DOI: 10.5604/01.3001.0013.7917

JAMME

Journal of Achievements in Materials and Manufacturing Engineering

International Scientific Journal published monthly by the World Academy of Materials and Manufacturing Engineering

Kernel function based regression approaches for estimating the oxygen transfer performance of plunging hollow jet aerator

M. Kumar *, N.K. Tiwari, S. Ranjan

Department of Civil Engineering, National Institute of Technology Kurukshetra, Kurukshetra, India

* Corresponding e-mail address: munishmehta123@gmail.com

ABSTRACT

Purpose: To evaluate the capability of various kernels employed with support vector regression (SVR) and Gaussian process regression (GPR) techniques in estimating the volumetric oxygen transfer coefficient of plunging hollow jets.

Design/methodology/approach: In this study, a data set of 81 observations is acquired from laboratory experiments of hollow jets plunging on the surface of water in the tank. The jet variables: jet velocity, jet thickness, jet length, and water depth are varied accordingly and the values of volumetric oxygen transfer coefficient is computed. An empirical relationship expressing the oxygenation performance of plunging hollow jet aerator in terms of jet variables is formulated using multiple nonlinear regression. The performance of this nonlinear relationship is compared with various kernel function based SVR and GPR models. Models developed with the training data set (51 observations) are checked on testing data set (24 observations) for performance comparison. Sensitivity analysis is carried out to examine the influence of jet variables in effecting the oxygen transfer capabilities of plunging hollow jet aerator.

Findings: The overall comparison of kernels yielded good estimation performance of Radial Basis Function kernel (RBF) and Pearson VII Function kernel (PUK) using the SVR technique which is followed by nonlinear regression, and other kernel function based regression models.

Research limitations/implications: The results of the study pertaining to the performance of kernels are based on the current experimental conditions and the estimation potential of the regression models may fluctuate beyond the selection of current data range due to data-dependant learning of the soft computing models.

Practical implications: Volumetric oxygen transfer coefficient of plunging hollow jets can be predicted precisely using SVR model by employing RBF as kernel function as compared to empirical correlation and other kernel function based regression models.

Originality/value: The comparative analysis of kernel functions is conducted in this study. In previous studies, the predictive modelling approaches are implemented in simulating the aeration properties of cylindrical solid jets only, while this paper simulates the volumetric oxygen transfer coefficient of diverging hollow jets with the jet variables by utilizing polynomial, normalized polynomial, PUK, and RBF kernels in SVR and GPR.

Keywords: Volumetric oxygen transfer coefficient, Multiple nonlinear regression, Gaussian Process Regression, Support Vector Regression

Reference to this paper should be given in the following way:

M. Kumar, N.K. Tiwari, S. Ranjan, Kernel function based regression approaches for estimating the oxygen transfer performance of plunging hollow jet aerator, Journal of Achievements in Materials and Manufacturing Engineering 95/2 (2019) 74-84.

ANALYSIS AND MODELLING

1. Introduction

Water jets plunging into receiving pool of water represent an effective and economical means of aeration. A scattering of air bubbles happens underneath the pool surface as a high speed liquid jet plunges on the liquid surface subsequent to being driven out from a specific fall height. The surrounding air around the plunging jet entrains at the impinging point generating a high level of disturbances due to jet impact. Exchange of oxygen mass occurs between the air bubbles and water pool which improves with the submergence time of air bubbles underneath the pool surface. The vertical dispersion and horizontal scattering of air bubbles enhances the bubble activity as well as the retention time underneath which is expected to be beneficial for oxygen transfer. The construction and operation of plunging jet equipment is simple, energetically attractive, and it produces aeration and mixing together with the jet impact [1-3]. Due to these inherent advantages, plunging jets have applications in fermentation, chemical, and treatment of wastewater [4].

Many studies exist in literature concerning the prominence of jet variables of plunging jets on the performance of oxygenation of water [5] discussed the impact of no. of jets plunging from a particular height into a water tank and showed improvement in oxygen transfer rate. In some recent work on oxygen transfer by plunging jets, the significance of jet length, jet velocity, water depth, no. of tandem jets and jet power is extensively studied and observed influential in a flowing channel of water [2,6]. Research is conducted concerning the effect of jet angle and jet length for plunging jets in water tank [1,7-9]. Some researchers performed aeration studies on the modified nozzle shapes of plunging jets [10,11]. [4,12] recognised the use of air holes on circular and venturi nozzles plunging into stagnant water and studied the aeration properties [3] and [13] investigated the oxygen transfer properties of hollow jets plunging into water tank. The impact of jet angle with the variation of jet thickness of plunging hollow jets on oxygen transfer properties is studied by [14].

Soft computing gained considerable attention in the area of hydraulics and water resources engineering. New artificial intelligence based methods have been introduced recently and successfully implemented in the simulation and estimation of the actual data. The utility of soft computing approaches in the field of aeration is documented in some studies [15,16] in their studies, success fully implemented ANFIS (adaptive neuro fuzzy inference system), LS-SVM (least square support vector machines) on the data sets of air entrainment rate and aeration efficiency observed from falling overfall jets from triangular weirs with input jet parameters as discharge, drop height, and sharp crest angle. The performance of theses modelling approaches were compared with multiple linear and multiple nonlinear regression based predictive equations. In an another study by [17], the capability of GEP (genetic expression programming) is tested in relating the triangular weir variables with the air entrainment rate as well as aeration efficiency and found to working well [18,19] achieved good predictive accuracy of oxygen mass transfer coefficient of multiple jets plunging into stagnant water pool by applying SVM (support vector machines) and GP (Gaussian process) regression approaches [20] predicted the air entrainment rate of plunging water jets with basic water jet properties using GEP and ANN modelling and compared the results with multiple linear and nonlinear regression equations. In a predictive study, [21] examined the performance of ANN, GEP, and multiple linear and nonlinear regression in predicting the penetration depth of water jets impinging on stagnant water pool [22], in their study based on the results of previous research, evaluated the capability of GEP in the estimation of air entrainment rate, bubble penetration depth and oxygen transfer efficiency of plunging water jets [23] experimented on plunging water jets with extended discharge and effectively estimated the penetration depth by using ANN and non-linear regression techniques [24] acknowledged the applicability of SVM and M5 tree regression techniques in predicting the oxygen mass transfer related to the jet thickness, jet velocity and jet angle of hollow jets plunging into a water tank. The performance of dimensional and non-dimensional data is compared along with predictive correlations yielded from nonlinear regression.

In this work, the performance comparison of various kernel functions (normalized polynomial, polynomial, RBF, and PUK) is checked for the accurate predictions of volumetric oxygen transfer coefficient using Support Vector Regression (SVR) and Gaussian Process Regression (GPR) techniques. Based on past works, the soft computing based data mining techniques are mostly utilized in field of cylindrical nozzle plunging jets for the estimation of aeration properties, while the present study highlights the performance of soft computing (SVR and GPR) based models on the experimental data collected from plunging hollow jet aerators. The oxygen mass transfer properties of plunging hollow jets is correlated with the four basic jet variables: jet thickness, jet velocity, jet length, and water depth in the pool. So the objective of the study is to formulate an empirical relationship considering volumetric oxygen mass transfer coefficient as output variable and compare the prediction capability with various kernel function based regression models.

2. Materials and methods

2.1. Experimental set up and collection of data

The sketch of experimental set up I given in Figure 1. An aeration tank of water holding capacity of (0.87× 0.87× 0.87) m^3 is connected to a water pump of 1HP using a GI pipe line of 2 inches diameter. An electromagnetic flow meter is used at the delivery pipe for the accurate measurement (±0.5% of reading) of discharge values of circulated water. At the delivery end of water discharging pipe, a cone having an angle 60° from the base, made of Perspex material, is fitted to form a water jet falling in hollow formation on the surface of water inside the aeration tank. The up and down movement of cone respective to the pipe delivery end alters the jet flow area and so is the thickness of water jet sheet. The fabrication of hollow jet aerator assembly and procedure adopted to calculate the jet parameters (jet thickness and jet velocity) is given in details in our previous paper [14]. A control valve is used to adjust and regulate the flow through pipe. The distance travelled by the impinging water jet in atmosphere is the termed as jet length (L_j) i.e. the height of delivery end of pipe from the surface of water. In this work, jet length is varied from 0.1 m to 0.4 m by maintaining a constant water depth of 0.6 m and water depth is varied from 0.4 m to 0.7 m by maintaining a constant jet length of 0.1 m .The jet lengths and water depths are varied in four increments of 0.1 m (±1 mm accuracy). The annular gap or jet thickness (5 mm-10 mm) is measured using a scale (±0.5 mm accuracy). The temperature of water is measured using a digital thermometer (±0.1°C accuracy). After filling a particular volume of water in the tank and adjusting a constant jet thickness and jet length, setting of a particular discharge is done. Thereafter, tank water is deoxygenated by mixing sodium sulphite and cobaltous chloride into it [11].

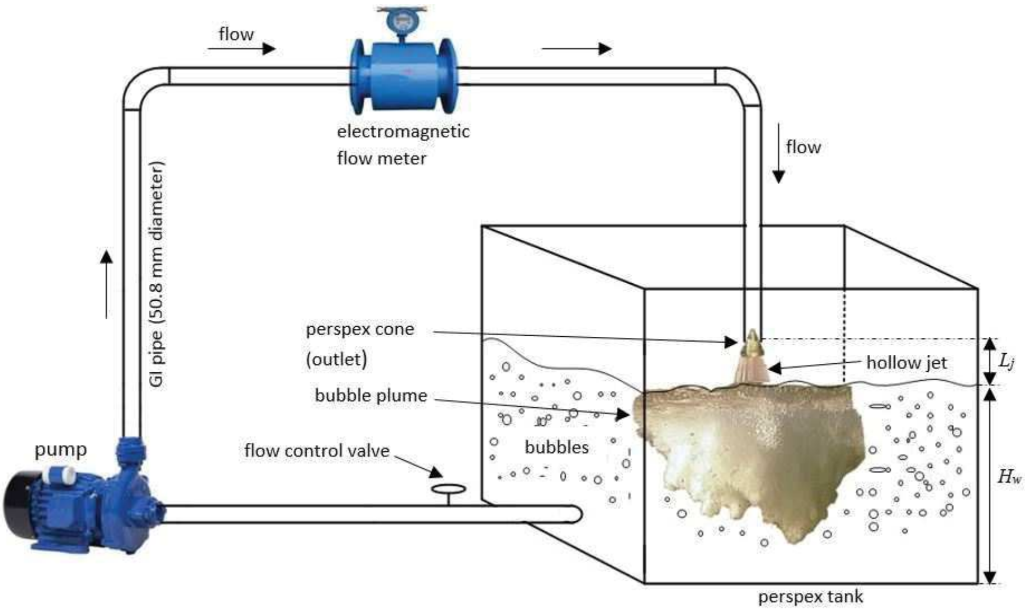

Fig. 1. Experimental set up

Thereafter, a water sample is collected from tank water for the observation of initial dissolved oxygen (C_o). After that, the pump is started and the water is allowed to circulate for a time period of 40s. Following this, a sample is collected for the observation of final dissolved oxygen (C_t). The procedure is repeated for the desired values of jet variables and representative samples of water are collected. Azide modification method [25] was used to measure the oxygen dissolved in initial and final samples of water. The value of volumetric oxygen transfer coefficient of water at test temperature T°C can be obtained from the following relation [5].

$$(K_L a)_T = \frac{1}{t}\ln\left[\frac{C_s - C_0}{C_s - C_t}\right] \quad (1)$$

where $(K_L a)_T$ is termed as volumetric oxygen transfer coefficient (s^{-1}) at test temperature T°C, $C_s$is the saturation oxygen concentration (mg L^{-1}), C_0 is the initial oxygen concentration before aeration (mg L^{-1}), C_t= final oxygen concentration after aeration (mg L^{-1})

For performance comparison, the value volumetric oxygen transfer coefficient at standard temperature of 20°C and pressure of 1 atmosphere can be calculated from the following relation [5].

$$K_L a = (K_L a)_T 1.024^{(20-T)} \quad (2)$$

2.2. Data set

For the execution of soft computing models, a total of 81 experimental observations of $K_L a$ (1/s) is measured by employing plunging hollow jet aerator with the variation of jet thickness (T_j), jet velocity (V_j), jet length (L_j) and water depth (H_w). Two data groups is formed from the total experimental observations for the training (57 readings) and testing (24 readings) of the modelling techniques. The grouping of the data was based on random selection from the total readings. The features of both group of data sets is mentioned in Table 1.

Table 1.
Features of the data set

Parameter	Training					Testing				
	$T_j(m)$	$L_j(m)$	$H_w(m)$	$V_j(m/s)$	$K_L a(1/s)$	$T_j(m)$	L_j (m)	$H_w(m)$	$V_j(m/s)$	$K_L a(1/s)$
Mean	0.008	0.19	0.57	3.14	0.0229	0.007	0.17	0.56	3.29	0.0277
Median	0.010	0.10	0.60	2.95	0.0154	0.005	0.10	0.60	3.26	0.0202
SD	0.003	0.11	0.09	1.41	0.0215	0.003	0.10	0.09	1.49	0.0253
Kurtosis	-2.062	-0.70	-0.15	-0.21	1.4049	-2.156	-0.07	-0.37	-1.08	-0.3021
Skewness	-0.108	0.91	-0.68	0.77	1.3628	0.179	1.15	-0.58	0.16	0.8958
Minimum	0.005	0.10	0.40	1.03	0.0018	0.005	0.10	0.40	1.10	0.0021
Maximum	0.010	0.40	0.70	6.19	0.0863	0.010	0.40	0.70	6.13	0.0837

2.3. Regression approaches

Multiple nonlinear regression (MNLR)

A multi-nonlinear relationship is considered by using $K_L a$ as dependant variable and the parameters, T_j, V_j, L_j, and H_w as explanatory variables. A nonlinear regression model is established using training data expressing the $K_L a$ in terms of jet parameters (Equation 4) based on following functional relationship.

$$K_L a = f(a\, V_j^{x_1} T_j^{x_2} L_j{}^{x_3} H_w{}^{x_4}) \quad (3)$$

where a is the constant, x_1, x_2, x_3, x_4 are the coefficients of the function and can be acquired by minimizing the sum of squares of error in approximation.

$$K_L a = 2.535\, V_j^{2.91} T_j^{1.73} L_j{}^{0.023_3} H_w{}^{-0.63} \quad (4)$$

Support Vector Regression (SVR)

This method is a regression and classification approach which originates from statically learning theory [26]. The SVMs classification techniques depend on the standard of ideal division of classes. In the event that the classes are divisible: this strategy chooses, from amongst the endless number of linear classifiers, the one with minimum generalization error, Along these lines, the chosen hyper plane will be one that leaves the most extreme edge between the two classes, where edge is characterized as the total of the separations of the hyper plane from the nearest purpose of the two classes. It very well may be accomplished by anticipating the first arrangement of factors into a higher dimensional element space and figuring a straight characterization issue in the element space [27,28].

Gaussian Processes Regression (GPR)

The Gaussian (GP) models depend on the presumption that nearby observations ought to pass on data about one another. They indicate an earlier specifically over function space. Therefore, the GP is a natural generalization distribution whose covariance is a matrix and mean is a vector. The Gaussian method is based upon the function whereas distribution relies upon the vector. Due to earlier information pertaining to the function, the validation is not necessary for speculation and Gaussian process regression model can comprehend the prescient distribution related to test input [29].

A Gaussian procedure is characterized as an accumulation of arbitrary factors, any limited number which has a joint multivariate Gaussian distribution. The *n* number of pairs ($x_i \times y_i$) have been made through the $(\chi \times \gamma)$ which indicates the input and output data domain, correspondingly. It is assumed that $y \subseteq R$, accordingly the GP on χ is uttered by mean function $\mu: \chi \rightarrow \mathcal{R}$ and covariance function $\kappa: \chi \times \chi \rightarrow \Re$.

2.4. Kernel functions and user-defined parameters

Implementation of SVR and GPR techniques requires the choice of suitable kernel function which works internally to map the given data to a high dimensional feature space for processing. Four type of kernel functions utilized in this study are:

1. Normalized polynomial kernel (norm poly):

$$K(x_i, x) = K(x_i, x)/\sqrt{K(x_i, x_i), K(x, x)}$$

2. Polynomial kernel (poly):

$$K(x_i, x) = ((x_i, x)+1)^d$$

3. Radial Basis Function kernel (RBF):

$$K(x_i, x) = e^{-\gamma|x_i-x|^2}$$

4. Pearson VII Function kernel (PUK):

$$K(x_i, x) = \left(1/\left[1+(2\sqrt{\|x_i-x\|^2}\sqrt{2^{(1/\omega)}-1}/\sigma)^2\right]^\omega\right)$$

After selecting the kernels, the next step requires selection of kernel specific parameters based on the model performance. The exponent (*d*) in normalized polynomial and polynomial kernel, kernel width (γ) in radial basis function (RBF) kernel, and parameters, σ (controls Pearson width) and ω (tailing factor of the peak) in Pearson VII Function kernel (PUK) need to be established based on the precision in prediction.

Three standard statistical measures, coefficient of determination (R^2), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are selected as principle measures to evaluate the accuracy of predictive modelling methods. Based on modelling performance, a common value of regularization parameter (C=13) in SVR and Gaussian noise (0.1) in GPR is established and chosen in all the four kernel functions to achieve a fair comparison of kernels. After that, the models are tuned for kernel specific parameters (d, γ, σ, ω) in both SVR and GPR techniques. The optimization of user defined parameters is executed by carrying out several runs with these parameters on the training data and investigating the performance of the developed models on testing data. Smaller values of RMSE and MAE deduce closer approximation of the experimental data by the modelling methods. Larger R^2 values agree to a stronger matching of trends in the experimental data by the model predictions. The established values of user defined parameters identified from various runs are provided in Table 2.

Table 2.
Model specific user-defined parameters

Method	Model specific parameters
SVR_norm poly	C =13, d =20.3
SVR_poly	C =13, d =2
SVR_rbf	C =13, γ = 2.6
SVR_puk	C =13, ω = 0.4, σ = 1.1
GPR_norm poly	Gaussian noise =0.1, d =9
GPR_poly	Gaussian noise =0.1, d =2
GPR_rbf	Gaussian noise =0.1, γ = 3.5
GPR_puk	Gaussian noise =0.1, ω = 0.7, σ = 0.7

3. Modelling results

3.1. Results of MNLR model

The obtained relationship using MNLR is plotted with the actual values of their respective data sets of training and testing and represented in Figure 2. The regression equation (4) plotted in Figure reveals nearby fitting of predicted data points of training and testing to the perfect agreement line and suggests good estimates by the developed functional relationship. The RMSE observed with training and testing of the models are 0.00575 (R^2=0.93) and 0.01030 (R^2=0.85), respectively (Tab. 3).

Table 3.
Performance of the models

Model	Training data			Testing data			Rank
	R^2	RMSE	MAE	R^2	RMSE	MAE	
MNLR	0.93	0.00575	0.00312	0.85	0.01030	0.00559	3
SVR_norm poly	0.74	0.01083	0.00334	0.50	0.01885	0.01321	9
SVR_poly	0.92	0.00596	0.00344	0.84	0.01078	0.00611	5
SVR_rbf	0.99	0.00236	0.00103	0.88	0.00993	0.00590	1
SVR_puk	1.00	0.00012	0.00011	0.86	0.01021	0.00622	2
GPR_norm poly	0.75	0.01065	0.00562	0.61	0.01752	0.01019	8
GPR_poly	0.84	0.00880	0.00649	0.79	0.01175	0.00823	7
GPR_rbf	0.99	0.00187	0.00134	0.86	0.01083	0.00677	6
GPR_puk	1.00	0.00057	0.00037	0.87	0.01046	0.00632	4

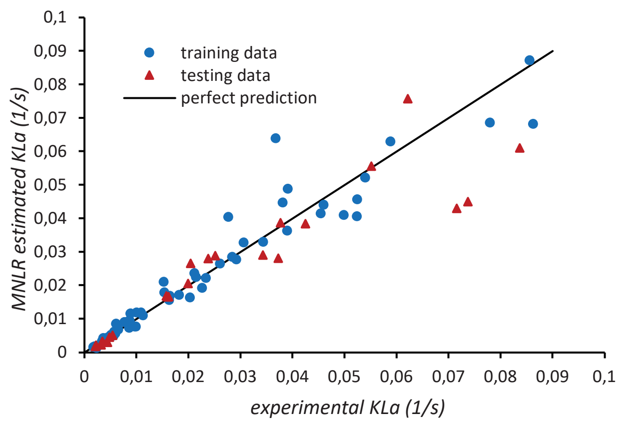

Fig. 2. Scattering of $K_L a$ using MNLR model (Equation 4)

3.2. Results of SVR models

The established user-defined parameters of the various kernel functions for support vector regression (SVR) are presented in Table 2. The results of the developed models are depicted as scattering plots for the training and testing phases of the data and represented with the line of perfect prediction (y=x) in order to apprehend the scattering.

Figure 3 and Figure 4 illustrate the plots between actual and predicted values of volumetric oxygen transfer coefficient ($K_L a$) by SVR models for the training and testing phases, respectively. Observing, Figures 3 and 4 indicate that the prediction of normalized polynomial kernel is worse in training as well as testing and the predicted $K_L a$ data points lie far from the line of perfect prediction as compared to other kernels. On the other hand, polynomial kernel has slightly superior performance than normalized polynomial kernel and the points are relatively near to the perfect prediction line. RBF and PUK kernels have good prediction performance and the estimated values of $K_L a$ in training and testing are closer to the perfect prediction line. All the kernel functions fall in performance during testing of the data and the predicted data is relatively distant from line of perfect prediction as compared to training data. The performance statistics for the SVR model is given in Table 2. From Table 2, the RMSE values of RBF kernel is lowest in testing stage, hence, RBF kernel based SVR performs best in approximating the actual values of $K_L a$ and outperforms normalized poly, polynomial, and PUK kernels using SVR.

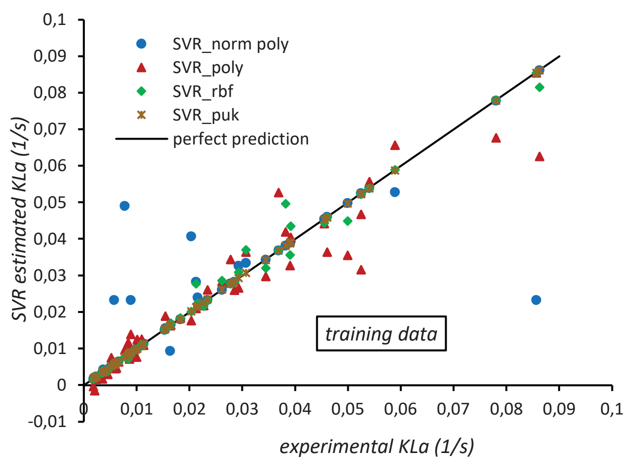

Fig. 3. Scattering of $K_L a$ using SVR models (Training)

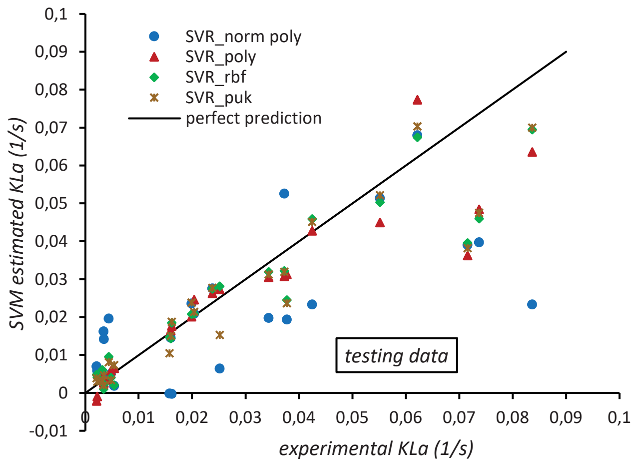

Fig. 4. Scattering of $K_L a$ using SVR models (Testing)

3.3. Results of GPR models

The established user-defined parameters of the various kernel functions for Gaussian Process Regression (GPR) are presented in Table 2. The results of the developed models are depicted as scattering plots for the training and testing phases of the data and represented with the line of perfect prediction (y=x) in order to apprehend the scattering. Figure 5 and Figure 6 illustrate the plots between actual and predicted values of volumetric oxygen transfer coefficient ($K_L a$) by GPR models for the training and testing phases, respectively. Similar to the results observed with the SVR models, normalized polynomial and polynomial kernels using GPR have poor generalization performance and the estimated results are inferior to the RBF and PUK kernel functions. The scattering of the predicted $K_L a$ values is higher (Figs. 5 and 6) as well as the error values in prediction of $K_L a$ with both the kernels as compared to other kernels (Tab. 3). Polynomial kernel has edge over Normalized Polynomial kernel as the RMSE and MAE values are lower in Polynomial GPR. (Tab. 3). In testing, the performance in prediction of $K_L a$ is obtained best with PUK based GPR (RMSE=0.01046, R^2= 0.87) as compared to RBF kernel based GPR (RMSE=0.01083, R^2= 0.86) and hence PUK based GPR performs best in approximating the actual values of $K_L a$ and outperforms normalized poly, polynomial, and RBF kernels using GPR.

3.4. Comparison of kernel functions

Evaluating the results from Table 3, normalized polynomial kernel based regression models using SVR and GPR have worse capabilities in estimation and the majority of predicted data values are beyond ±20% lines. The error values (RMSE and MAE) are highest with normalized polynomial kernel using both regression approaches. The archived error values of polynomial kernel using SVR (RMSE=0.01078, MAE=0.00611) and GPR (RMSE= 0.01175, MAE=0.00823) are lower to the normalized polynomial kernel. The RBF kernel based SVR (RMSE= 0.00993, MAE=0.00590) ranked 1st and the prediction is superior to PUK based SVR (RMSE=0.01021, MAE= 0.00622) which attains the 2nd rank. But in GPR modelling, PUK kernel ranked best and achieves RMSE and MAE values as 0.01046 and 0.00632, respectively as compared to RBF based GPR (RMSE=0.01083, MAE=0.00677).

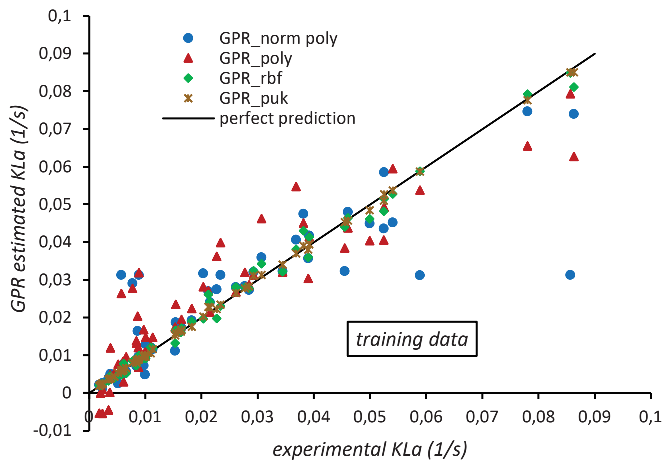

Fig. 5. Scattering of $K_L a$ using GPR models (Training)

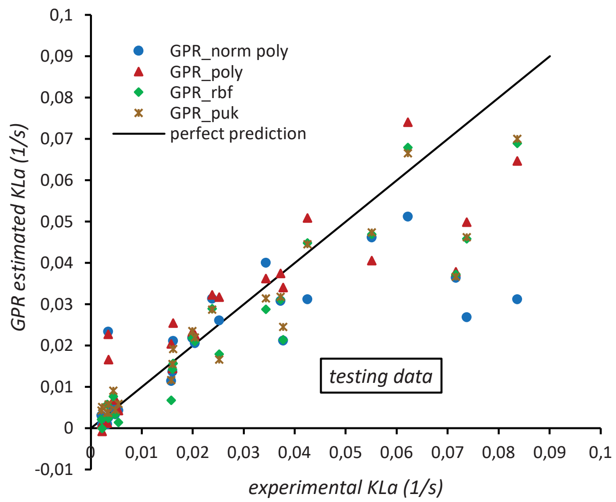

Fig. 6. Scattering of $K_L a$ using GPR models (Testing)

The performance of all the kernel functions improves significantly in SVR as compared to GPR except for the normalized polynomial kernel which shows good accuracy with GPR relative to SVR. So SVR models have better estimation capability than GPR models in this study. Based on testing of the models, the overall ranking of the regression models in prediction of $K_L a$ are as SVR_rbf (1st), SVR_puk (2nd), MNLR (3rd), GPR_puk (4th), SVR_poly (5th), GPR_rbf (6th), GPR_poly (7th), GPR_norm poly (8th), SVR_norm poly (9th). Figure 7 shows the scatter plot of testing data using MNLR, SVR and GPR. Most of the data predicted by RBF and PUK kernels in SVR, and MNLR (Equation 4) scatters almost within or nearer to ±20% lines than the other models.

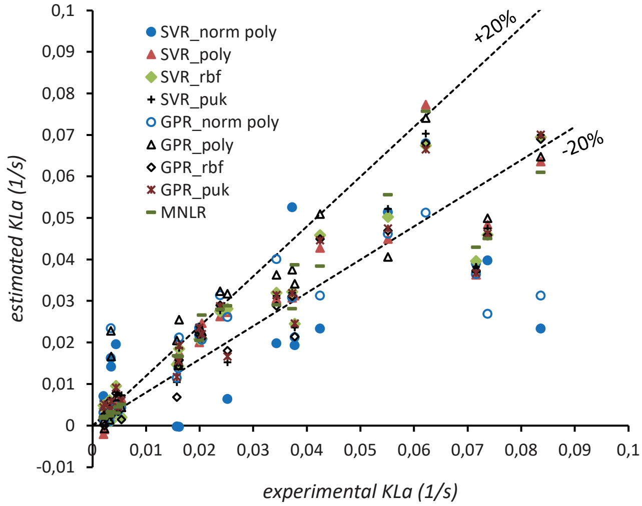

Fig. 7. Scatter plot (±20%) of prediction by SVR, GPR, and MNLR models (Testing)

The performance of SVR_rbf is best relative to other regression models, so the prediction potential of this model is compared with the experimental $K_L a$ as well as nonlinear regression equation (MNLR) predicted $K_L a$. Figure 8 and Figure 9 show the impact of jet velocity on $K_L a$ for jet thicknesses of 5 mm and 10 mm, respectively. As observed with the previous studies [3,14], $K_L a$ significantly increases with increase in jet velocity. With increase in jet velocity, the impact of the jet at the surface of pool water increases which leads to relatively deeper penetration as well as heavy turbulence with high velocity jets. The performance of both the modelling techniques is good in simulating the actual behaviour of oxygen mass transfer to water as the predicted $K_L a$ is close to the experimental data. SVR_rbf is working superior in precisely following the experimental values and the simulated points resides near to the actual experimental values comparative to the MNLR model (Fig. 8 and Fig. 9).

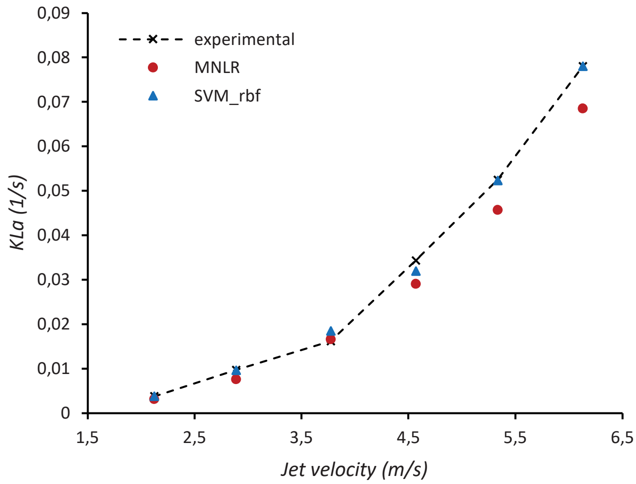

Fig. 8. Variation of experimental $K_L a$ with MNLR and RBF kernel based SVR showing the impact of jet velocity (T_j=0.005 m, L_j=0.2 m, H_w=0.6 m)

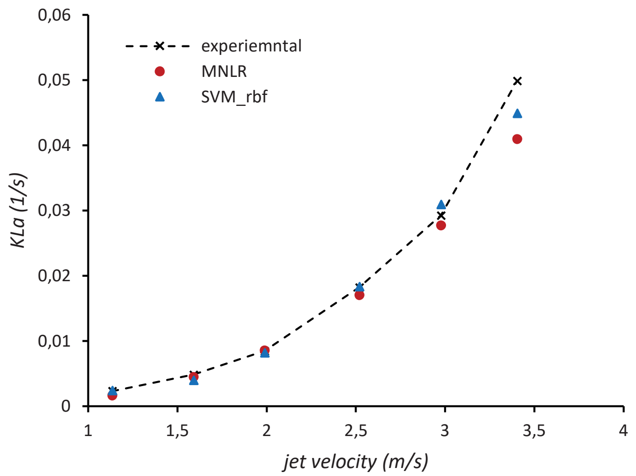

Fig. 9. Variation of experimental $K_L a$ with MNLR and RBF kernel based SVR showing the impact of jet velocity (T_j=0.01 m, L_j=0.2 m, H_w=0.6 m)

3.5. Sensitivity analysis of the jet variables

Sensitivity analysis was used to determine the most significant jet variable in the prediction of volumetric oxygen transfer coefficient of hollow jet aerator. For this, RBF kernel based SVR regression, performing best with the data set was used. Different set of training data was created by removing one input variable at a time and results were reported in terms of coefficient of determination (R^2) and root mean square error (RMSE). The degree of change produced in the RMSE and R^2 values depicts the dominance of the variable in influencing the oxygen mass transfer. Results from Table 4 suggests that the most dominating variable is the jet velocity and have significant role in predicting the volumetric oxygen transfer coefficient in comparison to other input parameters.

Relative change in RMSE and R^2 is considerably higher by removing jet velocity (Tab. 4). Jet thickness is the 2nd most influencing parameter. Jet length is the 3rd most influencing parameter after jet velocity and jet thickness, respectively.

Table 4.
Sensitivity analysis using RBF kernel based SVR regression

Input combination	Input parameter removed	R^2	RMSE (1/s)
$\boldsymbol{T_j, V_j, L_j, H_w}$	-	0.99	0.00236
$\boldsymbol{V_j, L_j, H_w}$	T_j	0.83	0.0088
$\boldsymbol{T_j, L_j, H_w}$	V_j	0.075	0.02147
$\boldsymbol{T_j, V_j, H_w}$	L_j	0.93	0.00559
$\boldsymbol{T_j, V_j, L_j}$	H_w	0.95	0.00479

4. Conclusions

In this study, experimental results of volumetric oxygen transfer coefficient ($K_L a$) observed with various jet variables by creating hollow water jet impinging on the surface of water in the tank is modelled. An empirical nonlinear relationship of $K_L a$ is developed by employing jet velocity, jet thickness, jet length, and water depth as input variables and the performance of this relationship is compared with Support Vector Regression (SVR) and Gaussian Process Regression (GPR). Four kernel functions, namely, normalized polynomial, polynomial, Radial Basis Function (RBF), and Pearson VII Function kernel (PUK) are implemented in SVR as well as GPR techniques.

All the kernel functions (except for normalized polynomial kernel) show improvement in prediction of $K_L a$ when employed with SVR technique as compared to GPR technique, so, SVR has good predictive accuracy relative to GPR.

In SVR technique, both the RBF as well as PUK kernels are working fine but due to slightly less error in estimation, RBF kernel ranked best out of the other kernel functions, while in GPR technique, PUK is performing best relative to other kernel functions.

The prediction capability of multiple nonlinear regression based empirical relationship is greater than all the developed GPR models but inferior to RBF and PUK based SVR models. Hence the relationship is simply useful and reasonable in determining the $K_L a$ close to the experimental values with the current data set.

Based on the outcomes from this study, SVR using RBF kernel is performing well and predicts the experimental $K_L a$ nearly within a scatter of ±20%.

The jet velocity comes out to be the most important parameter followed by jet thickness and jet length, respectively in influencing the $K_L a$ based on the sensitivity results using RBF based SVR model.

References

[1] E. Van de Sande, J.M. Smith, Mass transfer from plunging water jets, The Chemical Engineering Journal 10/2 (1975) 225-233, DOI: https://doi.org/10.1016/0300-9467(75)80041-4.

[2] K. Chipongo, M. Khiadani, Oxygen transfer by multiple vertical plunging jets in tandem, Journal of Environmental Engineering 143/1 (2016) 04016072, DOI: https://doi.org/10.1061/(asce)ee.1943-7870.0001145.

[3] S. Deswal, D.V.S. Verma, Performance evaluation and modeling of a conical plunging jet aerator, International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering 1/11 (2007) 616-620.

[4] M.E. Emiroglu, A. Baylar, Study of the influence of air holes along length of convergent-divergent passage

of a venturi device on aeration, Journal of Hydraulic Research 41/5 (2003) 513-520, DOI: https://doi.org/10.1080/00221680309499996.
[5] S. Deswal, Oxygen transfer by multiple inclined plunging water jets, International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering 2/3 (2008) 57-63.
[6] M.E. Jahromi, M. Khiadani, Experimental study on oxygen transfer capacity of water jets discharging into turbulent cross-flow, Journal of Environmental Engineering 143/6 (2017) 04017007, DOI: https://doi.org/10.1061/(asce)ee.1943-7870.0001194.
[7] A. Ohkawa, D. Kusabiraki, Y. Shiokawa, N. Sakai, M. Fujii, Flow and oxygen transfer in a plunging water jet system using inclined short nozzles and performance characteristics of its system in aerobic treatment of wastewater, Biotechnology and Bioengineering 28/12 (1986) 1845-1856, DOI: https://doi.org/10.1002/bit.260281212.
[8] K. Tojo, K. Miyanami, Oxygen transfer in jet mixers, The Chemical Engineering Journal 24/1 (1982) 89-97, DOI: https://doi.org/10.1016/0300-9467(82)80054-3.
[9] K. Tojo, N. Naruko, K. Miyanami, Oxygen transfer and liquid mixing characteristics of plunging Jet reactors, The Chemical Engineering Journal 25/1 (1982) 107-109, DOI: https://doi.org/10.1016/0300-9467(82)85027-2.
[10] S. Singh, S. Deswal, M. Pal, Performance analysis of plunging jets having different geometries, International Journal of Environmental Sciences 1/6 (2011) 1154-1167.
[11] T. Bagatur, A. Baylar, N. Sekerdag, The effect of nozzle type on air entrainment by plunging water jets, Water Quality Research Journal of Canada 37/3 (2002) 599-612, DOI: https://doi.org/10.2166/wqrj.2002.040.
[12] M.E. Emiroglu, A. Baylar, Role of nozzles with air holes in air entrainment by a water jet, Water Quality Research Journal Canada 38/4 (2003) 785-795, DOI: https://doi.org/10.2166/wqrj.2003.049.
[13] S. Ranjan, Hydraulics of Jet Aerators, Journal of the Institution of Engineers (India): Environmental Engineering Division 88 (2008) 29-32.
[14] M. Kumar, S. Ranjan, N.K. Tiwari, Oxygen transfer study and modeling of plunging hollow jets, Applied Water Science 8/5 (2018) 121, DOI: https://doi.org/10.1007/s13201-018-0740-8.
[15] A. Baylar, D. Hanbay, E. Ozpolat, An expert system for predicting aeration performance of weirs by using ANFIS, Expert Systems with Applications 35/3 (2008) 1214-1222, DOI: https://doi.org/10.1016/j.eswa.2007.08.019.
[16] Baylar, D. Hanbay, M. Batan, Application of least square support vector machines in the prediction of aeration performance of plunging overfall jets from weirs, Expert Systems with Applications 36/4 (2009) 8368-8374, DOI: https://doi.org/10.1016/j.eswa.2008.10.061.
[17] A. Baylar, M. Unsal, F. Ozkan, C. Kayadelen, Estimation of air entrainment and aeration efficiencies of weirs using genetic expression programming, KSCE Journal of Civil Engineering 18/6 (2014) 1632-1640, DOI: https://doi.org/10.1007/s12205-014-1058-1.
[18] S. Deswal, M. Pal, Comparison of polynomial and radial basis kernel functions based SVR and MLR in modeling mass transfer by vertical and inclined multiple plunging jets, International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering 9/9 (2015) 1236-1240.
[19] S. Deswal, Modeling oxygen-transfer by multiple plunging jets using support vector machines and gaussian process regression techniques, International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering 5/1 (2011) 1-6.
[20] T. Bagatur, F. Onen, A predictive model on air entrainment by plunging water jets using GEP and ANN, KSCE Journal of Civil Engineering 18/1 (2014) 304-314, DOI: https://doi.org/10.1007/s12205-013-0210-7.
[21] F. Onen, Prediction of penetration depth in a plunging water jet using soft computing approaches, Neural Computing and Applications 25/1 (2014) 217-227, DOI: https://doi.org/10.1007/s00521-013-1475-y.
[22] T. Bagatur, F. Onen, Prediction of flow and oxygen transfer by a plunging water jets with genetic expression programming (GEP) models, Arabian Journal for Science and Engineering 39/6 (2014) 4421-4432, DOI: https://doi.org/10.1007/s13369-014-1092-9.
[23] M. Kramer, S. Wieprecht, K. Terheiden, Penetration depth of plunging liquid jets – A data driven modelling approach, Experimental Thermal and Fluid Science 76 (2016) 109-117, DOI: https://doi.org/10.1016/j.expthermflusci.2016.03.007.
[24] M. Kumar, N.K. Tiwari, S. Ranjan, Prediction of oxygen mass transfer of plunging hollow jets using regression models, ISH Journal of Hydraulic

Engineering (2018) 1-8, DOI: https://doi.org/10.1080/09715010.2018.1435311.
[25] APHA, AWWA, WEF, Standard methods for the examination of water and waste water, APHA Washington DC, 2005, 4.38-4.140.
[26] Cortes, V. Vapnik, Support-vector networks, Machine Learning 20/3 (1995) 273-297, DOI: https://doi.org/10.1007/BF00994018.
[27] A.J. Smola, Regression estimation with support vector learning machines, Master's thesis, Technische Universität München, 1996.
[28] V.N. Vapnik, Statistical learning theory, Vol. 3, Wiley, New York, 1998.
[29] C.E. Rasmussen, Gaussian processes for machine learning, Cambridge, 2006.