

Prediction of discharge coefficient of side weir using adaptive neuro-fuzzy inference system

Abbas Parsaie¹ · Amir Hamzeh Haghiabi¹

Received: 25 June 2015 / Accepted: 29 February 2016 / Published online: 11 April 2016
© Springer International Publishing Switzerland 2016

Abstract Predicting the discharge coefficient of the hydraulic structures is one of the main subjects related to the hydro-system management. Weirs are the common hydraulic structure widely used in the water engineering projects. Side weir is the common type of hydraulic structure used in water engineering projects. Principal component analysis of the affective parameters on the side weir discharge coefficient leads to develop optimal structure for the empirical formulas and artificial intelligent models. In this paper, the principal component analysis (PCA) technique was used to define the most important affective parameters on the discharge coefficient of side weir ($C_{d_{sw}}$). The result of the PCA showed that the Froude number and ratio of the weir height to the upstream flow depth (P/h_1) are the most influential parameters affecting the $C_{d_{sw}}$. Developing the adaptive neuro-fuzzy inference system (ANFIS) based on the PCA result showed that the optimal ANFIS structure is related to consider the five and four Gaussian membership function for the Froude number and P/h_1 parameters, respectively. The correlation coefficient of the ANFIS model during the training and testing stage was found to be 0.96 and 0.86 correspondingly.

Keywords Principal component analysis · Optimal model structure · ANFIS · Side weir · Discharge coefficient

Introduction

Study on the hydraulic phenomena is based on the definition affective parameters. To this purpose, influence parameters such as fluid properties, hydraulic and geometric variables are collected together and using the dimensional analysis such as Buckingham π theorem the dimensionless parameters are derived (Dehdar-behbahani and Parsaie 2016; Chen 2015). Usually using the design of experiment (DOE) techniques, the influence of the independent parameters on the dependent parameter is defined. In this approach for defining the impact of the independent parameter on the dependent parameter during the experiments, other parameter remains constant (Antony 2014). Today by advancing the data mining approaches such as neural network models in almost all areas of water engineering fields especially in the water engineering studies (Azamathulla et al. 2016; Parsaie 2016a, b), researchers have attempted to use these techniques for predicting and modeling the hydraulic or hydrologic phenomena (Tayfur 2014). As clear from the name of the data mining approaches, developing these models are based on the data set; therefore, investigators for developing the types of the data mining models have tried to collect the related data set from the various reliable sources such as peer-reviewed article and handbooks and books, etc. (Araghinejad 2013). During the data collection process defining the most affective independent parameters sometimes becomes difficult therefore to this purpose several mathematical approached such as principal component analysis as multivariable analysis techniques, etc., have been proposed. Using these approaches leads to define the most affective parameter on the desired phenomenon (Remesan and Mathew 2014). Since the focus of this research is on the side weir discharge coefficient, so the most follow

✉ Abbas Parsaie
abbas_parsaie@yahoo.com

Amir Hamzeh Haghiabi
haghiabi@yahoo.com

¹ Department of Water Engineering, Lorestan University, Khorramabad, Iran

illustration is on this subject. Side weir is a type of weir which is set up on the side wall channel and most of the time installed parallel to the flow direction (Haghiabi 2012; Heidarpour et al. 2008). Side weir is used for removing the excess flow from the hydro-systems such as irrigation and drainage network, sewage, etc. (Bagheri et al. 2014; Haddadi and Rahimpour 2012; Parsaie et al. 2015a). Several studies such as experimental, analytical and artificial intelligent techniques have been used for calculating and predicting the side weir discharge coefficient (Vatankhah 2013a, b; Parsaie and Haghiabi 2014). In the experimental studies researchers have tried to improve the performance of the side weir, to this purpose various shapes have been proposed for the crest of side weirs which most of these categorized as nonlinear crest. In the field of numerical modeling using the computational fluid dynamic and artificial intelligent techniques can be stated (Aydin and Emiroglu 2013). In the computational hydraulic field, the water surface profile and flow properties were studied (Parsaie and Haghiabi 2015a, b). Side weir discharge coefficient was predicted and modeled by most types of neural network techniques such as multilayer perceptron (MLP) neural network, adaptive neuro-fuzzy inference system (ANFIS), and group method of data handling (GMDH) (Ebtehaj et al. 2015a; Emiroglu et al. 2011b; Kisi et al. 2012). Based on the reports the accuracy of these models are much more than the empirical formulas. Using the AI model together with numerical methods leads to increase the accuracy of the numerical simulation (Parsaie et al. 2015b; Parsaie and Haghiabi 2015a, c). Although the AI techniques have the ability to model complex systems, optimal structure of these models is an important subject which is discussed in the model development process. Several mathematical approaches such as gamma test, Monte Carlo simulation and principal component analysis

such as multivariable analysis have been proposed to this purpose (Martinez et al. 2010). In this paper using the PCA as the most important parameter on the side weir, discharge coefficient is derived and in the following by considering the PCA results, an evaluation is conducted on the performance of the empirical formulas which have been proposed for $C_{d_{sw}}$. At the end, the ANFIS model is developed based on the PCA results.

Method and materials

Discharge coefficient of side weir is proportional to the hydraulic and geometric parameters. Figure 1 shows a schematic shape of the side weir and the most important parameters in the subcritical flow condition.

As seen in Fig. 1, the most important parameters are the flow velocity (V_1), side weir length (L), diversion angle of the flow (ψ), weir height (P) and the longitudinal slope of the channel (s_0). Equation (1) collected the mentioned parameters.

$$C_{d_{sw}} = f(v_1, L, b, h_1, P, \psi, s_0) \quad (1)$$

Using the Buckingham theorem leads to derive dimensionless parameters which are basic parameters for developing the empirical formulas and AI models. The result of the Buckingham theory is given in the Eq. (2) (Emiroglu et al. 2011a).

$$C_{d_{sw}} = f_2 \left(Fr_1, \frac{L}{b}, \frac{L}{h_1}, \frac{P}{h_1} \right) \quad (2)$$

For calculating the $C_{d_{sw}}$ some of the most famous empirical formulas were collected and given in Table 1.

As mentioned in the past section, developing the AI models is based on the data set; therefore, about 477 data set-

Fig. 1 Sketch of side weir at subcritical flow condition

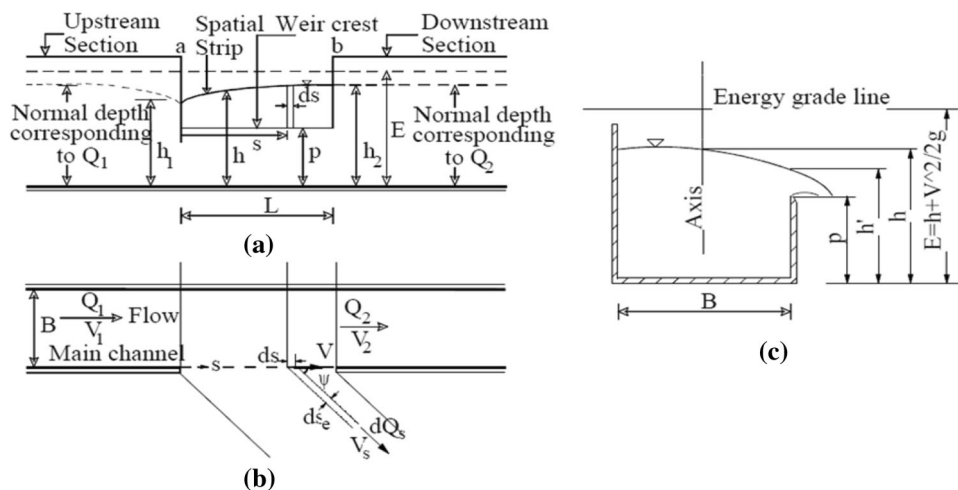


Table 1 Some empirical formulas to calculate the side weir discharge coefficient

Row	Author	Equation
1	Nandesamoorthy and Thomson (1972)	$C_d = 0.432 \left(\frac{2 - Fr_1^2}{1 - 2Fr_1^2} \right)^{0.5}$
2	Subramanya and Awasthy (1972)	$C_d = 0.432 \left(\frac{1 - Fr_1^2}{2 + Fr_1^2} \right)^{0.5}$
3	Yu-Tech (1972)	$C_d = 0.623 - 0.222Fr_1$
4	Ranga Raju et al. (1979)	$C_d = 0.81 - 0.6Fr_1$
5	Hager (1987)	$C_d = 0.485 \left(\frac{2 - Fr_1^2}{2 + 3Fr_1^2} \right)^{0.5}$
6	Cheong (1991)	$C_d = 0.45 - 0.221Fr_1$
7	Singh et al. (1994)	$C_d = 0.33 - 0.18Fr_1 + 0.49 \left(\frac{p}{h_1} \right)$
8	Jalili and Borghei (1996)	$C_d = 0.71 - 0.41Fr_1 + 0.22 \left(\frac{p}{h_1} \right)$
9	Borghei et al. (1999)	$C_d = 0.7 - 0.48Fr_1 + 0.3 \left(\frac{p}{h_1} \right) + 0.06 \left(\frac{L}{h_1} \right)$
10	Emiroglu et al. (2011a, b)	$C_d = \left[0.836 + \left(-0.035 + 0.39 \left(\frac{p}{h_1} \right)^{12.69} + 0.158 \left(\frac{L}{b} \right)^{0.59} + 0.049 \left(\frac{L}{h_1} \right)^{0.42} + 0.244Fr_1^{2.125} \right)^{3.018} \right]^{5.36}$

Table 2 Range of collected data related to the side weir discharge coefficient

Data range	Fr ₁	P/h ₁	L/b	L/h ₁	Cd _{sw}
Min	0.09	0.03	0.21	0.19	0.09
Max	0.84	2.28	3.00	10.71	1.75
AVG	0.43	0.76	1.13	3.87	0.50
STDEV	0.18	0.43	0.85	3.06	0.17

related parameters of Eq. (2) were collected from reliable peer-reviewed journals and their ranges are given in Table 2. To calculate the discharge coefficient of side weir using the empirical formulas with regard to Table 2, the values of related parameters used in each of the empirical formulas are derived and then taken into the empirical formula and then the discharge coefficient will be calculated.

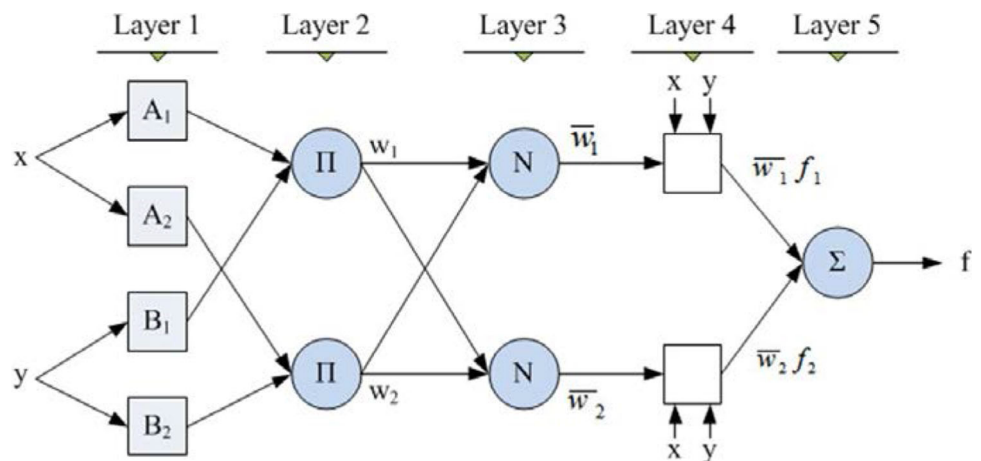
Principal component analysis (PCA)

The principal component analysis (PCA) is an advanced category in the factor analysis approaches and usually used for data reduction in the field of engineering. The main application of the PCA is in the compression and classification of data; the other main use of this approach is to reduce the dimensionality of a data set (sample) by finding a new set of variables, smaller than the original set of variables that nonetheless retains most of the sample’s information (Camacho et al. 2015, Martinez et al. 2010).

Adaptive neuro-fuzzy inference systems (ANFIS)

Adaptive neuro-fuzzy inference systems (ANFIS) is a powerful tool for modeling of complex system based on

Fig. 2 ANFIS model structure



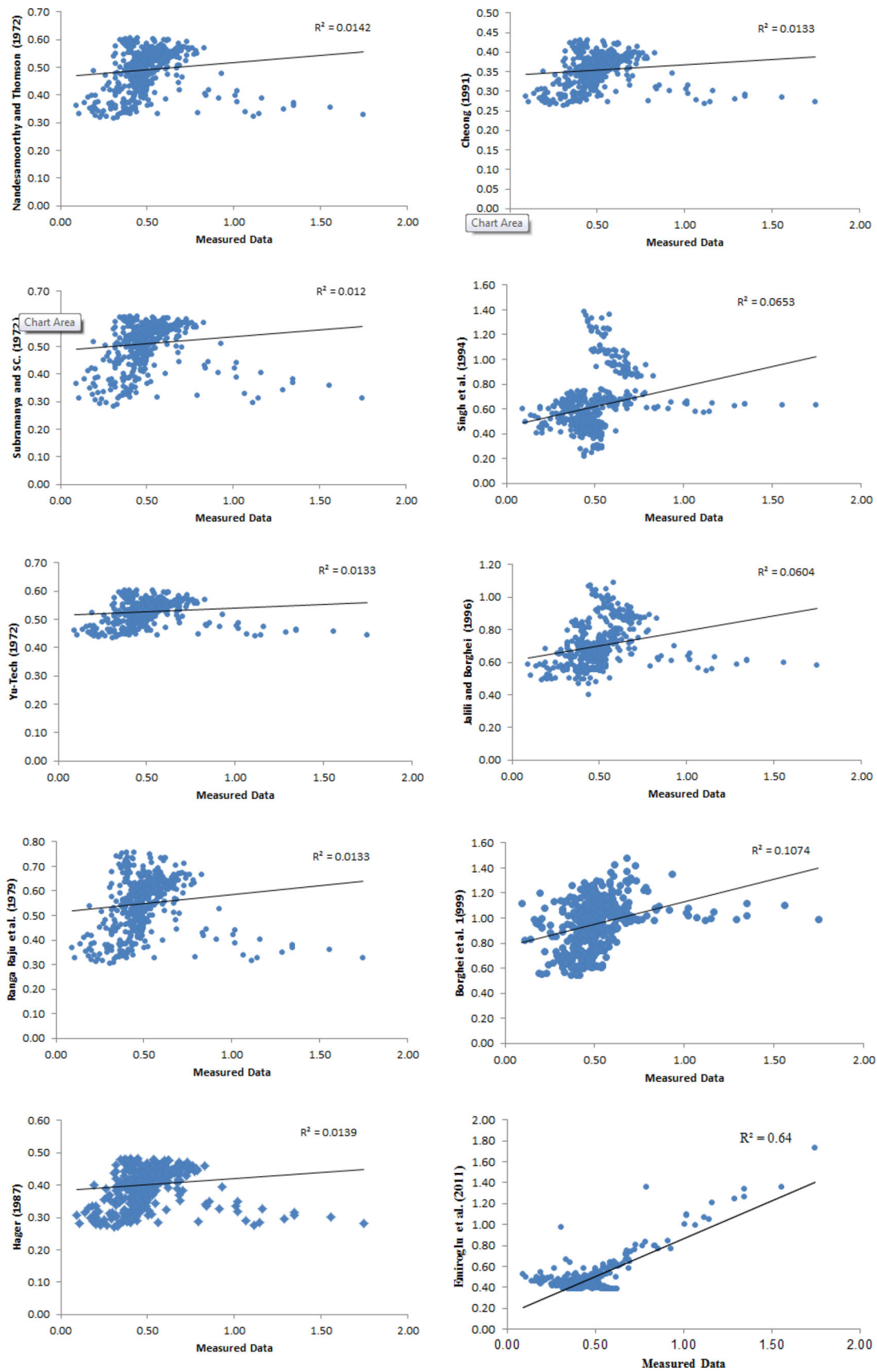


Fig. 3 Performance of the empirical formulas to calculate the Cd_{sw}

input and output data. ANFIS are realized by an appropriate combination of neural and fuzzy systems. This combination enables to use both the numeric power of intelligent systems. In fuzzy systems, different fuzzification and defuzzification strategies with different rules were considered for input parameters. For determining the effect of fuzzy logic on the input data, three stages should be considered. One-selecting the membership function for each input variable. In this stage, a Gaussian function for each of input variable maybe considered. Figure 2 shows a fuzzy reasoning process. For simplicity, illustrating a fuzzy system with two input variables and one output was considered. Suppose that the rule base containing two fuzzy if–then rules:

Rule1 : if x is A_1 and y is B_1 then $f_1 = p_1x + q_1y + r_1$

Rule2 : if x is A_2 and y is B_2 then $f_2 = p_2x + q_2y + r_2$,

where A_1 ; A_2 and B_1 ; B_2 are the MFs for inputs x and y ; respectively; p_1 ; q_1 ; r_1 and p_2 ; q_2 ; r_2 are the parameters of the output function. ANFIS architecture is presented in Fig. 2 as follows: in the first layer, all the input variables gave the grade membership with membership function; in layer 2, all the membership grades will be multiplies together; in layer 3, all the grades of member will be normalized; in layer 4, the contribution of all the rules will be computed; and in the last layer, output variable will be computed as weighted average of grade membership (Ri-ahi-Madvar et al. 2009).

Table 3 The performance of empirical formulas

Author	R^2	RSME
Nandesamoorthy and Thomson (1972)	0.01	0.00
Subramanya and Awasthy (1972)	0.01	0.00
Yu-Tech (1972)	0.01	0.00
Ranga Raju et al. (1979)	0.01	0.00
Hager (1987)	0.01	0.01
Cheong (1991)	0.01	0.01
Singh et al. (1994)	0.07	0.01
Jalili and Borghei (1996)	0.06	0.01
Borghei et al. (1999)	0.11	0.02
Emiroglu et al. (2011a, b)	0.64	0.03

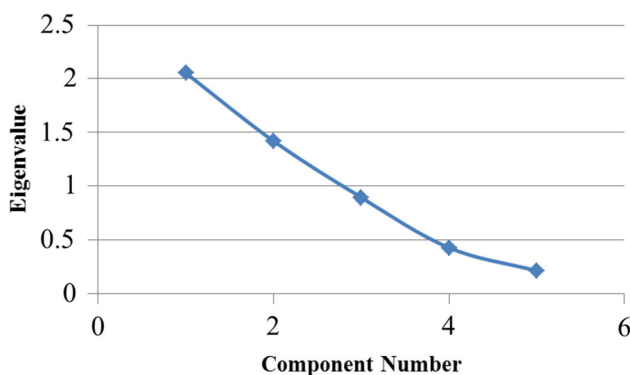


Fig. 4 The screw graph resulted from the PCA technique

Table 4 The table of component variance resulted from PCA technique

Component	Initial eigenvalue			Extraction sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	2.053	41.059	41.059	2.053	41.059	41.059
2	1.418	28.350	69.409	1.418	28.350	69.409
3	0.893	17.857	87.267			
4	0.426	8.518	95.785			
5	0.211	4.215	100.000			

Results and discussion

Empirical formula results

The performance of empirical formulas was evaluated by conducting a comparison with measured data. The results of the each empirical formula were plotted versus the measured data and are shown in Fig. 3. The standard error indices such as correlation coefficient (R^2) and root mean square of error (RMSE) were calculated for assessing the performance of the empirical formulas. The results of the error indices are given in Table 3. As clear from Fig. 3 and Table 3, the Emiroglu formula with correlation coefficient 0.64 and root mean square error 0.03 is accurate among the empirical approaches.

PCA result

To define the most affective parameters on the Cd_{sw} , the PCA technique was carried out on the collected data set, the ranges of which are given in Table 2. The results of the PCA are given in Fig. 4 and Table 4. As shown in Fig. 4, the Froude number and ratio of the weir height to the flow depth (P/h_1) are the most important parameters for predicting the Cd_{sw} . By paying attention to the PCA results and results of the empirical formulas obtained in Fig. 2, it

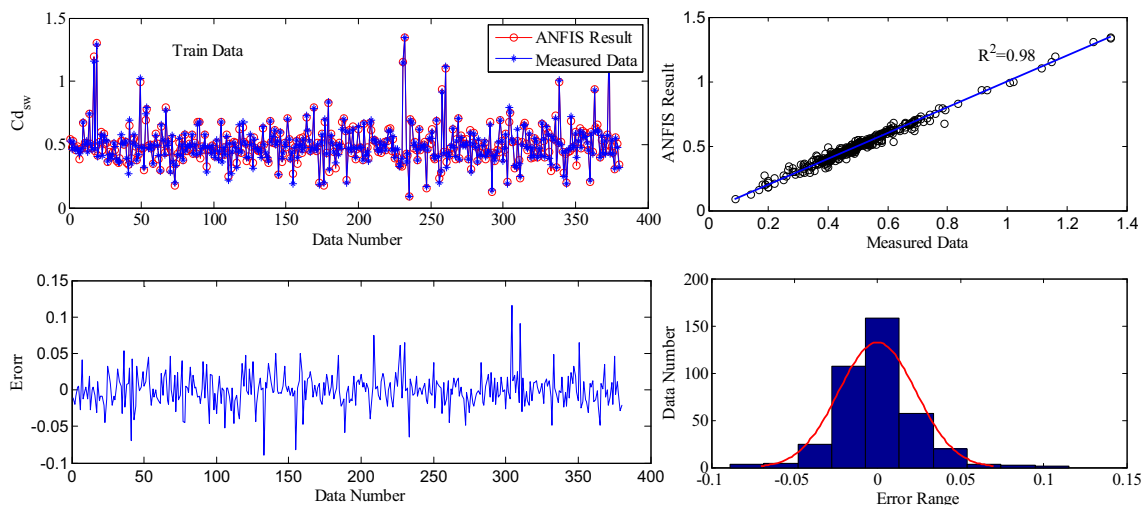


Fig. 5 The performance of ANFIS model during the training stage

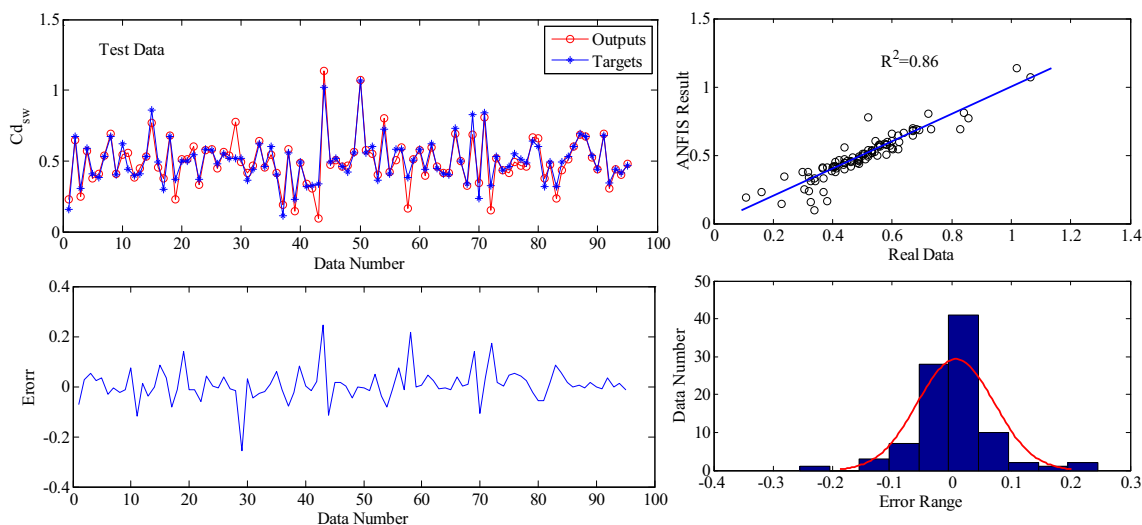


Fig. 6 The performance of ANFIS model during the testing stage

could be found that the empirical formulas which considered more weight for the parameters such as Fr_1 and especially P/h_1 , such as Emiroglu formula are more accurate when compared to other empirical formulas.

ANFIS models development

Developing ANFIS models similar to other neural network models is based on the data set. To this purpose, the data set ranges of these given in Table 2 were used and divided into two groups as training and testing. Choosing training and testing data sets was based on the randomized approach. Designing the structure of the ANFIS included the definition number of the membership function, hidden layer(s), activation function and learning algorithm. Choosing the

number of the hidden layers and other model structures almost is based on trial and error, but the experience of the designer and recommendations of the other investigators who conducted similar studies are useful. Another approach for developing an optimal structure for ANFIS model is using the mathematical approach such as PCA. Result of the PCA shows that the Fr_1 and P/h_1 are the most important parameters in the Cd_{sw} prediction. Developing the ANFIS model structure based on the PCA requires considering more number of neurons to the Fr_1 and P/h_1 . The ANFIS model has a main advantage when compared to other ANN models such as multilayer perceptron neural network (MLP) model in the utility of structure designing stage. This utility is related to specifying the number of the membership function to the input variables based on these influences on the output

Table 5 The structure and summary of the ANFIS structures

Parameter	NoN	MF	And method	OR method	Defuzz method	Agg method	Type	R^2		RMSE	
Fr_1	5	gaussmf	Prod	Max	wtaver	Max	Sugeno	Train	Test	Train	Test
P/h_1	4	gaussmf	Prod	Max	wtaver	Max					
L/b	1	gaussmf	Prod	Max	wtaver	Max		0.96	0.86	0.060	0.085
L/h_1	2	gaussmf	Prod	Max	wtaver	Max					

parameter. PCA results can be applied for developing the structure of ANFIS model, so the utility of ANFIS model leads to develop a model that is more optimal and has more reliability because each parameter which is more affective on the outputs can get more membership function. The results of the ANFIS model to predict the Cd_{sw} are shown in Figs. 5 and 6. As mentioned in the past, the data set is randomly divided in two groups as training and testing data set. Training data set is about 80 % of the total collected data and the remaining data set (20 %) was used for testing. The structural of the ANFIS which has best performance is given in Table 5; as shown in Table 5, the Gaussian function (gaussmf) was considered for the membership function and weight average (wtaver) approach was considered for defuzzification method. As shown in Figs. 5 and 6, the histogram and distribution of the errors are also plotted for assessing the performance of the ANFIS model in stage of the training and testing. As clear from Table 5, the Fr and P/h_1 have more neurons when compared to the other parameters. Overall, as shown in Figs. 5 and 6, the ANFIS model's ability is suitable for predicting the values of the Cd_{sw} in the training and testing stages and also this model has suitable performance to predict the maximum values of the Cd_{sw} . The results of this study uphold the results of Ebtehaj et al. (2015a) and Ebtehaj et al. (2015b). Ebtehaj et al. (2015a) stated that for prediction of discharge coefficient of side weir using the GMDH, Fr and P/h_1 are the most important parameters also reviewing the studies which were conducted by Ebtehaj et al. (2015b) and Emiroglu et al. (2011a, b) showed that they considered more weight for the both parameters during the model development.

Conclusion

In predicting the discharge coefficient of the weirs spatially, side weirs play a key role in the hydro-system management. Recently by advancing the neural network techniques in the water engineering studies, modeling of hydraulic phenomena is carried out more accurately. Although the ANN models have high ability for predicting the discharge coefficients of the hydraulic structures, especially side weirs, optimal designing of these structure is an important factor which leads to increase the reliability

of the ANN model. Using the mathematical techniques such as principal component analysis (PCA) helps to define the most important parameters which have influence on the desired phenomena. In this paper, using the PCA it was found that the Froude number and ratio of the weir height to the upstream flow depth (P/h_1) are the most influential parameters on the Cd_{sw} . Therefore, during the ANFIS model development more number of membership function was considered to these parameters. Overall, using the PCA results leads to preparing an optimal structure the ANFIS model.

References

- Antony J (2014) Design of experiments for engineers and scientists. Elsevier, London
- Araghinejad S (2013) Data-driven modeling: using MATLAB® in water resources and environmental engineering. Springer, Netherlands
- Aydin MC, Emiroglu ME (2013) Determination of capacity of labyrinth side weir by CFD. Flow Meas Instrum 29:1–8
- Azamathulla HM, Haghiabi AH, Parsaie A (2016) Prediction of side weir discharge coefficient by support vector machine technique. Water Sci Technol Water Supply. doi:10.2166/ws.2016.014
- Bagheri S, Kabiri-Samani AR, Heidarpour M (2014) Discharge coefficient of rectangular sharp-crested side weirs, part I: traditional weir equation. Flow Meas Instrum 35:109–115
- Borghei S, Jalili M, Ghodsian M (1999) Discharge coefficient for sharpcrested side weir in subcritical flow. J Hydraul Eng 125:1051–1056. doi:10.1061/(ASCE)0733-9429(1999)125:10(1051)
- Camacho J, Perez-Villegas A, Rodríguez-Gómez RA, Jimenez-Mañas E (2015) Multivariate exploratory data analysis (MEDA) toolbox for Matlab. Chemometr Intell Lab Syst 143:49–57
- Chen SH (2015) Hydraulic structures. Springer, Berlin, Heidelberg
- Cheong H (1991) Discharge coefficient of lateral diversion from trapezoidal channel. J Irrig Drain Eng 117:461–475. doi:10.1061/(ASCE)0733-9437(1991)117:4(461)
- Dehdar-behbahani S, Parsaie A (2016) Numerical modeling of flow pattern in dam spillway's guide wall. Case study: Balaroud dam, Iran. Alex Eng J 55(1):467–473
- Ebtehaj I, Bonakdari H, Zaji AH, Azimi H, Khoshbin F (2015a) GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs. Eng Sci Technol Int J 18(4):746–757
- Ebtehaj I, Bonakdari H, Zaji AH, Azimi H, Sharifi A (2015b) Gene expression programming to predict the discharge coefficient in rectangular side weirs. Appl Soft Comput 35:618–628

- Emiroglu ME, Agaccioglu H, Kaya N (2011a) Discharging capacity of rectangular side weirs in straight open channels. *Flow Meas Instrum* 22:319–330
- Emiroglu ME, Bilhan O, Kisi O (2011b) Neural networks for estimation of discharge capacity of triangular labyrinth side-weir located on a straight channel. *Expert Syst Appl* 38:867–874
- Haddadi H, Rahimpour M (2012) A discharge coefficient for a trapezoidal broad-crested side weir in subcritical flow. *Flow Meas Instrum* 26:63–67
- Hager W (1987) Lateral outflow over side weirs. *J Hydraul Eng* 113(4):491–504
- Haghiabi AH (2012) Hydraulic characteristics of circular crested weir based on Dressler theory. *Biosyst Eng* 112(4):328–334
- Heidarpour M, Habili JM, Haghiabi AH (2008) Application of potential flow to circular-crested weir. *J Hydraul Res* 46(5):699–702
- Jalili M, Borghei S (1996) Discussion: discharge coefficient of rectangular side weirs. *J Irrig Drain Eng* 122:132
- Kisi O, Emin Emiroglu M, Bilhan O, Guven A (2012) Prediction of lateral outflow over triangular labyrinth side weirs under subcritical conditions using soft computing approaches. *Expert Syst Appl* 39:3454–3460
- Martinez WL, Martinez AR, Solka J (2010) *Exploratory data analysis with MATLAB*, 2nd edn. CRC Press, Boca Raton
- Nandesamoorthy T, Thomson A (1972) Discussion of spatially varied flow over side weir. *J Hydrol Eng* 98(12):2234–2235
- Parsaie A (2016a) Analyzing the distribution of momentum and energy coefficients in compound open channel. *Model Earth Syst Environ* 2(1):1–5
- Parsaie A (2016b) Predictive modeling the side weir discharge coefficient using neural network. *Model Earth Syst Environ* 2(2):1–11
- Parsaie A, Haghiabi A (2014) Predicting the side weir discharge coefficient using the optimized neural network by genetic algorithm. *Sci J Pure Appl Sci* 3:103–112
- Parsaie A, Haghiabi AH (2015a) Computational modeling of pollution transmission in rivers. *Appl Water Sci*. doi:[10.1007/s13201-015-0319-6](https://doi.org/10.1007/s13201-015-0319-6)
- Parsaie A, Haghiabi A (2015b) The effect of predicting discharge coefficient by neural network on increasing the numerical modeling accuracy of flow over side weir. *Water Resour Manag* 29(4):973–985
- Parsaie A, Haghiabi A (2015c) Predicting the longitudinal dispersion coefficient by radial basis function neural network. *Model Earth Syst Environ* 1:1–8
- Parsaie A, Haghiabi A, Moradinejad A (2015a) CFD modeling of flow pattern in spillway's approach channel. *Sustain Water Resour Manag* 1:245–251
- Parsaie A, Yonesi H, Najafian S (2015b) Predictive modeling of discharge in compound open channel by support vector machine technique. *Model Earth Syst Environ* 1:1–6
- Ranga Raju KG, Prasad B, Grupta SK (1979) Side weir in rectangular channels. *J Hydraul Div* 105:547–554
- Remesan R, Mathew J (2014) *Hydrological data driven modelling: a case study approach*. Springer International Publishing, New York
- Riahi-Madvar H, Ayyoubzadeh SA, Khadangi E, Ebadzadeh MM (2009) An expert system for predicting longitudinal dispersion coefficient in natural streams by using ANFIS. *Expert Syst Appl* 36:8589–8596
- Singh R, Manivannan D, Satyanarayana T (1994) Discharge coefficient of rectangular side weirs. *J Irrig Drain Eng* 120:814–819. doi:[10.1061/\(ASCE\)0733-9437\(1994\)120:4\(814\)](https://doi.org/10.1061/(ASCE)0733-9437(1994)120:4(814))
- Subramanya K, Awasthy SC (1972) Spatially varied flow over sideweirs. *J Hydraul Div* 98:1–10
- Tayfur G (2014) *Soft computing in water resources engineering: artificial neural networks, fuzzy logic and genetic algorithms*. WIT Press, Southampton
- Vatankhah A (2013a) Water Surface profiles along a rectangular side weir in a U-shaped channel (analytical findings). *J Hydrol Eng* 18:595–602
- Vatankhah AR (2013b) Water surface profile along a side weir in a parabolic channel. *Flow Meas Instrum* 32:90–95
- Yu-Tech L (1972) Discussion of spatially varied flow over side weir. *J Hydraul Eng* 98(11):2046–2048