RESEARCH ARTICLE

Smoothing River Discharge Time Series Computed Using the Velocity-Area Method

Kim Dianne B. Ligue • **Joseph E. Acosta**

University of the Philippines Mindanao, PHILIPPINES

Abstract

Early warning systems for flood disaster mitigation involves hydrological and hydraulic models. These models require the discharge series as input and calibration. The indirect measurement of discharge, derived using the velocity-area method, presents errors, and thus, uncertainty. The associated uncertainty is mainly caused by the error in estimation since the direct measurement of discharge is neither feasible nor cost-efficient. Smoothing is employed to address the issue. Three smoothing techniques are proposed, i.e. Fourier smoothing, kernel smoothing using the Gaussian Density function, and LOESS Curve Fitting. Two river basins located in Davao del Sur, Lipadas and Padada, were evaluated. The Nash-Sutcliffe Efficiency (NSE) and RMSE-Observations Standard Deviation Ratio (RSR) were used to evaluate smoothing performance. Results showed that the Gaussian kernel smoothing technique outperformed both Fourier method and LOESS for both river discharge series. The values for both NSE and RSR indicated that the technique produced very good performances. The quality of smoothed discharge series was studied using two quality functions, Quality of Discharge (QOD) and BALANCE. Results showed that a more appropriate method would result in a better discharge quality regardless of the smoothing parameter chosen. Therefore, the smoothed discharge series is affected by the choice of smoothing technique and the method of choice is crucial. This study suggests that Gaussian kernel smoothing is a promising technique in smoothing discharge, with bandwidth at around 2 to 10. A very good quality of smoothed discharge is to be expected when the Gaussian kernel technique gives a very good smoothing performance.

Keywords: Smoothing • Fourier smoothing • Gaussian Kennel smoothing • LOESS Curve Fitting • NSE • RSR • discharge series • discharge quality

Correspondence: KDB Ligue. Department of Mathematics, Physics, and Computer Science, College of Science and Mathematics, University of the Philippines Mindanao, Mintal, Tugbok District, Davao City 8022, Philippines. Telephone: +63 82 293 0302. Email: kbligue@up.edu.ph

Author Contribution: KBDL, JEA: Conceptualization, Writing–Review and Editing, Validation; KBDL: Formal Analysis, Methodology, Software, Visualization, Writing–Original Draft; JEA: Methodology, Supervision, Funding Acquisition, Project Administration, Resources, Validation.

Editor-in-Charge: Dr. Emma Ruth V. Bayogan, University of the Philippines Mindanao

Received: 21 May 2020 **Accepted:** 22 September 2021 **Published:** 29 September 2021

Funding Source: This work is funded by the Department of Science and Technology - Philippine Council for Industry, Engineering, Engineering and Emerging Technology Research and Development and the University of the Philippines Mindanao.

Competing Interest: The authors have declared no competing interest.

Citation: Ligue KBD and Acosta JE. 2021. Smoothing river discharge time series computed using the velocity-area method. *Banwa* B 16: art049.

Smothing River Discharge Time Series Computed Using the Velocity-Area Method

Kim Dianne B. Ligue
Joseph E. Acosta
University of the Philippines Mindanao, PHILIPPINES

Introduction

Flood disasters inflict billions worth of infrastructural damages and millions of casualties. To cope with this, developing countries opt for nonstructural early warning systems in mitigating damages caused by flood disasters versus its expensive structural alternative (Jayawardena et al. 2014).

The basic technical component of an early warning system involves simulations formulated from hydrological and hydraulic mathematical models (Domeneghetti et al. 2012) that can be computed automatically, e.g. by use of software programs like HEC-HMS (HEC 2000) and HEC-RAS (HEC 2001). Discharge, the volume or flow of water passing a gauging station in a river at a particular time (Bowen and Pallister 2000), is important as input (boundary and initial conditions) and as calibration data for these kinds of models (Di Baldassarre 2012). However, directly measuring the continuous discharge record needed for these models is neither feasible nor cost-efficient (U.S. Geological Survey 2015).

An alternate practice is to make use of stage data, the height of the water surface above an established altitude. The direct measurement of water stage on a continuous basis is more feasible and cost-efficient than the direct measurement of discharge (ISO 1100–2 2010; Di Baldassarre 2012; U.S. Geological Survey 2015). Stage data is converted into discharge using a stage-discharge relationship, also known as the rating curve (Di Baldassarre 2012; U.S. Geological Survey 2015). The rating curve is commonly computed as:

$$Q = C(H + \alpha)^{\beta} \tag{1}$$

where Q is the discharge or flow rate, H is the water stage, C and β are the calibration coefficients and α is the offset value for stage-of-zero flow (ISO 1100–2 2010).

To empirically derive the rating curve, discharge is manually estimated using alternative techniques. The most used is the velocity area method given by:

$$Q'(x,t) = A(x,t)V(x,t) \tag{2}$$

where x is the river chainage, t is the sampling time, $Q'(x,t)$ is the measured river discharge, $A(x,t)$ is the cross sectional area, and $V(x,t)$ is the average flow velocity at that certain cross section (ISO 748 2007; U.S. Geological Survey 2015).

The expected associated uncertainty caused by the errors in the estimation are often not stated in the accompanied predictions even though only estimates of the unknown true discharge values are used (Herschy 2002). These uncertainties are quite alarming and are reported to be as high as 30% of the observed values (Di Baldassarre and Montanari 2009). Numerous studies were conducted to explore quantitative measurement of the uncertainty of these hydrological models (Di Baldassarre and Montanari 2009; Booij et al. 2011; Domeneghetti et al. 2012; Guerrero et al. 2012; Birgand et al. 2013). A significant portion of this uncertainty is caused by the error in velocity measurement. In total, river discharge measurement used to calibrate the rating curve at the 95% confidence level is affected by an uncertainty of about 5% (ISO 748 2007).

Smoothing is usually employed to address data with known measurement errors, also called noise. This method removes noise allowing important patterns to stand out so that more important information is extracted from the data set. While smoothing techniques were employed on hydrological flood models (Xiong 2005), no study is reported evaluating its effects on the rating curve (Di Baldassarre and Montanari 2009; Domeneghetti et al. 2012). This study evaluated three smoothing techniques, all varying in its complexity and power in extracting important information depending on type and the state of the target data.

Fourier Smoothing Method

The Fourier series decomposes a periodic function or signal into the sum of the set of trigonometric functions, sine and cosine. Applications of the Fourier series to smoothing extend to hydrological models (Xiong et al. 2005). One main justification for the use of Fourier smoothing is the periodicity of the target time series (Takezawa 2005; Pekarova et al. 2006; Esomba 2015). It is especially applied to discharge models because of the inherent periodicity of the discharge series (Takezawa 2005). The method is used to reduce the high frequency components of the time series in favor of the low frequency components so that the more important trend is reflected (Kimball 1974).

To smooth the data series using the Fourier series method, a Fourier series representation was utilized by using a property of the harmonic function given by:

$$Q_t^{FS} = a_0 + \sum_{j=1}^{p} \left[A_j \cos\cos\left(\frac{2\pi j}{N}\cdot t\right) + B_j \sin\sin\left(\frac{2\pi j}{N}\cdot t\right)\right] \quad (3)$$

where Q_t^{FS} is the smooth discharge for time t, N is the number of points in the data set smoothed, a_0 is the mean of the original discharge computation, A_j and B_j are the Fourier coefficient, j is the order of the harmonic, and p is the number of harmonics retained (Xiong et al. 2005). To estimate the values of a_0, A_j, and B_j, the following equations were used:

$$a_0 = \frac{1}{N}\sum_{t=1}^{N} Q_t \quad (4)$$

$$A_j = \frac{2}{N}\sum_{t=1}^{N} Q_t \cos\cos\left(\frac{2\pi j}{N}\cdot t\right), \quad (5)$$

$$B_j = \frac{2}{N}\sum_{t=1}^{N} Q_t \sin\sin\left(\frac{2\pi j}{N}\cdot t\right), \quad (6)$$

The values obtained from Q_t^{FS} is the new smoothed discharge series.

Kernel Smoothing Method

Kernel smoothing is a nonparametric statistical technique that uses noisy observations to represent an irregular data set as a smooth line or surface (Li and Racine 2007). There are numerous kernel methods used for smoothing and one of the popular ones is the Gaussian Kernel expressed as:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (7)$$

where $K(u)$ is the kernel function at point u (Vert et al. 2004).

Nadaraya-Watson estimator. In 1964, Nadaraya and Watson proposed, in separate studies, a linear smoother that aims to estimate a variable by locally weighted average using a kernel as a density function (Li and Racine 2007). The Nadaraya-Watson estimator, $\widehat{m}(X)$, is given by:

$$\widehat{m}(X) = \frac{\sum_{i=1}^{N} K\left(\frac{x-x_i}{v}\right)\cdot y_i}{\sum_{i=1}^{N} K\left(\frac{x-x_i}{v}\right)} = \sum_{i=1}^{n} W\left(\frac{x-x_i}{v}\right) y_i \quad (8)$$

where W is the weight function defined by:

$$W\left(\frac{x-x_i}{v}\right) = \frac{K\left(\frac{x-x_i}{v}\right)}{\sum_{i=1}^{N} K\left(\frac{x-x_i}{v}\right)} \quad (9)$$

and K is a kernel with a bandwidth, v, and (x_i,y_i) are the original points at $i = 1,2,\ldots,N$ where N is the total number of original points. $W\left(\frac{x-x_i}{v}\right)$ is the weighting term with sum equal to 1, that is $\sum_{i=1}^{n} W\left(\frac{x-x_i}{v}\right) = 1$ (Nadaraya 1994; Watson 1964).

The method is used in numerous smoothing problems, including hydrological models (Xiong et al. 2005). It is famous for its amenability in routine and automatic applications (Faucher et al. 2001).

In performing the Kernel smoothing method, the Nadaraya-Watson Estimator in Eq. [8] was used to smooth the discharge. The Gaussian kernel expressed in Eq. [7] was used, which represents how the weight among the neighboring data points is determined. Values for the smoothing parameter or bandwidth, v, were chosen based on the study by Xiong et al. (2005). The result obtained using Eq. [8] was recorded as the Gaussian Kernel smoothed discharge series.

LOESS Curve Fitting

LOESS, or local regression, is a later generalization of LOWESS, or locally weighted scatterplot smoothing, a non-parametric statistical smoothing technique that is based on a combination of multiple regression models in a

k-nearest-neighbor-based meta-model (Fox 2002).

The method is nonparametric because the linearity assumptions of conventional regression methods have been relaxed. Instead of estimating a general regression function in the form of $y=mx + c$, a low degree polynomial is fitted for a subset of the data. This subset comprises of some point x, whose response is being estimated, and other points nearest to x. An estimate is computed for every point in the data set. The local polynomial is fitted using weighted least squares, giving more weight to points near the response than the point far away. For this reason, the method is also called a locally weighted polynomial regression. The commonly used weighting function, which is used in this study, is the tri-cube weighting function given by:

$$w(x) = (1 - |x|^3)^3. \tag{10}$$

where $w(x)$ is the weighting function at x, the smoothing parameter, such that $|x| < 1$ (Altman 1992). A smoothing parameter, α, is predefined as input by the data analyst, which value ranges from $\frac{(\lambda+1)}{n}$ to 1, with λ denoting the degree of the polynomial. The bandwidth value, α, represents the proportion of data used in each fit.

The smaller the value of α, the closer the polynomial regression function conforms to the data. However, if values for α are too small, insufficient data near the point x are used for estimation, which results in a large variance. If the values for α used are too large, then there will be over smoothing, which causes bias. Hence, a tradeoff between bias and variance is realized by choosing the value for the smoothing parameter. Useful values typically lie in the range of 0.25 to 0.5 for most LOESS applications (Cleveland 1981).

Materials and Methods

Three techniques were implemented in this study to evaluate the effect of smoothing the computed discharge to the discharge series quality. The steps followed are summarized in a flow chart shown in Figure 1.

Data Gathering

River discharge series computed using the velocity-area method of two river basins, Padada and Lipadas, were gathered for this study. These are lifted from the Phil-LiDAR 1 program, an expansion of the DREAM program that aims to produce 3D flood and hazard maps for the 2/3 of the Philippine river systems (UP DREAM Program 2014). For each river basin, a discharge

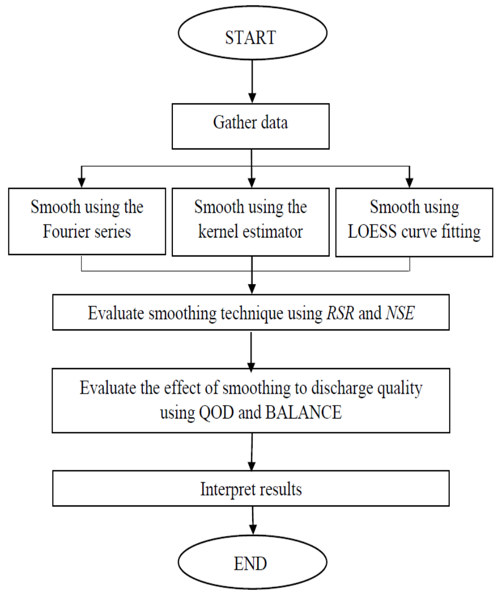

FIGURE 1 Flow chart of the research design

value in cubic meters per second was recorded every 10 minutes for a period of 15 days. The data series were derived from the product of the cross-sectional area and the average velocity gathered and pre-computed by the LiDAR 1 team.

The Lipadas River Basin is located in Toril, Davao City, while the Padada River Basin is located in Davao del Sur. For Lipadas River, data were gathered from 10 August 2015 at 16:30 to 25 August 2015 at 16:30. For Padada River, data were gathered from 24 September 2015 16:30 to 9 October 2015 at 3:00. A total of 2,161 data values were recorded for the Lipadas River Basin, and 2,073 data values for the Padada River Basin.

Evaluating the Smoothing Techniques

To evaluate the accuracy of our modeling techniques and to assess their predictive power as hydrological models, we use the RMSE-Observations Standard Deviation Ratio and

Nash-Sutcliffe Efficiency, respectively.

RMSE-Observations Standard Deviation Ratio (RSR**).** The root-mean-square error (RMSE), also known as the root-mean-square deviation (RMSD), is a widely used model accuracy measurement because of its technique that aggregates the magnitude of the errors in predictions (Willmott and Matsuura 2005). The RMSE has numerous applications and is used extensively especially in model calibration (Anderson and Woessner 1992; Nyarusanda 2011; Birgand et al. 2013).

In a method developed by Moriasi et al. (2007), the RMSE is standardized using the observations standard deviation ($STDEV_{obs}$). This method is the RMSE-observations standard deviation ratio (RSR) computed as the ratio of RMSE and $STDEV_{obs}$, and is given by the equation:

$$NSE = 1 - \frac{\sum_{t=1}^{T} \left(Q_0^t - Q_m^t\right)^2}{\sum_{t=1}^{T} \left(Q_0^t - \underline{Q}_0\right)^2} \tag{11}$$

where Y^{mean} is the mean of the observed discharges, Y_i^{sim} is the modeled (simulated) discharge, and Y_i^{obs} is the observed discharge at time i for $i = 1,2,3,\ldots,n$ and n is the number of data points.

Nash-Sutcliffe Efficiency (NSE)**.** The Nash-Sutcliffe efficiency (NSE) is a normalized statistic that determines relative magnitude of residual variance compared to measured data variances. It is defined as:

$$NSE = 1 - \frac{\sum_{t=1}^{T} \left(Q_0^t - Q_m^t\right)^2}{\sum_{t=1}^{T} \left(Q_0^t - \underline{Q}_0\right)^2} \tag{12}$$

where $\underline{Q}_0$ is the mean of the observed discharges, Q_m^t is the modeled (simulated) discharge, and Q_0^t is the observed discharge at time t (Nash and Sutcliffe 1970).

The Nash-Sutcliffe efficiency takes value of less than or equal 1, such that an efficiency of 1 denotes a perfect match of modeled to observed discharge. Essentially, the closer the value of NSE to 1, the better the model (Nash and Sutcliffe 1970). A negative value or a zero denotes an unacceptable performance. NSE is proposed for use mainly for assessing efficiency of discharge simulation models (Nash and Sutcliffe 1970; Xiong et al. 2005; Akhtar et al. 2009).

Table 1 summarizes the performance rating for the RSR and the NSE.

TABLE 1 General performance ratings for RSR and NSE (Lifted from Moriasi et al. 2007; UP DREAM Phil LIDAR 2014)

Performance rating	*RSR*	*NSE*
Very good (VG)	0.00 < *RSR* ≤ 0.50	0.75 < *NSE* ≤ 1.00
Good (G)	0.50 < *RSR* ≤ 0.60	0.65 < *NSE* ≤ 0.75
Satisfactory (S)	0.60 < *RSR* ≤ 0.70	0.50 < *NSE* ≤ 0.65
Unsatisfactory (U)	*RSR* < 0.70	*NSE* < 0.50

Use of QOD and BALANCE to Quantify Discharge Series Quality

To assess the quality of the smoothed discharge series, we used two quality functions, Quality of Discharge (QOD) and BALANCE (Booij et al. 2011).

QOD considers the slope of the hydrograph, a graph showing discharge rate versus time. This function is comparable to the NSE, and has strong similarity in the model (Booij et al. 2011).

The QOD is defined as:

$$QOD = 1 - \frac{\sum_{t=1}^{N} \left[Q_a(i) - Q_o(i)\right]^2}{\sum_{t=1}^{N} \left[Q_o(i) - \underline{Q}_o\right]^2} \tag{13}$$

where Q_a is the adapted discharge, Q_o is the original discharge, i is the time step and N is the total number of time steps. Like the NSE, the optimal value is 1 (Andréassian et al. 2001).

The second function, called BALANCE (Booij et al. 2011), considers the water BALANCE between the original and adapted discharge series. It is given by:

$$BALANCE = \frac{\sum_{i=1}^{N} \left[Q_a(i) - Q_o(i)\right]}{\sum_{i=1}^{N} \left[Q_o(i)\right]} \tag{14}$$

where Q_a is the adapted discharge, Q_o is the original discharge, i is the time step and N is the total number of time steps. The closer the value for BALANCE to zero is obtained, the better the model (Andréassian et al. 2001).

The quality functions QOD and BALANCE were applied to assess the quality of modeled discharge series and rainfall time series. It is usually compared to an objective function Y, defined to evaluate the power of the adapted model (Booij et al. 2011; Andréassian et al. 2001).

The objective function Y combines the NSE and the relative volume error (RVE), and defined as:

$$Y = \frac{NSE}{|RVE|} \tag{15}$$

where NSE is given in Eq. [12], and RVE is defined as:

$$RVE = \frac{\sum_{i=1}^{N} [Q_s(i) - Q_o(i)]}{\sum_{i=1}^{N} [Q_o(i)]} \tag{16}$$

where Q_s is the smoothed discharge at i, Q_o is the original discharge at i, and N is the total number of time steps (Akhtar et al. 2009).

The methods Fourier smoothing, Kernel smoothing, NSE, RSR, QOD and BALANCE were implemented using Scilab, an open source software for numerical computations (SAS 2015), while LOESS Curve Fitting was implemented using Paleontological Statistics, or PAST, an open source statistical software for scientific data analysis (Hammer 2015).

Results and Discussion

River Basins

The discharge values are relatively smaller in Lipadas than that of the Padada as shown on Figures 2 and 3. This is attributed to the difference in the size of the river basins with Lipadas river having a cross section area of 1.98 sq. m. while Padada is at 44.15 sq. m. The major hydrological characteristics of the two river basins are listed in Table 2.

Smoothing the Discharge Series

For the Fourier smoothing method, the number of harmonics p applied were 1, 2, 3, 4, 6, and 10. This was based on the study by Xiong et al. (2005) with a little adjustment of adding a p value of 10. The value of 10 was added to assess the sensitivity of the results to extreme values. Also based from the same study, a bandwidth v of 2, 5, 10, 20, 30, and 50 were applied for the Kernel smoothing using the Gaussian density function. The bandwidth v in units takes the number of ten-minute intervals accounted per run. Lastly, for the LOESS method, the values of α considered were 0.1, 0.25, 0.3, 0.4, 0.5, and 0.75. As a rule of thumb, favorable results are found from considering the values in the closed interval [0.25, 0.5] (Cleveland 1981), so the investigations were focused on that interval. To arrive with a more conclusive result, values of α outside that interval were also investigated. Tables 3 to 5 show the result of smoothing the discharge series of both rivers evaluated using NSE and RSR. Figures 4 to 9 display the plots of the smoothed discharge series.

Figure 4 and 5 shows that as the harmonics order p increases, the smooth curve produced by the Fourier Series becomes less smooth and the shape follows more closely to the shape of the original series. It seems that the local variations in the original series are reflected more as the order of the harmonics retained increases. This is accompanied with the increase in the value of NSE and decrease in the value of RSR, which indicates improving performance as seen in Table 3. This result coincides with that of the study by Xiong et al. (2005). However, Table 3 also shows that this method recorded unsatisfactory performances for the Lipadas dataset, only reaching a satisfactory performance when the harmonics p was set at $p=10$. These results indicate that the Fourier Smoothing may not be a favorable method for this type of dataset.

Figures 6 and 7 show that the smoothed discharge with a bandwidth v of lesser value follows more closely the shape of the original series. The graph of the smoothed series becomes smoother as the value of the bandwidth v increases. It follows that the increase in bandwidth v is accompanied with a gradual decrease in NSE and increase in RSR as seen in Table 4 indicating a decline in performance. The kernel method displayed better performance than the Fourier method. The curves resulting using the Gaussian kernel method imitate the original series more closely than the Fourier series method.

The LOESS Curve Fitting displayed the poorest performance when applied to the Lipadas dataset as seen in Figure 8. The values in Table 5 show unsatisfactory performances for all smoothing parameters explored, which indicate the poor choice of LOESS technique to smooth the Lipadas discharge series. However, the LOESS Curves for the Padada series shown in Figure 9 produced better results, even performing better

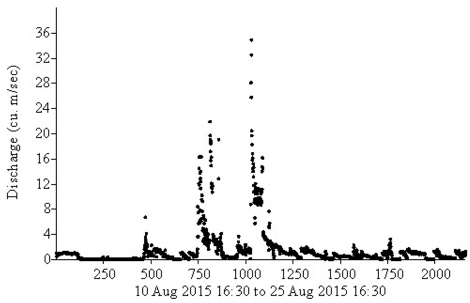

FIGURE 2 Discharge series of Lipadas River Basin

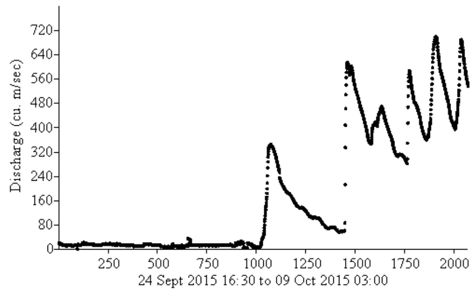

FIGURE 3 Discharge series of Padada River Basin

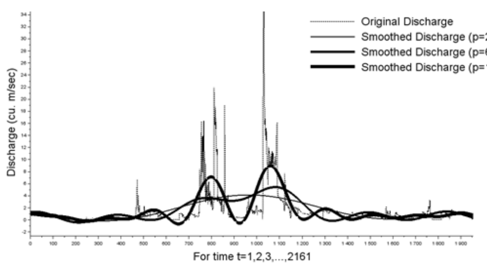

FIGURE 4 Smoothed discharge series of Lipadas River using Fourier Method

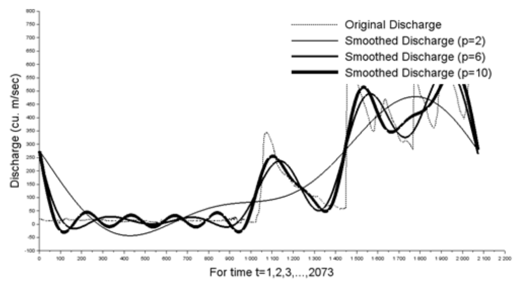

FIGURE 5 Smoothed discharge series of Padada River using Fourier Method

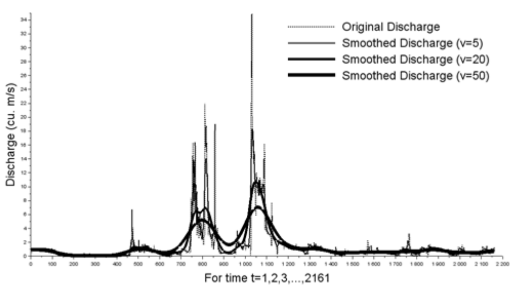

FIGURE 6 Smoothed discharge series of Lipadas River using Gaussian Kernel Method

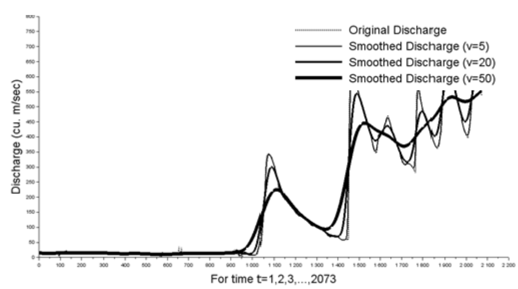

FIGURE 7 Smoothed discharge series of Padada River using Gaussian Kernel Method

TABLE 2 Summary of the hydrological data for Lipadas River Basin and Padada River Basin

River Basin	Rain Fall (mm/10 mins)	Velocity (m/sec)	Cross Section/ Area (sq.m)	Water Level/ Stage (m)	Discharge (cu. m/sec)
Lipadas	0.03165	0.61724	1.98304	2.30680	1.41302
Padada	0.05625	2.33302	44.14714	2.98311	175.98135

TABLE 3 Results of smoothing Lipadas River and Padada River discharge series using Fourier Smoothing method

Order p	Lipadas River		Padada River	
	NSE	*RSR*	*NSE*	*RSR*
1	0.14 (U)	0.93 (U)	0.58 (S)	0.65 (S)
2	0.23 (U)	0.88 (U)	0.67 (G)	0.57 (G)
3	0.24 (U)	0.87 (U)	0.70 (G)	0.55 (G)
4	0.25 (U)	0.87 (U)	0.73 (G)	0.52 (G)
6	0.30 (U)	0.84 (U)	0.84 (VG)	0.40 (VG)
10	0.54 (S)	0.68 (S)	0.87 (VG)	0.37 (VG)

TABLE 4 Results of smoothing Lipadas River and Padada River discharge series using Gaussian Kernel Smoothing method

Bandwidth v	Lipadas River		Padada River	
	NSE	*RSR*	*NSE*	*RSR*
2	0.93 (VG)	0.26 (VG)	0.9997 (VG)	0.02 (VG)
5	0.85 (VG)	0.39 (VG)	0.9979 (VG)	0.05 (VG)
10	0.75 (G)	0.50 (VG)	0.9918 (VG)	0.09 (VG)
20	0.63 (S)	0.61 (S)	0.9738 (VG)	0.16 (VG)
30	0.58 (S)	0.65 (S)	0.95199 (VG)	0.22 (VG)
50	0.48 (U)	0.72 (U)	0.9115 (VG)	0.30 (VG)

TABLE 5 Results of smoothing Lipadas River and Padada River discharge series using LOESS Curve Fitting method

Smoothing Parameter α	Lipadas River		Padada River	
	NSE	*RSR*	*NSE*	*RSR*
0.1	0.28 (U)	0.85 (U)	0.90 (VG)	0.32 (VG)
0.25	0.15 (U)	0.92 (U)	0.83 (VG)	0.42 (VG)
0.3	0.13 (U)	0.93 (U)	0.81 (VG)	0.44 (VG)
0.4	0.07 (U)	0.96 (U)	0.80 (VG)	0.45 (VG)
0.5	0.03 (U)	0.98 (U)	0.79 (VG)	0.45 (VG)
0.75	-0.02 (U)	1.01 (U)	0.80 (VG)	0.45 (VG)

than the Fourier smoothing technique based on the values of NSE and RSR as seen in Table 5. Still, the kernel smoothing technique resulted in better performance than both Fourier and LOESS techniques. The stark difference in the results between the two datasets indicate the sensitivity of this smoothing technique to the pattern in the dataset to be smoothed, which warrant attentiveness in using this method. In general, increasing the proportion of data used in each fit, as represented by the value of α, resulted in a decline in performance. As the value of α increases, the plot becomes smoother and eventually neglects the shape of the original series. After all, the polynomial regression function is closer to the data when α is smaller (Cleveland 1981).

Discharge Quality of the Smoothed Series

The three best performances for each method were evaluated. Figures 10 and 11 plot the result for the Lipadas and Padada discharge series, respectively. The ideal fit, taking the shape of "*", is plotted where the values of NSE and the objective function Y is set to 1, and the value of BALANCE is set to 0.

Figure 10 shows that the power of the smoothing technique greatly affects the discharge quality for the Lipadas Series. The extent of its effect is considerably lesser in the Padada Series where the results are clumped together as seen in Figure 11. It is, therefore, important to study the pattern of the data to better extract information from the series when applying a smoothing technique. Both Figures 10 and 11 indicate that models that give better performances also produce a higher quality of discharge and vice versa.

The values observed for the QOD are predictable and follow the general pattern of the previous observations. The value for BALANCE, however, deserves attention. Even though the Fourier method produced an unsatisfactory result when evaluated using NSE and QOD, excellent values of BALANCE are recorded to be as low as approximately 0%. This means that the deviation

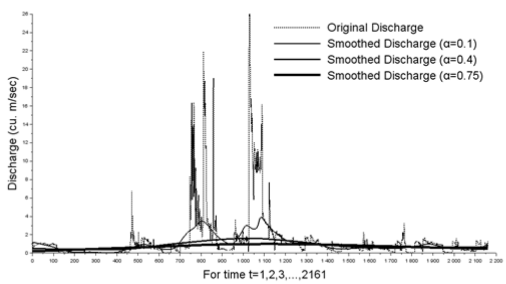

FIGURE 8 Smoothed discharge series of Lipadas River using LOESS Curve Fitting Method

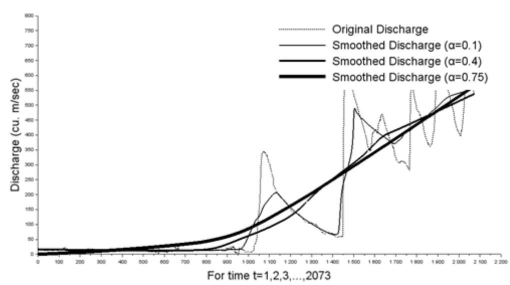

FIGURE 9 Smoothed discharge series of Padada River using LOESS Curve Fitting Method

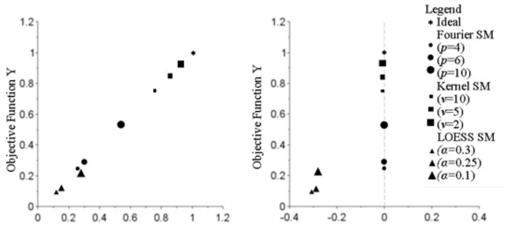

FIGURE 10 Relation between quality functions Quality of Discharge (QOD) and BALANCE and objective function Y for Lipadas discharge series

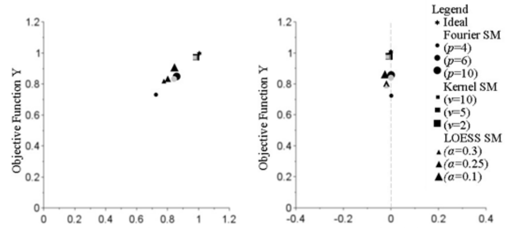

FIGURE 11 Relation between quality functions Quality of Discharge (QOD) and BALANCE and objective function Y to Padada discharge series

of the smoothed to the original discharge series relative to the volume of the original series is insignificant. This is reflected in most of the results, with the exception of LOESS Curve Fitting applied to Lipadas.

On another note, the values obtained using the same method generally fall closely together. This is more observable in Figure 10. Marks in this figure of the same shape clump together and are distinguishable from the other marks. This is evidence that even though changing the smoothing parameter increases the performance of the method, it is not enough to overpower a better fit method. Thus, choice of method is important. Specifically, it is obvious that the Gaussian Kernel method outdo the other methods, even when different extent of the parameters of the other methods are evaluated.

Comparing Smoothed Discharge with Discharge Computed using HEC-HMS

The HEC-HMS smoothed discharge series only accounted for the period from 13 August 2015 at 12:00 to 14 August 2015 at 23:20. Only a subset of the data previously evaluated were used amounting to 213 data points. The values of NSE and RSR were evaluated for the Gaussian Kernel method and were compared to the result obtained from the HEC-HMS smoothed method. Table 6 shows that the HEC-HMS smoothed discharge only gave a satisfactory performance when evaluated using NSE and RSR while the Gaussian Kernel smoothed series gave a very good performance. This outcome is also observable in Figure 12. The plot of the Gaussian Kernel smoothed series follows more closely to the original series than the HEC-HMS smoothed series. This result strengthens this study's conclusion on the strength of the Gaussian kernel smoothing method applied to the discharge series.

Conclusions and Recommendations

Improved performance is observed when there is an increase in the value of harmonics order p for the Fourier method, and a decrease in the value of bandwidth v for the Gaussian Kernel method, as well as the value of the smoothing parameter α in the LOESS method.

Table 6. Summary of the performance of HEC-HMS smoothed vs Gaussian Kernel smoothed series with evaluated using Nash-Sutcliffe Efficiency (NSE) and RMSE Observations Standard Deviation Ratio (RSR)

Smoothing method	Statistic	
	NSE	*RSR*
HEC-HMS	0.60 (S)	0.64 (S)
Gaussian Kernel with		
v = 2	0.83 (VG)	0.41 (VG)

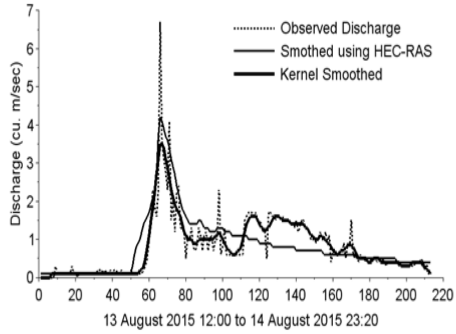

FIGURE 12 Smoothed discharge series obtained from HEC-HMS plotted against the Gaussian Kernel smoothed discharge with for Lipadas River

The models that gave very good performances would also produce a higher quality of discharge. The weaker method would result in an unacceptable quality of discharge series. The result of the two quality functions, QOD and BALANCE, indicate that simply changing the smoothing parameter of a smoothing method does not outperform another better fit smoothing method. Therefore, the choice of method used is important.

Consolidating the results of all evaluation techniques performed, the Gaussian Kernel method significantly outperforms the Fourier and LOESS method. The choice of bandwidth v at the ball park is around 2 to 10. If the Gaussian Kernel method performs excellently, a very good quality of smoothed discharge is expected.

However, having considered only two river basins for this study, it may be argued that further research on more river basins may be needed to explore the smoothing power of the Gaussian

Kernel. It would also be significant to define a technique in identifying a more specific value of the smoothing parameter or bandwidth ν that will not result in undersmoothing or over smoothing of the discharge series.

Acknowledgment

The author would like to thank UP Mindanao Phil – LiDAR 1 for the LiDAR DSM and LiDAR derived datasets, Department of Science and Technology - Philippine Council for Industry, Engineering and Emerging Technology Research and Development (DOST-PCIEERD) and University of the Philippines Mindanao for the facilities and funding.

References

AKHTAR M, AHMAD N, BOOIJ MJ. 2009. Use of regional climate model simulations as input for hydrological models for the Hindukush-Karakorum-Himalaya region. Hydrol. Earth Syst. Sci. 13: 1075–1089.

ALTMAN NS. 1992. An introduction to kernel and nearest neighbor nonparametric regression. Am. Stat. 46(3): 175–185.

ANDERSON MP, WOESSNER WW. 1992. Applied groundwater modeling: simulation of flow and advective transport. 2nd ed. Massachusetts, United States of America: Academic Press.

ANDRÉASSIAN V, PERRIN C, MICHEL C, USART-SANCHEZ I, LAVABRE J. 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. J. Hydrol. 250: 206–223.

BIRGAND F, LELLOUCHE G, APPELBOOM TW. 2013. Measuring flow in non-ideal conditions for short-term projects: uncertainties associated with the use of stage-discharge rating curves. J. Hydrol. 503: 186–195.

BOOIJ MJ, VAN DEN TILLAART SPM, KROL MS. 2011. Risks in hydrological modelling due to uncertainties in discharge determination. In: Bloschl G, Takeuchi K, Jain S, Farnleitner A, Schumann A, editors. Risk in water resources management. Proceedings of Symposium H03 at IUGG2011; 2011 28 June-7 July; Melbourne. Melbourne: IAHS Publ. 347. p. 95–100 .

BOWEN A, PALLISTER J. 2000. AS level geography. Portsmouth, United States of America: Heinemann.

CLEVELAND WS. 1981. LOWESS: A program for smoothing scatterplots by robust locally weighted regression. Am. Stat. 35(1): 54.

DI BALDASSARRE G, MONTANARI A. 2009. Uncertainty in river discharge observations: A quantitative analysis. Hydrol. Earth Syst. Sci. 13(6): 913–921.

DI BALDASSARRE G. 2012. Floods in changing climate: inundation modeling. Cambridge, United Kingdom: Cambridge University Press.

DOMENEGHETTI A, CASTELLARIN A, BRATH A. 2012. Assessing rating-curve uncertainty and its effects on hydraulic model calibration. Hydrol. Earth Syst. Sci. 16: 1191–1202.

ESOMBA R. 2015. Smoothing techniques using basis functions: Fourier basis. R-bloggers [Internet]. [accessed 2015 December 01]. http://www.r-bloggers.com/smoothing-techniques-using-basis-functions-fourier-basis/.

FAUCHER D, RASMUSSEN PF, BOBE´E B. 2001. A distribution function based bandwidth selection method for kernel quantile estimation. J. Hydrol. 250(1–4): 1–11.

FOX J., 2002. Nonparametric regression: Appendix to an R and S-PLUS companion to applied regression. California, United States of America: Sage Publications.

GUERRERO JL, WESTERBERG IK, HALLDIN S, XU CY, LUNDIN LC. 2012. Temporal variability in stage–discharge relationships. J. Hydrol. 446–447: 90–102.

HAMMER O. 2015. Past 3.x: the past of the future. PAST. [accessed 2015 November 19]. http://folk.uio.no/ohammer/past/.

HERSCHY W. 2002. The uncertainty in a current meter measurement. Flow Meas. Inst. 13: 281–284.

[HEC] HYDROLOGIC ENGINEERING CENTER. 2000. Hydrologic modeling system HEC-HMS technical reference manual. California, United States of America: U.S. Army Corps of Engineers.

[HEC] HYDROLOGIC ENGINEERING CENTER. 2001. HEC-RAS hydraulic reference manual. California, United States of America: US Army Corps of Engineers.

[ISO] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) 748. 2007. Hydrometry — Measurement of liquid flow in open channels using current-meters or floats. https://www.iso.org/standard/37573.html

[ISO] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) 1100–2. 2010. Hydrometry—measurement of liquid flow in open channels—part 2: determination of the stage-discharge relationship. https://www.iso.org/standard/42207.html

JAYAWARDENA AW, PERERA EDP, ZHU B, AMARASEKARA JD, VEREIVALU V. 2014. A comparative study of fuzzy logic systems approach for river discharge prediction. J. Hydrol. 514: 85–101.

KIMBALL BA. 1974. Smoothing data with Fourier transformations. Agro. J. 66: 259–262.

LI Q, RACINE JS. 2007. Nonparametric econometrics: theory and practice. New Jersey, United States of America: Princeton University Press.

MORIASI DN, ARNOLD JG, VAN LIEW MW, BINGNER RL, HARMEL RD, VEITH TL. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE. 50(3): 885–900.

NADARAYA EA. 1994. On estimating regression. Theory Probab. Its Appl. 9(1): 141–2.

NASH JE, SUTCLIFFE JV. 1970. River flow forecasting through conceptual models part I—a discussion of principles. J. Hydrol. 10(3): 282–290.

NYARUSANDA L. 2011. Calibration of LiDAR histogram databases using a non-linear mathematical model. Tanzania: University of Dar es Salaam.

PEKAROVA P, MIKLANEK P, PEKAR J. 2006. Long-term trends and runoff fluctuations of European rivers. Climate Variability and Change—Hydrological Impacts. Proceedings of the Fifth FRIEND World Conference; Havana, Cuba. Havana, Cuba: IAHS Publication 308. p. 520–525.

[SAS] SCILAB ENTERPRISES. 2015. About Scilab [Internet]. Scilab Enterprises . [accessed 2015 November 19]. http://www.scilab.org/scilab/about

TAKEZAWA K. 2005. Introduction to nonparametric regression. New Jersey, United States of America: John Wiley & Sons.

UP DREAM PROGRAM. 2014. Phil-LiDAR 1 [Internet]. Disaster Risk and Exposure Assessment for Mitigation (DREAM). [accessed 2015 November 15]. https://dream.upd.edu.ph/about-us/phil-lidar-1/

U.S. GEOLOGICAL SURVEY. 2015. How streamflow is measured, part 1: Mmeasuring stream stage [Internet]. U.S. Geological Survey. [accessed 2015 November 14]. http://water.usgs.gov/edu/streamflow1.html

VERT JP, TSUDA K, SCHÖLKOPF B. 2004. Kernel methods in computational biology. Massachusetts, United States of America: Massachusetts Institute of Technology (MIT) Press.

WATSON GS. 1964. Smooth regression analysis. Ind. J. Stat. 26(4): 359–372. https://www.jstor.org/stable/25049340

WILLMOTT C, MATSUURA K. 2005. Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in assessing average model performance. Clim. Res. 30: 79–82.

XIONG L, GUO S, O'CONNOR KM. 2005. Smoothing the seasonal means of rainfall and runoff in the Linear Perturbation Model (LPM) using the kernel estimator. J. Hydrol. 324: 266–282.