase was fitted by a linear equation, where the slope was less than 0.0005. Although the change points between two phases were different for different years, all occurred around the 15th day; for example, the change points were 16, 13, and 16 days for stations 02130900, 02131000, and 02135200, respectively.

CONCLUSIONS AND DISCUSSION

This article provides a new method for estimating the flow duration curve by the Tsallis entropy. For any watershed, researchers can apply this method with knowledge of the mean, minimum, and maximum streamflow values. With the steps discussed in this article, researchers can obtain the parameter M from the mean and maximum streamflow data, which can then be applied to estimate streamflow. In addition, mean and maximum streamflow values of ungauged stations can be estimated using the relationship between the mean or maximum streamflow values and drainage area. In this way, flow duration can be also estimated for ungauged stations.

The following conclusions are therefore drawn from this study: (1) With the entropy method, the parameters of flow duration curves can be determined in terms of specified constraints, which themselves are determined from observations. This obviates the need for fitting. (2) For all stations in this study, the computed flow duration curves were found to represent the observed duration curves satisfactorily. The Tsallis entropy-based predicted flow duration curves fit better than the Shannon entropy-based curves and have less estimation error. (3) The flow duration curves can be forecasted using the entropy method, and the prediction is consistent with actual flow recurrence intervals. (4) The proposed method also performs satisfactorily in forecasting the FDC at nearby stations using the power relationship between discharge and drainage area. (5) The sensitivity of

entropy was examined for different intervals, and the rate of entropy increase was consistent for each station.

The benefits of an empirical flow duration model are quite evident for water planners but extend further to include water quality planning. For example, the U.S. Army Corps of Engineers performs significant dredging operations to keep waterways free of sediment. A model of flow duration curves provides a path to modeling stream power, sediment input, and sediment movement based on various return periods of total flows over a year, facilitating planning of dredging needs and costs.

REFERENCES

Al-Hamdan, O. Z., & Cruise, J. F. (2010). Soil moisture profile development from surface observations by principle of maximum entropy. *J. Hydrol. Eng., 15*(5), 327-337. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000196.

Barbe, D. E., Cruise, J. F., & Singh, V. P. (1991). Solution of three-constraint entropy-based velocity distribution. *J. Hydraul. Eng., 117*(10), 1389-1396. http://dx.doi.org/10.1061/(ASCE)0733-9429(1991)117:10(1389).

Chiu, C. L. (1987). Entropy and probability concepts in hydraulics. *J. Hydraul. Eng., 113*(5), 583-600. http://dx.doi.org/10.1061/(ASCE)0733-9429(1987)113:5(583).

Cui, H., & Singh, V. P. (2012). On the cumulative distribution function for entropy-based hydrologic modeling. *Trans. ASABE, 55*(2), 429-438. http://dx.doi.org/10.13031/2013.41384.

Cui, H., & Singh, V. P. (2013a). One-dimensional velocity distribution in open channels using the Tsallis entropy. *J. Hydrol. Eng., 19*(2), 290-298. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000610.

Cui, H., & Singh, V. P. (2013b). Two-dimensional velocity distribution in open channels using the Tsallis entropy. *J. Hydrol. Eng., 18*(3), 331-339. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000610.

Foster, H. (1934). Duration curves. *Trans. ASCE, 99*, 1213-1267.

Gupta, V. K., Mantilla, R., Troutman, B., Dawdy, D., & Krajewski, W. F. (2010). Generalizing a nonlinear geophysical flood theory to medium-sized river networks. *Geophysical Res. Lett., 37*(11), L11402. http://dx.doi.org/10.1029/2009GL041540.

Hao, L., & Singh, V. P. (2011). Entropy theory for two-dimensional velocity distribution. *J. Hydrol. Eng., 16*(4), 303-315. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000319.

Hickox, G. H., & Wessenauer, G. O. (1933). Application of duration curves to hydroelectric studies. *Trans. ASCE, 98*, 1276-1308.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Rev., 106*(4), 620-630. http://dx.doi.org/10.1103/PhysRev.106.620.

Jaynes, E. (1982). On the rationale of maximum entropy methods. *Proc. IEEE, 70*(9), 939-952. http://dx.doi.org/10.1109/PROC.1982.12425.

Miller, C. (1951). Analysis of flow duration, sediment-rating curve method of computing sediment yield. Denver, Colo.: U.S. Department of Interior, Bureau of Reclamation.

Mitchell, W. D. (1957). Flow-duration of Illinois streams. Springfield, Ill.: Illinois Department of Public Works and Buildings, Division of Waterways.

Ogden, F. L., & Dawdy, D. R. (2003). Peak discharge scaling in small Hortonian watersheds. *J. Hydrol. Eng., 8*(2), 64-73. http://dx.doi.org/10.1061/(ASCE)1084-0699(2003)8:2(64).

Searcy, J. K. (1959). Flow-duration curves. USGS Water-Supply Paper 1542-A. Reston, Va.: U.S. Geological Survey.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Tech. J., 27*(3), 379-423. http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Singh, V. P. (1998). *Entropy-Based Parameter Estimation in Hydrology*. Boston, Mass.: Kluwer Academic. http://dx.doi.org/10.1007/978-94-017-1431-0.

Singh, V. P. (2010a). Entropy theory for derivation of infiltration equations. *Water Resources Res., 46*(3), 1-20. http://dx.doi.org/10.1029/2009WR008193.

Singh, V. P. (2010b). Entropy theory for movement of moisture in soils. *Water Resources Res., 46*(3), 1-12. http://dx.doi.org/10.1029/2009WR008288.

Singh, V. P. (2010c). Tsallis entropy theory for derivation of infiltration equations. *Trans. ASABE, 53*(2), 447-463. http://dx.doi.org/10.13031/2013.29585.

Singh, V. P. (2013). *Entropy Theory in Environmental and Water Engineering*. New York, N.Y.: John Wiley and Sons. http://dx.doi.org/10.1002/9781118428306.

Singh, V. P., & Hao, L. (2011). Entropy theory for distribution of one-dimensional velocity in open channels. *J. Hydrol. Eng., 16*(9), 725-735. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000363.

Singh, V., Byrd, A., & Cui, H. (2013). Flow duration curve using entropy theory. *J. Hydrol. Eng.* http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000930.

Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *J. Statistical Physics, 52*(1-2), 279-487.

U.S. EPA. (2007). An approach for using load duration curves in the development of TMDL. EPA-841-B-07-006. Washington, D.C.: U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds.