MODELING OF AERATION EFFICIENCY OF SMALL PARSHALL FLUME BY RANDOM FOREST REGRESSION

Amit Kumar[1], Dr. N. K. Tiwari[2], Dr. Subodh Ranjan[3]

[1]M.Tech Student,Department of Civil Engineering,National Institute of Technology, Kurukshetra

[2]Associate Professor,Department of Civil Engineering,National Institute of Technology, Kurukshetra

[3]Professor, Department of Civil Engineering,National Institute of Technology,Kurukshetra

ABSTRACT

This paper examines the potential of Random Forest regression (RF) in forecasting the aeration efficiency of Small Parshall flume. Data set consists of 60 observations which were obtained from the laboratory experiments. Out of 60 data set, arbitrarily selected 40 observations were used for training whereas residual 20 dataset were used for testing the models. Input data set consists of velocity, head, and sin θ, whereas aeration efficiency is considered as output. A comparison of results suggests that random forest works well and it could successfully use in modelling of aeration efficiency of small Parshall flume.

Keywords: -Aeration efficiency,Aeration performance, Dissolved oxygen, Parshall flume,Random Forest Regression.

I. INTRODUCTION

Dissolved oxygen (DO) is necessary to healthy streams and lakes.The existence of oxygen in water is a positive sign and the lack of oxygen is a sign of severe pollution. Many naturally occurring biological and chemical processes use oxygen, hence decreasing the concentration of DO in water. The physical process of transfer of oxygen or oxygen absorption from the atmosphere acts to replenishthe used oxygen. This process has been termed aeration or reaeration. A sufficient supply of dissolved oxygen is essential in natural rivers and in several water treatment processes. The concentration of dissolved oxygen can be improved by entraining air bubbles in a receiving pool. The transfer of oxygen across the air-water boundary at the hydraulic structure, such as Parshall flume,weir,gate and spillway may be an essential source of dissolved oxygen(DO) in the river system.A parshall flume can be consist of converging section, throat section and ends with diverging section. For a hydraulic engineer it is important to measure the flow in the hydraulic structures. Agricultural users do not able to use the proper use of water due to the absence of flow measuring devices.Without such measurement it is difficult to make a definite statement as how much amount of water actually used.One of the most cited water quality parameters in the fresh water hydrosphere(lakes,reservoirs and rivers) is dissolved oxygen (DO). The concentration of oxygen in surface waters is a main indicator of the quality of that water for human use as well

ISBN: 978-93-86171-36-8

as use by the aquatic biota. Many naturally occurring biological and chemical processes use oxygen, thus diminishing the DO concentration in the water.The growth of population and desire of man to raise living standards have resulted in higher and stronger volumes of waste matter which require treatment .The sewage coming from municipalities require an efficient and economic system of treatment.Dissolved oxygen concentration is very essential for aquatic life in the river and canal water.Hence in order to increasing the dissolved oxygen(DO) concentration in the channels having low slope a Parshall flume is essential for the increasing the DO.Some research work has been done by the researchers related to the performance of aeration of hydraulic structures.Gulliver et al., (1998) [1] estimated that the surfacearea of the bubbles in a 0.3m deep spillway flow can provide 500 times the surface area of the air water interface provided by a free surface. Dursun, O. Faruk, (2016) [2] completed the experimental study of the aeration performance of parshall flume. Chanson, H., (2009) [3] has done the work of turbulent air–water flows in hydraulic structures.Gulliver & Rindels, (1993) [4] done the measurement of air-water oxygen transfer at hydraulic structures.Chanson, H. (1997) [5], McKeogh & Ervine (1981) [6] have done some experimental work on reaeration at low-head hydraulic structures. Baylar, A., (2002) [7] Study on the effect of type selection of weir aeratorson oxygen transfer. Baylar, A., and Bagatur, T.,(2000) [8]done the work onaeration performance of weirs.

Soft computing techniques such as Random Forest regression (RF) are recently developed methods which can be used for prediction of aeration efficiency.RF techniques in developing prediction models for estimating the aeration efficiency. It learns features of the data set and adjust the system characterstics according to a given error criterion.

Fig.1 A longitudinal view of the water flume used

II.RANDOM FOREST REGRESSION

Random forest is a classification and regression process, consisting of a combination of tree predictors where every tree is produced utilizing a random vector sampled autonomously from the input vector. In regression, tree predictor proceeds on numerical values as arbitrary to class labels used by the random forest classifier (Breiman, 2001) [9]. Random forest regression used in present study consists of using a combination of variables or randomly selected variable at each node to maturea tree. Plan of a tree predictor required the decision of a variable choice measure and a pruning strategy. Various ways to deal with the choice of variable for tree

ISBN: 978-93-86171-36-8

acceptance is proposed in in literature and most approaches assign a quality measure directly to the variable. The most frequently used variable selection measures in tree induction are Information Gain Ratio criterion and Gini Index (Brieman et. al., 1984) [10]. The design of random forest regression permits a tree to develop to the maximum depth of new training data using the mixing of variables. These full-developed trees are not pruned back. This is one of the main advantages of random forest regression over other tree techniques like M5P tree model (Quinlan, 1993) [11]. Studies propose that the decision of the pruning process and not the variable determination, influence the performance of tree based algorithms (Pal and Mather, 2003) [12]. Breiman, (1999) [13] suggests that as the quantity of trees increases, the speculation error always converges even without pruning the tree and over fitting is not an issue in view of the Strong Law of large numbers (Feller, 1968) [14]. Number of variable used (*m*) at every node to grow a tree and the quantity of trees to be developed (*k*) are two user-defined parameters required for random forest regression (Breiman, 1999). At every node, just selected variables are found through for the best split. Thus, the random forest regression consists of *k* trees, where *k* is the quantity of trees to be developed which can be any value defined by the user. In the random forest based regression, the yield values arenumeric, thus the mean-squared generalization error for any numerical predictor is acquired. The random forest predictor is developed by taking the average of generalization error over *k* trees. In present study random forest and bagged random forest are used.

III.AERATION EFFICIENCY

The change in oxygen concentration over time in aparcel of water as the parcel travels through a hydraulicstructure can be expressed as:

$$\frac{dC}{dt} = K_L \frac{A}{V}(C_S - C)............(1)$$

WhereK_L = bulk liquid film coefficient. The values of Cs and C are the saturation concentration of oxygen in water at prevailing ambient conditions and the actualconcentration of oxygen in the water at time t difference being proportional to the concentration gradient. The term A is the air water contact area and V is the volume of water associated with the equation (1)does not consider sources and sinks of oxygen in thewater body because their rates are relatively slowcompared to the oxygen transfer that occurs at most hydraulic structures due to the increase in free surface turbulence and the large quantity of air that is normally entrained into the flow.The predictive relations assume that Cs is constant and determined by the water atmosphere partitioning. If that assumption is made, Cs is constant with respect to time, and the oxygen transfer efficiency (aeration efficiency), E may be defined as:

$$E = \frac{C_d - C_u}{C_s - C_u} .... (2)$$

Where u and d are the subscripts indicates the upstream and downstream locations, respectively. If the value of transfer efficiency is 1.0 then up to the saturation value full transfer has occurred at the structure. The value of saturation concentration may be obtained from charts. The aeration efficiency is generally normalized to a 20C

ISBN: 978-93-86171-36-8

standard for providing the uniform basis for the comparision of measurement results. Gulliver et al.,(1990) [15]suggested the equation which should be used to describe theinfluence of temperature:

$$1-E_{20} = (1-E)^{1/f} \quad \text{................ (3)}$$

Where E is the transfer efficiency at actual water temperature;E_{20} is the transfer efficiency for 20^0C; and f is the exponent described as:

$$f=1.0+2.1 \times 10^{-2}(T-20) + 8.26 \times 10^{-5}(T-20)^2 \quad \text{……….. (4)}$$

Where T is the temperature.

IV. METHODOLOGY AND DATA SET

4.1 Methodology

Four models of Parshall flume were used in this study.Each experimental run started by filling of tank with clean water in which sodium sulphite (Na_2So_3) and cobalt chloride (CoCl_2) were added to water in suitable amounts corresponding to the volume of available water and mixed thoroughly to reduce DO concentration to 0 mg/l. The water was fed into the flume by a centrifugal pump from tank and the water falling from the Parshall flume was allowed to fall into downstream water pool. Five runs were conducted for each Parshall flume model by varying the flow rate from 0.96 to 4.97 l/s in 1 l/s steps. In order to visualize the flow the side surface of open channel was made of glass. Water depth measurements were conducted using the point gauge scale at the converging, throat and diverging section of Parshall flume .Dissolved oxygen measurements were conducted by using azide-modification method and temperature measurement was also done simultaneously during sampling.

4.2 Data Set

Input Parameters	Units	Train Data			
		Min	Max	Mean	St. Dev.
Velocity	M/Sec	0.035	0.14	0.102	0.026
Depth	Cm	3.82	18.64	11.091	3.808
Sin Ɵ	Degree	0.351	0.758	0.579	0.155
Efficiency	%	0.112	0.476	0.234	0.09
		Test Data			
		Min	Max	Mean	St. Dev.
Velocity	M/Sec	0.04	0.15	0.102	0.028
Depth	Cm	4.73	15.48	10.816	3.608
Sin Ɵ	Degree	0.35	0.758	0.579	0.157
Efficiency	%	0.109	0.455	0.246	0.098

Data set consists of 60 observations were used and obtained from the laboratory experiments. Out of 60 observations arbitrarily selected 40 observations were used for training, whereas remaining data set (20) were used for testing the models. Input data set consists of velocity, head and sin Ɵ, whereas aeration efficiency was considered as output. The characteristics of experimental data are specified in Table 1.

Table 1. The characteristics of training and testing data set.

Fig. 2 Model of Small Parshall flume.

V. RESULTS AND DISCUSSION

To assess the effectiveness of Random Forest model approach in predicting the aeration efficiency. Correlation coefficient (C.C), mean absolute error (MAE) and root mean square error (RMSE) values obtained with test data set were used to the performance of Random Forest model.Table 2 provides the values of the correlation coefficient (C.C), mean absolute error (MAE) and root mean square error (RMSE) values provided by Random Forest predictive approach. The results suggest the improved performance by random forest regression in terms of CC, MAE, RMSE values used with this data set. A substantial improvement in predictive accuracy of random forest approach indicates that it can be effectively used in predicting of aeration efficiency.

Table 2.Detail of performance evaluation parameters using RF of training and testing data set

Modelling Approach	Training		
	C.C	MAE	RMSE (%)
Random Forest	0.9799	0.0144	0.0205
Testing			
	C.C	MAE	RMSE (%)
Random Forest	0.8799	0.0379	0.0477

ISBN: 978-93-86171-36-8

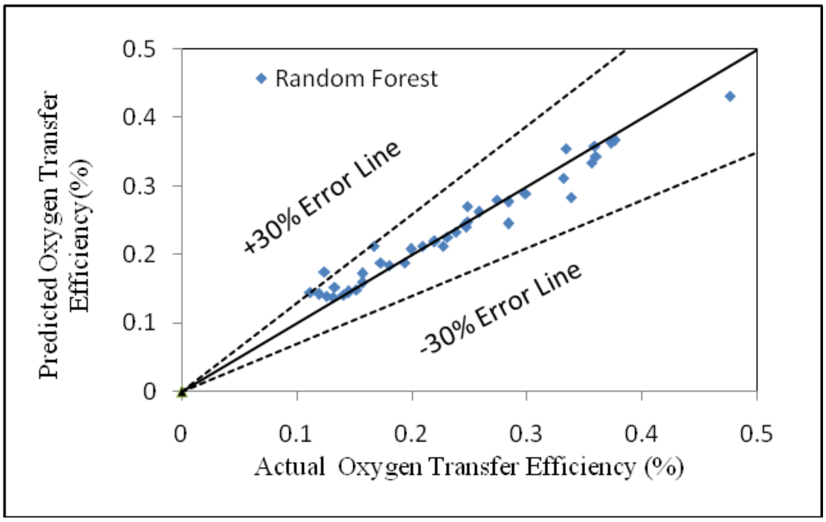

Fig. 3 Actual vs. Predicted values of oxygen transfer efficiency using Random Forest Model with training data.

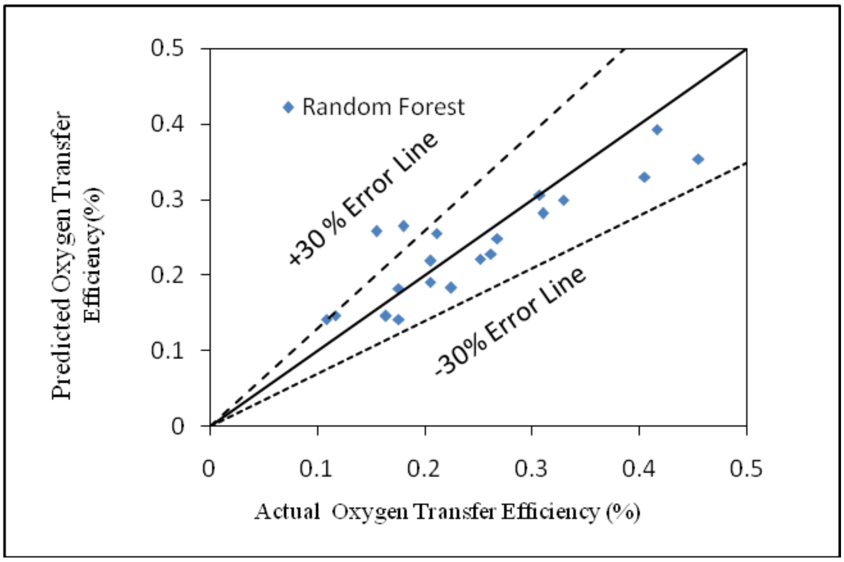

Fig. 4 Actual vs. Predicted values of oxygen transfer efficiency using Random Forest Model with testing data

Fig. 3 and 4 showing the agreement of actual and predicted oxygen transfer efficiency of Parshall flume for training and testing data set using Random Forest regression respectively. The values of correlation coefficient 0.9799, 0.8799, values of mean absolute error 0.0144, 0.0379 and values of Root mean square error 0.0205, 0.0477 were achieved for training and testing data set using Random Forest regression respectively. Fig. 5and 6 indicates that the performance of Random Forest is better in training and testing period. Random Forest estimated values are closer the actual values and follows the same path of actual values.

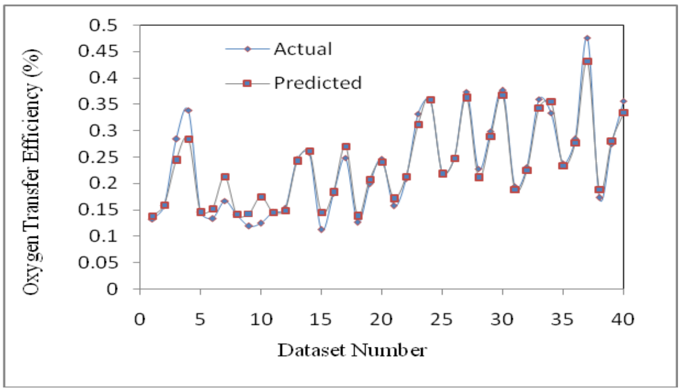

Fig. 5 Variation in predicting values of oxygen transfer efficiency using the Random Forest regression approach in comparison to actual values of oxygen transfer efficiency with training data.

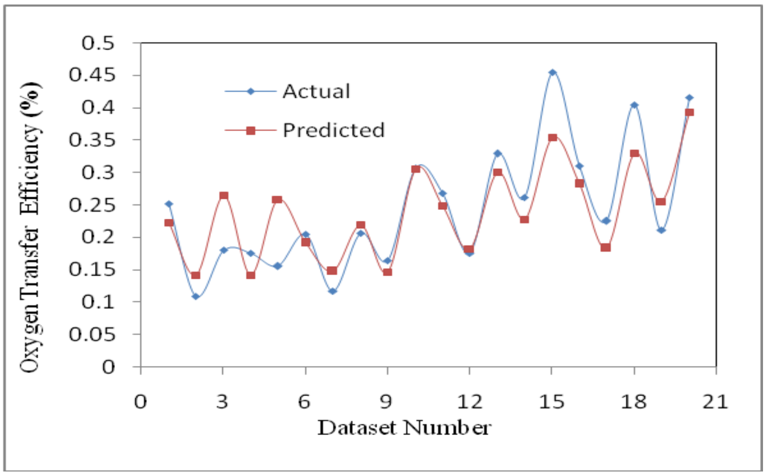

Fig. 6 Variation in predicting values of oxygen transfer efficiency using the Random Forest regression approach in comparison to actual values of oxygen transfer efficiency with testing data.

VI. CONCLUSION

This paper examines the potential of Random Forest approaches in predicting the aeration efficiency of Parshall flume. From the comparison of performance evaluation parameters, it has been concluded that Random Forest approach works well. It can be successfully used in estimation of aeration efficiency of small Parshall flume.

REFERENCES

[1] Gulliver, J. S., Wilhelms, S. C., & Parkhill, K. L. (1998). Predictive capabilities in oxygen transfer at hydraulic structures. *Journal of hydraulic engineering*, *124*(7), 664-671.

[2] Dursun, O. F. (2016). An experimental investigation of the aeration performance of parshall flume and venturi flumes. *KSCE Journal of Civil Engineering*, *20*(2), 943-950.

[3] Chanson, H. (2009). Turbulent air–water flows in hydraulic structures: dynamic similarity and scale effects. *Environmental Fluid Mechanics*, *9*(2), 125-142.

[4] Gulliver, J. S., & Rindels, A. J. (1993). Measurement of air-water oxygen transfer at hydraulic structures. *Journal of Hydraulic Engineering*, *119*(3), 327-349.

[5] Chanson, H. (1997). Air entrainment in the developing flow region of plunging jets—part 1: theoretical development. *Journal of Fluids Engineering*, *119*, 597.

[6] McKeogh, E. J., & Ervine, D. A. (1981). Air entrainment rate and diffusion pattern of plunging liquid jets. *Chemical Engineering Science*, *36*(7), 1161-1172.

[7] Baylar, A. (2002). *Study on the effect of type selection of weir aerators on oxygen transfer* (Doctoral dissertation, Ph. D. thesis, Firat University, Elazig, Turkey (in Turkish).

[8] Baylar, A., &Bagatur, T. (2000). Aeration performance of weirs. *WATER SA-PRETORIA-*, *26*(4), 521-526.

[9] Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

[10] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

[11] Quinlan, J. R. (1993). Combining instance-based and model-based learning. In *Proceedings of the Tenth International Conference on Machine Learning*, 236-243.

[12] Pal, M., & Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote sensing of environment*, *86*(4), 554-565.

[13] Breiman, L. (1999). Prediction games and arcing algorithms. *Neural computation*, *11*(7), 1493-1517.

[14] Feller, W. (1968). *An introduction to probability theory and its applications: volume I* (Vol. 3). New York: John Wiley & Sons.

[15] Gulliver, J. S., Thene, J. R., & Rindels, A. J. (1990). Indexing gas transfer in self-aerated flows. *Journal of environmental engineering*, *116*(3), 503-523.