Journal of Hydrology 564 (2018) 748–757

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

On the generation of stream rating curves

John D. Fenton

Institute of Hydraulic and Water Resources Engineering, Vienna University of Technology, Karlsplatz 13/222, 1040 Vienna, Austria

ARTICLE INFO

This manuscript was handled by Marco Borga, Editor-in-Chief, with the assistance of Yasuto Tachikawa, Associate Editor

Keywords:
Discharge measurement
Streamgauging
Rating curves
Rivers
Velocity measurements

ABSTRACT

Traditional methods for the calculation of rating curves from measurements of water level and discharge are criticised as being limited and complicated to implement, such that manual methods are still often used. Two methods for automatic computation are developed using least-squares approximation, one based on polynomials and the other on piecewise-continuous splines. Computational problems are investigated and procedures recommended to overcome them. Both methods are found to work well and once the parameters for a gauging station have been determined, rating data can be processed automatically. For some streams, ephemeral changes of resistance may be important, evidenced by scattered or loopy data. For such cases, the approximation methods can be used to generate a rating envelope as well, allowing the routine calculation also of maximum and minimum expected flows. Criticism is made of current shift curve practices. Finally, the approximation methods allow the specification of weights for the data points, enabling the filtering of data, especially decreasing the importance of points with age and allowing the computation of a rating curve for any time in the past or present.

1. Introduction

A rating curve is a relationship between the discharge Q of a stream and h, the stage or surface elevation, so that when routine measurements of stage at a gauging station are made, the flow can be estimated. The curve is calculated from a number of (h, Q) rating data points from that station, using relatively infrequent measurements of the velocity distribution, cross-section, and stage of the stream.

The problem of the automatic calculation of rating curves has received relatively little research attention. The main problem seems to be the perceived success and almost universal use of the power function

$$Q = C(h - h_0)^{\mu}, \tag{1}$$

where C, h_0 and μ are constants, and which is a straight line on $(\log Q, \log(h - h_0))$ axes. The reasons for it being a problem include:

- On one hand it is too simple, with only three parameters, and is limited in its accuracy and generality.
- On the other hand, it is too complicated, such that the three parameters occur nonlinearly and solving for them is difficult such that manual methods are often used.

The power function, and its representation as a straight line on logarithmic axes appears ubiquitously in books, standards, and lecture notes. Whereas it is sometimes a convenient approximation to the relationship $Q(h)$ over the whole range of data, in general it is not. It is an over-simplification of the real hydraulics at many gauging stations. Such a formula is valid for an infinitely-wide weir in infinitely-deep water or for uniform flow in an infinitely-wide rectangular channel. There is no reason for a real rating curve to follow such a function closely. Insufficient knowledge of hydraulics has led to a too-great belief in the power function, on one hand by practitioners, and on the other by theoreticians in related disciplines. This has led to complicated procedures in some organisations where sequences of power functions are used, and a great deal of trouble goes into the laborious manual fitting of piecewise-continuous straight lines on logarithmic axes by adjusting the offsets h_0 for each on different vertical $\log(h - h_0)$ axes.

The more general representation of Q by a polynomial of higher degree M has been in the background for some time:

$$Q = a_0 + a_1 h + a_2 h^2 + \ldots + a_M h^M = \sum_{m=0}^{M} a_m h^m, \tag{2}$$

where $a_0,\ a_1,\ \ldots, a_M$ are coefficients. It was presented by Herschy in the first edition of his book in 1985, most recently in Herschy (2009, p195), in International Standard 7066-2 (1988), and in Morgenschweis (2010, p384). Standard linear least-squares methods can be used to determine the coefficients. Mirza (2003) used it successfully with just $M = 3$, and in that scholarly work gave considerable attention to statistical matters.

Reading those sources and water industry websites, but also reading between the lines, it seems that the approximation by polynomials,

E-mail address: JohnDFenton@gmail.com.
URL: http://johndfenton.com.

https://doi.org/10.1016/j.jhydrol.2018.07.025
Received 16 September 2017; Received in revised form 18 May 2018; Accepted 11 July 2018

despite its promise, has not often been adopted, and usually only implemented to low degree. Herschy wrote in the first edition of his book in 1985, and 24 years later again in the third edition, (Herschy, 2009, p195): "however some user experience is still required with this method before it is accepted as an alternative to the existing methods", implying that its use has been languishing.

Fenton and Keller (2001, Section 6.3.2), suggested writing the polynomial for Q raised to the power ν, specified *a priori*:

$$Q^{\nu} = a_0 + a_1 h + a_2 h^2 + \ldots + a_M h^M = \sum_{m=0}^{M} a_m h^m, \tag{3}$$

which is actually a simple generalisation of the power function, Eq. (1), written in the form $Q = (a_0 + a_1 h)^{1/\nu}$, to $Q = (a_0 + a_1 h + a_2 h^2 + \ldots)^{1/\nu}$. They recommended a value of $\nu = \frac{1}{2}$, on the basis of that being the mean value in the hydraulic discharge formulae for a sequence of weir and channel cross-sections that modelled local and channel control (Fenton, 2001). The use of such a fractional value of ν has two effects:

1. For small flows, h and Q small, the data usually is such that
$$Q^{\nu} = a_0 + a_1 h, \tag{4}$$
with $\nu = \frac{1}{2}$, is a surprisingly good approximation when compared with power function approximations in which ν is a free parameter, as shown in Fenton (2015b, Section 3.4). In this small flow limit the more general polynomial approximation just has to simulate nearly-linear variation, which it can easily do.
2. For larger flows, when the higher degree terms in Eq. (3) become important, the use of Q^{ν} means that the magnitude of the dependent variable to be approximated is numerically much smaller, so that, instead of a range of say, $Q \approx 1$ to $10^4\ \mathrm{m^3\,s^{-1}}$, for $\nu = \frac{1}{2}$ a numerical range $Q^{1/2} \approx 1$ to 10^2 has to be approximated. This is a simple version of a *power transformation* used in more formal data analysis applications to stabilise variance and to make the data more normal distribution-like.

In recent years there have been a number of papers with a quite different way of looking at the problem, using Bayesian statistics. Le Coz et al. (2014) provided an excellent survey both of that field and the rating curve problem generally. However, all the papers they referred to used either a single power function or two or more of them, each in the belief that they were following hydraulic principles. It is the assertion here that there is little fundamental about the power function or the application of hydraulic theory, and here a rather different path will be followed, treating the problem as one of data approximation.

In that spirit, Coxon et al. (2015) used LOWESS (LOcally WeightEd Scatterplot Smoothing) to obtain rating curves for a huge number of sites. The method considered each stage-discharge measurement as the central point in a subset of the data points. The estimate of the discharge for the data point and its variance was generated by fitting a weighted linear regression to the selected data. Weights were dependent upon the differences in stage and gave most weight to data closest to the central measurement. To account for outlier points, two passes were made, then a weight function was used to weight each data point according to how far the point was from the first fitted line, reducing the impact of those furthest from it. The procedure could be used to satisfy the goal in this work, of developing methods for practical automatic computation. It seems good in principle, but there are a number of adjustable parameters and the reduction of importance of outlying points might deny the importance of some causative processes and trends at work. It functioned well for the demanding application that Coxon et al. required of it, where the main thrust was the quantification of uncertainty rather than the generation of approximations.

Fenton (2015b), hereafter referred to as Report I, considered several aspects of the problem of the automatic generation of rating curves. The present paper is based on that report, which contains more detail. Here first, the application of polynomial approximation methods is treated at length. Several mathematical reasons for problems associated with them are given, with physical explanations and methods for overcoming them. It is considered imperative to use series of Chebyshev polynomials rather than the simple polynomials shown above which are series of monomials h^m. Also it is desirable to approximate, not values of discharge Q, but Q^{ν}, where ν is a fractional exponent, as in Eq. (3). It is usually able to be taken to be $\frac{1}{2}$, but in extreme cases can be calculated by a method that is presented. Other than ν, the degree M of the polynomial series is the only free parameter. It is possible to use large values of M but if the data has gaps there will usually be one degree beyond which large fluctuations occur in between accurate approximation of the data points. To overcome that problem, an alternative approximation method is developed using piecewise-continuous splines, in which case the parameters of the problem are the number and stage values of knot points between which simple quadratic or cubic spline functions are used. A simple automatic method is suggested for the placing of those knots, just requiring there to be the same number of data points in each interval. This usually works well. Otherwise, in difficult cases the values of stage for the knot points can be specified. Results for both the polynomial method and the approximating spline method are presented. They are both found to perform well and have the potential to be standard procedures for rating curve generation. Then possible reasons for scatter of rating points are discussed. For such data, the methods can be modified to calculate additionally a rating envelope, giving likely maximum and minimum flow rating curves. For discrepant points it is suggested that current use of shift curves should be re-examined. Finally, the approximation methods are simply modified to allow the importance of data points to decrease with age. This allows the generation of a rating curve on any date in the past also, thereby determining any relatively slow long term changes in the stream.

2. Polynomial approximation

Eq. (3) is now generalised by considering the approximating function to be made up, not of a series of monomials h^m, but of more general functions $f_m(h)$:

$$Q^{\nu} = \sum_{m=0}^{M} a_m f_m(h) = a_0 f_0(h) + a_1 f_1(h) + \ldots + a_M f_M(h). \tag{5}$$

In application, the functions are specified *a priori*, and the unknown coefficients a_m found by least-squares fitting to data points (h_n, Q_n^{ν}) for $n = 1, \ldots, N$. Each of the functions applies over the whole data range of h, and so methods using them are *global* ones, as distinct from those in Section 3 below where piecewise-continuous *local* functions are used. We will consider the functions $f_m(h)$ each to be a polynomial of degree m so that the sum of such polynomials in Eq. (5), including the last one at $m = M$, is also a polynomial of degree M, and we can refer to methods using them as *polynomial approximation*. It will be found that Chebyshev polynomials for the $f_m(h)$ are particularly useful.

There are three problems here with the approximation: the rapid variation of data at the low flow end, the large range of discharge Q, and the ability of the approximating functions to describe arbitrary variations. These problems will be overcome, as is now described.

2.1. Exponent ν

2.1.1. Usual adequacy of $\nu = 1/2$

Traditionally, the power function has often been required to model all the data. By writing it in the form of Eq. (4), $Q^{\nu} = a_0 + a_1 h$, while it incorporates the usual rapid variation and large curvature on (Q, h) axes at low-flows, it is obvious that it is a limited approximation to the whole rating curve problem. Concerning the actual value of ν to use, Report I (Fig. 2) showed that for each of seven different stations,