DEVELOPING A NEW TOTAL SEDIMENT TRANSPORT FORMULA VIA POLYNOMIAL BEST SUBSET REGRESSION

[1]DAVUT OKCU, [2]EBUBEKIR SEZGIN, [3]HATICE OKCU

University for Business and Technology, Kosovo
E-mail: [1]davut.okcu@ubt-uni.net, [2]ebubekirsezgin34@gmail.com, [3]htcklc43@gmail.com

Abstract - The aim of this study is to derive a new total sediment load formula which is more accurate and which has less application constraints than the well-known formulae of the literature. There are many sediment transport formulas in the literature but most of them are derived from small data sets of natural rivers or experimental designs. Each formulation has its own restrictions, which depends on the original dataset of that study. None of them have gained universal acceptance. To achieve this aim, a wide range of dataset is compiled which includes both experimental lab(flume) data and Natural river data so that a very large range (of parameters) has been achieved. Then this dataset is used to generate a new formula.Five most known total sediment transport formulae, which are approved by American Society of Civil Engineers (ASCE) are used for benchmarking. The dimensionless parameters of these widely used formulae are used as inputs in a new regression approach. The new approach is called Polynomial Best subset regression (PBSR) analysis. The aim of the PBRS analysis is fitting and testing all possible combinations of the input variables and selecting the best subset. All the input variables with their second and third powers are included in the regression to test the possible relation between the explanatory variables and the dependent variable. While selecting the best subset a multistep approach is used that depends on significance values and the Multicollinearity degrees of inputs. According to proposed formulae, the sediment transport phenomenon is deeply related to velocity of water (U), slope (S), depth (H), shear velocity (U*), sediment size (d50) and specific gravity (Gs). The new formula is compared to others in a holdout dataset and detailed performance investigations are conducted for field and lab datasets within this holdout data. Different goodness of fit statistics are used as they represent different perspectives of the model accuracy. After the detailed comparisons are figure out, PBSR is the most accurate equation that is also applicable on both flume and river data. Especially, on field dataset the prediction performance of the proposed formula outperformed the benchmark formulations.

Index Terms - River Hydrology, Sediment Transport, Total Sediment Load, Polynomial Best Subset Regression (PBSR).

I. INTRODUCTION

Sedimentation refer to the motion of solid particles, called sediment. The natural processes of erosion, transportation and sedimentation have been active throughout geological time and have shaped the present landscape of our world. Today, they can cause severe engineering and environmental problems [1]. In alluvial hydraulics, sediment has been defined as rock, mineral particles such as clay, sand, silt, gravel, and boulders transported or moved by the flowing water. The main sources of sediment production are the weathering of rocks, erosion by the flow of water over soil surfaces, channel bed erosion, and bank caving. Traditionally, sediment is decomposed into wash load or river load depending on the particles origin. Another classification is done as suspended sediment load or bed load and these definitions are related with the movement of the particles.Sediment transport prediction is vital to water engineering problems like dam management, scoring and bridge problems, irrigation channel sedimentation, economical life of reservoirs etc. It is important to possess a reasonable measurement of sediment transport rate due to its importance in the planning, design and management of hydraulic and hydrological projects [2]. By the perspective of environmental engineering, the moving sediments also affect the water quality of river systems and act as a transport mechanism for materials such as pesticides, heavy metals, nutrient, decomposable organics, and bacteria. Hence, the phenomenon of sediment transport is of great engineering importance, and together with its related problems, governs a large number of situations that are of major concern to the civilized man. Another source of difficulty in the development and assessment of sediment load predictors is the existence of uncertainty and inaccuracy in sediment transport data. This is due to limitations in measurement equipment and the large temporal and spatial variations of concentration and transport rates of sediment particles in natural channels. Usually a limited number of samples are taken from the channel, which may not be truly representative of the entire flow. In general the samples are taken from the same river for the same cross section within a time period. This data creates a time series (daily, monthly and annually). In literature there are many studies in time domain [3].

II. LITERATURE REVIEW

In 1950, Reference[4] introduced his bed-load function, which is based on the probability concept. He estimated bed load for different size of sediments found on the bed. Reference [5] proposed a relationship that give both quantity and quality of total, suspended and bed loads as functions of stream and sediment characteristics. Reference [6] developed graphical solutions for total load based on laboratory and field data. Reference [7] evaluated works of [6] on

Amazon, Orinoco, and Mississippi Rivers. In their analysis, they found reference [6] equation overestimated the unit sediment discharge for the rivers under study. The total sediment load includes the wash load and the bed-material load. The bed-material load consists of bed-load and suspended load. Generally, two approaches are available for predicting the bed-material load in a river. One is to estimate the bed-load and suspended load in separate calculations. This is based on the fact that the hydrodynamics of each mode of sediment transport is different. The methods developed by [4]-[8]. Reference [9] has approached the sediment transport phenomenon from the point of view of the rate of potential energy expenditure of the flow. Reference [9] defined the unit stream power, which is the product of mean flow velocity and energy slope (VS), as the rate of expenditure of flow potential energy per unit weight of water. From the analysis of a massive data bank, Reference [9] found that Unit stream power is the best dominant variable, which could be related to the sediment concentration. Reference [10] introduced the power concept and similarity principle to obtain sediment transport function. Reference [10] formula was developed based on flume data, using sediment size of bed material as input. Reference [11] had introduced sediment transport functions in terms of three dimensionless groups namely, size, mobility and transport rate of sediment. His functions are based on flume data carried out with uniform or near uniform sediment with flow depths up to 0.4 meters. Reference [12] tested 14 formulas using a compendium of sediment transport data from the laboratory and field records. He concluded that Reference [10]-[11]-[12] formulas are acceptable.

Most total load equations are actually total bed material equations. A sediment transport equation based on universal steam power is presented for the prediction of bed material concentration of the rivers or streams by [13].Reference [14] later modified his unit stream power formula for the computation of total bed-material load in a sediment-laden river with a high concentration of fine suspended material. Reference [14] presented a formula to estimate total bed material load in a sediment laden river with high concentration of fine material. They derived their sediment concentration functions using multiple linear regression analysis of laboratory data, based on unit stream power theory. Reference [15] had proposed the use of energy concept in the development of sediment transport equation based on universal stream power by [9]. According to them, this has advantage in eliminating the energy slope as a parameter. They had confirmed that the relationships derived from flume experiments of shallow flows should not be universally applied to large rivers with deep flows. Reference [16] recommended six total sediment load formulas; [9], [10], [11], [12], [15], [17].

Reference [18] developed a mathematical model, based on the kinematic wave theory that predict the evolution and movement of bed profiles in alluvial channels under the equilibrium conditions. In order to discretization the equations, the explicit finite difference method was used. To test the model, flume and field data was used. Then, they improved the model, for non-equilibrium conditions. Reference [19] developed a total bed material formula by using multiple linear regression model. The authors focused on high gradient river sediment transportby using Regression models [20].

III. RESEARCH METHODOLOGY

The dataset is compiled from many different sources those includes observational river data and the outputs of many experimental studies. Reference [12] produced one of the most comprehensive compilations of existing flume and field data in this area, and the present work is based on those data, which consist of both field and flume type data. Initially, each data set has complete records of the flow discharge (Q-m3/s), channel width (B-m), flow depth (H-m), hydraulic slope (S), median sediment size d50-m), sediment gradation (σ), specific gravity of sediment (G) and temperature (T-C). Bed material load transport - concentration (C-ppm) is a function of these parameters.

$$C = f\left(Q, B, H, S, d_{50}, \sigma, G, T\right)$$

All parameters are derived by using these attributes. In the model many non-dimensional variables are used that is produced by using Buckingham pi theorem. In Table 1, descriptive statistics of river/lab dimensions and some sediment parameters are presented. The descriptive statistics of the dataset indicate that the range of the river discharges are very high. As shown in the field part of Table 1 the widest river is about 1100 m and the narrowest river is only 0.35 m width. So the range is very high same as the discharge values. On the other side within the experimental designs the widest and the narrowest designs are 2.44 and 0.31, respectively, the range of sediment concentration (C_ppm) and d50_mm are very large for flume data. All the variables are included in the analysis with their first, second and third order force therefore 30 parameters are used for best subset selection process. While taking the powers (of the parameters in using them in regression as inputs) increases the nonlinear estimation capability of the model, this process increases the collinearity problem in the developed models.

		Valid N	Mean	Median	Min.	Max.	Range	Perc. %10	Perc. %90	Std. Dev.	Skewness	Kurtosis
FIELD	Q-(m^3/s)	1173	1505	134.78	0.001	28826	28826	2.86	4899	3824	4.00	18.52
	B-(m)	1173	181.01	82.60	0.35	1109	1109	13.72	491	235	2.23	4.90
	H-(m)	1173	3.001	1.703	0.034	17.282	17.2	0.316	9.327	3.607	1.697	2.012
	S	1173	0.685	0.720	0.006	6.690	6.684	0.042	1.500	0.685	2.156	11.59
	d50_(mm)	1173	0.740	0.323	0.083	3.400	3.317	0.161	2.204	0.821	1.402	0.312
	C-(ppm)	1173	522.5	200.24	5.61	5830	5824	44.00	1420	847.7	3.14	11.25
LAB	Q-(m^3/s)	927	0.13	0.05	0.001	2.08	2.07	0.01	0.36	0.19	3.79	24.60
	B-(m)	927	1.12	0.91	0.31	2.44	2.13	0.49	2.44	0.64	1.08	-0.04
	H-(m)	927	0.147	0.125	0.032	0.585	0.553	0.064	0.287	0.083	1.174	1.521
	S	927	0.002	0.002	0.000	0.011	0.011	0.001	0.003	0.001	2.214	7.065
	d50_(mm)	927	0.416	0.375	0.100	1.500	1.400	0.150	0.930	0.263	1.368	2.191
	C-(ppm)	927	875.76	249.60	2.90	12900	12897	26.00	2480	1552	3.57	16.66
TOTAL	Q-(m^3/s)	2100	840.50	3.84	0.001	28826	28826	0.02	2129	2953	5.47	35.35
	B-(m)	2100	101.60	19.20	0.31	1109	1109	0.70	390	197	3.05	10.13
	H-(m)	2100	1.7408	0.3290	0.0323	17.28	17.2	0.076	6.233	3.046	2.521	5.96
	S	2100	0.3832	0.0438	0.0002	6.6900	6.690	0.001	1.290	0.614	2.655	13.99
	d50_(mm)	2100	0.5969	0.3435	0.0830	3.4000	3.317	0.150	2.204	0.658	2.065	3.11
	C-(ppm)	2100	678.44	213.05	2.90	12900	12897	33.91	1829	1223	4.04	23.21

Table I: Descriptive statistics of raw dataset.

Like any data mining project, before developing new models the dataset needs to be prepared for the analysis. In this manner, data partition and data transformations are performed. Before the model development process dataset is divided into two sets as training (70%) and validation (30%). The validation part is not used in model development. The models are compared in validation part considering the lab data (44%) and field data (56%) partition. Total number of data set is 2100, [lab data (927) and field data (1173)].

After data partition, Best subset regression models are modified and used to generate a new formula. As known Regression approach is an alternative to estimate the sediment load when there is no sediment data, such as particle size, to use a predictive sediment transport formula.

Reference [21] used possible-subset regression with stepwise regression. The authors pointed that there is a limitation in stepwise regression search approach which is it presumes there is a single "best" subset of X variables and seeks to identify it. Nevertheless, there is often no unique "best" subset. Therefore, for huge input numbers best subset solution might give the most parsimonious model if the comparison parameter is sensitive to input number.

The Best subset regression approach is adopted to fit and test (significance, F- Anavo) all possible combinations of the input variables in a regression equation and to select the best solution. The new approach is called Polynomial Best Subset Regression Model (PBSR). The aim of the polynomial design is to investigate if there is any nonlinear relationship between model inputs and the output by creating the nth power of the attributes and using them in the model considering the correlation between them. So n is taken as 3 in this study. In this study 10 non-dimensional variables are used as a starting point. In fact, for thirty parameters,

potentially 2^{30} models (1073741824) could be developed. Before initiating the permutation process of the algorithm a constraint is applied to decrease the possible number (potentially possible) of the models. In the study well known collinearity diagnostic Variance inflation factor (VIF) is used to eliminate the models initially (VIF >3000). After this elimination possible 53009101 variations (different input combinations) are tried and compared to each other. In this process, the adjusted R^2 is used to select the most parsimonious and the most accurate models. But the selected models still could have many collinear parameters. To overcome this problem the algorithm is modified to select the most accurate model while model input parameter selection is suppressed to decrease multicollinearity problem. The model algorithm is repressively adopted to eliminate the combinations (input sets) which includes very high VIF (>10) valued parameters by optimizing delta parameter that is used for sweeping operator in computation of reverse matrix. Different values of sweeping operators are tried between 10 and 1. These variables have been selected within the non-dimensional parameters of the 5well-known formula in the literature. These formulas are 1- Engelund and Hansen's (1967) (EH), 2- Ackers and White's (1973) (AW), 3- Yang's (1973) (YANG), 4- Karim's (1998) (KARIM) formula, 5-Molinas and Wu (2001) (MW) formulas.

$$C = f\left(S, \frac{H}{d_{50}}, \frac{U^3}{g.H.w}, \frac{U^* d_{50}}{v}, \frac{HS}{(G_s-1)d_{50}}, \frac{U}{w}, \frac{w}{U^*}, \frac{U}{\sqrt{(G_s-1)gd_{50}}}, \frac{wd_{50}}{v}, \frac{US}{w}\right)$$

Where represent of the symbols are; sediment concentration: (C-ppm), energy slope (S), flow depth (H-m), median sediment particle size (d_{50} -m), submerged specific gravity (G_s -1),velocity of water (U –m/s), shear velocity (U* -m/s), gravitational acceleration (g- m/s^2), settling velocity for a particle (w – m/s), water kinematic viscosity (v -m^2 /s).

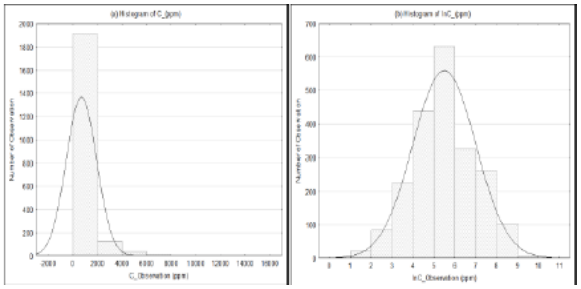


Figure 1 : Histogram of target Concentration(ppm), a- Raw data, b-Transformed data.

To ensure the normality assumption of regression models, the dependent variable (C-ppm) is transformed by natural logarithm. The histogram of the concentration of carried sediment is presented in Figure 4.1. As shown in the figure after transformation the probability density function become close fit to the normal curve, indicating the accordance for the use in regression models.

IV. RESULTSAND COMPARISION OF THE NEW FORMULA

A. New Proposed Formula (PBSR)

The output formula of the PBSR model is (1);

$$C = 34,45\frac{P^{3,239}J^{0,005}}{L^{0,066}R^{0,146}} \quad (1)$$

Where represent of the parameters in the equation (1) are C: total bed-material concentration in parts per million by weight, P:dimensionless particle parameter, J:dimensionless energy slope parameter, L: dimensionless length parameter, and R: dimensionless particle Reynold parameter, which are defined as (2):

$$\begin{cases} P = \dfrac{U}{\sqrt{(G_s-1)gd_{50}}} \\ lnJ = (lnS)^3 \Rightarrow J = e^{(LnS)^3} \\ lnL = (ln\dfrac{H}{d_{50}})^2 \Rightarrow L = e^{(\ln(H/d_{50}))^2} \\ R = \dfrac{U_* d_{50}}{\nu} \end{cases} \quad (2)$$

Where represent of the parameters in the equation (2) are U: velocity of water, S: energy slope, Gs-1: the submerged specific gravity,d_{50} : the median size of particle diameter, H: height of water, U*: shear velocity, that is calculated equation (3);

$$U_* = \sqrt{grS} \quad (3)$$

Where represent of the parameters in the equation (3) are g: gravitational acceleration, r: hydraulic radius, S: slope. So it can be concluded that this formula (1) can be used for the rivers which have a slope between 0.0002 and 6.69, Particle diameter-d50 with 0.083- 3.4 mm, concentration between 2.9 - 12900 ppm.

B. Model Evaluation and Comparison Criteria

In literature, there are many statistics to compare the models. In general these statistics are known as "Goodness of fit statistics". These statistics are useful to understand how model capture the real world or at what points the model has a bias. Therefore each of them have different focus points. All the models are compared by using the validation (not used) data. According to Reference [22], model validation is the process of demonstrating that a given site-specific model is capable of making "sufficiently accurate" simulations, although "sufficiently accurate" can vary based on project goals [23]. On the other hand, classical scatter diagrams are very useful to see where model underestimates or overestimates the real data (observational data). In this study beside scatter plots; correlation coefficients (r), Nash-Sutcliffe efficiency (NSE), and Logarithmic transformation variable (e) and Adjusted R2 statistics are used to compare the model accuracies. Additionally the error statistics Mean absolute error (MAE) and Percent bias (PBIAS) are used to examine the error perspective. All used statistics have different perspectives and focus on a partial side of evaluation. By using the determined model fit statistics, new proposed equation (1) and the benchmark model outputs are compared in validation dataset. The formula restrictions of the models are used in comparison process. Additionally, predictive capabilities of the models for field and lab data are investigated separately.

C. Validation whole data set comparisons

For the validation dataset the determined model performance statistics are given in Table II. According to table EH, AW and YANG formulas have bad performance in the sense of all criteria. Negative big values of NSE indicate that these three models are worse than average base models. The scatter plots of these models are given in Fig.1. As understood from Maximum Absolute Error statistics, these models produce very high inaccurate predictions so these unsignificant prediction decrease the performance of model sharply. According to the table 5.2, by the e and PBIAS the PBSR model tends to make underestimation while the MW tends to make overestimation.

For mixed validation data (Field+Flume) the best two models are PBSR and MW. When these two are compared, by all of the performance indices, the proposed model PBSR outperforms the others and the MW is relatively successful.

	PBSR	Karim	MW	EH	AW	Yang
Correlation	0.92	0.36	0.90	0.01	-0.02	0.02
R. er. mean	0.54	14.02	4.59	2775	20651	167
e	-0.01	0.81	0.53	1.99	0.88	1.30
NSE	0.82	-43.76	0.71	-23856	-3976	-4834
PBIAS	0.68	-556	-9.49	-35666	-5430	-5795
Adj. R2	0.846	0.119	0.810	-0.014	-0.008	-0.012
Max AE	3989	53607	1726	98831	22080	83189
Median AE	91	1703	393	637769	691	45673
Mean AE	287	4464	448	46625	41594	66418

Table II: Model comparison statistics for validation dataset.

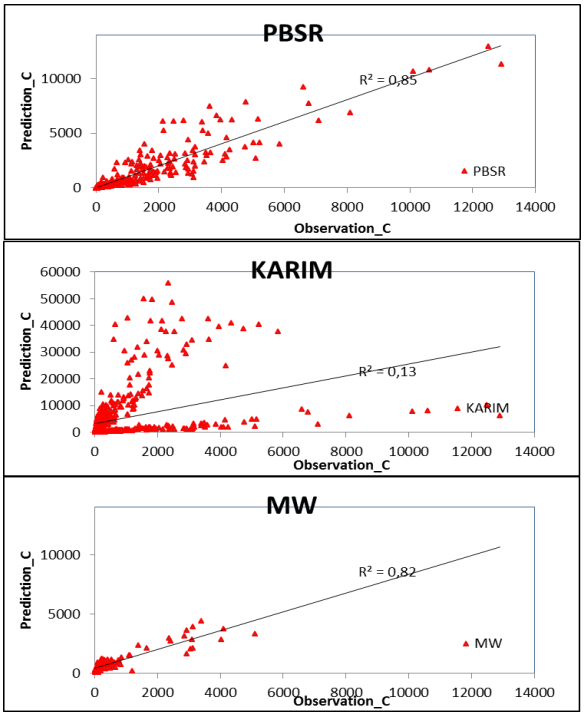


Figure 1: Scatter plot of PBSR-KARIM-MW for validation.

The models MW and PBSR give high correlation coefficients and the formula of the KARIM is relatively good. As seen that the output of the KARIM has a two way spread like the others (EH, AW and YANG).

D. Model comparisons for Field dataset.

The models are compared with their field data (within validation set) predictive capacities. Table III, represents the determined statistics of the models for field data. In comparison process only validation data is used and the specific restrictions of the models are considered.

The field data performances of the EH, AW and YANG are very bad estimation. The NSE statistics of these three are all negative and the correlation coefficients are almost zero. Mean relative errors are huge when compared to the other models (PBSR, KARIM and MW). For field data best two models are PBSR and KARIM.

	PBSR	Karim	MW	EH	AW	YANG
Correlation	0.86	0.84	0.28	0.30	-0.02	0.41
R.Er.mean	0.54	24.05	2.97	4661	37043	302
e	0.01	1.32	0.56	3.27	1.41	2.26
NSE	0.51	-217	-0.54	-611	-1942983	-2313
PBIAS	-13.3	-1470	-4.67	-9071	-1414486	-1424
Adj. R2	0.74	0.71	-0.25	0.07	-0.01	0.15
Max AE	3989	53607	567	9883	22080659	83189
Median AE	81.7	3665	453	8578	6350	95710
Mean AE	250	7756	426	7845	7460815	11759

Table III: Model comparison statistics for field dataset.

The correlation coefficients and the Adjusted R2 values of PBSR and KARIM are close to each other. Strangely the NSE value of KARIM is negative and high. This means that the model is worse than the base average model. KARIM gives very high overestimations for most field data. But the variations

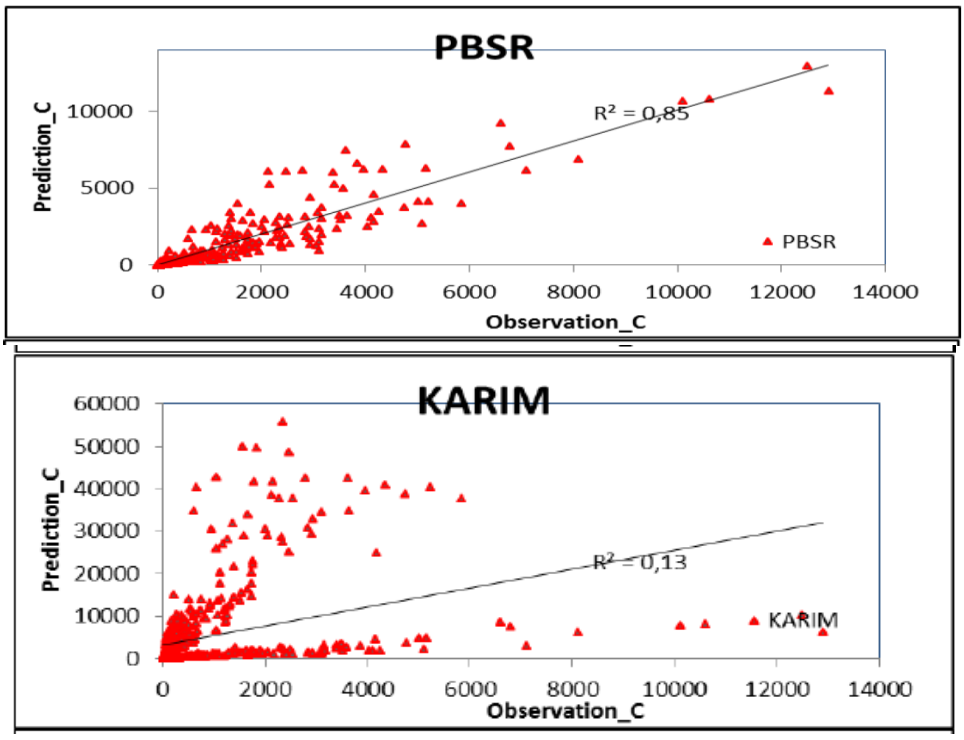
of KARIMs' predictions are synchronous with observations. This is why the correlation coefficients and the Adjusted R2 values are high. However, NSE and some other model comparison criteria detect this critical point. The PBIAS value is very high and negative (-1470) indicating that the model overestimates, the "e" (1.32) is positive and close to the "e" value of AW (1.41). So after detailed examinations it can be concluded that for field data set the proposed PBSR formula outperforms all the others.

E. Model comparisons for Flume (Lab) dataset.

The performance statistics of PBSR, KARIM and MW are close to each other. The ranking for all fit statistics of table IV is PBSR, KARIM and MW. As seen from the figures KARIM and MW formulas give high deviations for large values of sediment concentration. For Flume dataset PBSR, KARIM and MW give better performance than others (EH, AW, YANG) (Table IV), (Figure 2). By correlation coefficients and Adjusted R2 values these three models (EH, AW, YANG) seems they cannot capture almost nothing in the variation of the observed lab data. But this immediate result may not be completely true, since these statistics (like many others) are very sensitive to extreme data. For example after the restrictions of EH formula in validation dataset there exist only 154 data and for only 7 data the EH model produces extremely huge values. Because of this the correlation coefficient and the Adjusted R^2 value are approaching zero. If these 7 data points are eliminated from the dataset, then the correlation coefficients rise to 0.87 (from 0.03). However, these seven points are completely concordant with the restrictions of the formula; all the attributions of these 7 cases are within the restrictive boundaries of the formula.

	PBSR	KARIM	MW	EH	AW	YANG
Correlation	0.95	0.93	0.90	-0.03	0.07	0.005
R. Er. mean	0.55	1.49	4.87	174.30	25.93	14.06
e	-0.03	0.18	0.53	0.22	0.22	0.21
NSE	0.89	0.84	0.76	-8866	-561	-229
PBIAS	9.40	12.03	-12.48	-1398	-529	-538
Adj. R2	0.90	0.87	0.81	-0.03	-0.01	-0.03
Max AE	3107	6643	1726	953505	665133	172068
Median AE	118	152	373	154	185	205
Mean AE	333	352	452	27259	5708	8416

Table IV: Model comparison statistics for flume dataset.



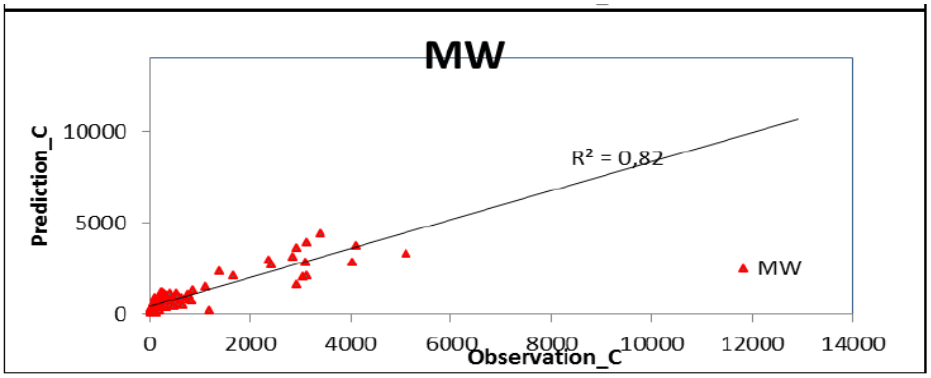


Figure 2: Scatter plots of flume data in validation, (PBSR-KARIM-MW).

CONCLUSION

In this study, 10 non-dimensional parameters of the sediment transport are used with their second and third order powers to capture nonlinear relationships. Since the parameters are generated from their powers there is a high correlation within the input dataset. The new formula is compared with 5 well known sediment formulae in the literature. The proposed formula outperforms the compared models in terms of model fit statistics

As a conclusion, the results show significant successful estimation of Polynomial Best subset regression (PBSR) for total sediment load. It has high prediction accuracy. Therefore, the proposed formula, PBSR model, can be used in sedimentation engineering applications.

ACKNOWLEDGMENTS

This study is based on the PhD thesis, 'Developing a New Total Sediment Transport Formula by Using River Parameters', of the first author in Istanbul Technical University (ITU).

REFERENCES

[1] Julien, P. Y. (2010). Erosion and sedimentation. Cambridge University Press
[2] Roushangar, K., Mehrabani, F. V., & Shiri, J. (2014). Modeling river total bed material load discharge using artificial intelligence approaches (based on conceptual inputs). Journal of Hydrology, 514, 114-122.
[3] Iadanza, C., & Napolitano, F. (2006). Sediment transport time series in the Tiber River. Physics and Chemistry of the Earth, Parts A/B/C, 31(18), 1212-1227.
[4] Einstein, H.A., (1950). The Bedload Function for Sediment Transportation in Openchannel Flows. In: Soil Conservation Servicen Technical Bulletin, No. 1026. U.S Dep. of Agriculture.
[5] Laursen, E. M. (1958). The total sediment load of streams. Journal of the Hydraulics Division, 84(1), 1-36.
[6] Colby, B. R. (1964). Discharge of sands and mean-velocity relationships in sand-bed streams (No. 462-A).
[7] Posada-G, L., & Nordin, C. F. (1993). Total Sediment Loads of Tropical Rivers.Hydraulics Engineering, 93, 258-262.
[8] Van Rijn, L. C. (1984a) "Sediment transport. Part I: Bed load transport." J. Hydraul. Eng., 110_10_, 1431–1456.
[9] Yang, C, T (1973). "Incipient Motion and Sediment Transport." Journal of the Hydraulics Division, ASCE, Vol. 99, No. HY10, pp. 1679-1704.
[10] Engelund, F., Hansen, E., (1967). A Monograph on Sediment Transport in Alluvial Streams. Teknish Forlag Technical Press, Copenhagen, Denmark, p 62.
[11] Ackers, P., White, W.R., (1973). Sediment transport: new approach and analysis. ASCE J. Hydraulic Eng. 99 (11), 2041.
[12] Brownlie, W. R. (1981). "Prediction of flow depth and sediment discharge in open channels." Rep. No. KH-R-43B, Lab. of Hydraulic Research, California Institute of Technology, Pasadena, Calif.
[13] Yang, C. T., & Wan, S. (1991). Comparisons of selected bed-material load formulas. Journal of Hydraulic Engineering, 117(8), 973-989.
[14] Yang, C. T., Molinas, A., & Wu, B. (1996). Sediment transport in the Yellow River. Journal of Hydraulic Engineering, 122(5), 237-244.
[15] Molinas, A., and Wu, B. (2001). "Transport of sediment in large sand-bed rivers." J. Hydraul. Res., 39_2_, 135–146.
[16] Garcia, M. H. (2008). Sedimentation engineering: processes, measurements, modeling, and practice. ASCE Manuals and Reports on Engineering Practice, (110).
[17] Karim, F. (1998). Bed material discharge prediction for nonuniform bed sediments. Journal of Hydraulic Engineering, 124(6), 597-604.
[18] Tayfur, G., & Singh, V. P. (2006). ANN and fuzzy logic models for simulating event-based rainfall-runoff. Journal of Hydraulic Engineering, 132(12), 1321-1330.
[19] Sinnakaudan, S., Ab Ghani, A., Ahmad, M.S.S., Zakaria, N.A., (2006). Multiplelinear regression model for total bed material load prediction. Journal of Hydraulic Engineering 132 (5), 521e528.
[20] Sinnakaudan, S. K., Sulaiman, M. S., & Teoh, S. H. (2010). Total bed material load equation for high gradient rivers. Journal of Hydro-environment Research,4 (3), 243-251.
[21] Neter, J., Wasserman, W., & Kutner, M. H. (1985). Applied linear statistical models (2nd ed.). Homewood, IL: Richard D.
[22] Refsgaard, J. C. (1997). Parameterisation, calibration and validation of distributed hydrological models. Journal of hydrology, 198(1), 69-97.
[23] Moriasi, D. N. et al. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. Asabe, 50(3), 885-900.

★★★