Water Resour Manage (2015) 29:2805–2818
DOI 10.1007/s11269-015-0971-5

Quantile Regression-Based Probabilistic Estimation Scheme for Daily and Annual Suspended Sediment Loads

Jenq-Tzong Shiau · Ting-Ju Chen

Received: 28 August 2014 / Accepted: 9 March 2015 /
Published online: 19 March 2015

Abstract Assessing suspended sediment loads in rivers is important since it affects water quality, hydraulic-facility design, and many other sediment-induced problems. Sediment-load estimation heavily depends upon empirical approaches such as a sediment rating curve, which is the empirical relationship between sediment load and river discharge. However, the sediment rating curve is insufficient to describe the inevitable scatter between sediment and discharge. This study aims to develop a probabilistic estimation scheme for daily and annual suspended sediment loads using quantile regression. All recorded daily suspended sediment load and discharge data are employed to construct quantile-dependent sediment rating curves. The empirical probability distribution of daily suspended sediment load is then built by integrating the conditional estimations associated with the corresponding quantiles for a given discharge. The probability distribution of a cumulative sediment load over a longer period can also be derived by the obtained daily sediment-load probability distributions and convolution theorem. The proposed approach is applied to the Laonung station located in southern Taiwan. The results indicate that the proposed approach provides not only the probabilistic description for daily and annual suspended sediment loads, but also the single estimations including the mean, median, and mode of the derived probability distribution. For the 1,110 recorded data of Laonung station during the 1959–2008 period, the proposed mean and median estimation schemes outperform the traditional sediment-rating-curve approach for less mean absolute errors.

Keywords Suspended sediment load · Sediment rating curve · Quantile regression

1 Introduction

Estimating sediments carried and transported by rivers is an important and essential issue in water-resources planning, design, and management since the information is required for reservoir sedimentation evaluation, hydraulic structures design, water-quality criterion and treatment costs assessment, ecological and recreation consideration, efficacy of watershed management, stable channels for navigation, and many other sediment-induced problems.

J.-T. Shiau (✉) · T.-J. Chen
Department of Hydraulic and Ocean Engineering, National Cheng Kung University, Tainan 701 Taiwan, Republic of China
e-mail: jtshiau@mail.ncku.edu.tw

Infrequent or periodically sampling of sediments, which is frequently met in most river basins, is insufficient to define a continuous record of sediment load and fail to determine the cumulative sediment load for a longer period. Sediment-load estimation thus heavily depends upon empirical approaches such as a sediment rating curve, which describes the relationship between sediment load and river discharge, to achieve these purposes. Although some researchers (e.g., Walling, 1977; McBean and Al-Nassri, 1988; Cohn et al., 1992; Kişi, 2005) had indicated that the regression and curve-fitting techniques are not adequate to describe the high degree of scatter existed between sediment load and discharge data, the sediment rating curve remains a popular and widely used approach to estimate sediment loads due to available streamflow data and calculation simplicity.

Many works had suggested various methods to improve predictive accuracy of the sediment rating curve because of its poor predictive performances. These methods include bias adjustment or bias correction factor (Ferguson, 1986; Cohn et al., 1989; Phillips et al., 1999; Asselman, 2000), multivariate rating curve (Cohn et al., 1992; Wang and Linker, 2008; Wang and Tian 2013), and different regression techniques (Hicks et al., 2000; Krishnaswamy et al., 2001; Tarras-Wahlberg and Lane, 2003). More recently, a considerable amount of studies have applied newly developed methods to construct the empirical relationship between sediment and discharge. These approaches include artificial neural network (Jain, 2001; Nagy et al., 2002; Alp and Cigizoglu, 2007; Rai and Mathur, 2008; Afan et al., 2015), genetic programming (Aytek and Kişi, 2008; Guven and Kişi, 2011; Kitsikoudis et al., 2014), fuzzy logic (Kişi, 2004; Lohani et al., 2007), wavelet−neural network (Partal and Cigizoglu, 2008), and support vector machines (Çimen, 2008; Lafdani et al., 2013).

Inherent complexity in the sediment transport process induces an inevitable scatter between sediment and discharge. Due to few samplings during high-flow periods, suspended sediment load estimation provided by a sediment rating curve for such periods has greater deviations generally. Besides, estimation errors may propagate from a short-period (daily) estimation to determine the cumulative load for a longer period (e.g., a year) using the sediment-rating-curve approach. Providing uncertainty information is as important as accurate sediment load estimation in water resources planning and management. Although a vast literature had proposed various approaches to improve estimation accuracy, uncertainty in sediment load estimation from the sediment rating curve is rarely considered. The uncertainty-related studies include McBean and Al-Nassri (1988) had attempted to establish a confidence interval of the mean suspended sediment concentration versus discharge. Clarke (1990a, 1990b) had discussed statistical characteristics such as bias and variance under the assumption of bivariate lognormal distribution of suspended sediment concentration and mean daily discharge. Rustomji and Wilkinson (2008) had used the bootstrap associated with Monte Carlo resampling techniques for estimating suspended sediment loads in order to quantify the uncertainty in the shape of the sediment rating curve. Wang et al. (2011) had adopted a generalized rating-curve approach with additional predictors to provide reliable load estimates with minimal bias and an associated measure of uncertainty. Vigiak and Bende-Michl (2013) had used bootstrap and Bayesian inference in estimating prediction intervals for daily and monthly constituent using improved eight-parameter rating curve.

This study aims to develop a probabilistic estimation scheme for daily and annual suspended sediment loads using quantile regression. Instead of constructing a single sediment rating curve at a specific quantile, a novel approach is proposed to derive a probabilistic distribution of daily suspended sediment loads through dense quantile-dependent sediment rating curves. Deriving probabilistic distribution of the cumulative suspended sediment load for a longer period such as a year is thus made possible by the derived daily suspended-sediment-loads distributions and utilization of convolution theorem. Advances of quantile regression for deriving probabilistic distributions of daily and annual suspended sediment loads are still novelties in literature. The proposed approach is applied to the Laonung station located in southern Taiwan for demonstration.

The remainder of this paper is organized as follows. Section 2 presents an introduction of sediment rating curve, quantile regression, and the approach for constructing probability distributions of daily and annual suspended sediment loads. Section 3 summaries the recorded daily data at the Laonung station located in southern Taiwan for the 1959–2008 period. Results and discussions are given in the Section 4, which is followed by the Section 5 of summaries and conclusions.

2 Methodology

2.1 Sediment Rating Curve

The sediment rating curve denotes the empirical relationship between measured sediment load and river discharge. The power law function is commonly adopted to model this relationship, which is expressed as

$$Q_s = aQ^b \tag{1}$$

where Q_s denotes the suspended sediment load; Q denotes the river discharge; a and b are constants which are estimated from measured data. The nonlinear least squares regression (Asselman, 2000) is employed in this study to obtain sediment rating curve constants in order to avoid the need of applying a bias correction factor to the estimated loads (Rustomji and Wilkinson, 2008).

The obtained suspended sediment rating curve transforms the recorded daily streamflow series into a sequence of daily suspended sediment load estimations. The cumulative suspended sediment load over a longer period such as a month or a year is thus calculated as the sum of daily suspended sediment load over such period. That is,

$$L_n = \sum_{i=1}^{n} \widehat{Q}_{s,\,i} = \sum_{i=1}^{n} aQ_i^b \tag{2}$$

where L_n denotes the cumulative suspended sediment load over an n-day period; $\widehat{Q}_{s,\,i}$ denotes the estimated suspended sediment load of the ith day through sediment rating curve; Q_i is the ith-day discharge; n denotes the number of days.

2.2 Quantile Regression

Quantile regression, initiated by Koenker and Basset (1978), is a statistical technique for regression on various quantiles rather than regression only on the mean. Since quantile regression can capture the variability of all parts of data distribution, it recently has been used in a broad range of applications such as economics and finances (Meligkotsidou et al., 2009; Gaglianone et al., 2011; Alagidede and Panagiotidis, 2012), biology and ecology (Cade and Noon, 2003; Austin, 2007; Cozzoli et al., 2013), and environmental science (Jagger and Elsner, 2009; Hirschi et al., 2011; Barbosa et al., 2011; Shiau and Huang, 2015). Detailed description of the theory of quantile regression can be found in Koenker (2005).

The quantile regression-based power-law sediment rating curves for various quantiles are expressed as

$$Q_{s,\,q} = a_q Q^{b_q} \tag{3}$$

where q is quantile which is constrained between 0 and 1; $Q_{s,\,q}$ is a conditional suspended sediment load at quantile q given a discharge of Q; a_q and b_q are quantile-dependent regression coefficients, which are determined by the selected quantile q and minimizing the sum of asymmetrically weighted absolute deviations, i.e.,

$$\min \sum_{i:\,Q_{s,\,i} \geq a_q Q_i^{bq}} q\left|Q_{s,\,i} - a_q Q_i^{b_q}\right| + \sum_{i:\,Q_{s,\,i} < a_q Q_i^{bq}} (1-q)\left|Q_{s,\,i} - a_q Q_i^{b_q}\right| \tag{4}$$

There is no analytic solution for the quantile-dependent regression coefficients of a_q and b_q. The minimization problem can be formulated as a linear program and solved by an algorithm developed by Koenker and D'Orey (1987). Calculations in this study are carried out using a free-access software *quantreg* (Koenker, 2014). Instead of obtaining a single estimation of suspended sediment load for a given discharge, quantile regression-based model estimates suspended sediment loads at various quantiles for a given discharge. Examining entire distribution of suspended sediment load for a given discharge is thus made possible using quantile regression.

2.3 Probability Distribution of Daily Suspended Sediment Loads

All the paired sediment-discharge recorded data are used to construct suspended sediment rating curves at various quantiles (quantiles of 0.95, 0.9, 0.7, 0.5, 0.3, and 0.1 demonstrated in Fig. 1a). These dense quantile-dependent sediment rating curves are employed in order to have a finer resolution in probability distribution. Given a daily discharge of Q, the conditional suspended sediment $\widehat{Q}_{s,\,q}$ for each quantile-dependent sediment rating curve is calculated (various discharges of Q_1 and Q_2 demonstrated in Fig. 1a). These estimated conditional $\widehat{Q}_{s,\,q}$ associated with the corresponding quantile q constitute a rough estimation of the empirical cumulative distribution function (CDF) of the daily suspended sediment load (CDFs of Q_1 and Q_2 demonstrated in Fig. 1b). In this study, suspended sediment load is considered as a discrete random variable, the empirical probability mass function (PMF) is then constructed using the obtained CDF (PMFs of Q_1 and Q_2 demonstrated in Fig. 1c). Figure 1 illustrates the entire procedure for deriving empirical CDFs and PMFs of daily suspended sediment load given various discharges from estimated quantile regression-based sediment rating curves. The obtained PMF describe the probabilities of different suspended-sediment-load intervals for a given discharge, which offers rich-information than the single estimation obtained from the traditional sediment rating curve described by equation (1). The derived quantile regression-based suspended sediment rating curves thus transform the recorded daily streamflow series into a sequence of PMFs of daily suspended sediment load.

2.4 Probability Distribution of Annual Suspended Sediment Loads

The probabilistic description of cumulative suspended sediment loads for a year period can be constructed by the obtained PMFs of daily suspended sediment load and convolution theorem, which is described below.

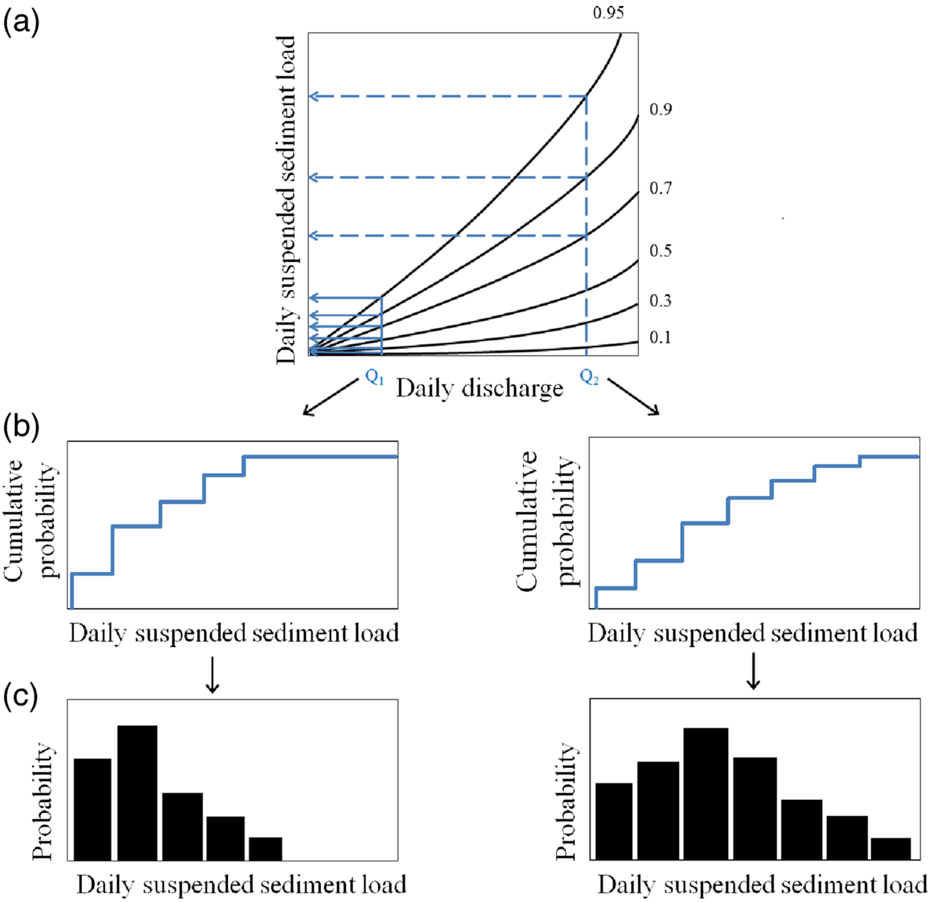

Fig. 1 Schematic illustration of constructing probabilistic distributions of daily suspended sediment loads, (**a**) quantile-dependent suspended sediment rating curves, (**b**) empirical cumulative distribution function (CDF) of various discharge Q_1 and Q_2, and (**c**) empirical probability mass function (PDF) of discharge Q_1 and Q_2

Let S_1 and S_2 be the discrete random variables which represent the daily suspended sediment loads for day 1 and day 2, respectively. The derived PMFs of S_1 and S_2 are denoted by $f_{S_1}(s_1)$ and $f_{S_2}(s_2)$, respectively. The cumulative suspended sediment load of these two successive days L_2 is the sum of S_1 and S_2, i.e., $L_2=S_1+S_2$. Under the assumption of independence between S_1 and S_2, the PMF of L_2 can be derived by the convolution theorem (Ross, 2007). That is,

$$f_{L_2}(l_2)=\sum_{s_1=0}^{\infty} f_{S_1}(s_1) f_{S_2}(l_2-s_1) \tag{5}$$

where $f_{L_2}(l_2)$ denotes the PMF of L_2.

Repeating the procedure again, the PMF of cumulative suspended sediment load for three successive days can be derived. Let $L_3=S_1+S_2+S_3=L_2+S_3$. After deriving the PMF of L_2, convolution is implemented between L_2 and S_3 to obtain the PMF of L_3. Considering the recursive formula of $L_n=L_{n-1}+S_n$, where $n=2, 3, \ldots, N$, the above procedure can be used to construct the PMF of the annual suspended sediment load when $N=365$.

3 Data Used

The recorded daily suspended sediment load and river discharge data of Laonung station located in southern Taiwan are used in this study to demonstrate the proposed approach. River discharge and suspended sediment load data have been measured by Water Resource Agency in Taiwan. The daily streamflow sequence analyzed in this study has been collected between years 1959 and 2008. With a drainage area of 812 km^2 and mean annual rainfall of approximate 3,600 mm, the mean daily streamflow of Laonung station of the 1959−2008 period is 85.2 m^3/s, which is ranged between the maximum daily streamflow of 4,090 m^3/s and the minimum one of 0.35 m^3/s. However, infrequent sampling of daily suspended sediment load leading to the number of recorded suspended sediment load is approximate 6 % of recorded discharge data. A total of 1,110 paired sediment-discharge measured data are used to construct the empirical relationship between sediment load and river discharge, which are shown in Fig. 2.

Approximate 75 % of these 1,110 paired data are collected below mean daily discharge of 85.2 m^3/s, while only 17 daily suspended sediment loads are collected during the greater-than-1,000 m^3/s period. The mean value of these recorded daily suspended sediment load is 7.32×10^4 ton, which is ranged between 9.02 million ton (for measured discharge of 1,820 m^3/s) and 0.65 ton (for measured discharge of 10.8 m^3/s). Table 1 summaries the basic statistics of these recorded data of Laonung station. Although the recorded data shown in Fig. 2 illustrate a positive correlation, considerable scatter at the high-flow stage between recorded data cannot be ignored.

4 Results and Discussions

4.1 Estimated Daily and Annual Suspended Sediment Loads by Sediment Rating Curve

The suspended sediment rating curve of Laonung station is derived for the measured 1,110 paired sediment-discharge daily data of the 1959−2008 period. The obtained power-law rating curve is described below and also shown in Fig. 2.

$$\widehat{Q}_s = 2.22 \times 10^{-4} \times Q^{1.8819} \tag{6}$$

where $\widehat{Q}_s$ denotes the estimated daily suspended sediment load in 10^4 ton per day; Q denotes the daily river discharge in m^3/s.

Inherent scatter between paired sediment-discharge measured data shown in Fig. 2 induces estimation errors when the derived sediment rating curve is employed to estimate daily suspended sediment load for a given discharge. The mean absolute error between this set of 1,110 measured and estimated daily suspended sediment load is approximate 5.1×10^4 ton. However, the discharge of 1,820 m^3/s occurred in Aug. 9, 1994 leads to the maximum estimation error of 599.2×10^4 ton, which is caused by the difference between measured value of 902.2×10^4 ton and estimation of 303.0×10^4 ton. Estimation error is generally declined with decreasing discharge. For example, a mean absolute error of 1.6×10^4 ton is obtained when river discharge is below 500 m^3/s, while this value is increased to 56.8×10^4 ton for the condition of discharge exceeding 500 m^3/s. It is worth to note that the discharge of 1,840 m^3/s occurred in Jul. 31, 1982 carries a suspended sediment load of 192.9×10^4 ton and has an estimation of 309.3×10^4 ton using the derived sediment rating curve. This overestimation of 116.4×10^4 ton for 1,840 m^3/s associated with the underestimation of 599.2×10^4 ton caused by

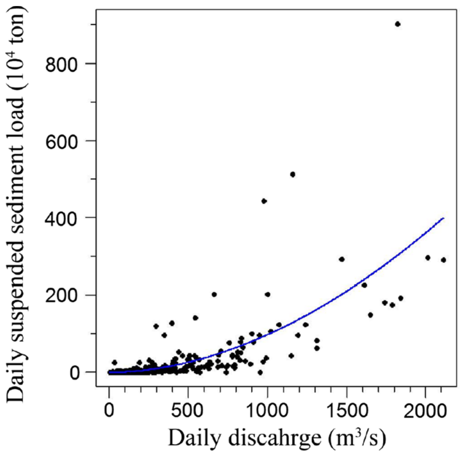

Fig. 2 Scatter plot of recorded daily suspended sediment load and discharge data of Laonung station, Taiwan for the 1959−2008 period and the corresponding suspended sediment rating curve

1,820 m^3/s reveal that an estimation error is inevitable and great uncertainty is existed in estimations even though the very close discharges.

The derived sediment rating curve is used to estimate the sequence of 18,263 daily suspended sediment load for the 1959–2008 period through recorded daily discharge. The yearly cumulative suspended sediment load is thus obtained by summing the estimated daily suspended sediment loads over the yearly period. Figure 3 illustrates the estimated yearly suspended sediment load and the mean daily discharge of Laonung station for the 1959–2008 period. Generally greater annual streamflow results in greater annual suspended sediment load. However, the maximum annual suspended sediment load and the maximum mean daily discharge do not occur in the same year. This phenomenon is attributed to the nonlinear power-law feature of the empirical sediment rating curve. The maximum annual suspended sediment load of 26.3 million ton occurred in 1960 is caused by three daily discharges

Table 1 Summaries of basic statistics of recorded daily suspended sediment load and discharge data of Laonung station, Taiwan for the 1959−2008 period

Station		Laonung
Drainage area (km^2)		812
Time period		1959−2008
Number of discharge data		18,263
Number of suspended sediment data		1,110
Daily discharge (m^3/s)	Mean	85.2
	Standard deviation	163.7
	Max.	4,090
	Min.	0.35
Daily suspended sediment load (10^4 ton)	Mean	7.32
	Standard deviation	42.1
	Max.	902.2
	Min.	6.5×10^{-5}

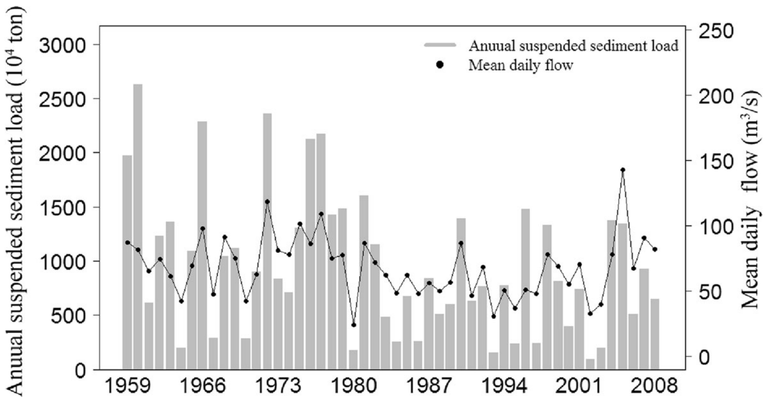

Fig. 3 Annual suspended sediment load estimated by sediment rating curve and the time series plot of mean daily discharge of Laonung station, Taiwan for the 1959−2008 period

exceeding 1,000 m^3/s. The mean daily streamflow of 143.1 m^3/s occurred in 2005 is ranked first during the 1959–2008 period. However, daily streamflow all below 500 m^3/s in 2005 leads to an annual suspended sediment load of 13.4 million ton, which is ranked 14th within the 50-year period. It is hard to quantify the uncertainty of the estimated yearly suspended sediment loads without measured data, although the uncertainty actually exists.

4.2 Probabilistic Estimation of Daily Suspended Sediment Loads

To quantify the estimation uncertainty of daily suspended sediment load, a probabilistic estimation scheme using quantile regression is proposed in this study. The regression constants of power-law quantile regression model are estimated first at dense quantiles which range from 0.01 to 0.99 with an increment of 0.01. Figure 4 shows the measured daily sediment-discharge data associated with the derived quantile-dependent sediment rating curves at quantiles of 0.9,

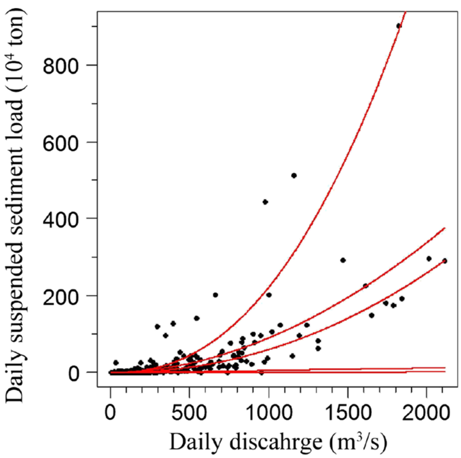

Fig. 4 Recorded daily suspended sediment load and discharge data of Laonung station, Taiwan for the 1959−2008 period associated with quantile-dependent suspended sediment curves at quantiles of 0.9, 0.7, 0.5, 0.3, and 0.1 (from top to bottom)

0.7, 0.5, 0.3, and 0.1 from top to bottom for illustration. For a given discharge, the conditional suspended sediment loads are calculated for each quantile and then integrated to obtain the empirical CDF and PMF for that discharge. The PMF is derived in a fine interval of 1×10^4 ton in order to have a finer resolution in probability distribution.

Left panel of Fig. 5 shows the obtained CDFs of daily suspended sediment loads given that discharge respectively equaling to 250, 496, and 998 m^3/s for demonstration. The corresponding PMFs for these three discharges are shown in the right panel of Fig. 5. Probabilistic description of various sediment intervals are evidently provided by the obtained PMFs. For instance, daily suspended sediment load for the discharge of 250 m^3/s ranges between 0 and 60×10^4 ton, which is shown in upper right panel of Fig. 5. However, a highly uneven and right-skewed PMF is observed in this figure. Daily suspended sediment load with a cumulative probability of 0.94 is distributed between 0 and 13×10^4 ton and the highest probability of 0.27 occurs in 0.5×10^4 ton (interval of $0-1\times10^4$ ton). This obtained PMF of daily suspended sediment load is useful to explain why the recorded daily suspended sediment loads for 250 m^3/s are different in three different days. That is, 32.4, 1.8, and 4.9×10^4 ton respectively occur in Jul. 14, 1961, Aug. 10, 1961, and Aug. 12, 1969. The PMFs of daily suspended sediment load for 496 and 998 m^3/s, respectively occurred in Jun. 12, 1995 and Aug. 6, 1962, also behave a right skewness.

The box-plot showing the minimum (first percentile), first quartile, medium (second quartile), third quartile, and the maximum (99th percentile) estimations of daily suspended sediment load for a given discharge is a useful approach to illustrate the probability distribution and compare the range of distribution for various discharges. Figure 6 demonstrates the box-plot of daily suspended sediment load for various discharges. It is worth to note that these distributions are highly right-skewed, which are in line with the quantile-dependent sediment rating curves shown in Fig. 4. That is, lower-quantiles sediment rating curves are denser than the higher-quantiles sediment rating curves. The increasing range and interquartile range with increasing discharge exhibit greater uncertainty for higher streamflow. For example, the interquartile range increases from 23.4 to 88.0×10^4 ton when discharge increases from 500 to 1000 m^3/s. This value further increases to 338.7×10^4 ton when discharge reaches 2000 m^3/s. The greater range of daily suspended sediment load for higher discharge also indicates a

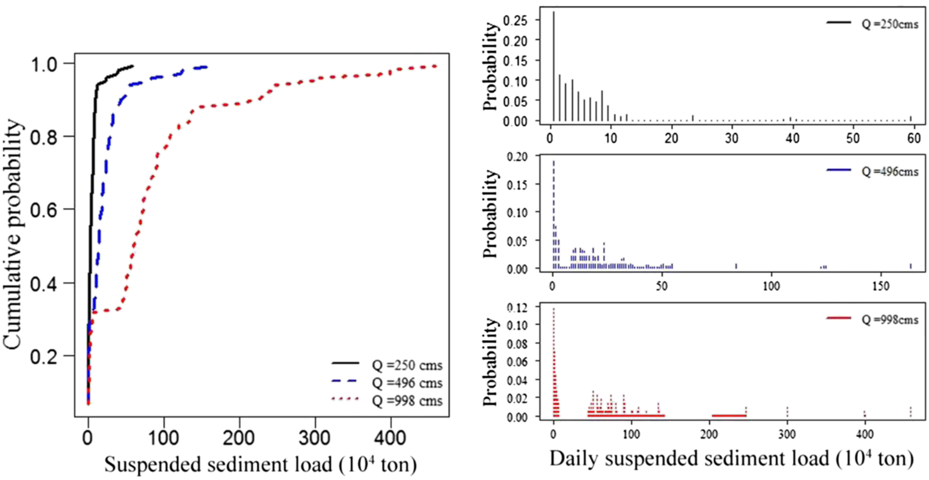

Fig. 5 Derived cumulative distribution function (CDF) and probability mass function (PMF) for the given discharge of 250, 496, and 998 m^3/s

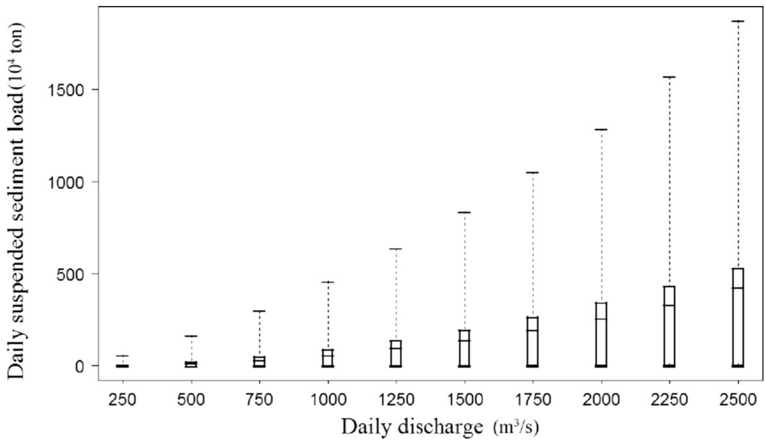

Fig. 6 Box-plot of daily suspended sediment load given various discharges for Laonung station

possibility of extreme events. For instance, the maximum daily suspended sediment load for 2,000 m^3/s is 1,286 × 10^4 ton. Although the extreme value is associated with a small probability, it is a potential threat to damage hydraulic-facility operation. Incorporating the risk of such extremes into water-resources planning and management is essential, but this information is not provided by the traditional sediment-rating-curve approach.

Since the PMF of daily suspended sediment load is derived, the single estimation such as the popular central-tendency measures of mean, median, and mode can also be derived. For instance, the mean, median, and mode of daily suspended sediment load for the discharge of 250 m^3/s are 5.8, 2.9, and 0.5 × 10^4 ton, respectively, which are less than the single estimation of 7.2 × 10^4 ton obtained by the traditional sediment rating curve. Repeating the same procedure to construct the CDF and PMF from the quantile-dependent rating curves for each recorded daily discharge, the single estimations are calculated and compared with the recorded daily suspended sediment loads. The mean absolute errors of mean, median, and mode for these 1,110 recorded data are 4.8, 4.3, and 6.9 × 10^4 ton, respectively. The proposed mean and median estimation schemes derived from the quantile regression-based model evidently outperform the traditional sediment rating curve, which induces the mean absolute error of 5.1 × 10^4 ton. Since the median estimation scheme has the minimum mean absolute error for the Laonung station, this method is suggested if the single estimation is needed for planning or management purpose.

The quantile regression-based approach provides not only the probabilistic description of daily suspended sediment load, but also a single estimation. Besides, the single estimations of mean and median have less mean absolute errors than the traditional sediment-rating-curve approach for Laonung station. Furthermore, probabilities of extreme events threatening hydraulic-facility operation are also provided by the proposed approach for planners and managers to assess risk of such extreme events.

4.3 Probabilistic Estimation of Annual Suspended Sediment Loads

The continuous records of daily discharge for the 1959–2008 period are employed to construct the sequence of PMF of suspended sediment load for that period. The PMF of cumulative suspended sediment load for a yearly period can be constructed by the derived daily PMFs and convolution theorem.

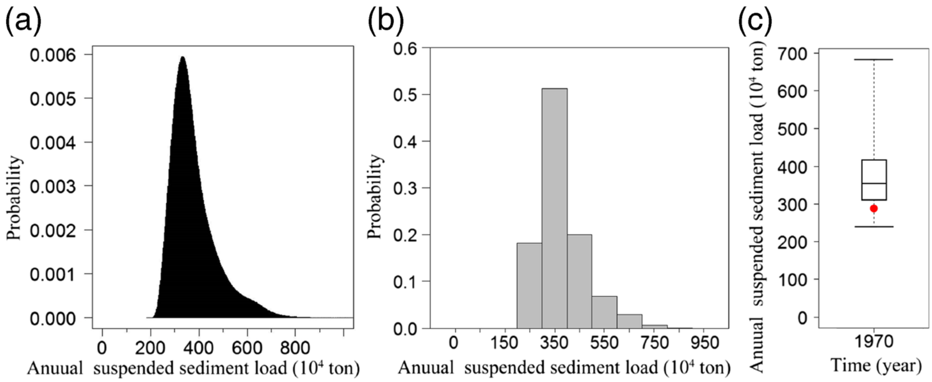

Fig. 7 Derived probability mass function (PMF) of annual suspended sediment load of 1970 for interval width of (**a**) 1×10^4 ton, (**b**) 100×10^4 ton, and (**c**) the corresponding box plot (*red dot* denoting the value estimated from the traditional sediment rating curve)

Figure 7a shows the PMF of annual suspended sediment load of 1970, which is roughly ranged between 200 and 800×10^4 ton. The fine interval of 1×10^4 ton replaced by a course interval of 100×10^4 ton leads to the histogram shown in Fig. 7b. This figure evidently shows that annual suspended sediment load of 350×10^4 ton (interval of $300-400\times10^4$ ton) has the highest probability of 0.51. In order to easily compare PMF of annual suspended sediment load for each year, Fig. 7a is transferred into a box-plot and shows the minimum (first percentile), first quartile, medium, third quartile, and the maximum (99th percentile) values, which is shown in Fig. 7c. This figure shows that the interquartile range ranges between 312 and 418×10^4 ton, which illustrates the most likely sediment load with a probability of 0.5 to occur within this range. The single estimations of annual suspended sediment load for the mean, median, and mode in 1970 are 378, 355, and 333×10^4 ton, respectively, which are greater than the value of 287×10^4 ton estimated by the traditional sediment rating curve.

Figure 8 illustrates the box-plot of annual suspended sediment load of Laonung station over the 1959–2008 period. Comparing with the single annual suspended sediment load estimated from the traditional sediment rating curve shown in Fig. 3 (also shown in Fig. 8 as red dots),

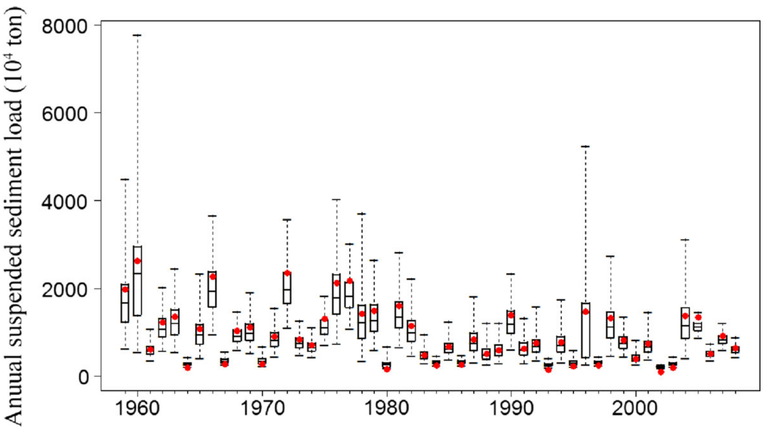

Fig. 8 Box-plot of annual suspended sediment load of Laonung station over the 1959−2008 period (*red dot* denoting the value estimated from the traditional sediment rating curve)

this proposed scheme provide a uncertainty information. For instance, the annual sediment load in 1990 roughly ranges between 623 and $2{,}352\times10^4$ ton, while the most likely 50 % of annual sediment load ranges between 990 and $1{,}501\times10^4$ ton, and the estimated mean, median, and mode of annual sediment load are 1,270, 1,203, and $1{,}069\times10^4$ ton, respectively. Although no recorded annual sediment loads can be used to check which single estimation has less error, the median estimation of annual sediment load is suggested because of the minimum mean absolute error obtained by this scheme for daily suspended sediment load.

The proposed approach provides probabilistic description of daily and annual suspended sediment loads. These derived empirical PMFs offers water-resources planners and managers useful information to assess risk or uncertainty of sediment-related problems. For example, useful life of reservoirs determined by sedimentation can be expressed in terms of probability distribution using the derived sequence of PMFs of annual suspended sediment load.

5 Summaries and Conclusions

Providing accurate suspended sediment load estimation together with uncertainty information is an important and essential issue in water-resources planning, design, and management. Quantile-dependent sediment rating curves associated with a probabilistic description scheme for daily and annual suspended sediment loads using quantile regression are proposed in this study to achieve this purpose.

Paired sediment-discharge recorded data are firstly used to construct dense quantile-dependent sediment rating curves. The estimated conditional suspended sediment loads at various quantiles associated with the corresponding quantiles for a given discharge are then integrated to build the empirical probability mass function (PMF) of daily suspended sediment loads. Derivation of the PMF of annual suspended sediment loads is made possible by the derived PMFs of daily suspended sediment loads and the convolution theorem.

A total of 1,110 recorded daily suspended sediment load and discharge data during the 1959–2008 period of the Laonung station located in southern Taiwan are used for illustrating the proposed novel scheme. The derived empirical PMFs of daily suspended sediment load for discharge equaling to 250, 486, and 998 m^3/s are shown in Fig. 5 for demonstration. The proposed probabilistic description for daily suspended sediment load evidently describes the inherent scatter between sediment and discharge data. The results of single estimations indicate that mean and median estimation schemes have mean absolute errors of 4.8 and 4.3×10^4 ton, respectively, which outperform the traditional sediment-rating-curve approach with a mean absolute error of 5.1×10^4 ton. The empirical PMFs in terms of box-plot of annual suspended sediment loads over the 1958–2008 period are shown in Fig. 8, which offer rich-information than the single estimation obtained from the traditional sediment-rating-curve approach.

Quantile regression-based approach provides not only the probabilistic description of daily and annual suspended sediment loads, but also a single estimation as well as the extremes. The obtained results offer water-resources planners and managers useful information to assess risk or uncertainty of sediment-induced problems such as the risk of water quality failing to meet the sediment-concentration criterion or the probabilistic description of useful life of reservoirs determined by sedimentation. These sediment-related problems are not considered in this work and remain as future topics for extending this research.

Acknowledgments Financial support for this study was graciously provided by the National Science Council, Taiwan, ROC (Grant No. NSC102-2221-E006-187). Valuable comments from anonymous referees and associate editor for improving presentation are greatly appreciated.

References

Afan HA, El-Shafie A, Yaseen ZM, Hameed MM, Wan Mohar WHM, Hussain A (2015) ANN based sediment prediction model utilizing different input scenarios. Water Resour Manag 29(4):1231–1245

Alagidede P, Panagiotidis T (2012) Stock returns and inflation: evidence from quantile regressions. Econ Lett 117(1):283–286

Alp M, Cigizoglu HK (2007) Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environ Model Softw 22(1):2–13

Asselman NEM (2000) Fitting and interpretation of sediment rating curve. J Hydrol 234:228–248

Austin M (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecol Model 200(1–2):1–19

Aytek A, Kişi Ö (2008) A genetic programming approach to suspended sediment modelling. J Hydrol 351:288–298

Barbosa SM, Scotto MG, Alonso AM (2011) Summarising changes in air temperature over Central Europe by quantile regression and clustering. Nat Hazards Earth Syst Sci 11(12):3227–3233

Cade BS, Noon BR (2003) A gentle introduction to quantile regression for ecologists. Front Ecol Environ 1(8): 412–420

Çimen M (2008) Estimation of daily suspended sediments using support vector machines. Hydrol Sci J 53(3): 656–666

Clarke RT (1990a) Statistical characteristics of some estimators of sediment and nutrient loadings. Water Resour Res 26(9):2229–2233

Clarke RT (1990b) Bias and variance of some estimators of suspended sediment load. Hydrol Sci J 35(3):253–261

Cohn TA, DeLong LL, Gilroy EJ, Hirsch RM, Wells DK (1989) Estimating constituent loads. Water Resour Res 25(5):937–942

Cohn TA, Caulder DL, Gilroy EJ, Zynjuk LD, Summers RM (1992) The validity of a simple statistical model for estimating fluvial constituent loads: an empirical study involving nutrient loads entering Chesapeake Bay. Water Resour Res 28(9):2353–2363

Cozzoli F, Bouma TJ, Ysebaert T, Herman PMJ (2013) Application of non-linear quantile regression to macrozoobenthic species distribution modelling: comparing two contrasting basins. Marine Ecol Process Ser 475:119–133

Ferguson R (1986) River loads underestimated by rating curves. Water Resour Res 22(1):74–76

Gaglianone WP, Lima LR, Linton O, Smith DR (2011) Evaluating value-at-risk models via quantile regression. J Bus Econ Stat 29(1):150–160

Guven A, Kişi Ö (2011) Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming. Water Resour Manag 25(2):691–704

Hicks DM, Gomez B, Trustrum NA (2000) Erosion thresholds and suspended sediment yields, Waipaoa River Basin, New Zealand. Water Resour Res 36(4):1129–1142

Hirschi M, Seneviratne SI, Alexandrov V, Boberg F, Boroneant C, Christensen OB, Formayer H, Orlowsky B, Stepanek P (2011) Observational evidence for soil-moisture impact on hot extremes in southwestern Europe. Nat Geosci 4(1):17–21

Jagger TH, Elsner JB (2009) Modeling tropical cyclone intensity with quantile regression. Int J Climatol 29(10): 1351–1361

Jain SK (2001) Development of integrated sediment rating curves using ANNs. J Hydraul Eng 127(1):30–37

Kişi Ö (2004) Daily suspended sediment modeling using a fuzzy-differential evolution approach. Hydrol Sci J 49(1):183–197

Kişi Ö (2005) Suspended sediment estimation using neuro-fuzzy and neural network approaches. Hydrol Sci J 50(4):683–696

Kitsikoudis V, Sidiropoulos E, Hrissanthou V (2014) Machine learning utilization for bed load transport in gravel-bed rivers. Water Resour Manag 28(11):3727–3743

Koenker R (2005) Quantile regression. Cambridge University Press, Cambridge

Koenker R (2014) Quantreg: Quantile Regression. R Package Version 4.76. <http://CRAN.R-project.org/package=quantreg>

Koenker R, Basset G (1978) Regression quantiles. Econometrica 46(1):33–50

Koenker R, D'Orey V (1987) Computing regression quantiles. Appl Stat 36(3):383–393

Krishnaswamy J, Richter DD, Halpin PN, Hofmockel MS (2001) Spatial patterns of suspended sediment yields in a humid tropical watershed in Costa Rica. Hydrol Process 15(12):2237–2257

Lafdani EK, Nia AM, Ahmadi A (2013) Daily suspended sediment load prediction using artificial neural networks and support vector machines. J Hydrol 478:50–62

Lohani AK, Goel NK, Bhatia KKS (2007) Deriving stage-discharge-sediment concentration relationship using fuzzy logic. Hydrol Sci J 52(4):793–807

McBean EA, Al-Nassri S (1988) Uncertainty in suspended sediment transport curves. J Hydraul Eng 114(1):63–74

Meligkotsidou L, Vrontos ID, Vrontos SD (2009) Quantile regression analysis of hedge fund strategies. J Emperic Fin 16(2):264–279

Nagy HM, Watanabe K, Hirano M (2002) Prediction of sediment load concentration in rivers using artificial neural network model. J Hydraul Eng 128(6):588–595

Partal T, Cigizoglu HK (2008) Estimation and forecasting of daily suspended sediment data using wavelet−neural network. J Hydrol 358:317–331

Phillips JM, Webb BW, Walling DE, Leeks GJL (1999) Estimating the suspended sediment loads of rivers in the LOIS study area using infrequent samples. Hydrol Process 13(7):1035–1050

Rai RK, Mathur BS (2008) Event-based sediment yield modeling using artificial neural network. Water Resour Manag 22(4):423–441

Ross SM (2007) Introduction to probability models, 9th edn. Academic, Burlington

Rustomji P, Wilkinson SN (2008) Applying bootstrap resampling to quantify uncertainty in fluvial suspended sediment loads estimated using rating curves. Water Resour Res 44(9), W09434. doi:10.1029/2007WR006088

Shiau JT, Huang WH (2015) Detecting distributional changes of annual rainfall indices in Taiwan using quantile regression. Journal of Hydro-environment Research, http://dx.doi.org/10.1016/j.jher.2014.07.006

Tarras-Wahlberg NH, Lane SN (2003) Suspended sediment yield and metal contamination in a river catchment affected by El Niño events and gold mining activities: the Puyango river basin, southern Ecuador. Hydrol Process 17(15):3101–3123

Vigiak O, Bende-Michl U (2013) Estimating bootstrap and Bayesian prediction intervals for constituent load rating curve. Water Resources Research 49(12), doi:10.1029/2013WR013559

Walling DE (1977) Assessing the accuracy of suspended sediment rating curves for a small basin. Water Resour Res 13(3):531–538

Wang P, Linker LC (2008) Improvement of regression simulation in fluvial sediment loads. J Hydraul Eng 134(10):1527–1531

Wang YG, Tian T (2013) Sediment concentration prediction and statistical evaluation for annual load estimation. J Hydrol 482:69–78

Wang YG, Kuhnert P, Henderson B (2011) Load estimation with uncertainties from opportunistic sampling data −a semiparametric approach. J Hydrol 396:148–157