Derivation of Rating Curves Using Entropy Theory

V. P. Singh

Abstract. *Using the entropy theory, this study derives the stage-discharge relation, often called the rating curve, which is based on two simple constraints: (1) the total probability and (2) the mean logarithmic discharge. Parameters of the derived curves are determined with the use of these two constraints. The entropy theory permits a probabilistic characterization of the rating curve and hence the probability density function underlying the curve. It also permits a quantitative assessment of the uncertainty of the rating curve. The derived rating curves are tested using field data and are found to be in agreement with the curves obtained by the least square method.*

Keywords. *Entropy, Principle of maximum entropy, Rating curve, Shannon entropy, Stage-discharge relation.*

The relation between stage and discharge is often referred to as the stage-discharge rating curve or simply the rating curve. It is a kinematic relation (Singh, 1993, 1996). In principle, a rating curve is a relation between flux (usually volumetric) and concentration related to a river or stream. In the case of a stage-discharge relation, the volumetric flux is represented by flow discharge (cubic meters per second) and its corresponding flow concentration by flow depth or stage above a datum. In the case of a sediment rating curve, the concentration can be represented by suspended sediment concentration. Thus, the relation between suspended sediment concentration and discharge defines a suspended sediment rating curve. Similarly, if concentration is represented by pollutant concentration, then the relation between pollutant concentration and discharge corresponds to the pollutant rating curve. Thus, there are different types of rating curves used in hydrology. Since the rating curves are of similar form from an algebraic viewpoint, and fundamental to most rating curves is the estimation of discharge, this study focuses on the stage-discharge rating curve only. Rating curves are employed for a variety of purposes, including the determination of discharge for measured stage, calibration of physically based hydraulic and hydrologic models, evaluation of flood inundation, and damage assessment. They are also used for constructing continuous records of discharge, continuous time series of sediment discharge or sediment concentration, continuous pollutant graphs, floodplain mapping, storage variation, hydraulic design, catchment routing, damage assessment, and so on (WMO, 1994).

A common method used to construct a rating curve for a gauging site is to plot observed discharge and stage data on graph paper and fit an equation using regression or a least square method. Depending on the river gauging site and the underlying geomorphic controls, three types of rating curves, which are parabolic or power type, have been derived and are employed in practice (Corbett, 1962; Lambie, 1978). Parameters of these curves are determined either graphically or by using a mathematical or statistical method, such as least square method, maximum likelihood, pseudo-maximum likelihood, segmentation, artificial neural networks, simulated annealing, and genetic programming (Lee et al., 2010).

Ideally, the rating curve should be a smooth curve of parabolic shape, without reversals in curvature, so that a unique relation between stage and discharge is easily established. In the absence of an abrupt change in the slope of the rating curve, the rate of increase in stage corresponding to a specified increase in discharge should be reasonably consistent throughout the stage. In the case of the relation for a channel, the channel must be capable of regulating or stabilizing the flow past the gauge such that for a given stage or height of the water surface, the discharge past the gauge must remain unaltered. The relation is controlled by a section or reach of the channel below the gauge, known as the section control. It eliminates the influence of all other boundary conditions on the velocity of flow at the stage. A control may be complete or partial. A complete control, as the name suggests, governs the stage-discharge relation for the full range of stage and is independent of all downstream conditions. On the other hand, a partial control governs the relation for only part of the range in stage. It may act in concert with other controls and be an essential part of the complete control (Corbett, 1962; Petersen-Overleir and Reitan, 2005).

However, a rating curve is often subject to a number of uncertainties, as discussed by Clarke (1999): (1) There are inherent errors in discharge measurements due to the computation of average velocity, relative accuracy of the flow meter equipment, and the experimenters (Sauer and Meyer, 1992). (2) Since discharge is a product of cross-sectional area and flow velocity and cross-sectional area is computed using depth and width, determining the discharge at a place where velocity variations are kept to a minimum or selecting a stable river cross-section, which is usually associated with a high depth to width ratio and relatively little erosion and deposition of sediment, is not always easy. (3) The section control must be capable of maintaining a fairly stable

Submitted for review in April 2010 as manuscript number SW 8543; approved for publication by the Soil & Water Division of ASABE in September 2010.

The author is **Vijay P. Singh, ASABE Member Engineer,** Caroline and William N. Lehrer Distinguished Chair in Water Engineering, and Professor, Department of Biological and Agricultural Engineering, and Professor, Department of Civil and Environmental Engineering, Texas A&M University, 321 Scoates Hall, 2117 TAMU, College Station, Texas 77843-2117; phone: 979-845-3931; e-mail: vsingh@tamu.edu.

  ISSN 2151-0032

relation between discharge and water stage at the selected point above it. If the channel control is comprised of channel slope, resistance, and dimensions over a considerable distance, which varies inversely with slope and increases with increasing stage, and if the distance increases farther downstream and includes controls of new downstream features, then the curvature of the rating curve may exhibit reversals. If a channel has a flat slope, then the control at high stages may extend so far downstream that it may involve backwater effects, which may not occur at lower stages. Abrupt changes in controls and submergence of controls most likely cause irregularities in the slope of the stage-discharge relations. (4) In order for a stage-discharge relation to be stable for a given discharge, both relations of slope to stage and slope to discharge must remain unaltered. This will correspond to a complete control in its effectiveness. These relations would be constant for steady-state conditions, which do not occur often. Nevertheless, the sites for the position of gauges and controls must be selected such that the variation in discharge for a given stage, due to variations in slope, velocity, or channel conditions, is small during the period of time involved. (5) A permanent control ensures a permanent stage-discharge relation at all times as long as slope remains the same. For a permanent control, the position with respect to the datum of the gauge, its distance downstream from the gauge, and the condition of the streambed between the gauge and the part of the channel controlling the stage-discharge relation must remain unchanged. However, in real life, these conditions are seldom met, and a permanent control and consequent permanent stage-discharge relation do not remain unchanged. Even if the control may seem permanent, the stage-discharge relation may change. (6) There can be more than one control for high and low flows (Yoo and Park, 2010).

Herschy (1995) investigated errors in discharge due to errors in velocity and depth measurements. Considering channel instability, DeGagne et al. (1996) developed a decision support system. Because there can a change in control from low flow to high flow, a segmentation method, which is developed in segments, has been used to represent a rating curve (Torsten et al., 2002; Schmidt and Yen, 2002; Petersen-Overleir, 2004; Petersen-Overleir and Reiten, 2005; Sivapragasam and Mutill, 2005; Overlier, 2006). This discussion shows that there is an element of randomness in the stage-discharge curve, and it will therefore be reasonable to argue that discharge can be treated as a random variable. Although significant temporal variability in discharge has been recognized, little effort has been made to account for its probabilistic characteristics when establishing rating curves and to quantify uncertainty in a rating curve. One way to accomplish the twin objectives of defining the probability distribution and the uncertainty of a rating curve is to use the entropy theory. This theory has an advantage over other methods in several respects. First, it takes account of the information available on the rating curve, such as moments (mean, variance, etc.) of discharge. These moments are more stable in time than individual measurements. Second, it permits us to quantify the information or uncertainty associated with the curve. Third, it paves the way to determine data sampling or the number of measurements needed to determine a robust rating curve. Fourth, it obviates the need for estimating the rating curve parameters empirically or by curve fitting. Fifth, since the parameters estimated by the entropy theory are expressed in terms of the specified constraints, they have physical meaning or they can be interpreted in terms of the given information. These considerations motivated the use of the entropy theory.

The objective of this study therefore is to (1) derive, using the entropy theory, the stage-discharge rating curve dominated by friction control and its three special forms, (2) determine the rating curve parameters from the specified information expressed as constraints, (3) derive the probability density function associated with the three rating curves, (4) determine the entropy associated with these curves, and (5) test the three rating curves with field measurements.

Forms of Rating Curves

A rating curve for a gauge in a channel dominated by friction is normally expressed in a power form (Corbett, 1962; Kennedy, 1964) as:

$$Q = a(y - y_0)^b + c \tag{1}$$

where Q is the discharge (L^3/T, e.g., $ft^3\ s^{-1}$ or $m^3\ s^{-1}$); y is the stage or height of water surface (L, e.g., ft or m); y_0 is the height (L) when discharge is negligible and is usually taken as a constant value or is sometimes used as a fitting parameter; b is an exponent; and a (L^{3-b}/T) and c (L^3/T) are parameters, where L is the length dimension and T is the time dimension. Equation 1 is a general form and specializes into three popular types that are commonly employed (Corbett, 1962). Although the three forms differ from each other in their parameters, these forms have been popularly used and reported separately in the hydraulic literature and stem from river morphological characteristics. Therefore, they are described as such.

Type 1: In this case, $y_0 = 0$ and $c = 0$. Equation 1 then becomes:

$$Q = ay^b \tag{2a}$$

or in logarithmic form:

$$\log Q = \log a + b \log y \tag{2b}$$

Type 2: In this case, $c = 0$. Equation 1 then becomes:

$$Q = a(y - y_0)^b \tag{3a}$$

or in logarithmic form:

$$\log Q = \log a + b \log(y - y_0) \tag{3b}$$

Type 3: In this case, $y_0 = 0$. Equation 1 then becomes:

$$Q = ay^b + c \tag{4a}$$

or in logarithmic form:

$$\log(Q - c) = \log a + b \log y \tag{4b}$$

It should be noted that values of parameters a, b, and c will vary from one relation to another. In the hydraulic literature, equations 2 to 4 have been applied. The objective here is to derive these relations using the entropy theory.

Derivation of Rating Curves Using Entropy Theory

The procedure for deriving rating curves, based on the entropy theory, is comprised of (1) Shannon entropy, (2) specification of constraints, (3) maximization of entropy in concert with the principle of maximum entropy, (4) derivation of the probability distribution of discharge, (5) maximum entropy, (6) determination of Lagrange multipliers, and (7) derivation of the rating curve. It is assumed that temporally averaged discharge Q is a random variable with a probability density function (PDF) denoted as $f(Q)$. Each of these steps is discussed in the following sections.

Shannon Entropy

The Shannon entropy (Shannon, 1948; Shannon and Weaver, 1949) of discharge Q or of $f(Q)$, $H(Q)$, can be expressed as:

$$H(u) = -\int_{Q_0}^{Q_D} f(Q)\,\ln f(Q)\,dQ \tag{5}$$

where Q_0 and Q_D represent the lower and upper limits of discharge for integration. Equation 5 expresses a measure of uncertainty about $f(Q)$ or the average information content of sampled Q. The objective here is to derive the least-biased $f(Q)$, which can be accomplished by maximizing $H(Q)$, subject to specified constraints, in accordance with the principle of maximum entropy (POME) (Jaynes, 1957, 1982). Maximizing $H(Q)$ is equivalent to maximizing $f(Q)\ln f(Q)$. In order to determine $f(Q)$ that is least biased toward what is not known as regards discharge, the principle of maximum entropy (POME) developed by Jaynes (1957, 1982) is invoked, which requires specification of certain information, expressed in terms of what is called constraints, on discharge. According to POME, the most appropriate probability distribution is the one that has the maximum entropy or uncertainty, subject to these constraints.

Specification of Constraints

For deriving the stage-discharge relation, following Singh (1998) the constraints to be specified are the total probability law, which must always be satisfied by the probability density function of discharge written as:

$$C_1 = \int_{Q_0}^{Q_D} f(Q)dQ = 1 \tag{6}$$

and

$$C_2 = \int_{Q_0}^{Q_D} \ln Q f(Q)dQ = \overline{\ln Q} \tag{7}$$

Equation 7 is the mean of the logarithmic discharge values. Equation 6 is the first constraint defining the total probability law (C_1), and equation 7 is the second constraint (C_2) defining the mean of the logarithm of discharge.

Maximization of Entropy

In order to obtain the least biased probability density function of Q, $f(Q)$, the Shannon entropy, given by equation 5, is maximized following POME, subject to equations 6 and 7. To that end, the method of Lagrange multipliers is employed. The Lagrangian function then becomes:

$$L = -\int_{Q_0}^{Q_D} f(Q)\ln f(Q)dQ - (\lambda_0 - 1)\left(\int_{Q_0}^{Q_D} f(Q)dQ - C_1\right) - \lambda_1\left(\int_{Q_0}^{Q_D} \ln Q f(Q)dQ - C_2\right) \tag{8}$$

where λ_0 and λ_1 are the Lagrange multipliers. Recalling the Euler-Lagrange equation of calculus of variation and differentiating equation 8 with respect to f, noting that f is variable and Q is a parameter, and equating the derivative to zero, one obtains:

$$\frac{\partial L}{\partial f} = 0 = -[\ln f(Q) + 1] - (\lambda_0 - 1) - \lambda_1 \ln Q \tag{9}$$

Derivation of Probability Distribution

Equation 9 leads to the entropy-based probability density function (PDF) of velocity as:

$$f(Q) = \exp[-\lambda_0 - \lambda_1 \ln Q]$$
$$\text{or } f(Q) = \exp(-\lambda_0) Q^{-\lambda_1} \tag{10}$$

The PDF of Q contains the Lagrange multipliers λ_0 and λ_1, which can be determined using equations 6 and 7. The cumulative probability distribution function of Q can be obtained by integrating equation 10 as:

$$F(Q) = \frac{\exp(-\lambda_0)}{-\lambda_1 + 1}\left[Q^{-\lambda_1+1} - Q_0^{-\lambda_1+1}\right] \tag{11}$$

Note that when $Q = Q_0$:

$$f(Q)\big|_{Q=Q_0} = \exp(-\lambda_0) Q_0^{-\lambda_1} \tag{12}$$

Maximum Entropy

Substitution of equation 10 in equation 5 yields the maximum entropy or uncertainty of discharge:

$$H(Q) = \lambda_0 + \lambda_1 \overline{\ln Q} \tag{13}$$

Determination of Lagrange Multipliers

Substitution of equation10 in equation 6 leads to:

$$\lambda_0 = -\ln(-\lambda_1 + 1) + \ln\left[Q_D^{-\lambda_1+1} - Q_0^{-\lambda_1+1}\right] \tag{14}$$

Differentiating equation 14 with respect to λ_1 produces:

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{1}{-\lambda_1 + 1} + \frac{-[\ln Q_D]Q_D^{-\lambda_1+1} + [\ln Q_0]Q_0^{-\lambda_1+1}}{Q_D^{-\lambda_1+1} - Q_0^{-\lambda_1+1}} \tag{15}$$

On the other hand, substitution of equation 10 in equation 6 can also be written as:

$$\lambda_0 = \ln \int_{Q_0}^{Q_D} Q^{-\lambda_1} dQ \tag{16}$$

Differentiating equation 16 with respect to λ_1, one gets:

$$\frac{\partial\lambda_0}{\partial\lambda_1} = -\frac{\int_{Q_0}^{Q_D}(\ln Q)Q^{-\lambda_1}\,dQ}{\int_{Q_0}^{Q_D}Q^{-\lambda_1}\,dQ} \tag{17}$$

Multiplying and dividing equation 17 by exp(-λ_0), and using equation 6, the result is:

$$\frac{\partial\lambda_0}{\partial\lambda_1} = -\frac{\int_{Q_0}^{Q_D}(\ln Q)\exp(-\lambda_0)Q^{-\lambda_1}\,dQ}{\int_{Q_0}^{Q_D}\exp(-\lambda_0)Q^{-\lambda_1}\,dQ} = -\overline{\ln Q} \tag{18}$$

Equating equation 15 to equation 18, an expression for λ_1 is obtained in terms of the constraint and the limits of integration of Q as:

$$\frac{1}{-\lambda_1+1} + \frac{-[\ln Q_D]Q_D^{-\lambda_1+1} + [\ln Q_0]Q_0^{-\lambda_1+1}}{Q_D^{-\lambda_1+1} - Q_0^{-\lambda_1+1}} = -\overline{\ln Q} \tag{19}$$

Derivation of Rating Curve

Let the maximum stage (channel flow depth) be denoted D. It is then assumed that all values of stage y measured from the bed to any point between 0 and D are equally likely. In reality this is not highly unlikely because at different times different values of stage do occur. This is also consistent with the Laplacian principle of insufficient reason. The cumulative probability distribution of discharge can then be expressed as the ratio of the stage to the point where discharge is to be considered and the stage up to the maximum water surface. The probability of discharge being equal to or less than a given value of Q is y/D. At any stage (measured from bed) less than a given value y, the discharge is less than a given value, say, Q. Thus, the cumulative distribution function of discharge, $F(Q) = P$(discharge $\leq$ a given value of Q), where P = probability, can be expressed as:

$$F(Q) = \frac{y}{D} \tag{20}$$

$F(Q)$ denotes the cumulative distribution function (CDF), and Q is discharge (m^3 s^{-1}). It should be noted that on the left side of equation 20, in the argument of function F, the variable is Q, whereas on the right side the variable is y. The CDF of Q is not linear in terms of Q unless Q and y are linearly related. Of course, it is plausible that $F(Q)$ might have a different form. Since equation 20 constitutes the fundamental hypothesis employed here for deriving the stage-discharge relation using entropy, it will be useful to evaluate its validity. This hypothesis (i.e., the relation between the cumulative probability $F(Q)$ and the ratio y/D) should be tested for a large number of natural rivers. This hypothesis was tested for several sets of data on stage and discharge, as shown for a sample data set in figure 1, which shows that the hypothesis expressed by equation 20 is approximately valid. It may also be noted that a similar hypothesis has been employed when using the entropy theory for deriving infiltration equations by Singh (2010a, 2010b), soil moisture profiles by Singh (2010c), and velocity distributions by Chiu (1987). It may also be emphasized that even if the above hypothesis is not strictly valid, it will not greatly influence the results because it merely allows the entropy theory to lead to the rating curves that are desired.

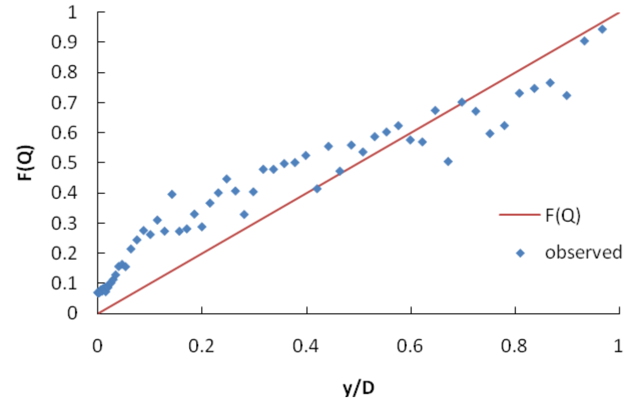

Figure 1. Relation between the cumulative probability *F*(*Q*) and the ratio *y*/*D*.

The probability density function is obtained by differentiating equation 20 with respect to Q:

$$f(Q) = \frac{dF(Q)}{dQ} = \frac{1}{D}\frac{dy}{dQ}$$

$$\text{or}\quad f(Q) = \left(D\frac{dQ}{dy}\right)^{-1} \tag{21}$$

The term $f(Q)\,dQ = F(Q + dQ) - F(Q)$ denotes the probability of velocity being between Q and $Q + dQ$.

Substituting equation 10 in equation 14, one gets:

$$\frac{\exp(-\lambda_0)}{-\lambda_1+1}Q^{-\lambda_1+1}\Big|_{Q_0}^{Q} = \frac{1}{D}y\Big|_{y_0}^{y} \tag{22}$$

Equation 22 yields:

$$Q = \left[Q_0^{-\lambda_1+1} + \frac{\exp(\lambda_0)(-\lambda_1+1)}{D}(y-y_0)\right]^{\frac{1}{-\lambda_1+1}} \tag{23}$$

Equation 23 can be considered as a general rating curve, for it encompasses all three types of rating curves outlined earlier.

Rating Curve Type 1

If $y_0 = 0$ and $Q_0 = 0$ in equation 23, then the result is:

$$Q = \left[\frac{\exp(\lambda_0)(-\lambda_1+1)}{D}\,y)\right]^{\frac{1}{-\lambda_1+1}} \tag{24}$$

$$\text{Let } a = \left[\frac{\exp(\lambda_0)}{D}(-\lambda_1+1)\right]^{\frac{1}{-\lambda_1+1}} \tag{25}$$

$$\text{and } b = \frac{1}{-\lambda_1 + 1} \tag{26}$$

Then, equation 24 yields the rating curve given by equation 2a. This rating curve is used in practice. When plotted on log-log paper, equation 24 will be straight line. It may now be interesting to evaluate the Lagrange multipliers for this simple case and hence parameters *a* and *b*.

Determination of Lagrange Multipliers

Substitution of equation 10 in equation 6 yields:

$$\exp(\lambda_0) = \frac{Q_D^{-\lambda_1+1}}{-\lambda_1+1}$$

$$\text{or } \lambda_0 = -\ln(-\lambda_1+1) + (-\lambda_1+1)\ln(Q_D) \tag{27}$$

where Q_D is the discharge at $y = D$.

Differentiating equation 27 with respect to λ_1, one obtains:

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\ln Q_D + \frac{1}{-\lambda_1+1} \tag{28}$$

One can also write from equations 10 and 6:

$$\lambda_0 = \ln \int_0^{Q_D} Q^{-\lambda_1} dQ \tag{29}$$

Differentiating equation 29 with respect to λ_1 and simplifying, one obtains:

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\overline{\ln Q} \tag{30}$$

Equating equation 28 to equation 30 leads to an estimate of λ_1:

$$\lambda_1 = 1 - \frac{1}{\ln Q_D - \overline{\ln Q}} \tag{31}$$

Therefore, exponent *b* of the power form rating curve becomes:

$$b = \ln Q_D - \overline{\ln Q} \tag{32}$$

Equation 32 shows that exponent *b* of the power form rating curve can be estimated from the values of the logarithm of maximum discharge at the water surface covering the channel fully and the average of the logarithmic values of discharge. The higher the difference between these logarithm values, the higher will be the exponent.

The Lagrange multiplier λ_0 can now be expressed as:

$$\lambda_0 = \frac{\ln Q_D}{\ln Q_D - \overline{\ln Q}} + \ln(\ln Q_D - \overline{\ln Q}) \tag{33a}$$

Therefore

$$a = \frac{Q_D}{D^{\ln Q_D - \overline{\ln Q}}} \tag{33b}$$

The PDF of *Q* can be expressed as:

$$f(Q) = \exp(-\lambda_0) Q^{\frac{1}{b}-1} = \frac{(Q_D)^{1/(\ln Q_D - \overline{\ln Q})}}{\ln Q_D - \overline{\ln Q}} Q^{\frac{1}{b}-1} \tag{34}$$

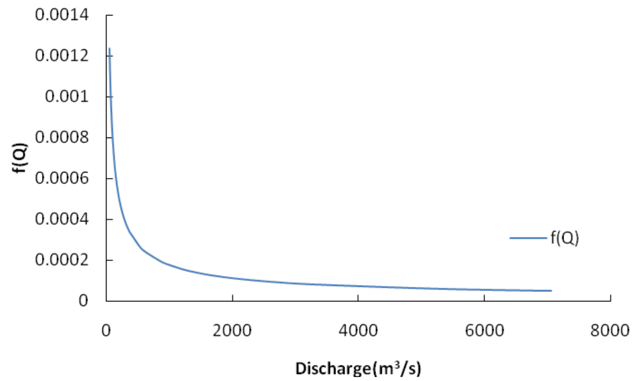

Figure 2. Probability density function of discharge for USGS gauging station 08082000-Salt-FK-Brazos.

and the CDF as:

$$F(Q) = b\exp(-\lambda_0) Q^{\frac{1}{b}} = (Q_D)^{1/(\ln Q_D - \overline{\ln Q})} Q^{\frac{1}{b}} \tag{35}$$

For $b < 1$, the probability density function (PDF) monotonically increases from 0 to $\exp(-\lambda_0) Q_D^{\,(1/b)-1}$. Figure 2 shows a PDF of discharge values observed at a USGS gauging station (08082000-Salt-FK-Brazos) on the Brazos River.

The entropy (in Napiers) of the discharge distribution can be obtained by substituting equation 34 in equation 5:

$$H = \lambda_0 - \left(\frac{1}{b} - 1\right)\overline{\ln Q} = -\frac{\ln Q_D}{\ln Q_D - \overline{\ln Q}} + \ln[\ln Q_D - \overline{\ln Q}] - \left(\frac{1}{b} - 1\right)\overline{\ln Q} \tag{36}$$

Rating Curve Type 2

In this case, $Q_0 = 0$ at $y = y_0$. Therefore, equation 23 becomes:

$$Q = \left[\frac{\exp(\lambda_0)(-\lambda_1+1)}{D}\right]^{\frac{1}{-\lambda_1+1}} (y - y_0)^{\frac{1}{-\lambda_1+1}} \tag{37}$$

Using equation 25 and 26, equation 37 becomes:

$$Q = a(y - y_0)^b \tag{38}$$

This type of rating curve is also used in practice. The PDF, CDF, and entropy associated with this curve remain the same as for rating curve type 1. Here, parameters *a* and *b* will have the same definitions but with *y* replaced by $y - y_0$ as in the case of type 1.

Rating Curve Type 3

In this case, let $q = Q - Q_0$, where Q_0 is some small value. It is assumed that $q = 0$ at $y = 0$. Then the derivation in the case of rating curve type 1 will hold, and equation 23 will be become:

$$Q = ay^b + c \tag{39}$$

which includes equation 24 as a special case. Here, parameters *a* and *b* will have the same definitions but with *Q* replaced by *q* as in the case of type 1 and $c = Q_0$.

Table 1. Stream gauging stations on the Brazos River, Texas.

USGS Number	Location	County	Latitude	Longitude	Hydrologic Unit	Drainage Area, mi^2 (km^2)	Contributing Area, mi^2 (km^2)	Gauge Datum above Sea Level NGVD, ft (m)
08079600	Justiceburg	Garza	33°02′18″	101°11′50″	12050004	1,466 (4,497.40)	244 (748.54)	2,222.47 (677.41)
08080500	Aspermont	Stonewall	33°00′29″	100°10′49″	12050004	8,796 (26,984.37)	1,864 (5718.38)	1,624.79 (494.26)
08082000	Salt-FK	Stonewall	33°20′02″	100°14′16″	12050007	5,130 (15,737.81)	2,496 (7657.23)	1,588.70 (484.24)
08082500	Seymour	Baylor	33°34′51″	99°16′02″	12060101	15,538 (47,667.48)	5,972 (18320.90)	1,238.97 (377.64)
08083100	Clear-FK	Fisher County	32°47′15″	100°23′18″	12060102	228 (699.46)	228 (699.46)	1,885.09 (577.58)

Evaluation

The objective here is to evaluate if entropy-based parameters and hence rating curves yield acceptable discharge values for given stage or flow depth, but not to advocate the preference of one rating curve over others. To that end, five stream gauging stations located on the Brazos River, Texas, were selected. These stations are located in four different counties and are operated by the U.S. Geological Survey (USGS). Table 1 gives relevant information on these stations. The drainage areas at these stations vary from 228 to 15,538 square miles (699.46 to 47,667.48 km^2) and represent a broad range of flow conditions.

Parameters were computed using equations for all five stations and are shown in table 2. For comparison, parameters of the three rating curves were also determined using the least square method, and values of these parameters are also given in table 2. The entropy values obtained were also determined for the five stations and are given in table 2.

In order to quantitatively evaluate the goodness of fit of the three rating curves, the following statistical measures were employed:

Coefficient of correlation:

$$r^2 = \frac{\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(Q_o(i)-\overline{Q}_o\right)\left(Q_c(i)-\overline{Q}_c\right)}}{\left[\frac{1}{N}\sum_{i=1}^{N}\left(Q_o(i)-\overline{Q}_o\right)\right]^{0.5}\left[\frac{1}{N-1}\sum_{i=1}^{N}\left(Q_c(i)-\overline{Q}_c\right)^{0.5}\right]} \tag{40}$$

Average bias:

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}[Q_o(i)-Q_c(i)] \tag{41}$$

Root mean square (RMS):

$$\text{RMS} = \sqrt{\frac{\sum [Q_o(i)-Q_c(i)]^2}{N-1}} \tag{42}$$

Table 2. Parameters *a* and *b* of rating curves for five gauging stations (for period of record 2006-2009).

Name of Station and USGS Number	Rating Curve	Entropy-Based Parameters		LS-Based Parameters		y_0 (m)	Q_0 (m^3 s^{-1})	H (bits)
		a	*b*	*a*	*b*			
DMF-Brazos 08079600	Type 1	26.35	2.34	68.30	1.67	--	--	1.85
	Type 2	26.35	2.34	79.04	2.27	-2.5	--	1.85
	Type 3	26.35	2.34	186.8	0.95	--	64.00	1.85
DMF-Brazos 08080500	Type 1	57.5	3.14	104.58	2.45	--	--	2.14
	Type 2	57.5	3.14	0.82	5.09	-2.5	--	2.14
	Type 3	57.5	3.14	23.1	3.59	--	64.0	2.14
Salt-FK-Brazos 08082000 (2006-09)	Type 1	3.89	3.66	0.03	6.66	--		2.30
	Type 2	3.89	3.66	265.1	2.1	2.8		2.30
	Type 3	16.59	2.95	1.05	4.64		190.0	2.08
Brazos 08082500	Type 1	67.5	3.16	104.58	2.44	--	--	2.42
	Type 2	67.5	3.16	0.38	5.44	-1.7	--	2.42
	Type 3	10.44	4.12	459.4	0.25	--	276.0	2.42
Clear-FK-Brazos 08083100	Type 1	0.017	4.21	0.07	3.83	--	--	2.44
	Type 2	0.07	4.21	0.001	5.08	-1.9	--	2.00
	Type 3	1.71	2.72	0.88	3.1	--	0.67	2.00

Table 3. Statistical measures for evaluating the performance of entropy and least square methods.

Name of Station and USGS Number	Rating Curve	Entropy Method r^2	Entropy Method Bias ($m^3\ s^{-1}$)	Entropy Method RMS ($m^3\ s^{-1}$)	Entropy Method NSE	Least Square Method r^2	Least Square Method Bias ($m^3\ s^{-1}$)	Least Square Method RMS ($m^3\ s^{-1}$)	Least Square Method NSE
DMF-Brazos 080779600	Type 1	0.916	43.75	167.69	0.91	0.938	43.60	167.48	0.91
	Type 2	0.906	-196.43	838.18	0.884	0.816	-77.69	403.67	0.881
	Type 3	0.914	36.74	165.30	0.913	0.969	-13.58	37.05	0.996
DMF-Brazos 08080500	Type 1	0.951	83.31	254.00	0.987	0.80	271.18	1021.75	0.792
	Type 2	0.824	-261.92	1270.40	0.872	0.821	-54.560	741.86	0.841
	Type 3	0.985	63.31	248.04	0.988	0.811	-275.21	1696.95	0.811
Salt-FK-Brazos 08082000	Type 1	0.412	120.75	1005.90	0.472	0.845	-190.07	2410.54	0.844
	Type 2	0.667	179.41	1035.1	0.674	0.856	-64.88	1680.82	0.856
	Type 3	0.476	470.75	1105.30	0.629	0.843	-104.94	2074.00	0.843
Brazos 8082500	Type 1	0.943	84.95	252.8	0.975	0.792	276.65	1037.7	0.792
	Type 2	0.827	-280.48	1252.01	0.827	0.879	-128.59	1067.37	0.888
	Type 3	0.98	262.84	538.16	0.985	0.808	-96.33	2351.38	0.808
Clear-FK-Brazos 08083100	Type 1	0.982	33.99	135.57	0.987	0.963	-78.01	334.27	0.963
	Type 2	0.925	-19.77	162.71	0.981	0.967	29.37	196.54	0.973
	Type 3	0.988	33.49	135.45	0.987	0.92	-77.514	335.15	0.921

Nash-Sutcliffe efficiency:

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N} [Q_o(i) - Q_c(i)]^2}{\sum_{i=1}^{N} [Q_o(i) - \overline{Q}_c(i)]^2} \tag{43}$$

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. These measures were computed for all three curve types and for all stations and are given in table 3.

In general, both the entropy-based curves and the curves determined with the least square method were comparable, but the agreement between computed and observed curves was sometimes better for the entropy-based method than for the least square method, and sometimes the reverse was true. Furthermore, type 1 curve better represented the observed stage-discharge relation than type 2 and type 3 curves. In some cases, type 2 as well as type 3 curves were satisfactory. This is also reflected in the statistical measures given in table 3.

For illustrative purposes, computed and observed rating curves for two stations, station 2 (08080500-DMF-Brazos) and station 5 (08083100-Clear-FK-Brazos), are shown in figures 3 to 8. For station 2, a rating curve of type 1 fitted the data reasonably well, as shown in figure 3. This is also seen from the high correlation value of the coefficient of determination (r^2) of 0.951 and the Nash-Sutcliffe efficiency (NSE) of 0.987. For this station, the average bias produced by the type 1 curve was small (83.31 $m^3\ s^{-1}$), and so was the root mean square error (RMS) (254.0 $m^3\ s^{-1}$). These measures for the least square method were: r^2 = 0.8, bias = 271.18 $m^3\ s^{-1}$, RMS = 1021 $m^3\ s^{-1}$, and NSE = 0.792. The goodness of fit is remarkable, given that no fitting was involved and the parameters were based solely on the constraints or the information specified beforehand. For this station, entropy-based rating curves of type 2 and type 3 were also found suitable. For the type curve 2, the statistical measures for the entropy-based method were: r^2 = 0.824, bias = -261.923 $m^3\ s^{-1}$, RMS = 1270.4 $m^3\ s^{-1}$, and NSE = 0.872, and for the least square method they were: r^2 = 0.821, bias = -54.56 $m^3\ s^{-1}$, RMS = 741.86 $m^3\ s^{-1}$, and NSE = 0.841. For the type 3 curve, the statistical measures for the entropy-based method were: r^2 = 0.985, bias = 63.31 $m^3\ s^{-1}$, RMS = 248.04 $m^3\ s^{-1}$, and NSE = 0.998, and for the least square method they were: r^2 = 0.811, bias = -275.21 $m^3\ s^{-1}$, RMS = 1696.95 $m^3\ s^{-1}$, and NSE = 0.811. Both methods were comparable, as shown in figures 4 and 5.

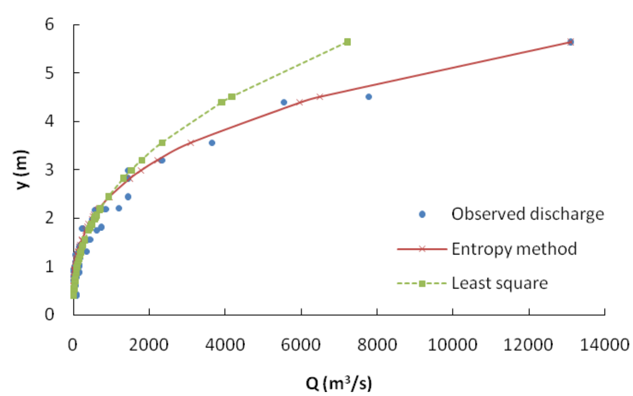

Figure 3. Entropy method-based and least square method-based rating type 1 curve and observed stage-discharge relation for station 2 (08080500-DMF-Brazos).

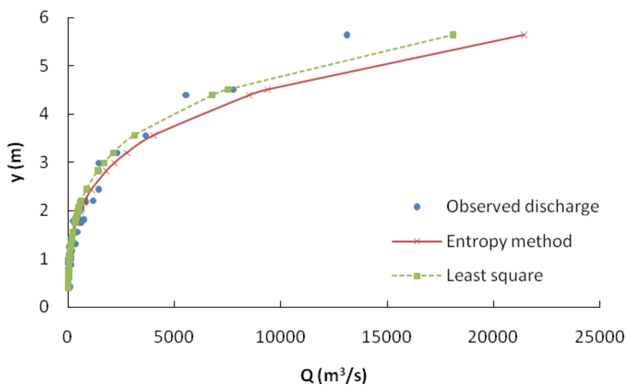

Figure 4. Entropy method-based and least square method-based type 2 curve and observed stage-discharge relation for station 2 (08080500-DMF-Brazos).

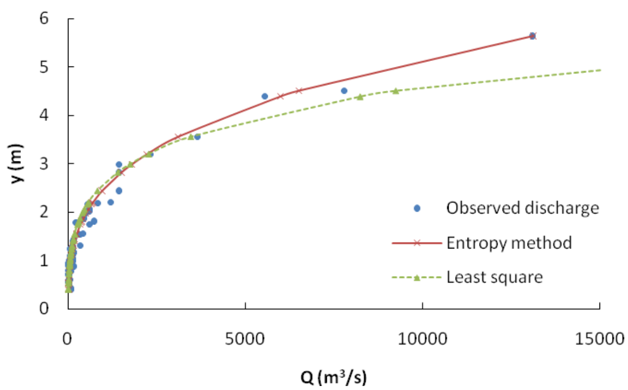

Figure 5. Entropy method-based and least square method-based type 3 curve and observed stage-discharge relation for station 2 (08080500-DMF-Brazos).

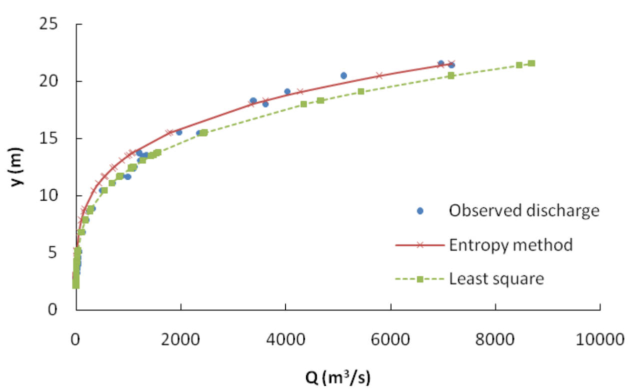

Figure 6. Entropy method-based and least square method-based type 1 curve and observed stage-discharge relation for station 5 (08083100-Clear-FK-Brazos).

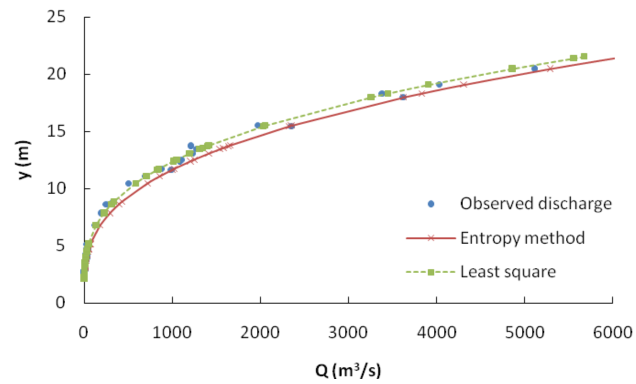

Figure 7. Entropy method-based and least square method-based type 2 curve and observed stage-discharge relation for station 5 (08083100-Clear-FK-Brazos).

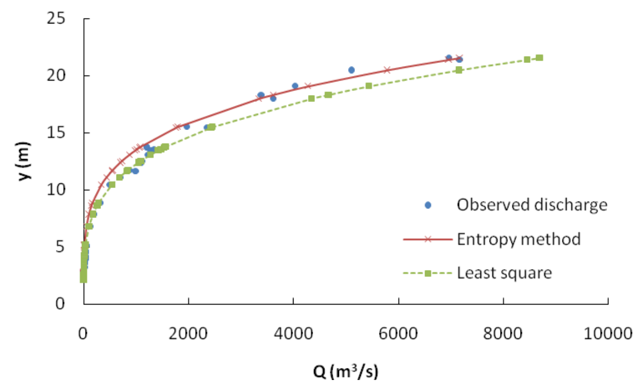

Figure 8. Entropy method-based and least square method-based type 3 curve and observed and stage-discharge relation for station 5 (08083100-Clear-FK-Brazos).

For station 5, the entropy-based type 1 rating curve fitted the data well, as shown in figure 6. For this curve, the statistical measures for the entropy-based method were: r^2 = 0.982, bias = -33.99 $m^3 s^{-1}$, RMS = 135.57 $m^3 s^{-1}$, and NSE = 0.987, and for the least square method they were: r^2 = 0.963, bias = -78.01 $m^3 s^{-1}$, RMS = 334.27 $m^3 s^{-1}$, and NSE = 0.963. The entropy-based method and the least squares method were comparable, but the least square method slightly under-estimated the discharge values for specified flow depths. For the type 2 curve, the least square method had a slight edge, as seen in figure 7 and the statistical measures. For the entropy-based method, these statistical measures were: r^2 = 0.925, bias = -19.77 $m^3 s^{-1}$, RMS = 162.71 $m^3 s^{-1}$, and NSE = 0.981, and for the least square method they were: r^2 = 0.967, bias = 29.37 $m^3 s^{-1}$, RMS = 196.54 $m^3 s^{-1}$, and NSE = 0.973. The least square method had a slight edge.

The entropy-based type 3 rating curve fitted the observed data reasonably well, as seen in figure 8 and the statistical measures. The least square method also fitted the data well. For the entropy-based method, the statistical measures were: r^2 = 0.988, bias = 33.49 $m^3 s^{-1}$, RMS = 135.45 $m^3 s^{-1}$, and NSE = 0.987, and for the least square method they were: r^2 = 0.92, bias = -77.51 $m^3 s^{-1}$, RMS = 334.15 $m^3 s^{-1}$, and NSE = 0.921. The entropy method had a slight edge.

For all five stations, the entropy method-based rating curves compared well with the observed rating curves and the least square method-based curves. The entropy values reflect relative goodness of fit. The entropy values were quite comparable, and agreements between observed and computed rating 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. The lower values reflect less uncertainty, meaning a better fit, and this was observed somewhat.

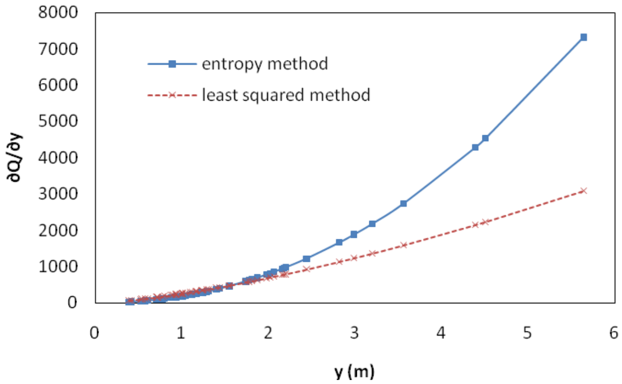

Figure 9. Change in discharge with the change in parameter *y* as a function of *y* for rating curve type 1 for USGS gauging station 08080500-DMF-Brazos.

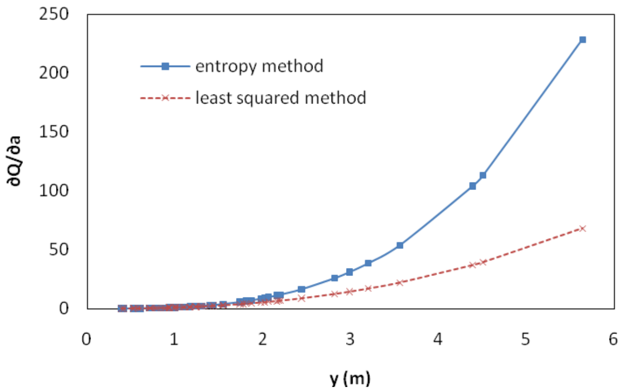

Figure 10. Change in discharge with the change in parameter *a* as a function of *y* for rating curve type 1 for USGS gauging station 08080500-DMF-Brazos.

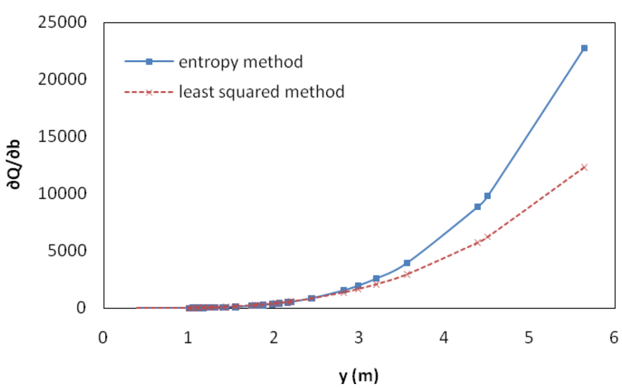

Figure 11. Change in discharge with the change in parameter *b* as a function of *y* for rating curve type 1 for USGS gauging station 08080500-DMF-Brazos.

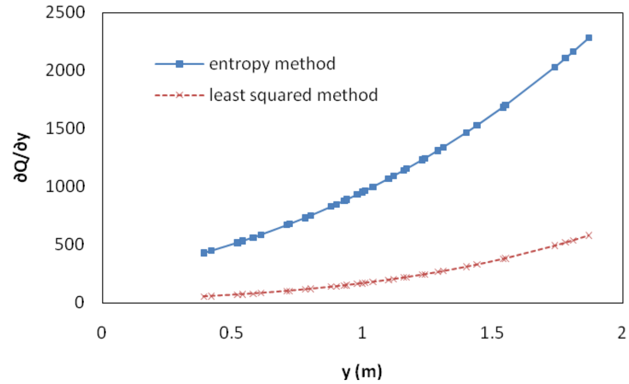

Figure 12. Change in discharge with the change in parameter *y* as a function of *y* for rating curve type 2 for USGS gauging station 08080500-DMF-Brazos.

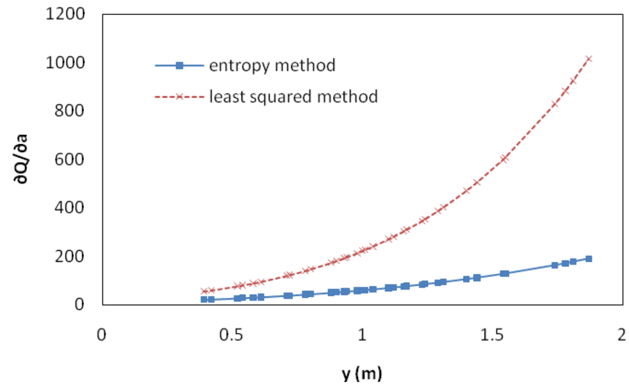

Figure 13. Change in discharge with the change in parameter *a* as a function of *y* for rating curve type 2 for USGS gauging station 08080500-DMF-Brazos.

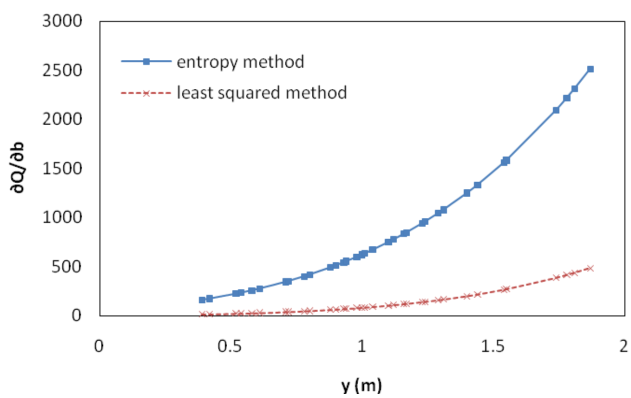

Figure 14. Change in discharge with the change in parameter *b* as a function of *y* for rating curve type 2 for USGS gauging station 08080500-DMF-Brazos.

Sensitivity Analysis

The rating curve has one dependent variable (Q), one independent variable (y), and two or three parameters depending on the type of curve. For type 1, the two parameters are a and b. For type 2, the parameters are a, b, and y_0. For type 3, the parameters are a, b, and c. It would be interesting to determine the sensitivity of rating curves to their parameters. To that end, one can write:

For type 1 curves:

$$\frac{\partial Q}{\partial y} = aby^{b-1}; \ \frac{\partial Q}{\partial a} = 1; \ \frac{\partial Q}{\partial b} = y^b \ln y \qquad (44)$$

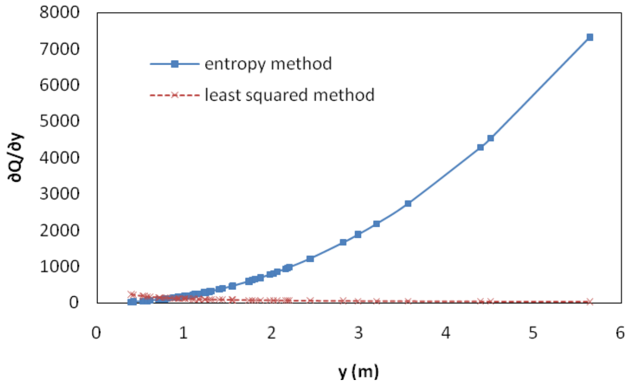

Figure 15. Change in discharge with the change in parameter *y* as a function of *y* for rating curve type 3 for USGS gauging station 08080500-DMF-Brazos.

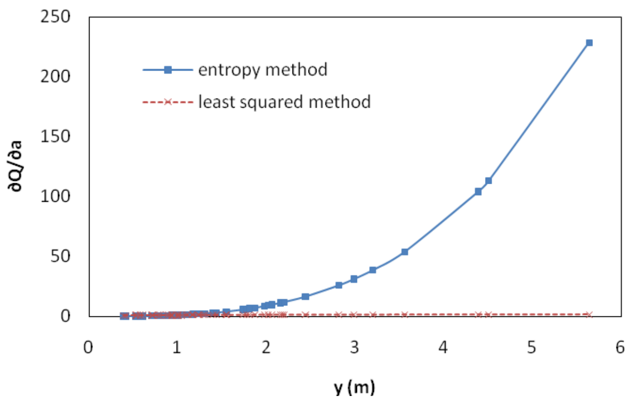

Figure 16. Change in discharge with the change in parameter *a* as a function of *y* for rating curve type 3 for USGS gauging station 08080500-DMF-Brazos.

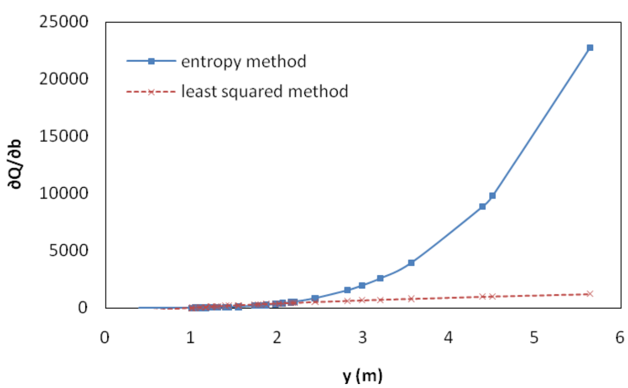

Figure 17. Change in discharge with the change in parameter *a* for rating curve type 3 for USGS gauging station 08080500-DMF-Brazos.

For type 2 curves:

$$\frac{\partial Q}{\partial y} = ab(y - y_0)^{b-1}; \ \frac{\partial Q}{\partial a} = 1;$$

$$\frac{\partial Q}{\partial b} = (y - y_0)^b \ln(y - y_0) \quad (45)$$

For type 3 curves:

$$\frac{\partial Q}{\partial y} = aby^{b-1}; \ \frac{\partial Q}{\partial a} = 1;$$

$$\frac{\partial Q}{\partial Q_0} = 1; \ \frac{\partial Q}{\partial b} = y^b \ln y \quad (46)$$

Each derivative of Q with respect to a given parameter can be called a sensitivity coefficient. The values of the coefficients help assess the effect of parameters on the rating curve and also reflect their significance. Using equations 44 to 46, the coefficients were computed for all curves for all five stations using both the entropy method and the least square method. For economy of space, these coefficients are plotted for the three curves for only station 2 (08080500-DMF-Brazos), as shown in figures 9 to 17. Curves would be similar but not the same for other stations. For type 1 curves, as shown in figures 9 to 11, the entropy-based sensitivity coefficient for input parameter y and parameters a and b as functions of flow depth are significantly higher as flow depth exceeds 2 m and become even higher for higher flow depths. This is even truer for type 2 curves, as shown in figures 12 to 14. However, for type 3 curves, this is true only for the entropy method, as shown in figures 14 to 17; for the least square method, they remain almost independent of flow depth.

Conclusions

The following conclusions are drawn from this study: (1) With the entropy method, parameters of rating curves can be determined in terms of specified constraints, which themselves are determined from observations. This obviates the need for fitting. (2) All three rating curves (types 1, 2, and 3) are found to represent observed rating curves quite well. (3) The rating curves computed using the entropy method compare well with those computed using the least square method.

Acknowledgements

Ms. Huijuan Cui, Graduate Student, Watershed Management and Hydrologic Science Program, at Texas A&M University helped with computations and graphs. Her help is gratefully acknowledged.

References

Chiu, C. L. 1987. Entropy and probability concepts in hydraulics. *J. Hydraulic Eng.* 113(5): 583-600.

Corbett, D. M. 1962. Stream-gaging procedure: A manual describing methods and practices of the Geological Survey. USGS Water Supply Paper 888. Washington, D.C.: U.S. Geological Survey.

Clarke, R. T. 1999. Uncertainty in the estimation of mean annual flood due to rating-curve indefinition. *J. Hydrol.* 222: 185-190.

DeGagne, M. P. J., G. G. Douglas, H. R. Hudson, and S. P. Simonovic. 1996. A decision support system for the analysis and use of stage-discharge rating curves. *J. Hydrol.* 184(3-4): 225-241.

Herschy, R. W. 1995. *Streamflow Measurements.* 2nd ed. London, U.K.: Chapman and Hall.

Jaynes, E. T. 1957. Information theory and statistical mechanics. *Physical Review* 106(4): 620-630.
Jaynes, E. T. 1982. On the rationale of maximum-entropy methods. *Proc. IEEE* 70(9): 939-952.
Kennedy, E. J. 1964. Chapter A10: Discharge rating at gaging stations. In *Applications of Hydraulics*, Book 3. Washington, D.C.: U.S. Geological Survey.
Lambie, J. C. 1978. Chapter 1: Measurement of flow-velocity-area-methods. In *Hydrometry: Principles and Practices*. R. W. Herschy, ed. Chichester, U.K.: Wiley.
Lee, W. S., K. S. Lee, S. U. Kim, and E. S. Chung. 2010. The development of rating curve considering variance function using pseudo-likelihood estimation method. *Water Resources Mgmt.* 24(2): 321-348.
Overlier, P. 2006. Modeling stage-discharge relationships affected by hysteresis using the Jones formula and nonlinear regression. *Hydrol. Sci.* 51(3): 365-388.
Petersen-Overleir, A. 2004. Accounting for heteroscedasticity in rating curve estimates. *J. Hydrol.* 292: 173-181.
Petersen-Overleir, A., and T. Reiten. 2005. Objective segmentation in compound rating curves. *J. Hydrol.* 311: 188-201.
Sauer, V. B., and R. W. Meyer. 1992. Determination of error in individual discharge measurements. USGS Open-File Report 92-144. Washington, D.C.: U.S. Geological Survey.
Schmidt, A. R., and B. C. Yen. 2002. Stage-discharge ratings revisited. Hydraulic measurements and experimental methods. *Proc. EWRI and IAHR Joint Conf.*, Colorado. Reston, Va.: ASCE Press.
Shannon, C. E. 1948. A mathematical theory of communications, I and II. *Bell System Tech. J.* 27: 379-443.
Shannon, C.,E. and Weaver, W. 1949. *The Mathematical Theory of Comunication.* Urbana, Ill.: University of Illionis Press.
Sivapragasam, C., and N. Mutill. 2005. Discharge curve extension: A new approach. *Water Resources Mgmt.* 19(5): 505-520.
Singh, V. P. 1993. *Elementary Hydrology*. Englewood Cliffs, N.J.: Prentice Hall.
Singh, V. P. 1996. *Kinematic Wave Modeling in Water Resources: Surface Water Hydrology*. New York, N.Y.: John Wiley and Sons.
Singh, V. P. 1998. *Entropy-Based Parameter Estimation in Hydrology.* Boston, Mass.: Kluwer Academic.
Singh, V. P. 2010a. Entropy theory for derivation of infiltration equations. *Water Resources Res.* 46: W03527, doi:10-1029/2009WR008193.
Singh, V. P. 2010b. Tsallis entropy theory for derivation of infiltration equations. *Trans. ASABE* 53(2): 447-463.
Singh, V. P. 2010c. Entropy theory for movement of moisture in soils. *Water Resources Res.* 46: W03516, doi:10-1029/2009WR008288.
Torsten, D., M. Gerd, and S. Torsten. 2002. Extrapolating stage-discharge relationships by numerical modeling. *Proc. Intl. Conf. on Hydraulic Eng*., Warsaw, Poland.
WMO. 1994. *Guide to Hydrological Practices*. 5th ed. WMO-No. 168. Geneva, Switzerland: World Meteorological Organization.
Yoo, C., and J. Park. 2010. A mixture-density-network based approach for finding rating curves: Facing multimodality and unbalanced data distribution. *KSCE J. Civil Eng.* 14(2): 243-250.