



# Mathematical and Computer Modelling of Dynamical Systems

Methods, Tools and Applications in Engineering and Related Sciences

ISSN: (Print) (Online) Journal homepage: [www.tandfonline.com/journals/nmcm20](http://www.tandfonline.com/journals/nmcm20)

## An iterative approach for deriving and solving an accurate regression equation

Manaye Getu Tsige, Andreas Malcherek & Ivo Baselt

**To cite this article:** Manaye Getu Tsige, Andreas Malcherek & Ivo Baselt (2024) An iterative approach for deriving and solving an accurate regression equation, *Mathematical and Computer Modelling of Dynamical Systems*, 30:1, 73-90, DOI: [10.1080/13873954.2024.2313014](https://doi.org/10.1080/13873954.2024.2313014)

**To link to this article:** <https://doi.org/10.1080/13873954.2024.2313014>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 03 Mar 2024.



[Submit your article to this journal](#)



Article views: 330



[View related articles](#)



[View Crossmark data](#)

# An iterative approach for deriving and solving an accurate regression equation

Manaye Getu Tsige<sup>a</sup>, Andreas Malcherek<sup>b</sup> and Ivo Baselt<sup>b</sup>

<sup>a</sup>Department of Water Resources Engineering, Adama Science and Technology University, Adama, Ethiopia;

<sup>b</sup>Department of Civil Engineering and Environmental Sciences, Universität der Bundeswehr München, Neubiberg, Germany

## ABSTRACT

This paper introduces a method for deriving an accurate regression equation based on a set of any paired data, and a technique for solving the equation. For a practical example, we used five hundred seventy-one pairs of sediment concentration and river flow data to derive an accurate sediment rating equation. The graphs of the measured and predicted sediment concentrations matched each other, and data correlation showed Nash–Sutcliffe efficiency (NSE) of 0.9999860, coefficient of determination ( $R^2$ ) of 0.99998679, root mean square error (RMSE) of 0.0345, mean average error (MAE) of 0.0067, volume error (VE) of 1, and sum of square error (SSE) of 0.678631. To explain the technique of deriving and solving the accurate regression equation, separate files of video presentation and excel spreadsheet are provided as supplementary materials. In general, the method can be used to model any processes, and any calibration and validation processes can be addressed.

## ARTICLE HISTORY

Received 4 September 2023  
Accepted 25 January 2024

## KEYWORDS

Regression equation;  
polynomial function;  
sediment rating curve

## Introduction

The relationship between independent and dependent variables is governed by accepted scientific laws (Seber and Wild 2003) or it is expressed by mathematical, statistical, empirical, analytical, or numerical models. To find the best fit model to the measured data, parameters of the model can be estimated either through calibration or by regression analysis. The performance of the model is evaluated by using different statistical indicators.

Regression analysis, a technique for finding the relation among variables, is important to all scientific work where interpretations need to be drawn from measured data sets (Wu and Yen 1992). Authors (Seal 1967; Finney 1996; Barnes 1998; Galton 2001) highlighted the history related to the regression analysis, and authors (Fernandez-Delgado et al. 2019) provided an extensive experimental survey of regression methods.

If the relationship between dependent and independent variables is known or their relationship is defined by a chosen model, parameters of the model can be determined by

**CONTACT** Manaye Getu Tsige  [manaye.tsige@unibw.de](mailto:manaye.tsige@unibw.de)  Department of Water Resources Engineering, Adama Science and Technology University, Adama, Ethiopia

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/13873954.2024.2313014>

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

the parametric regression method. The results of analysing data using a parametric model may heavily depend on the chosen model for regression and variance functions, moreover also on a possibly underlying preliminary transformation of the variables (Bunke et al. 1999). Non-parametric regression methods, on the other hand, have in general a slower rate of convergence, but need no explicit specification of the form of the regression function (Glad 1998); the resulting curve is hence completely determined by the data themselves (Glad 1998). Different types of parametric or non-parametric regression methods, and their descriptions or applications are given in (Linnet 1998; Seber and Wild 2003; Qian and Reckhow 2005; Li and Yin 2009; Lolli and Gasperini 2012; Wang and Du 2014; Yong 2014; Özsoy and Örkücü 2016; Fernandez-Delgado et al. 2019).

The regression methods which can be used for both parametric and non-parametric regression analysis are artificial neural network (Specht 1991; Wu and Yen 1992; Zhang et al. 1998) and fuzzy regression method (Bárdossy et al. 1993; Yang and Lin 2002; Hao and Chiang 2008). Compared to other regression approaches, the artificial neural network is more appropriate than other approaches (Wiese and Schaper 1993; Pao 2008; Rahman and Asadujjaman 2021). The artificial neural network was designed to study the behaviour of real, nonlinear, complex systems, and they are particularly effective in solving problems where the correlations between the dependent and independent variables are well-known (Kopal et al. 2022). However, their precise description by classical mathematical methods is too complicated, too simplified, or impossible (Du and Swamy 2014; Kopal et al. 2022) and they also embody much uncertainty and difficulty (Masters and Land 1997; Zhang et al. 1998; Tomandl and Schober 2001; Morala et al. 2021). The neural network model could be a more useful nonlinear regression tool if it successfully incorporates human knowledge (heuristics) and other regression techniques (Wang 1999).

In actual modelling, the underlying processes are generally complex and not well understood, this means that we have little or no idea about the form of the relationship (Seber and Wild 2003). For example, different authors indicate that the power function is a commonly used nonlinear regression approach to model the sediment rating curve (eg., (Asselman 2000; Heng and Suetsugi 2014; Hapsari et al. 2019)). However, the error of the regression equation is very large. Therefore, finding an accurate regression method to derive an accurate regression equation based on a set of any paired data becomes important for the best accurate representation of any process. In this paper, we provide a procedure to derive a complex equation expressing the relationship between dependent and independent variables based on a set of any paired data.

## 2. Methodology

### 2.1. An iterative approach to derive an accurate regression equation

To arrive at iteration steps, let us begin from the following definition.

**Definition 2.1.** For given values of paired variables  $S$  and  $Q$ , variables  $x$  and  $y$  are defined by

$$y = i(bS)^{1/u} + jQ \quad (1)$$

$$x = k(hQ)^{1/w} + t \quad (2)$$

where  $i, b, u, j, k, h, w$  and  $t$  are constants.

Let  $y \approx y(x)$ , where  $y(x)$  is the function.

Since a polynomial function can accommodate and generate negative value, positive value, or both negative and positive values, let us consider a polynomial function  $y(x)$

$$y(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + c \quad (3)$$

Therefore,

$$y = y(x) + e \quad (4)$$

where  $e$  is the error value

Substitute equation 1 into 4

$$i(bS)^{1/u} + jQ = y(x) + e \quad (5)$$

Rearrange equation 5

$$S = \frac{1}{bi^u} (y(x) - jQ + e)^u \quad (6)$$

In equation 6, variables  $y(x)$ ,  $Q$  and  $e$  are connected by plus and minus sign. It shows that values of variables  $y(x)$ ,  $Q$  or  $e$  have an individual effect on a value of variable  $S$  (i.e. if we use  $y(x)$  in place of  $Q$  and vice versa, or if we use  $Q$  in the place of  $e$  and vice versa, a value of  $S$  will be different). This is a reason why we defined  $x$  and  $y$  in the above way to arrive at equation 6.

Let

$$e = e_1 + e_2 + e_3 + \dots + e_{p-1} + e_p \quad (7)$$

Substitute equation 7 into 6

$$S = \frac{1}{bi^u} (y(x) - jQ + e_1 + e_2 + e_3 + \dots + e_{p-1} + e_p)^u \quad (8)$$

Equation 8 represents an actual value of variable  $S$ . In equation 8, if a value of error  $e_p$  is the minimum tolerable error that could be ignored, then the sum of error values  $e_1, e_2, e_3, \dots, e_{p-1}$  represents an approximate value of the total error  $e$ . Therefore, the predicted value of variable  $S$  (let us say  $S_p$ ) is given by

$$S_p = \frac{1}{bi^u} (y(x) - jQ + e_1 + e_2 + e_3 + \dots + e_{p-1})^u \quad (9)$$

Therefore, the difference between  $S$  and  $S_p$  is an error, which is equal to  $E_p$  (i.e.  $E_p = S - S_p$ ), where  $p - 1$  refers to the number values of error  $e$  should be required to derive an accurate regression equation at  $p$  number of iteration steps. If there are  $p - 1$  number of values of error  $e$  (i.e.  $e_1, e_2, e_3, \dots, e_{p-1}$ ), there are also  $p - 1$  number of values of corresponding error  $E$  (i.e.  $E_1, E_2, E_3, \dots, E_{p-1}$ ).

Logic is now if we are able to express error  $e$  as a function of error  $E$ , we can derive an accurate regression equation. This is because of both errors (i.e.  $e$  and  $E$ ) is the function of variables  $Q$  and  $S$ . Therefore, we define an iterative procedure to approximate a value of error  $e$  based on a value of the corresponding error  $E$ . Let an approximate value of error  $e_1, e_2, e_3, \dots, e_{p-1}$  be equal to  $r_1, r_2, r_3, \dots, r_{p-1}$  respectively. Therefore, the following iteration steps are defined based on equation 9 and the explanations above.

For the first iteration step ( $p = 1$ ),  $e_0 = 0$ ,  $E_0 = 0$  and  $r_0 = 0$ . Therefore, the first predicted value of variable  $S$  (i.e.  $S_1$ ) is determined by

$$S_1 = \frac{1}{bi^u} (y(x) - jQ)^u \quad (10)$$

If  $S_1 \approx S$ , no need to proceed to the next iteration step. If  $S_1 \not\approx S$ , we proceed to the next iteration step.

For the second iteration step ( $p = 2$ ),  $e_1, E_1$  and  $r_1$  is determined by

$$e_1 = y - y(x) \quad (11)$$

$$E_1 = S - S_1 \quad (12)$$

$$r_1 = f_1(E_1) \quad (13)$$

where  $f_1$  is the polynomial regression function between the values of  $e_1$  and  $E_1$ . Therefore, at the second iteration step, the second predicted value of variable  $S$  (i.e.  $S_2$ ) is determined by

$$S_2 = \frac{1}{bi^u} (y(x) - jQ + r_1)^u \quad (14)$$

If  $S_2 \approx S$ , no need to proceed to the next iteration step. If  $S_2 \not\approx S$ , we proceed to the next iteration step.

For the third iteration step ( $p = 3$ ),  $e_2, E_2$  and  $r_2$  is determined by

$$e_2 = y - (y(x) + r_1) \quad (15)$$

$$E_2 = S - S_2 \quad (16)$$

$$r_2 = f_2(E_2), \quad (17)$$

where  $f_2$  is the polynomial regression function between the values of  $e_2$  and  $E_2$ . Therefore, at the third iteration step, the third predicted value of variable  $S$  (i.e.  $S_3$ ) is determined by

$$S_3 = \frac{1}{bi^u} (y(x) - jQ + r_1 + r_2)^u \quad (18)$$

If  $S_3 \approx S$ , no need to proceed to the next iteration step. If  $S_3 \not\approx S$ , we proceed to the next iteration step.

For the fourth iteration steps ( $p = 4$ ),  $e_3, E_3$  and  $r_3$  is determined by

$$e_3 = y - (y(x) + r_1 + r_2) \quad (19)$$

$$E_3 = S - S_3 \quad (20)$$

$$r_3 = f_3(E_3), \quad (21)$$

where  $f_3$  is the polynomial regression function between the values of  $e_3$  and  $E_3$ . Therefore, at the fourth iteration step, the fourth predicted value of variable  $S$  (i.e.  $S_4$ ) is determined by

$$S_4 = \frac{1}{bi^u} (y(x) - jQ + r_1 + r_2 + r_3)^u \quad (22)$$

If  $S_4 \approx S$ , no need to proceed to the next iteration step. If  $S_4 \not\approx S$ , we proceed to the next iteration step.

For the  $(p - 1)$ th iteration step,  $e_{p-2}$ ,  $E_{p-2}$  and  $r_{p-2}$  is determined by

$$e_{p-2} = y - (y(x) + r_1 + r_2 + r_3 + \dots + r_{p-3}) \quad (23)$$

$$E_{p-2} = S - S_{p-2} \quad (24)$$

$$r_{p-2} = f_{p-2}(E_{p-2}), \quad (25)$$

where  $f_{p-2}$  is the polynomial regression function between the values of  $e_{p-2}$  and  $E_{p-2}$ . Therefore, at the  $(p - 1)$ th iteration step, the  $(p - 1)$ th predicted value of variable  $S$  (i.e.  $S_{p-1}$ ) is determined by

$$S_{p-1} = \frac{1}{bi^u} (y(x) - jQ + r_1 + r_2 + r_3 + \dots + r_{p-2})^u \quad (26)$$

For the  $p$ th iteration step,  $e_{p-1}$ ,  $E_{p-1}$  and  $r_{p-1}$  is determined by

$$e_{p-1} = y - (y(x) + r_1 + r_2 + r_3 + \dots + r_{p-2}) \quad (27)$$

$$E_{p-1} = S - S_{p-1} \quad (28)$$

$$r_{p-1} = f_{p-1}(E_{p-1}), \quad (29)$$

where  $f_{p-1}$  is the polynomial regression function between the values of  $e_{p-1}$  and  $E_{p-1}$ . Therefore, at the  $p$ th iteration step, the  $p$ th predicted value of variable  $S$  (i.e.  $S_p$ ) is determined by

$$S_p = \frac{1}{bi^u} (y(x) - jQ + r_1 + r_2 + r_3 + \dots + r_{p-2} + r_{p-1})^u \quad (30)$$

## 2.2. Determining the final form of the accurate regression equation

Suppose at the  $p$ th iteration step,  $S \approx S_p$ . Then, the final form of an accurate regression equation is obtained through substitutions.

Substitute equation 10 into 12

$$E_1 = S - \frac{1}{bi^u} (y(x) - jQ)^u \quad (31)$$

Substitute equation 14 into 16

$$E_2 = S - \frac{1}{bi^u} (y(x) - jQ + r_1)^u \quad (32)$$

Substitute equation 18 into 20

$$E_3 = S - \frac{1}{bi^u} (y(x) - jQ + r_1 + r_2)^u \quad (33)$$

Substitute equation 26 into 28

$$E_{p-1} = S - \frac{1}{bi^u} (y(x) - jQ + r_1 + r_2 + r_3 \dots + r_{p-2})^u \quad (34)$$

Substitute equation 31 into 13

$$r_1 = f_1 \left( S - \frac{1}{bi^u} (y(x) - jQ)^u \right) \quad (35)$$

Substitute equation 32 into 17

$$r_2 = f_2 \left( S - \frac{1}{bi^u} (y(x) - jQ + r_1)^u \right) \quad (36)$$

Substitute equation 33 into 21

$$r_3 = f_2 \left( S - \frac{1}{bi^u} (y(x) - jQ + r_1 + r_2)^u \right) \quad (37)$$

Substitute equation 34 into 29

$$r_{p-1} = f_{p-1} \left( S - \frac{1}{bi^u} (y(x) - jQ + r_1 + r_2 + r_3 \dots + r_{p-2})^u \right) \quad (38)$$

Substitute equation 35 into 36; equations 35 and 36 into 37; equations 35, 36 and 37 into 38 and so on. After all substitutions have been done one after the other, then the final resulting equation is very long. But, we can see that  $r_1, r_2, r_3 \dots r_p$  is the function of variables  $Q$  and  $S$ . For given values of paired variables,  $i, b, u, j, k, h, w$  and  $t$  are all constants. Therefore,

$$r_1 + r_2 + r_3 + \dots + r_{p-1} = f_1(Q, S) + f_2(Q, S) + f_3(Q, S) \dots + f_{p-1}(Q, S) \quad (39)$$

Substitute equation 39 into 30

$$S_p = \frac{1}{bi^u} (y(x) - jQ + f_1(Q, S) + f_2(Q, S) + f_3(Q, S) \dots + f_{p-1}(Q, S))^u \quad (40)$$

Substitute equation 2 into 3

$$y(x) = a_n (k(hQ)^{1/w} + t)^n + a_{n-1} (k(hQ)^{1/w} + t)^{n-1} + a_{n-2} (k(hQ)^{1/w} + t)^{n-2} + \dots + c \quad (41)$$

From equation 41,  $y(x)$  is the function of variable  $Q$ . Therefore,

$$y(x) = f(Q) \quad (42)$$

Substitute equation 42 into 40

$$S_p = \frac{1}{b_i^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + f_3(Q, S) \dots + f_{p-1}(Q, S))^u \quad (43)$$

Suppose at the  $p$ th iteration step,  $S \approx S_p$ . Therefore, equation 43 is given by

$$S \approx \frac{1}{b_i^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + \dots + f_{p-1}(Q, S))^u \quad (44)$$

Equation 44 is the shorthand form of a very long equation. The substituting equations' power constants  $u$ ,  $w$ , and  $n$  make the equation complex and difficult to simplify. However, the substituting equations that form the complex equation are easily interconnected in an Excel spreadsheet or programmed in Matlab. As we can see from equation 44, there are only two variables  $Q$  and  $S$ . Therefore, we can solve this equation for a given value of  $Q$  or  $S$ . A procedure to solve the equation is provided in [Section 2.5](#).

### **2.3. Determining initial values for deriving an accurate regression equation**

In [Sections 2.1](#) and [2.2](#), we showed the steps to derive and determine the final form of the accurate regression equation based on values of paired variables  $S$  and  $Q$ . To start deriving the equation based on the values of the paired variables  $S$  and  $Q$ , we should have to first determine the constants (see equations 1 and 2). The polynomial function (see equation 3) directly describes the relationship between variables  $x$  and  $y$ , but it indirectly describes the relationship between variables  $S$  and  $Q$ . Therefore, for given values of paired variables  $S$  and  $Q$ , we find values of constants  $i$ ,  $b$ ,  $u$ ,  $j$ ,  $k$ ,  $h$ ,  $w$ , and  $t$  for equations 1 and 2 such that plots of  $x$  versus  $y$  yield a smooth curve of a polynomial function. Accordingly, once all values of constants are known, the initial and final values of variables will be determined by following the iteration steps above.

### **2.4. Deriving an accurate sediment rating equation**

In above sections, we indicated the general directions showing how to derive and determine the final form of the accurate regression equation, and we also indicated the direction showing how to determine the initial values to start deriving the equation. For a practical example, we use sediment concentration and corresponding river or stream-flow data (see [Table 1](#) and [2](#)) to derive an accurate sediment rating equation. In the table, suspended sediment concentration data is represented by variable  $S$ , whereas flow data is represented by variable  $Q$ .

To make it clear, we use the following steps to derive an accurate sediment rating equation based on the above pairs of sediment concentration and river or streamflow data.

**Table 1.** Sediment concentration versus river or streamflow data.

S	Type										
	Q	S	Q	S	Q	S	Q	S	Q	S	Q
0.14	4.45	0.38	20.12	2.36	94.00	1.39	0.17	1.49	31.02	0.69	199.26
1.28	6.25	1.32	15.30	0.32	17.68	1.03	1.53	0.55	25.32	0.90	189.28
0.16	2.21	12.49	149.06	3.93	42.80	0.44	21.83	0.77	17.64	0.57	198.42
0.24	37.89	0.28	15.04	4.07	119.92	1.12	93.23	0.44	3.15	0.53	319.65
0.21	33.19	14.66	37.81	2.42	6.09	1.30	74.07	1.33	12.06	0.77	110.63
0.14	30.44	1.98	167.03	0.74	0.49	0.92	1.60	1.26	14.15	0.67	104.23
0.19	37.09	4.62	114.72	2.67	57.20	1.18	24.21	0.74	5.39	0.51	153.16
0.26	40.82	2.58	59.48	1.00	56.88	0.62	17.47	1.01	53.83	0.35	159.87
0.15	1.51	3.67	152.82	1.29	52.23	4.95	144.55	4.32	39.80	0.25	135.77
0.24	2.51	1.15	238.53	10.35	57.74	5.58	108.34	1.63	13.08	0.17	3.12
2.14	54.81	1.91	58.37	25.17	49.32	0.81	9.69	24.41	2.40	0.18	29.89
0.67	23.65	0.89	4.79	1.81	108.29	0.76	9.72	28.09	0.63	0.23	3.64
1.18	30.74	0.61	5.10	36.73	159.62	11.86	15.58	2.37	39.32	0.76	12.60
0.43	32.12	1.41	32.37	51.81	179.98	1.40	93.23	1.88	17.10	0.53	113.22
0.75	38.05	0.58	24.74	90.85	69.49	5.28	105.25	2.81	25.48	0.43	97.20
0.32	20.44	2.24	11.55	4.96	128.92	1.16	104.70	19.72	1.30	0.80	50.24
3.53	55.26	1.42	13.48	4.11	104.01	0.48	36.19	23.65	1.51	0.17	1.28
0.41	21.89	13.49	40.42	2.22	82.63	1.37	38.07	19.66	0.99	0.37	0.31
2.49	35.16	6.76	43.20	0.42	5.31	1.81	42.83	0.16	0.29	10.07	37.64
1.21	63.67	2.70	75.92	0.33	6.18	0.56	37.75	0.52	2.53	0.54	0.45
0.65	37.58	1.33	130.30	0.33	5.18	1.49	37.75	17.00	5.25	0.29	35.88
0.41	22.34	11.75	176.14	0.20	2.21	0.42	48.70	0.30	0.23	3.11	117.10
2.83	44.08	5.42	111.82	0.16	2.48	0.43	48.70	0.80	0.31	5.37	207.80
1.48	24.09	2.94	98.13	0.28	3.08	0.42	35.76	23.20	32.44	5.04	95.13
0.63	22.51	3.83	99.30	0.95	5.64	0.78	48.92	0.43	0.43	3.36	146.50
6.73	133.89	1.35	168.55	0.68	6.99	7.12	95.55	31.05	50.11	3.87	152.71
0.45	31.28	1.24	20.22	0.34	3.72	0.60	54.25	20.31	10.16	2.63	50.12
0.48	31.03	1.78	62.43	0.19	5.43	0.78	56.13	0.32	0.22	5.95	62.99
0.44	57.53	3.43	71.05	1.38	53.34	2.79	13.58	0.36	0.44	3.26	73.70
0.20	34.61	2.63	38.20	1.31	67.62	3.78	10.14	14.60	228.98	3.23	129.79
0.28	32.44	3.13	39.30	2.18	8.69	3.30	5.27	30.02	127.80	2.07	122.71
0.37	26.92	0.93	48.68	3.16	44.75	0.65	41.22	7.70	10.93	0.65	73.37
0.20	2.70	0.38	28.72	0.38	12.68	2.85	47.26	33.64	28.16	3.04	129.32
2.58	93.17	0.36	34.10	4.97	31.55	0.39	41.71	27.86	30.54	2.19	152.27
0.15	1.21	2.87	151.57	3.24	44.48	0.30	39.25	2.58	2.54	4.74	180.53
0.38	33.75	7.17	33.49	4.64	45.20	1.27	65.49	9.56	12.85	1.55	118.12
0.34	30.32	0.57	16.89	39.48	56.21	3.44	271.87	3.57	21.47	0.17	3.19
0.19	3.98	4.71	37.33	0.34	6.30	1.74	208.40	4.75	16.66	7.24	225.41
0.36	4.61	3.49	40.91	59.70	61.38	0.39	29.66	1.63	2.89	5.13	176.56
1.39	102.17	1.34	33.79	13.36	148.11	1.80	85.73	37.06	22.19	5.27	221.30
1.11	59.16	3.86	74.78	0.63	4.54	3.33	108.97	0.33	0.23	2.47	171.11
0.52	33.87	0.48	24.69	0.40	0.57	0.85	69.59	0.27	1.72	3.94	276.35
0.27	152.47	7.52	75.00	0.30	0.13	5.34	127.63	0.47	1.56	1.08	110.00
18.53	78.29	1.28	16.81	0.19	0.48	2.48	77.18	0.36	0.75	2.52	137.09
0.38	2.89	0.97	27.20	0.63	0.60	12.29	98.50	37.66	36.81	0.28	17.58
0.27	11.05	1.08	59.65	1.08	1.01	3.42	110.31	11.92	214.26	0.21	17.16
0.54	14.02	2.94	66.77	1.25	23.21	4.41	70.46	23.07	67.11	0.24	19.31

- (1) For given values of paired variables  $S$  and  $Q$ , estimate constants  $i, b, u, j, k, h, w$ , and  $t$  such that plots of  $x$  versus  $y$  yields a smooth curve of polynomial function (refer to equations 1 and 2)
- (2) Choose a polynomial regression function that fits the plots of  $x$  versus  $y$
- (3) From the regression equation in step 2, find the constants of equation 3
- (4) Calculate  $y$  by using equation 1

**Table 2.** Sediment concentration versus river or streamflow data.

Type											
S	Q	S	Q	S	Q	S	Q	S	Q	S	Q
0.79	19.02	1.57	66.74	0.35	14.29	8.77	108.26	2.89	6.26	0.63	101.23
1.29	58.66	0.23	32.58	5.50	31.05	0.29	46.20	9.81	9.92	2.68	274.87
0.60	4.56	2.93	66.22	0.36	1.30	0.69	35.27	0.39	0.17	2.01	140.60
0.23	24.35	3.61	23.95	0.24	1.83	6.14	60.28	0.29	1.85	0.94	148.94
0.30	29.23	0.58	15.38	0.27	1.83	7.78	82.47	3.20	7.15	3.40	179.02
0.36	23.58	4.46	119.18	2.61	46.23	5.45	72.37	0.33	1.27	4.83	181.19
0.73	226.44	1.53	16.09	2.30	107.37	1.80	72.65	0.50	0.88	2.76	175.82
1.04	231.54	4.27	30.15	5.80	355.23	6.28	144.50	8.06	6.52	4.44	214.76
0.71	173.97	0.54	30.39	18.50	133.38	0.76	13.54	0.48	1.13	1.76	115.29
1.59	50.64	0.51	26.49	4.67	148.95	0.72	12.26	19.58	85.06	4.95	177.35
4.82	99.40	0.28	12.27	1.87	132.59	0.82	11.28	0.30	0.72	1.48	101.10
2.82	18.86	0.33	61.50	3.57	113.26	0.80	12.24	1.16	1.21	6.10	145.92
0.61	72.30	0.32	38.59	6.87	118.23	0.73	9.70	0.39	1.14	6.09	152.61
1.17	84.22	1.28	23.90	5.27	121.66	0.81	10.55	0.36	0.64	4.34	142.45
0.41	12.03	0.24	49.53	0.25	1.50	0.75	10.38	0.46	0.36	3.20	153.48
0.46	16.91	8.39	232.91	0.27	1.80	0.86	11.69	0.23	0.64	6.96	263.48
1.50	16.62	0.20	0.42	0.31	0.78	0.76	11.85	0.33	1.38	1.77	127.29
0.62	17.65	0.18	6.47	2.19	0.76	55.67	103.16	0.26	0.97	2.00	88.42
0.42	10.38	0.76	11.92	11.84	0.77	40.87	56.90	0.35	1.63	3.36	108.43
0.38	62.83	0.30	10.18	4.26	0.74	20.17	25.76	11.58	18.41	4.12	109.92
0.48	44.38	0.54	4.93	5.77	0.32	11.20	16.43	0.21	1.97	0.91	93.72
0.29	25.15	31.76	55.98	4.50	0.44	9.71	19.69	14.76	12.43	2.84	94.94
0.27	2.72	7.66	197.55	6.59	0.43	3.03	37.75	45.57	9.86	4.59	129.44
1.62	11.78	1.38	134.05	0.92	11.76	11.27	26.97	8.05	3.37	2.51	139.24
5.35	53.47	1.17	22.75	31.40	1.60	69.93	34.78	13.44	5.32	2.58	144.18
0.29	4.23	1.19	195.35	8.40	0.93	6.13	21.76	15.25	23.31	5.72	171.62
7.07	200.33	0.97	193.68	5.20	0.44	39.35	117.00	34.20	16.55	3.96	207.69
0.57	7.86	1.37	201.84	12.68	0.44	27.25	37.75	15.01	16.64	2.13	146.80
1.05	9.04	1.69	205.97	0.71	9.93	21.58	30.66	18.28	24.10	3.31	148.29
0.68	53.32	1.11	139.99	4.39	0.72	18.56	23.01	8.58	10.44	4.87	187.62
0.45	66.91	1.30	168.31	36.95	0.72	11.80	19.02	0.21	0.43	2.62	118.57
0.60	51.73	2.25	264.60	0.71	13.37	6.45	22.07	14.17	10.15	1.38	94.30
0.35	5.62	1.14	233.50	10.20	361.08	3.67	97.29	0.88	0.65	7.70	208.42
0.53	6.71	1.71	18.52	2.47	254.38	1.12	86.98	25.65	29.16	5.50	112.99
0.25	8.90	2.14	236.63	2.19	245.57	2.98	83.68	12.38	10.89	5.11	184.17
7.06	18.77	3.22	17.85	2.87	51.24	5.70	119.69	35.03	6.75	5.17	154.67
1.19	50.21	0.76	17.13	8.82	154.82	3.23	124.85	2.94	1.35	3.95	127.96
1.58	111.47	2.74	13.87	2.64	235.94	4.67	126.00	5.16	4.97	4.78	176.35
3.37	76.86	25.59	97.91	3.74	289.75	4.78	140.53	2.11	1.16	6.53	212.73
0.99	62.46	2.24	66.73	2.03	246.12	5.83	84.36	0.18	9.62	1.82	139.40
0.30	24.61	0.98	98.90	0.86	43.60	2.74	41.96	0.15	114.51	3.48	200.64
0.34	26.30	1.00	73.00	14.64	43.91	0.45	38.08	0.12	3.72	5.06	153.94
3.59	87.21	3.23	296.16	11.77	75.13	0.42	42.77	0.28	3.37	6.85	190.77
0.49	24.87	2.43	285.66	4.13	413.23	1.03	41.25	0.28	3.37	2.29	116.25
3.57	29.98	0.56	5.42	2.65	170.43	0.43	24.19	0.73	4.92	7.65	259.50
0.88	30.47	0.59	5.76	3.10	410.41	1.94	90.56	0.20	6.88	2.97	153.25
4.95	64.42	1.60	0.37	0.30	1.91	1.32	108.50	0.31	2.99	2.53	122.99
1.30	28.83	0.73	27.55	0.29	0.95	3.50	38.88	0.17	2.30	2.32	161.80
										5.10	258.24

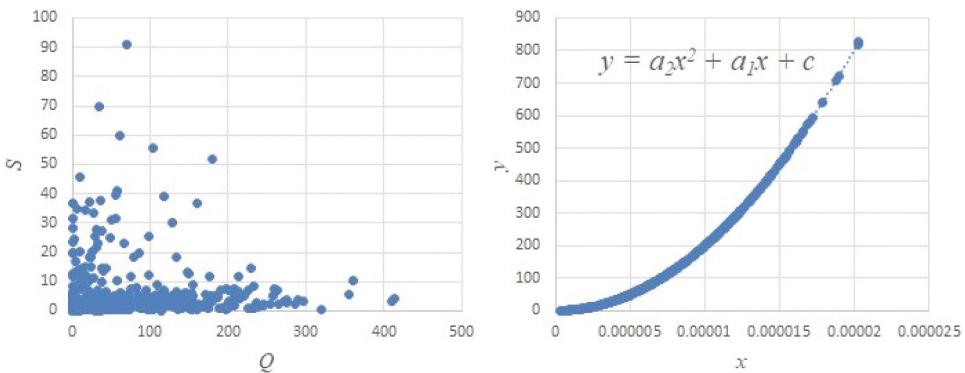
- (5) Calculate  $x$  by using equation 2
- (6) Calculate  $y(x)$  based on steps 3 and 5
- (7) Calculate  $S_1$  by using equation 10, where  $S_1$  represents the first predicted value.  
Plot graphs of measured ( $S$ ) and predicted ( $S_1$ ) values. If the graphs do not match each other, then we proceed to the next iteration step.
- (8) Calculate  $e_1$  by using equation 11

- (9) Calculate  $E_1$  by using equation 12
- (10) Consider a polynomial regression function to correlate  $e_1$  and  $E_1$
- (11) Calculate  $r_1$  by using the regression equation from step 10 (i.e. refer to equation 35)
- (12) Replace the calculated value of  $r_1$  from step 11 in equation 14
- (13) Calculate  $S_2$  by using equation 14, where  $S_2$  represents the second predicted value. Plot graphs of measured ( $S$ ) and predicted ( $S_2$ ) values. If the graphs do not match each other, then proceed to the next iteration step.
- (14) Replace the calculated value of  $r_1$  from step 11 in equation 15
- (15) Then, calculate  $e_2$  by using equation 15
- (16) Calculate  $E_2$  by using equation 16
- (17) Consider a polynomial regression function to correlate  $e_2$  and  $E_2$
- (18) Calculate  $r_2$  by using the regression equation from step 17 (i.e. refer to equation 36)
- (19) Replace the calculated value of  $r_1$  from step 11, and the calculated value of  $r_2$  from step 18 in equation 18
- (20) Then, calculate  $S_3$  by using equation 18, where  $S_3$  represents the third predicted value. Plot graphs of measured ( $S$ ) and predicted ( $S_3$ ) values. If the graphs do not match each other, then we proceed to the next iteration step, and so on.

We repeat the same procedure to calculate a value of  $S_p$  by using equation 30, where subscript  $p$  stands for number of iteration steps. During each iteration step, we plot graphs of the measured ( $S$ ) and predicted ( $S_p$ ) values. Our iteration procedure ends when the graphs almost match each other.

Based on the paired data given in the table 0, the values of the required constants (i.e.  $i, b, u, j, k, h, w,$  and  $t$ ) and variables (i.e.  $y(x), r_1, r_2, r_3, r_4, \dots, r_{14}$ ) had been determined by following the above steps. The values of these constants and variables are given below. **Figure 1** shows the graph of the original river or streamflow ( $Q$ ) versus sediment concentration ( $S$ ) data, and the graph of the transformed data ( $x$  versus  $y$ ) (see Section 2.3).

$$k = 0.000001 \tag{45}$$



**Figure 1.** The original and transformed data according to step one.

$$h = 1 \quad (46)$$

$$w = 2 \quad (47)$$

$$t = 0 \quad (48)$$

$$i = 0.0000111 \quad (49)$$

$$b = 0.0022 \quad (50)$$

$$u = 2 \quad (51)$$

$$j = 2 \quad (52)$$

$$y(x) = 1999999995091.56x^2 + 0.10499849x + 0.0000001 \quad (53)$$

$$r_1 = (-0.00008E_1^2 + 1.011E_1 + 0.004 - E_1) * 10^{-5} \quad (54)$$

$$r_2 = (1.052E_2 + 0.006 - E_2) * 10^{-6} \quad (55)$$

$$r_3 = (0.002E_3^2 + 1.089E_3 - 0.018 - E_3) * 10^{-6} \quad (56)$$

$$r_4 = (1.042E_4 - E_4) * 10^{-6} \quad (57)$$

$$r_5 = (0.007E_5^2 + 1.096E_5 - 0.006 - E_5) * 10^{-6} \quad (58)$$

$$r_6 = (0.001E_6^2 + 1.047E_6 - E_6) * 10^{-6} \quad (59)$$

$$r_7 = (0.02E_7^2 + 1.143E_7 - 0.004 - E_7) * 10^{-6} \quad (60)$$

$$r_8 = (0.005E_8^2 + 1.066E_8 - E_8) * 10^{-6} \quad (61)$$

$$r_9 = (0.013E_9^2 + 1.075E_9 - E_9) * 10^{-6} \quad (62)$$

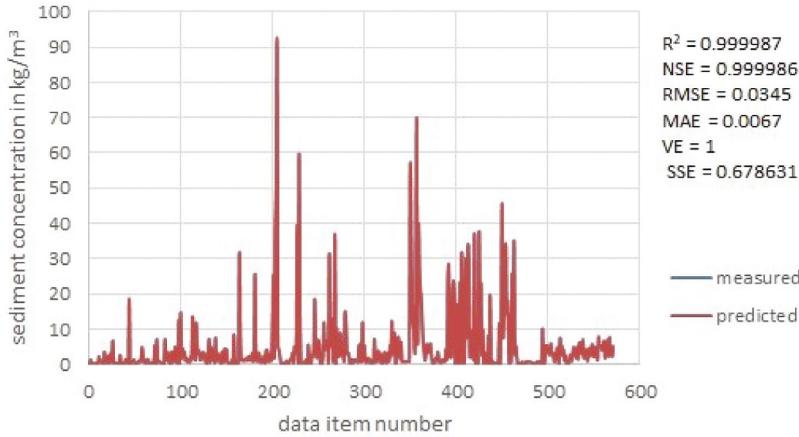
$$r_{10} = (0.026E_{10}^2 + 1.09E_{10} - 0.001 - E_{10}) * 10^{-6} \quad (63)$$

$$r_{11} = (0.026E_{11}^2 + 1.102E_{11} - E_{11}) * 10^{-6} \quad (64)$$

$$r_{12} = (-0.001E_{12}^2 + 1.04E_{12} - E_{12}) * 10^{-6} \quad (65)$$

$$r_{13} = (0.115E_{13}^2 + 1.139E_{13} - E_{13}) * 10^{-6} \quad (66)$$

$$r_{14} = (0.007E_{14}^2 + 1.048E_{14} - E_{14}) * 10^{-6} \quad (67)$$



**Figure 2.** Graphs of measured ( $S$ ) and predicted ( $S_p$ ) sediment concentration.

As the values of the above constants and variables were already determined, the final form of the accurate regression equation is obtained by direct substitutions (refer to Section 2.2). Therefore, the final form of the accurate sediment rating equation is given by

$$S \approx \frac{1}{b_i^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + \dots + f_{14}(Q, S))^u \tag{68}$$

For the final form of the equation, the graphs of measured ( $S$ ) and predicted sediment concentration ( $S_p$ ) matched each other (see Figure 2).

Since the final form of the equation is a very large and complex equation, the above values of the variables are easily interconnected in an excel spreadsheet or programmed in Matlab. Separate files of video presentation and Excel spreadsheet are provided as supplementary files.

**2.5. Solving the accurate sediment rating equation**

In the above section, we showed the procedures to derive the accurate sediment rating equation. For the paired suspended sediment concentration ( $S$ ) and flow data ( $Q$ ), we calculated each value of  $E_1, E_2, E_3, \dots, E_{14}$ , and the corresponding value of  $e_1, e_2, e_3, \dots, e_{14}$ , respectively. At the fifteenth iteration step, at the values of  $E_{14}$  and  $e_{14}$ , we found that  $S_{15} \approx S$ . Therefore, the last remaining errors are  $E_{15}$  and  $e_{15}$ . According to the steps above or Section 2.1, a value of  $E_{15}$  is determined by

$$E_{15} = S - S_{15} \tag{69}$$

Based on equation 43

$$S_{15} = \frac{1}{b_i^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + f_3(Q, S) \dots + f_{14}(Q, S))^u \tag{70}$$

Therefore,

$$E_{15} = S - \frac{1}{b_i^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + f_3(Q, S) \dots + f_{14}(Q, S))^u \tag{71}$$

Based on equations 1, 27 and 39

$$e_{15} = f_0(Q, S) - (f(Q) + f_1(Q, S) + f_2(Q, S) + f_3(Q, S) + \dots + f_{14}(Q, S)) \quad (72)$$

For each paired values of  $S$  and  $Q$ , there are corresponding values of  $E_{15}$  and  $e_{15}$ . Now, we take the values of  $E_{15}$  and  $e_{15}$  as paired input data to derive another equation that relates  $E_{15}$  and  $e_{15}$  by following the above steps, and so on. To derive the equation based on paired values of  $E_{15}$  and  $e_{15}$ , we calculate another values of  $E_p$  and  $e_p$  (see the steps above). To avoid confusion, let us express these other values of  $E_p$  and  $e_p$  in terms of  $E_p^*$  and  $e_p^*$ , respectively. Therefore, we define the following relationship.

$$\text{If } \lim_{E_p^* \rightarrow 0} (E_1^* + E_p^*) = E_1^*, \text{ then } E_{14} + E_p^* \approx E_{14} \quad (73)$$

For the given paired data ( $S$  and  $Q$ ), at the value of  $E_{14}$ , we found that  $S_{15} \approx S$ . According to steps above or [section 2.1](#), the value of  $E_{14}$  is determined by

$$E_{14} = S - S_{14} \quad (74)$$

Consider equation 43

$$S_{14} = \frac{1}{bi^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + f_3(Q, S) \dots + f_{13}(Q, S))^u \quad (75)$$

Therefore,

$$E_{14} = S - \frac{1}{bi^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + f_3(Q, S) \dots + f_{13}(Q, S))^u \quad (76)$$

For each paired value of  $S$  and  $Q$ , there is a corresponding value of  $E_{14}$ , which is a unique value. It is to mean that, for a given value of  $Q$ , there is only one value of  $S$  which results in a corresponding value of  $E_{14}$ , minimum value or zero value of  $E_{14}$  (i.e. there is no possibility to have two different values of  $E_{14}$  for the same values of paired data). From the relationship (refer to equation 73) to approximate a value of  $E_{14}$  for an unknown value of suspended sediment concentration or flow data, we keep on deriving a series of equations until a value of  $E_p^*$  is approximately zero or it is far apart from a value of  $E_1^*$ . In this case, the value of  $E_p^*$  determines the accuracy of the approximation. Therefore, to estimate an unknown value of suspended sediment concentration for a given value of flow data, a value of suspended sediment concentration that results in the minimum value of  $E_p^*$  is the solution.

Since the systems of equations forming the complex equation are very long, the separate files of the video presentation and Excel spreadsheet on deriving and solving the accurate sediment rating equation are provided as the supplementary files.

### 3. Results

The iterative approach for deriving an accurate regression equation based on values of paired variables is given in [Section 2.1](#). The procedures to determine the final form of the accurate regression equation are given in [Section 2.2](#). Accordingly, the shorthand form of the final accurate regression equation is given by

$$S \approx \frac{1}{bi^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + \dots + f_{p-1}(Q, S))^u, \tag{77}$$

where,  $S$  and  $Q$  are variables,  $b, i, u$  and  $j$  are constants for given values of paired data.

The accurate sediment rating equation which was derived based on five hundred seventy-one number of records of suspended sediment concentration and flow data is given by

$$S \approx \frac{1}{bi^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + \dots + f_{14}(Q, S))^u \tag{78}$$

The graphs of measured and predicted suspended sediment concentration matched each other (see Figure 2), and statistical measures for the data correlation are given in Figure 2. The procedures to solve the accurate regression equation are given in Section 2.5.

### 4. Discussions

The relationship between the sediment concentration and flow was given by the complex equation (it was not polynomial or other kinds of known function). This equation may reflect the complex relationship between the dynamic behaviour of flow and sediment transport.

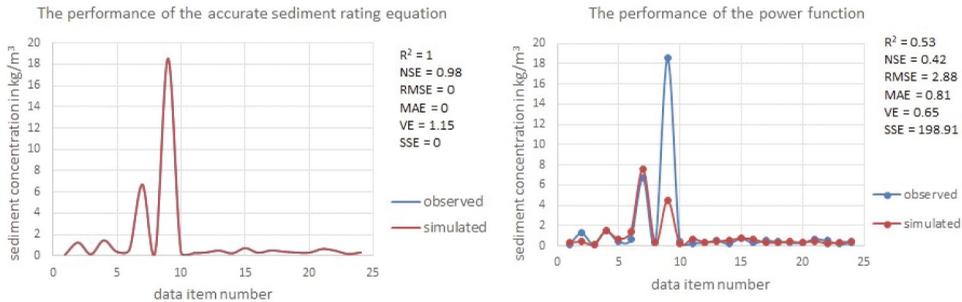


Figure 3. Comparison of sediment prediction accuracy of the proposed regression equation and the power function ( $S = 0.069Q^{0.9576}$ ) for the hombole watershed in Ethiopia.

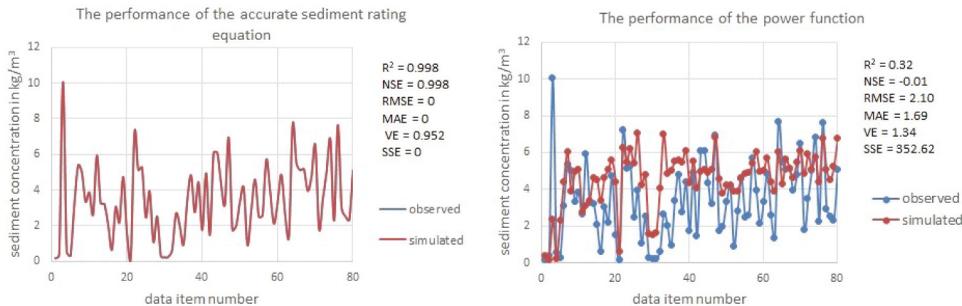
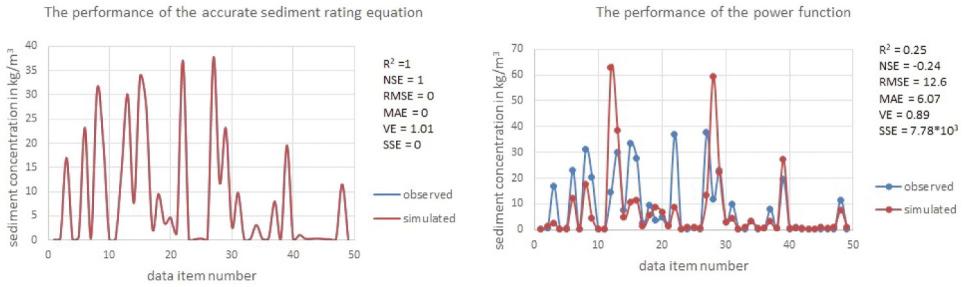
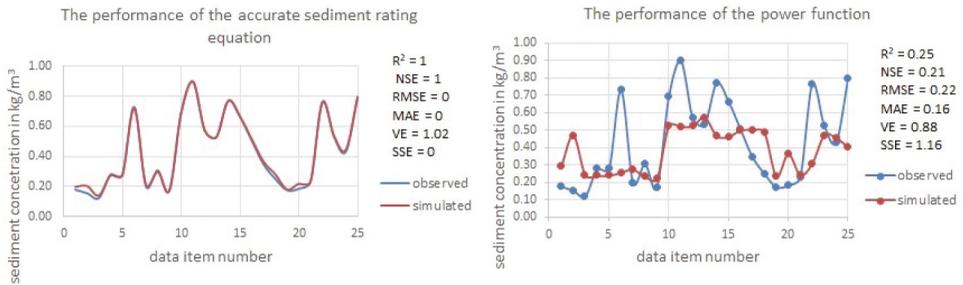


Figure 4. Comparison of sediment prediction accuracy of the proposed regression equation and the power function ( $S = 0.2036Q^{0.5475}$ ) for the Gumera watershed in Ethiopia, provided that all data records were taken into account without any preconditions.



**Figure 5.** Comparison of sediment prediction accuracy of proposed regression equation and the power function ( $S = 0.659Q^{0.839}$ ) for the mojo watershed in Ethiopia, provided that all data records were taken into account without any preconditions.



**Figure 6.** Comparison of sediment prediction accuracy of the proposed regression equation and the power function ( $S = 0.1901Q^{0.1916}$ ) for the Gilgel Gibe 1 watershed in Ethiopia, provided that all data records were taken into account without any preconditions.

A power function is a commonly used non-linear regression approach for predicting sediment from a given flow data. However, a regression error is very large. The comparison of sediment prediction accuracy of the proposed regression equation and power function are given in Figures 3–6. The proposed regression equation is very accurate. We can minimize a regression error as small as possible by increasing iteration steps.

Model calibration and validation are challenging tasks to apply a model for a particular purpose, even for further improvement of the model. For example, if we consider the Modified Universal Soil Loss Equation (MUSLE) or the improved MUSLE, finding the coefficient, soil erodibility, cover, and conservation practice factors of the MUSLE or the improved MUSLE through calibration is not a feasible approach (Tsige et al. 2022a, 2022b). This is because only a product effect of the coefficient and these factors is reflected in the MUSLE or the improved MUSLE rather than their individual effect during the calibration of sediment yield (Tsige et al. 2022a, 2022b). Therefore, the individual effect of model variables rather than their product effect on the engaged physical processes is important. Therefore, expressing the relationship between model variables in such a way that their individual effects can be seen on the engaged physical process is essential. The proposed regression method may play a significant role in this regard.

## 5. Conclusions

The accurate sediment rating equation was derived by following the proposed iteration steps. For the paired values of suspended sediment concentration ( $S$ ) and flow ( $Q$ ) data, the shorthand form of the final accurate sediment rating equation is given by

$$S \approx \frac{1}{b_i^u} (f(Q) - jQ + f_1(Q, S) + f_2(Q, S) + \dots + f_{14}(Q, S))^u, \quad (79)$$

where,  $b$ ,  $i$ ,  $u$  and  $j$  are constants for given values of paired data

In this paper, the polynomial regression functions were considered to derive very long and complex accurate regression equation. However, we can use any other known functions. And also, variables  $x$  and  $y$  were defined in such a way that individual effects of other variables can reflect on variable  $S$  (refer to [section 2.1](#)). However, we can define variables  $x$  and  $y$  in another way, and we follow the proposed iterative approach to derive an accurate regression equation.

The proposed iterative approach can be used to derive an accurate regression equation based on given values of paired variables. Therefore, the iterative approach can be used to model any processes, and any calibration and validation processes can be addressed.

In this paper, the iterative procedure is provided to solve the accurate regression equation. For further research, the analytical solution of the equation is recommended.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This research was funded by the German Academic Exchange Service and Universität der Bundeswehr München.

## References

- Asselman NEM. 2000. Fitting and interpretation of sediment rating curves. *J Hydrol.* 234(3):228–248. doi:[10.1016/S0022-1694\(00\)00253-5](https://doi.org/10.1016/S0022-1694(00)00253-5).
- Bárdossy A, Bogárdi I, Duckstein L. 1993. Theory and methodology: fuzzy nonlinear regression analysis of dose-response relationships. *Eur J Oper Res.* 66(1):36–51. doi:[10.1016/0377-2217\(93\)90204-Z](https://doi.org/10.1016/0377-2217(93)90204-Z).
- Barnes TJ. 1998. A history of regression: actors, networks, machines, and numbers. *Envir & Plan.* 30(2):203–223. doi:[10.1068/a300203](https://doi.org/10.1068/a300203).
- Bunke O, Droge B, Polzehl J. 1999. Model selection, transformations and variance estimation in nonlinear regression. *Stat: A J Theo & Appl Stat.* 33(3):197–240. doi:[10.1080/02331889908802692](https://doi.org/10.1080/02331889908802692).
- Du KL, Swamy MNS. 2014. *Neural networks and statistical learning*. London, Springer. doi:[10.1007/978-1-4471-5571-3](https://doi.org/10.1007/978-1-4471-5571-3).
- Fernández-Delgado M.S, Sirsat M, Cernadas E, et al. 2019. An extensive experimental survey of regression methods. *Neural Networks.* 111:11–34. doi:[10.1016/j.neunet.2018.12.010](https://doi.org/10.1016/j.neunet.2018.12.010)
- Finney DJ. 1996. A note on the history of regression. *J Appl Stat.* 23(5):555–557. doi:[10.1080/02664769624099](https://doi.org/10.1080/02664769624099).

- Galton SJM. 2001. Pearson, and the Peas: a brief history of linear regression for statistics instructors. *J Stat Educ*, 9.
- Glad IK. 1998. Parametrically guided non-parametric regression. *Scand J Stat.* 25(4):649–668. Available from <https://www.jstor.org/stable/4616530>.
- Hao P, Chiang J. 2008. Fuzzy regression analysis by support vector learning approach. *IEEE Trans Fuzzy Syst.* 16(2):428–441. doi:10.1109/TFUZZ.2007.896359.
- Hapsari D, Onishi T, Imaizumi F, et al. 2019. The use of sediment rating curve under its limitations to estimate the suspended load. *Rev Agric Sci.* 7:88–101. doi:10.7831/ras.7.0\_88
- Heng S, Suetsugi T. 2014. Comparison of regionalization approaches in parameterizing sediment rating curve in ungauged catchments for subsequent instantaneous sediment yield prediction. *J Hydrol.* 512:240–253. doi:10.1016/j.jhydrol.2014.03.003.
- Kopal I, Labaj I, Vrškova J, et al. 2022. A generalized regression neural network model for predicting the curing characteristics of carbon black-filled rubber blends. *Polymers.* 14(4):653. doi:10.3390/polym14040653
- Li H, Yin G. 2009. Generalized method of moments estimation for linear regression with clustered failure time data. *Biometrika.* 96(2):293–306. doi:10.1093/biomet/asp005.
- Linnert K. 1998. Performance of deming regression analysis in case of misspecified analytical error ratio in method comparison studies. *Clin Chem.* 44(5):1024–1031. doi:10.1093/clinchem/44.5.1024.
- Lolli B, Gasperini P. 2012. A comparison among general orthogonal regression methods applied to earthquake magnitude conversions. *Geophy J Int.* 190(2):1135–1151. doi:10.1111/j.1365-246X.2012.05530.x.
- Masters T, Land W *A new training algorithm for the general regression neural network, 1997 IEEE International Conference on Systems, Man, and Cybernetics, Orlando, USA, 1997.*
- Morala P, Cifuentes JA, Lillo RE, et al. 2021. Towards a mathematical framework to inform neural network modelling via polynomial regression. *Neural Networks.* 142:57–72. doi:10.1016/j.neunet.2021.04.036.
- Özsoy VS, Örkçü HH. 2016. Estimating the parameters of nonlinear regression models through Particle Swarm optimization. *Gazi Univ J Sci.* 29:187–199.
- Pao H. 2008. A comparison of neural network and multiple regression analysis in modeling capital structure. *Expert Syst Appl.* 35(3):720–727. doi:10.1016/j.eswa.2007.07.018.
- Qian SS, Reckhow KH. 2005. Nonlinear regression modeling of nutrient loads in streams: a Bayesian approach. *Water Resour Res.* 41.
- Rahman M, Asadujjaman M Implementation of artificial neural network on regression analysis, 2021 5th annual systems modelling conference, Canberra, Australia. 2021; doi:10.1109/SMC53803.2021.9569881.
- Seal HL. 1967. Studies in the history of probability and statistics. XV: the historical development of the gauss linear model. *Biometrika.* 54(1):1–24.
- Seber GAF, Wild CJ. 2003. *Nonlinear regression.* John Wiley & Sons, Inc., Hoboken, New Jersey.
- Specht DF. 1991. A general regression neural network. *IEEE Trans Neural Net.* 2(6):568–576. doi:10.1109/72.97934.
- Tomandl D, Schober A. 2001. A modified general regression neural network (MGRNN) with new, efficient training algorithms as a robust ‘black box’-tool for data analysis. *Neural Networks.* 14(8):1023–1034. doi:10.1016/S0893-6080(01)00051-X.
- Tsige MG, Malcherek A, Seleshi Y. 2022a. Estimating the best exponent and the best combination of the exponent and topographic factor of the modified universal soil loss equation under the hydro-climatic conditions of Ethiopia. *Water.* 14(9):1501. Available from <https://www.mdpi.com/2073-4441/14/9/1501>.
- Tsige MG, Malcherek A, Seleshi Y. 2022b. Improving the modified universal soil loss equation by physical interpretation of its factors. *Water.* 14(9):1450. Available from <https://www.mdpi.com/2073-4441/14/9/1450>.
- Wang F, Du T. 2014. Implementing support vector regression with differential evolution to forecast motherboard shipments. *J Amer Math Soc.* 41:3850–3855.

- Wang S. 1999. Nonlinear regression: a hybrid model. *Comput Oper Res.* 26(8):799–817. doi:10.1016/S0305-0548(98)00088-4.
- Wiese M, Schaper KJ. 1993. Application of neural networks in the QSAR analysis of percent effect biological data: comparison with adaptive least squares and nonlinear regression analysis. *SAR and QSAR in Envir Res.* 1(2–3):137–152. doi:10.1080/10629369308028825.
- Wu FY, Yen KK. 1992. Applications of neural network in regression analysis. *Comput Ind Eng.* 23(1–4):93–95. doi:10.1016/0360-8352(92)90071-Q.
- Yang M, Lin T. 2002. Fuzzy least-squares linear regression analysis for fuzzy input–output data. *Fuzzy Sets Syst.* 126(3):389–399. doi:10.1016/S0165-0114(01)00066-5.
- Yong L. 2014. Novel global harmony search algorithm for least absolute deviation. *J Appl Math.* 2014:1–6. doi:10.1155/2014/632975.
- Zhang G, Patuwo BE, Hu MY. 1998. Forecasting with artificial neural networks: the state of the art. *Int J Forecasting.* 14(1):35–62. doi:10.1016/S0169-2070(97)00044-7.