Modeling air concentration over macro roughness conditions by Artificial Intelligence techniques

T Roshni\textsuperscript{1} and S Pagliara\textsuperscript{2}

\textsuperscript{1}Assistant Professor, National Institute of Technology Patna, Bihar, India-800005
\textsuperscript{2}Professor, University of Pisa, Via Gabba 22, Pisa, Italy- 56122
Email: roshni@nitp.ac.in, s.pagliara@unipi.it

Abstract. Aeration is improved in rivers by the turbulence created in the flow over macro and intermediate roughness conditions. Macro and intermediate roughness flow conditions are generated by flows over block ramps or rock chutes. The measurements are taken in uniform flow region. Efficacy of soft computing methods in modeling hydraulic parameters are not common so far. In this study, modeling efficiencies of MPMR model and FFNN model are found for estimating the air concentration over block ramps under macro roughness conditions. The experimental data are used for training and testing phases. Potential capability of MPMR and FFNN model in estimating air concentration are proved through this study.

1. Introduction
Roughness occurs in nature in various situations such as flows in hilly streams, gravel beds, etc. The natural flows are very dynamic in nature showing a very large-scale variation in different time scales. Aeration happens due to the turbulent flow over macro and intermediate roughness conditions. Aerated flows or two-phase flows is recreated artificially in rivers by flows over block ramps or rock chutes [1]. Natural rocky rivers act as river training structures and the human intervention leads to regenerate a similar situation in the natural rivers as river training structure. These structures offer an alternative form of drop structures, fish ladders etc when the protection of the morphology and ecology of water course is required [2]. Air entrained in the flow leads to a positive effect in terms of flow reaeration and dissolved oxygen [3] representing a valid methodology to improve the fish habitat and to reduce the pollutant content of rivers or water bodies. Aeration studies over block ramps and stepped spillways have been carried out experimentally by [1], [2], [4], [5] etc. A close view of the aerated flows over macro roughness conditions is shown in Figure 1.

Figure 1. Aerated flow over ramps in River Maglioggio (Italy).[6].
Recent literature surveys show that a few investigators attempted to investigate the potential of artificial intelligent techniques in aeration studies [7]. This paper aims to develop models to estimate the air concentration over block ramps by Feedforward Neural Network (FFNN) and Minimax Probability Machine Regression (MPMR) models. Models are developed based on the experimental results of [1], [2]. This paper also emphasizes the applicability of regression models in the estimation of hydraulic parameter.

2. Theory of Models

2.1. Feedforward Neural Network (FFNN)
Artificial Neural Network acts as a powerful computational model and can applied to all fields of engineering. Previous literature work proves the efficacy of conventional model FFNN with Back Propagation algorithm. In FFNN, two phases involved are training and testing phase. Correspondingly, 70% and 30% of the data are divided for training and testing. In the three-layer system, 1st layer denotes the input layer and nodes in the input layer corresponds to number of inputs. Second layer is the hidden layer and the selection of nodes are by trial and error method and third is the output layer and the node corresponds to the output required. Model development proceeds until the error generated between the output and the target is within the desirable range. Transfer function used in this study is sigmoid function [8].

2.2. Minimax Probability Machine Regression (MPMR)
In recent studies only, the potential applicability of MPMR is noted and it acts a good alternative for Support Vector Machine (SVM) [9]. Related work of SVM can be seen in [10]. MPMR is introduced by Strohmann and Grudic [11]. Recent literature work of MPMR in prediction of seismic ultrasonic attenuation [12], prediction of fast fading channel [13] forecasting evaporative loss [9] and a few more reveals the potentiality of this regression model. MPMR follows a regression model for y as

\[ y = [\sum_{i=1}^{N} \beta_i K(x_i, x) + b] \pm \varepsilon \] (1)

In eqn. 1, x and y represents the inputs and corresponding outputs, \( K(x_i, x) \) is kernel function and generally, Radial Basis Function (RBF) is used as the kernel function. The other variables \( \beta, b \) are outputs obtained by MPMR algorithm and \( \varepsilon \) shows the limits of error fluctuations. The MPMR algorithm is established in MATLAB 2010.

3. Experimental Facilities and Data Collection
The source of data for the modeling analysis is taken from the study of [1], [2], [5], [14]. The experiment is conducted in a 7m long tilting flume with an effective width of 0.3m. Bed surface is filled with irregular boulders to simulate the conditions of roughness. Inflow to the flume is measured by digital flowmeter, point gauge to measure the depth of flow and single tip conductivity probe to measure the air concentration. Both the point gauge and the conductivity probe are connected to the trolley assembly, such that measurements can be taken at all possible locations. The single tip conductivity probe is connected to a data logger. The experiments vary for different slopes, discharges and different bed materials [1], [2]. The measurements were taken only in the uniform flow region. The experimental range of the data are shown in Table 1.

| Table 1. Experimental data range |
|-------------------------------|-----------------|----------------|-----------------|-----------------|-----------------|
|                              | \( d_e/D_{84} \) | Slope | \( Q \) (m³/s) | \( k_c/D_{84} \) | \( C_{mean} \) |
| Minimum                      | 0.52            | 0.136 | 0.006           | 0.64            | 0.13            |
| Maximum                      | 1.81            | 0.46  | 0.03            | 2.12            | 0.42            |
Where \( d_e \) is the effective depth of flow, \( D_{84} \) is for the bed material used, \( Q \) is the discharge, \( k_c \) is the critical flow depth, \( \bar{C}_{\text{mean}} \) is the average air concentration in the uniform flow region.

4. Performance Measures

As a performance measure, different indicators are used in this study. The indicators are Root Mean Square Error (RMSE), Coefficient of Correlation (R), Nash–Sutcliffe model efficiency (NSE), and Standard Deviation (STD).

\[
\text{RMSE} = \left( \frac{1}{N} \sum_{i=1}^{N} (h_{oi} - h_{ci})^2 \right)^{1/2}
\]

\( NSE = 1 - \frac{\sum_{i=1}^{N} ((h_{oi} - h_{ci})^2)}{\sum_{i=1}^{N} ((h_{oi} - \bar{h}_{oi})^2)} \)

\[
R = \frac{\sum_{i=1}^{N} (h_{oi} \cdot h_{ci})}{\sqrt{\sum_{i=1}^{N} h_{oi}^2 \sum_{i=1}^{N} h_{ci}^2}}
\]

Where, \( h_{oi} = \) observed value and \( h_{ci} = \) predicted value, and \( N = \) number of observations, \( P \) is the input quantity and \( \bar{h}_{oi} \) and \( \bar{h}_{ci} \) are the mean of the observed values and the calculated values [1].

5. Results and Discussion

The experimental results were modelled with FFNN model and MPMR model. The input parameters considered in this study for modeling are \( d_e/D_{84} \), Slope, \( Q \) and \( k_c/D_{84} \). The output obtained is air concentration. For the model development, 70\% of the data is taken for training and 30\% of the data for testing. For FFNN model, different architectures are developed with different hidden nodes and the best architecture is selected for lowest RMSE value of the model. Similarly, MPMR model is also developed by varying the internal parameters till optimum model is developed. Visual performance of the observed experimental data and the MPMR model and FFNN model data is shown in Figure 2a and Figure 3a. A plot of confidence interval (+95\% to -95\%) band and prediction interval band (+95\% to -95\%) is also plotted for MPMR and FFNN model results to show the performance of the developed models (Figure. 2b and 3b). Figures (2-3) display a very good correlation for both the models and it is difficult to infer a best model out of two. Overall performance of training and testing phase is verified by plotting Regression Error Characteristics (REC) Curve (shown in Figure 4). REC curve gives a plot of accuracy in terms of absolute deviation. In this curve also, both models show potential ability of prediction with a slightly more weightage to FFNN.

A novel representation of performance measure indicator is Taylor diagram. In this, Standard deviation, Correlation coefficient and RMSE (blue dotted line) (shown in Figure 5) values are considered together to find the best model. Based on the training result, FFNN is showing better result, for testing model MPMR is more towards the observed values. Overall performance shows that, there was a good agreement between the experimental and the model results for FFNN model. The performance indicators for training and testing phase is shown separately in Table 2. This is also in line with the above results.
Figure 2. (a) Visual comparison of MPMR model result with that of observed data (b) Data within the confidence interval band and prediction interval band.
Figure 3. (a) Visual comparison of FFNN model result with that of observed data (b) Data within the confidence interval band and prediction interval band.

Figure 4. REC Curve for MPMR and FFNN model result
Figure 5. Taylor diagram plotted for MPMR and FFNN Models.

Table 2. Model Performances Indices for FFNN and MPMR Models

|       | FFNN   | MPMR  | FFNN   | MPMR  |
|-------|--------|-------|--------|-------|
|       | Training |       | Testing |       |
| R     | 0.9595  | 0.94  | 0.8942  | 0.96  |
| RMSE  | 0.022   | 0.0267| 0.0445  | 0.02  |
| NSE   | 0.9178  | 0.88  | 0.66    | 0.85  |
| STD   | 0.0787  | 0.0743| 0.0918  | 0.0681|

6. Conclusions
Aeration concentration varies with the macro roughness and the slopes. Block ramps or rock chutes are used in the experimental study to recreate the macroroughness conditions. Through this work, the potential efficacy of the soft computing techniques in the field of hydraulic engineering is studied. For that, MPMR and FFNN models are used in estimation of air concentration over block ramps. Both the models show equal potential in modeling efficacy. A slight better performance was noticed for FFNN model than MPMR model. Overall, both MPMR and FFNN models proved their capability in modeling aeration concentration over block ramps.

Further investigations in this field can be extended to the application of mathematical models in other fields of hydraulic engineering. Similar type of work can be used to study the effect of aeration over different roughness conditions in improving dissolved oxygen concentration, effect on pollutant loads, etc.

7. References
[1] Pagliara S, Carnacina I and Roshni T 2010 Air-water flows in presence of staggered and row boulders under macro-roughness condition Water Resources Research 46(8),11pp
[2] Pagliara S and Chiavaccini P 2006 Energy dissipation on block ramps J. of Hyd. Engg., 132(1), 41-48.
[3] Moog D B and Jirka G H 1999 Stream Reaeration in Nonuniform Flow: Macroroughness Enhancement J. of Hyd. Engg., 125(1), 11-16.
[4] Chanson H 1995 Air bubble entrainment in free-surface turbulent flows. Report CH46/95, Department of Civil Engineering, The University of Queensland, Brisbane.
[5] Pagliara S, Carnacina I and Roshni T 2010 Self-aeration and friction over rock chutes in uniform flow conditions. J. of Hyd. Engg. 136(11), 959-964.
[6] Roshni T 2011 Two phase flows over macro and intermediate roughness. PhD Thesis. University of Pisa, Italy
[7] Khalifa A, Bayoumi S and Monayeri O EI 2011 Mathematical modeling of aeration efficiency and dissolved oxygen provided by stepped cascade aeration Water Science and Technology. 1-9.
[8] Sahoo S and Jha M K 2013 Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment Hydrogeol. J. 21, 1865–1887.

[9] Ravinesh C D and Samui P 2017 Forecasting Evaporative Loss by Least-Square Support-Vector Regression and Evaluation with Genetic Programming, Gaussian Process, and Minimax Probability Machine Regression: Case Study of Brisbane J. Hydrol. Eng., 22(6), 1-15.

[10] Nurcihan C 2014 Application of support vector machines and relevance vector machines in predicting uniaxial compressive strength of volcanic rocks J. of African Earth Sciences. 100, 634-644

[11] Strohmann T and Grudic G Z 2002 A formulation for minimax probability machine regression Proc., Advances in Neural Information Processing System, MIT Press, Cambridge, MA, 769–776.

[12] Manoj K, Mittal M and Samui P 2013 Performance assessment of genetic programming and minimax probability machine regression for prediction of seismic ultrasonic attenuation Earthq. Sci., 26(2), 147-150.

[13] Liming Y, Wang L, Sun Y and Zhang R 2010 Simultaneous feature selection and classification via Minimax Probability Machine Int J Comput Intell Syst. 3(6), 754-760.

[14] Pagliara S, Carnacina I and Roshni T 2011 Inception point and air entrainment on flows under macro-roughness conditions J. of Environ. Engg.137(7), 629-638.