

Research Paper

Investigation and Comparison Performance Three Data Mining Techniques ANN, GEP and RF in Estimating Flow Discharge Coefficients of Trapezoidal Broad-Crested Weirs

Bahram Nourani¹,Hadi Arvanaghi^{2*},Farzin Salmasi³

¹ Ph. D Candidate of Hydraulic Structures, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran, Email: nourani.bahram@tabrizu.ac.ir

² Associated Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran, Email: arvanaghi.hadi@gmail.com

³ Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran, Email: salmasi@tabrizu.ac.ir



10.22125/TWE.2023.173240

Received:

March 24, 2021

Accepted:

November 3, 2021

Available online:

June 25, 2023**Keywords:****Broad-crested weirs,
Intelligence models,
Discharge coefficient,
Finite volume method,
ANSYS FLUENT.****Abstract**

Accurate determination of the discharge coefficients plays an important role in estimating the discharge of the weirs. As a result, it is significant to estimate the discharge coefficients correctly. The purpose of this study is simulation and estimation the discharge coefficients of the broad-crested weirs. Hence, numerical simulation of hydraulic characteristics of these weirs were performed by ANSYS FLUENT software and results were obtained. Then, three intelligent models of ANN (MLP), GEP and RF were used to determine the discharge coefficients and the results of these models were compared. Assessment of the results were carried out using the statistical indices as R^2 , RMSE, KGE, and graphical diagrams. The amounts of R^2 , RMSE and KGE for ANN (MLP) were 0.906, 0.016, 0.927, for GEP 0.790, 0.025, and 0.780, and for RF 0.898, 0.013, and 0.841 were obtained, respectively. Due to the statistical criteria results and the range of RE% ($\mp 5\%$), the ANN (MLP) was selected as the superior model. Violin plot illustrates a great and close agreement between the probability distribution of estimated data with ANN (MLP) and the obtained data from the numerical simulation using the finite volume method. Furthermore, the probability density function plot of residuals in ANN (MLP) model was closer to the normal distribution function, compared with the other models in this study.

1. Introduction

Weirs are simple hydraulic structures that are used to control, regulate, and measure flow discharge in irrigation canals.

* **Corresponding Author:** Hadi Arvanaghi

Address: Department of Water Engineering,
Faculty of Agriculture, University of Tabriz, Tabriz,
Iran

Email: arvanaghi.hadi@gmail.com
Tel: 09144195085

There have been numerous studies on the hydraulic performance of broad-crested weirs, including study by Bos (1985), Hager and Schwalt (1994), Hager and Fritz (1998), Saker and Rhodes (2003) and etc. Sargison and Percy (2009) conducted experiments on trapezoidal broad-crested weirs with different upstream and downstream slopes. The results showed that the upstream face slope has the greatest effect on increasing the flow discharge, and reducing the upstream slope leads to an increase in the discharge coefficient (C_d). Using the experimental data of previous researchers, equations were developed for the discharge coefficient (C_d) of broad-crested weirs with sloping crests and weirs with sloping faces upstream, downstream, or both under free flow conditions. Azimi and Rajaratnam (2009) integrated their previous studies with available laboratory data to extract equations for the C_d . The results showed that for long broad-crested weirs, a positive slope for the crest increases the flow discharge compared to a horizontal crest. However, for broad-crested weirs, the slope of the crest has no effect on the C_d . It was also determined that for broad-crested and narrow-crested weirs, a negative slope for the crest increases the C_d .

Based on previous studies, estimating and determining the discharge coefficients (C_d) of trapezoidal broad-crested weirs has been limited and mostly carried out experimentally. Based on the literature review, numerical investigation of these type of weirs have received less attention by researchers. On the other hand, in recent years, intelligent models in estimating the C_d have been used by water industry researchers. Therefore, in the current study, after extracting results from numerical simulation using the ANSYS FLUENT software which solves the problems by the finite volume method (FVM), the performance of various data-mining methods, such as artificial neural networks-multilayer perceptron (ANN (MLP)), gene expression programming (GEP), and random forest (RF) models, were investigated in estimating the discharge coefficients (C_d) for these types of the broad-crested weirs.

2. Materials and Methods

In the present study, a flume with a length, width, and height of 5, 1, and 0.5 meters, respectively, was defined as the simulation domain in the software. The weir was considered in all simulated models at a distance of about 2.5 meters from the inlet of the flume. The weir height is constant in all models and equal to 20 centimetres (Nourani et al., 2021). To investigate the hydraulic characteristics of flow in broad-crested weirs, 164 models with different combinations of upstream face slope (β_1), downstream face slope (β_2), and crest slope (θ) were simulated at a specific ratio of the hydraulic head to weir length (h/L_r).

After numerical simulation and generation of data, the three intelligent models, namely ANN (MLP), GEP, and RF were used to estimate the value of the discharge coefficients (C_d) by changing 5 parameters of h/P , h/L_r , β_1 , β_2 , and θ . For this purpose, simulated data generated by the numerical model, was used in the present study. To achieve a program that can generalize the relations between the inputs and outputs of the model, the program needs to be trained. For this purpose, input and output parameters were divided into two groups. In this study the parameters of h/P , h/L_r , β_1 , β_2 , and θ were used as inputs and the parameter of C_d as output. Also, in all models, 75% of the total available data (123 numbers) were used for the training phase and 25% (41 numbers) were used for the testing phase.

In this study, several statistical indices besides graphical plots have been used to evaluate the accuracy of intelligent models in estimating the discharge coefficients (C_d). These indices include R^2 , RMSE, RE%, and KGE (Kling et al. 2012) and are used to assess the accuracy of the intelligent models in estimating the C_d .

3. Results

After defining the necessary parameters for each of the selected models, the estimated data from each model was used to evaluate their performance during the testing phase. Based on the results of statistical indices for each developed model, it was observed that the KGE value for all three intelligent models is in the very good range ($0.7 < KGE \leq 1.0$), indicating that these models have a good quality in estimating the discharge coefficient (C_d). However, the ANN (MLP) model has a relatively higher accuracy compared to GEP and RF models. Therefore, it can be said that selecting an optimal structure with a proper activation function reduces errors and increases the accuracy of the estimation of model. The ANN (MLP) provides much better results than other models in estimating the C_d and its results are closer to the observed results (extracted from the numerical simulation in this study).

5) Discussion and Conclusion

In this study, the performance of intelligent models including ANN, GEP, and RF in estimating the discharge coefficients (C_d) in broad-crested weirs were investigated using parameters such as h/P , h/L_r , β_1 , and β_2 . For this purpose, simulated data using the finite volume method with ANSYS FLUENT software was used.

- The results were evaluated using statistical metrics including R^2 , RMSE, and KGE, as well as graphical diagrams. For the ANN (MLP) model, the values of R^2 , RMSE, and KGE were 0.906, 0.016, and 0.927, respectively. For the GEP model, these values were 0.790, 0.025, and 0.780, respectively. And for the RF model, they were 0.898, 0.013, and 0.841, respectively.
- The statistical metrics indicate the superiority of the ANN (MLP) model over the other models, with a relative error (RE%) range of less than $\pm 5\%$.
- The violin plot shows a very close similarity between the probability distribution of the estimated data using the ANN (MLP) model and the data obtained from the finite volume method (FVM).
- Additionally, the probability distribution function of the residuals in the ANN (MLP) model is more symmetric and closer to normal distribution than in the other models examined.

6) Six important references

- 1) Azimi, A. H., and Rajaratnam, N. 2009. Discharge characteristics of weirs of finite crest length. *J. Hydraul. Eng.*, 135(12): 1081– 1085.
- 2) Bos, M. G., 1989. Discharge Measurements structures. *J. Irrig. Drain Eng.*, 120(1): 13-26.
- 3) Hager, H., and Schwalt. M. 1989. Broad crested weir. Intl. Institute for Land Reclamation and Improvement, Wageningen NL.
- 4) Hager, H., and Fritz, M. 1998. Hydraulics of Embankment weirs. *J. Hydr. Eng. ASCE*, 124(9).
- 5) Sargison, J. E., and Percy, A. 2009. Hydraulics of broad-crested weirs with varying side slopes, *Journal of irrigation and drainage engineering*, 135(1): 115-118.
- 6) Nourani B., Arvanaghi H., Salmasi F., 2021. Effects of different configurations of sloping crests and upstream and downstream ramps on the discharge coefficient for broad-crested weirs, *Journal of Hydrology*, Volume 603, Part B, 126940.

Conflict of Interest

Authors declared no conflict of interest.

بررسی و مقایسه عملکرد سه روش داده کاوی ANN، GEP و RF در تخمین ضریب دبی جریان در سرریزهای لبه پهن با مقطع طولی ذوزنقه‌ای

بهرام نورانی^۱، هادی ارونقی^۲، فرزین سلماسی^۳

تاریخ ارسال: ۱۴۰۰/۰۴/۰۱

تاریخ پذیرش: ۱۴۰۰/۰۸/۱۲

مقاله پژوهشی

چکیده

تعیین دقیق ضریب دبی نقش بسیار مهمی را در برآورد مقدار جریان از روی سرریزها دارد. لذا تخمین آن با روش‌های دقیق مهم می‌باشد. هدف از این تحقیق، مدل‌سازی و تخمین ضریب دبی در سرریزهای لبه پهن با مقطع طولی ذوزنقه‌ای می‌باشد. برای این منظور پس از شبیه‌سازی عددی جریان با نرم‌افزار ANSYS FLUENT و استخراج نتایج، از سه مدل هوشمند ANN(MLP)، GEP و RF در تخمین ضریب دبی استفاده شد و نتایج آن‌ها مورد مقایسه قرار گرفت. ارزیابی نتایج با استفاده از معیارهای آماری R^2 ، RMSE و KGE و همچنین دیاگرام‌های گرافیکی انجام گردید. مقادیر معیارهای آماری R^2 ، RMSE و KGE به ترتیب ۰/۹۰۶، ۰/۰۱۶ و ۰/۹۲۷ برای مدل ANN(MLP)، ۰/۷۹۰، ۰/۰۲۵ و ۰/۷۸۰ برای مدل GEP و ۰/۸۹۸، ۰/۰۱۳ و ۰/۸۴۱ برای مدل RF به دست آمد. نتایج معیارهای آماری، برتری ANN(MLP) نسبت به سایر روش‌ها را نشان می‌دهد. محدوده درصد خطای نسبی (%RE) برای مدل برتر کمتر از حدود ۵٪ \pm حاصل گردید. دیاگرام ویولنی نشان داد که توزیع احتمال داده‌های تخمینی با مدل ANN(MLP) با داده‌های حاصل از شبیه‌سازی عددی به روش حجم محدود تطابق بسیار نزدیک و مشابهی دارد. همچنین نمودار تابع چگالی احتمال باقی‌مانده‌ها در مدل ANN(MLP) نسبت به سایر مدل‌های مورد بررسی در تحقیق حاضر به توزیع نرمال نزدیکتر می‌باشد.

واژه‌های کلیدی: سرریز لبه پهن، ضریب دبی جریان، روش حجم محدود، مدل هوشمند، ANSYS FLUENT

^۱ دانشجوی دکتری سازه‌های آبی، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران، تلفن تماس: ۰۹۱۴۷۸۲۲۶۱۹، پست الکترونیکی: nourani.bahram@tabrizu.ac.ir

^۲ دانشیار، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران، تلفن تماس: ۰۹۱۴۴۱۹۵۰۸۵، پست الکترونیکی: arvanaghi.hadi@gmail.com (مسئول مکاتبه)

^۳ استاد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران، تلفن تماس: ۰۹۱۴۳۱۷۸۳۳۲، پست الکترونیکی: salmasi@tabrizu.ac.ir



مقدمه

سرریزها سازه‌های هیدرولیکی ساده‌ای هستند که به منظور کنترل، تنظیم و اندازه‌گیری جریان آب در کانال‌های آبیاری مورد استفاده قرار می‌گیرند. سرریزهای لبه‌پهن استاندارد با دارا بودن مزیت حساسیت کم نسبت به استغراق، مشخصات هندسی ساده و هزینه کم در اجرا و ساخت از متداول‌ترین این سازه‌ها در شبکه‌های آبیاری می‌باشند (فرهودی و همکاران، ۱۳۸۶).

در مورد عملکرد هیدرولیکی سرریزهای لبه‌پهن مطالعات بسیار زیادی صورت گرفته است. از جمله می‌توان به تحقیقات (Bos (1985), Hager and Schwalt (1994), Saker and Rhodes (2003), Hager and Fritz (1998), Sargison and Gogus et al. (2006) و ... اشاره نمود. Percy (2009) جریان روی سرریزهای لبه‌پهن با مقطع طولی دوزنقه‌ای را در شیب‌های مختلف بالادست و پایین‌دست به صورت آزمایشگاهی مورد بررسی قرار دادند. نتایج نشان داد که شیب وجه بالادست بیشترین اثر را در افزایش دبی جریان دارد و با کاهش شیب بالادست، ضریب دبی جریان افزایش می‌یابد. هم‌چنین ایشان در بررسی خود مشاهده کردند که شیب پایین‌دست تاثیری در مقدار ضریب دبی جریان ندارد. فرهودی و همکاران (۱۳۸۶) به بررسی تاثیر شیب‌دار کردن وجه بالادست سرریزهای لبه‌پهن مستطیلی بر ضریب جریان و مشخصات جریان در این سازه‌ها پرداختند و به این نتیجه رسیدند که ضریب دبی جریان و در نتیجه ظرفیت تخلیه سرریز، متناسب با کاهش شیب بالادست نسبت به سرریز لبه‌پهن استاندارد افزایش می‌یابد، که این افزایش در ضریب دبی جریان در سرریز با شیب ۱۵ درجه تا ۲۰٪ نیز می‌رسد. عباسپور و همکاران (۱۳۹۲) به مقایسه نتایج حاصل از شبیه‌سازی عددی جریان بر روی سرریزهای لبه‌پهن مستطیلی با وجوه شیب‌دار به وسیله نرم‌افزار فلوئنت^۱ با نتایج آزمایشگاهی پرداختند و نتایج نشان داد که مدل آشفتگی $k - \epsilon$ نتایج مطلوبی را در پیش‌بینی پروفیل سطح آب، پروفیل سرعت، ضریب دبی جریان و ضخامت لایه مرزی دارد. حداکثر خطا

در مطالعات آن‌ها برای بار آبی و ضریب دبی جریان به ترتیب ۴/۰۶ و ۶/۵۴ به دست آمد. (Madadi et al. (2014) به بررسی خصوصیات هیدرولیکی جریان در سرریزهای لبه‌پهن دوزنقه‌ای پرداختند. نتایج نشان داد با کاهش شیب وجه بالادست تا ۲۱ درجه، ضریب دبی جریان در حدود ۱۰ درصد افزایش می‌یابد و طول و ارتفاع ناحیه جداشدگی در تاج به ترتیب تا ۸۰ و ۹۵ درصد کاهش می‌یابد. (Azimi et al. (2013) با استفاده از داده‌های آزمایشگاهی محققان قبلی روابطی را برای ضریب دبی سرریزهای لبه‌پهن با تاج شیب‌دار و هم‌چنین سرریزهای با وجوه شیب‌دار در بالادست، پایین‌دست و یا هر دو در شرایط جریان آزاد به دست آوردند. (Azimi and Rajaratnam (2009) با ادغام مطالعات قبلی خود و داده‌های آزمایشگاهی در دسترس روابطی را برای ضریب دبی استخراج نمودند. نتایج نشان داد که برای سرریزهای لبه‌پهن طولانی، یک شیب مثبت برای تاج سرریز، دبی جریان را در مقایسه با حالت سرریز با تاج افقی افزایش می‌دهد. این در حالی است که برای سرریزهای لبه‌پهن، شیب تاج تاثیری در ضریب دبی جریان ندارد. هم‌چنین تعیین گردید که برای سرریزهای لبه‌پهن و لبه‌باریک، یک شیب منفی برای تاج سرریز، ضریب دبی را افزایش می‌دهد. (Aksoy and Doğan (2016) به بررسی آزمایشگاهی اثر زاویه نزدیک‌شونده بر بازده دبی جریان در سرریزهای لبه‌پهن پرداختند. آن‌ها آزمایشات خود برای چهار شیب (در شرایطی که شیب بالادست و پایین‌دست یکسان باشد) و دو طول تاج را به ازای سه دبی مختلف در یک فلوم آزمایشگاهی به طول ۸ متر و عرض ۱۵ سانتی‌متری انجام داده و به این نتیجه رسیدند که ضریب دبی جریان فقط تابعی از طول تاج سرریز نبود، بلکه تابعی از زاویه وجه بالادست نیز می‌باشد. (Lie et al. (2018) ضریب دبی جریان و ناحیه جداشدگی جریان بر روی سرریزهای لبه‌پهن با وجه شیب‌دار بالادست را با استفاده از نرم‌افزار OpenFOAM مورد بررسی قرار دادند. نتایج نشان داد که مدل عددی مذکور به خوبی پروفیل سطح جریان و ضریب

¹ FLUENT

این نوع سرریزها کمتر مورد توجه محققان بوده و ضمن اینکه بیشتر تحقیقات در این زمینه به گونه‌ای بوده که تاج سرریز به صورت افقی در نظر گرفته شده است. بنابراین در تحقیق حاضر ابتدا شرایط هیدرولیکی جریان بر روی سرریزهای لبه‌پهن با مقطع طولی ذوزنقه‌ای در دو حالت بدون تاج شیب‌دار (تاج افقی) و تاج شیب‌دار (تاج با شیب مثبت و تاج با شیب منفی) به روش عددی با شرایط مرزی معین با استفاده از روش حجم محدود^۵ و با به کارگیری نرم‌افزار ANSYS FLUENT مورد شبیه‌سازی قرار گرفت. سپس عملکرد روش‌های مختلف داده کاوی نظیر مدل ANN(MLP)، مدل GEP و مدل RF^۶ در تخمین ضریب دبی این نوع از سرریزها مورد بررسی قرار گرفت.

مواد و روش‌ها

معادله حاکم

برای یک سرریز لبه‌پهن و هم عرض کانال، رابطه دبی - اشل به صورت رابطه (۱) تعریف می‌گردد:

$$Q = C_d B \sqrt{2gH^3} \quad (1)$$

که در آن Q دبی جریان (m^3/s)، C_d ضریب دبی جریان B ، عرض فلوم یا سرریز (m)، g شتاب ثقل (m/s^2) و H بار آبی کل ($H = h + Q^2 / (2gB^2(h + P)^2)$)، h ارتفاع آب بالای تاج سرریز (m) و P ارتفاع سرریز (m) می‌باشند (Horton and Murphy, 1906). از آن جایی که اندازه‌گیری بار آبی کل در محل امکان‌پذیر نیست، بار آبی کل (H) با بار آبی اندازه‌گیری شده در بالادست سرریز (h) جایگزین شده و تاثیر این امر نیز در ضریب دبی جریان لحاظ می‌گردد.

$$Q = C_d B \sqrt{2gh^3} \quad (2)$$

آنالیز ابعادی پارامترهای موثر بر ضریب دبی جریان

دبی را پیش‌بینی می‌کند. هم‌چنین مشاهده شد که با افزایش زاویه شیب بالادست، ضریب دبی جریان کاهش می‌یابد. هم‌چنین در تحقیقات ایشان مشاهده گردید که ناحیه جدا شدگی در تاج سرریز در زوایای بیش از ۶۰ درجه اتفاق می‌افتد و با افزایش این زاویه، ابعاد ناحیه جدا شدگی افزایش می‌یابد.

در سال‌های اخیر استفاده از مدل‌های هوشمند در تخمین ضریب دبی جریان مورد استقبال محققان صنعت آب قرار گرفته که در این میان می‌توان به تحقیقات صورت گرفته توسط Roushangar et al. (2017) در تخمین ضریب دبی جریان در سرریزهای کنگره‌ای با مدل‌های ANFIS^۱ و GEP^۲، تخمین ضریب دبی جریان در سرریزهای کنگره‌ای و کنگره‌ای قوس‌دار با مدل SVM^۳ به وسیله Roushangar et al. (2018) و استفاده از الگوریتم‌های SVM و GEP برای افزایش دقت پیش‌بینی رابطه دبی - اشل در سرریزهای کنگره‌ای و لبه‌تیز مستغرق توسط فولادی‌پناه و همکاران (۱۳۹۹) اشاره نمود. هم‌چنین نورانی و همکاران (۲۰۱۹) به عملکرد مدل‌های هوشمند ANN^۴، GEP در تخمین پارامتر بی بعد h/p رابطه دبی - اشل سرریزهای لبه‌تیز پلان مثلثی پرداختند. ارزیابی نتایج به دست آمده برتری مدل ANN نسبت به مدل GEP را نشان داد و مشاهده گردید مدل ANN نتایج بسیار مشابهی با داده‌های نتایج مشاهداتی را می‌تواند برآورد نماید. Norouzi et al. (2019) عملکرد مدل‌های هوشمند ANNs و SVM را در پیش‌بینی ضریب دبی سرریزهای غیرخطی را بررسی نمودند. نتایج آن‌ها نشان داد که عملکرد هر دو مدل در پیش‌بینی ضریب دبی سرریز غیرخطی قابل قبول است اما نتایج حاصل از مدل ANNs به داده‌های آزمایشگاهی نزدیکتر است.

همان‌طور که مشاهده گردید، برآورد و تعیین ضریب دبی سرریزهای لبه‌پهن ذوزنقه‌ای به صورت محدود و بیشتر به صورت آزمایشگاهی صورت گرفته شده است. با توجه به بررسی منابع صورت گرفته مشاهده می‌شود بررسی عددی

⁴ Artificial Neural Network

⁵ Finite Volume Method

⁶ Random Forest

¹ Adoptive Neuro-Fuzzy Inference System

² Gene Expression Programming

³ Support Vector Machine



محدوده محاسباتی و پارامترهای مورد بررسی در تحقیق حاضر

در تحقیق حاضر یک فلومی به طول، عرض و ارتفاع به ترتیب ۵، ۱ و ۰/۵ متر به عنوان محدوده شبیه‌سازی در نرم‌افزار تعریف گردید. تاج سرریز در کل مدل‌های شبیه‌سازی شده در فاصله حدود ۲/۵ متری از ورودی فلوم در نظر گرفته شد. ارتفاع سرریز در تمامی مدل‌ها ثابت و برابر ۲۰ سانتی‌متر می‌باشد. به منظور بررسی خصوصیات هیدرولیکی جریان در سرریزهای لبه‌پهن، ۱۶۴ مدل با ترکیب‌های مختلفی از شیب وجه بالادست (β_1)، شیب وجه پایین‌دست (β_2) و شیب تاج (θ) در نسبت بار آبی به طول تاج سرریز (h/L_r) معین مورد شبیه‌سازی قرار گرفت. این شبیه‌سازی برای ۵ حالت برای هر کدام از شیب وجه بالادست (۹۰، ۷۵/۹۶، ۶۳/۴۳، ۴۵ و ۲۶/۵۶) و شیب وجه پایین‌دست (۹۰، ۷۵/۹۶، ۶۳/۴۳، ۴۵ و ۲۶/۵۶) در حالت تاج افقی (شیب صفر) و ۴ حالت برای هر کدام از شیب وجه بالادست و پایین‌دست (۷۵/۹۶، ۶۳/۴۳، ۴۵ و ۲۶/۵۶) با زاویه یکسان برای هر دو شیب وجه بالادست و پایین‌دست در حالت تاج با شیب مثبت (۹/۴۶ و ۴/۷۶) و تاج با شیب منفی (۹/۴۶ - و ۴/۷۶ -) در ۴ حالت بار آبی معین برای محدوده سرریزهای لبه‌پهن حقیقی ($0.1 < h/L_r \leq 1/4$) انجام گردید.

معادلات حاکم بر جریان در شبیه‌سازی عددی

پایه و اساس تمام روش‌های عددی در مهندسی هیدرولیک حل معادلات حاکم بر جریان شامل معادلات پیوستگی و معادلات مومنتم است که اصطلاحاً به معادلات ناویر-استوکس^۲ معروف می‌باشند. برای یک جریان تراکم ناپذیر با ویسکوزیته ثابت معادلات مذکور به در حالت تانسوری به ترتیب به فرم روابط (۷) و (۸) نوشته می‌شوند:

$$\frac{\partial}{\partial x_i} (\bar{u}_i) = 0 \quad (7)$$

در سرریزهای لبه‌پهن ضریب دبی جریان تابعی از پارامترهای هندسی، شرایط هیدرولیکی و خصوصیات سیال است که بر اساس این سه پارمتر رابطه دبی-اشل در چنین سرریزهایی می‌تواند با رابطه (۳) بیان گردد (شکل ۱).

$$F_1(q, h, L_r, P, \beta_1, \beta_2, \theta, g, \rho, \mu, \sigma) = 0 \quad (3)$$

در رابطه (۳)، q دبی در واحد عرض جریان عبوری از روی تاج سرریز (m^2/s)، h بار آبی روی تاج سرریز (m)، L_r طول تاج سرریز در جهت جریان (m)، P ارتفاع سرریز (m)، β_1 و β_2 زاویه شیب وجه بالادست و پایین‌دست، θ زاویه شیب تاج، g شتاب ثقل (m/s^2)، ρ جرم مخصوص آب (Kg/m^3)، μ لزجت دینامیکی آب (N/m^2) و σ کشش سطحی آب (N/m) می‌باشند. بر اساس تئوری باکینگهام^۱ و با در نظر گرفتن ρ ، g و h به عنوان پارامترهای تکرارشونده و اعمال یکسری عملیات ریاضی مجاز (ضرب عدد بی‌بعد در یک عدد ثابت) پس از استخراج پارامترهای بی‌بعد حاصل از رابطه (۳)، پارامترهای بی‌بعد نهایی می‌توانند به صورت رابطه (۴) تعریف گردند:

$$\frac{q}{\sqrt{2gh^3}} = F_2\left(\frac{\sigma}{\rho g h^2}, \frac{\mu}{\rho h^2 \sqrt{g}}, \frac{h}{L_r}, \frac{h}{P}, \beta_1, \beta_2, \theta\right) \quad (4)$$

در رابطه (۴)، دو پارامتر بی‌بعد اول حالتی از عدد وبر (We) و عدد رینولدز (Re) می‌باشند که به ترتیب تاثیر کشش سطحی و لزوجت سیال را نشان می‌دهند. Hager and Schwalt (1994) پیشنهاد کردند که عمق جریان بر روی تاج سرریز نباید کمتر از ۵۰ میلی‌متر باشد. بنابراین در تحقیق حاضر عمق‌های کمتر از ۵۰ میلی‌متر در آنالیز در نظر گرفته نشده‌اند. در صورتی که تاثیر کشش سطحی و لزوجت نادیده گرفته شود رابطه (۵) می‌تواند به صورت رابطه (۶) باز نویسی گردد.

$$\frac{q}{\sqrt{2gh^3}} = F_3\left(Re, We, \frac{h}{L_r}, \frac{h}{P}, \beta_1, \beta_2, \theta\right) \quad (5)$$

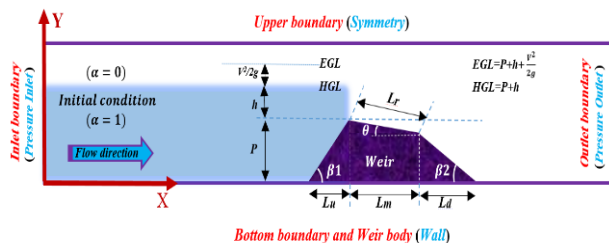
$$C_d = \frac{q}{\sqrt{2gh^3}} = F_4\left(\frac{h}{L_r}, \frac{h}{P}, \beta_1, \beta_2, \theta\right) \quad (6)$$

² Navier- Stokes

¹ Buckingham

اختصاص می‌دهد که تحت این شرایط المان دارای سطح آزاد سیال می‌باشد.

مقدار Tora et al. (2017) و Bayon et al. (2016) عددی α برابر $0/5$ ($\alpha=0/5$) را به عنوان معرف سطح آزاد سیال معرفی نمودند. هم‌چنین لازم به ذکر است برای گسسته‌سازی عبارت فشار از روش PISO^۳ و برای گسسته‌سازی عبارت ممتم از روش مرتبه دوم بالادست^۳ استفاده شد. در شکل ۱ موقعیت و نوع شرایط مرزی^۴ و هم‌چنین شرط اولیه^۵ برای محدوده محاسباتی در حضور هندسه سرریز نشان داده شده است.



شکل (۱): نحوه اعمال شرایط مرزی و اولیه در شبیه‌سازی عددی در تحقیق حاضر

همان‌طور که در شکل (۱) مشاهده می‌گردد، شرایط مرزی برای جریان ورودی به صورت فشار جریان ورودی (Pressure outlet)، برای خروجی به صورت فشار خروجی (Pressure Inlet)، برای کف کانال و بدنه سرریز لبه پهن داخل کانال به صورت دیوار (Wall) و برای بالای کانال به صورت متقارن (Symmetry) در حالت دو بعدی لحاظ شده است. شبیه‌سازی عددی جریان به صورت غیردائمی^۶ با گام زمانی^۷ $0/001$ ثانیه جهت برقراری در پایداری حل عددی انجام گردید و تا رسیدن به حالت دائمی ادامه یافت.

برای شبکه‌بندی^۸ محدوده محاسباتی در تحقیق حاضر از المان چهار وجهی و سه وجهی منظم با کیفیت مناسب

$$\rho \left(\frac{\partial \bar{u}_i}{\partial t} + u_j \frac{\partial \bar{u}_i}{\partial x_j} \right) = f_i - \frac{\partial \bar{P}}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\mu \frac{\partial \bar{u}_i}{\partial x_j} - \rho \overline{u_i' u_j'} \right) \quad (8)$$

در روابط فوق u_i و u_j مولفه بردار سرعت در راستای فضایی i و j ، f_i نیروی حجمی در راستای i ، P بیانگر فشار، μ ویسکوزیته سیال، ρ چگالی سیال و $\overline{u_i' u_j'}$ تنش رینولدزی می‌باشد که برای تعیین آن از مدل‌های آشفتگی استفاده می‌شود. مدل‌های آشفتگی متنوعی در شبیه‌سازی‌های عددی مورد استفاده قرار می‌گیرند که در تحقیق حاضر از مدل آشفتگی $k-\epsilon$ (RNG) بخاطر کاربرد گسترده و معروف بودن در ارائه نتایج دقیق در جریان‌های با شدت آشفتگی کم، پس از بررسی و صحت‌سنجی استفاده شد. استفاده از مدل آشفتگی $k-\epsilon$ (RNG) در بسیاری از شبیه‌سازی‌های هیدرولیکی نتایج بسیار مطلوبی را ارائه کرده است که در این بین می‌توان به شبیه‌سازی عددی جریان در سرریزهای استوانه‌ای توسط ورجاروند و همکاران (۲۰۰۹)، شبیه‌سازی جریان در سرریزهای منحنی‌الخط توسط Simsek et al. (2016) شبیه‌سازی جریان در سرریزهای پلان مثلثی توسط نورانی و همکاران (۲۰۱۹) اشاره نمود. هم‌چنین برای شبیه‌سازی جریان دو فازی در محدوده محاسباتی در تحقیق حاضر از روش حجم سیال^۱ (VOF) که توسط Hirt and Nichols (1981) ارائه شده است، استفاده گردید که رابطه انتقالی کسر سیال در آن با رابطه (۹) بیان می‌گردد:

$$\frac{\partial}{\partial t}(\alpha) + u_i \frac{\partial}{\partial x_i}(\alpha) + u_j \frac{\partial}{\partial x_j}(\alpha) = 0 \quad (9)$$

در شبیه‌سازی جریان دو فازی شامل آب و هوا، در صورتی که سلولی پر از هوا باشد مقدار نسبت حجمی صفر ($\alpha=0$)، در صورتی که سلول کاملاً پر از آب باشد، مقدار یک ($\alpha=1$) و در صورتی که سلولی هر دو سیال را داشته باشد، مقدار عددی بین صفر و یک ($0 < \alpha < 1$) را به خود

⁵ Initial Value

⁶ Transient

⁷ Time Step

⁸ Mesh Generation

¹ Volume of Fluid (VOF) Model

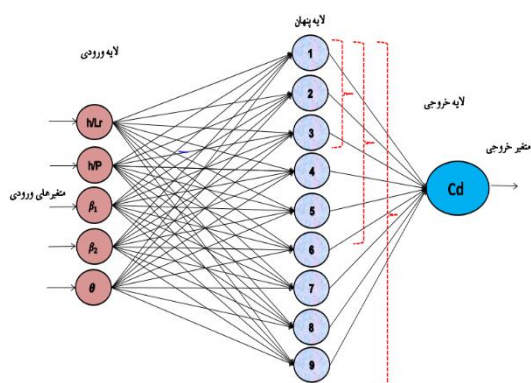
² Pressure-Implicit with Splitting of Operator

³ Second order upwards

⁴ Boundary Conditions



ورودی در خلال شبکه و در مسیری رو به جلو به صورت لایه لایه منتشر می‌گردد. لازم به ذکر است که معماری بهینه^۵ و نوع تابع محرک^۶ برای لایه ورودی و لایه خروجی پس از ارزیابی شرایط مختلف تعیین گردید. در این راستا مدل تخمینی با ANN(MLP) در محیط نرم‌افزار Statistica 12 به گونه‌ای تنظیم گردید که داده‌ها به صورت تصادفی^۷ برای دو مرحله آموزش^۸ و آزمون^۹ انتخاب گردد تا این ترکیب از داده‌ها برای مرحله آموزش و آزمون در سایر مدل‌ها نیز مورد استفاده قرار گیرد. شکل (۳) ساختار کلی مدل ANN(MLP) مورد استفاده در تحقیق حاضر با تعداد نرون‌های ۳، ۶ و ۹ در لایه پنهان را نشان می‌دهد.

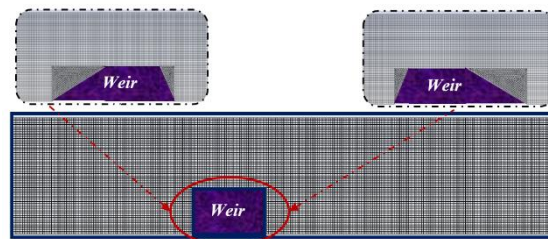


شکل (۳): ساختار کلی شبکه عصبی مصنوعی (ANN) در تحقیق حاضر

برنامه‌ریزی بیان ژن

مدل برنامه‌ریزی بیان ژن (GEP) تعمیم یافته الگوریتم ژنتیک بوده و یک تکنیکی با برنامه‌ریزی خودکار می‌باشد که راه‌حل مسئله را می‌تواند با استفاده از برنامه کامپیوتری ارائه کند. در این روش در ابتدای فرآیند هیچ‌گونه رابطه تابعی در نظر گرفته نشده و این روش قادر به بهینه‌سازی ساختار مدل و مولفه‌های آن می‌باشد. برنامه‌ریزی بیان ژن بر خلاف الگوریتم ژنتیک روی ساختار درختی فرمول‌ها

($Orthogonal\ Quality < 0.7$) استفاده شده است. همچنین یک آزمون مستقل از شبکه‌بندی در جهت کنترل عدم تاثیر اندازه شبکه‌ها بر نتایج حاصل از شبیه‌سازی عددی انجام گرفت. نتایج ارزیابی نشان داد با انتخاب مدل $k - \epsilon(RNG)$ و تعداد شبکه در حدود ۲۴۰۰۰، مدل عددی می‌تواند دقت بسیار مشابهی را به نتایج آزمایشگاهی (Azimi et al. (2014)) ارائه نماید. در شکل (۲) خصوصیات شبکه‌بندی محدوده محاسباتی در تحقیق حاضر برای سرریز لبه‌پهن استاندارد و سرریز لبه‌پهن با وجوه شیب‌دار در شرایطی که شیب وجه بالادست و پایین‌دست به ترتیب $1V:2H$ و $1V:0.5H$ و حالت عکس آن باشد، نشان داده شده است.



شکل (۲): نحوه اعمال شبکه‌بندی برای محدوده محاسباتی در شبیه‌سازی عددی در تحقیق حاضر

شبکه عصبی مصنوعی

شبکه عصبی مصنوعی (ANN) متشکل از مجموعه‌ای از نرون‌ها با ارتباطات داخلی بین یکدیگر می‌باشد که قادر است بر اساس اطلاعات و داده‌های ورودی، جواب‌های خروجی را تهیه نماید (Norouzi et al., 2019). یکی از مهمترین ساختارها در شبکه عصبی مصنوعی، شبکه چند لایه پیشخور می‌باشد. به طور معمول این شبکه‌ها شامل مجموعه‌ای از نرون‌های پایه می‌باشند که تشکیل دهنده لایه ورودی^۱، یک یا چند لایه پنهان^۲ و یک لایه خروجی^۳ می‌باشند. در تحقیق حاضر از شبکه با عنوان پرسپترون چند لایه^۴ (MLP) استفاده گردید که در آن سیگنال

⁶ Active Function

⁷ Random Sample

⁸ Train

⁹ Test

¹ Input Layer

² Hidden Layer

³ Output Layer

⁴ Multi-layer Perceptron

⁵ Optimum Structure

تشکیل هر درخت دسته متفاوتی از الگوهای موجود با در نظر گرفتن جایگزینی دوباره هر الگوی منتخب انتخاب می-شوند. اندازه هر دسته برابر تعداد الگوهای موجود خواهد بود. الگوریتم RF توسط (Bermin, 2001) به عنوان یک روشی که از توسعه جدید درخت‌های تصمیم‌گیری ایجاد شده ارائه گردید که پیش‌بینی چندین الگوریتم منفرد را باهم با استفاده از قوانین مبتنی ترکیب می‌کند. فرایند کلی الگوریتم RF به صورت ساده در شکل (۴) نشان داده شده است. در این الگوریتم ابتدا T نمونه خود راه انداز از داده‌های آموزش بیرون کشیده شده و سپس از هر نمونه خود راه انداز یک درخت طبقه‌بندی و رگرسیونی (CART¹) هرس نشده ایجاد می‌شود که برای انشعاب در هر گره CART تنها یک متغیر انتخاب شده به صورت تصادفی استفاده می‌شود. در نهایت خروجی طبقه‌بندی بر اساس یک نتیجه میانگین از پیش‌بینی‌های تمام درخت‌های منفرد آموزش دیده به دست می‌آید. لازم به ذکر است که یک مجموعه داده خود راه انداز مجموعه‌ای از نقاط انتخابی به طور تصادفی است که از مجموعه داده آموزشی بیرون کشیده شده است. برای اینکه اندازه نمونه آموزشی اولیه ثابت بماند، مجموعه داده خود راه انداز یک نسخه کپی از نقاط را به کار می‌برد. همچنین این الگوریتم به دلیل ارزیابی درونی نتایج هر طبقه‌بندی درختی که در داخل خود دارد و با وزن‌دهی به نتایج هر درخت می‌تواند نتایج صحیحی را تولید نماید.

بجای سلسله ارقام دودویی عمل می‌کند. ساختارهای درختی از مجموعه توابع (عملگرهای ریاضی مورد استفاده در فرمول‌ها) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) ایجاد می‌شوند.

مراحل برنامه‌ریزی بیان ژن به صورت زیر است:

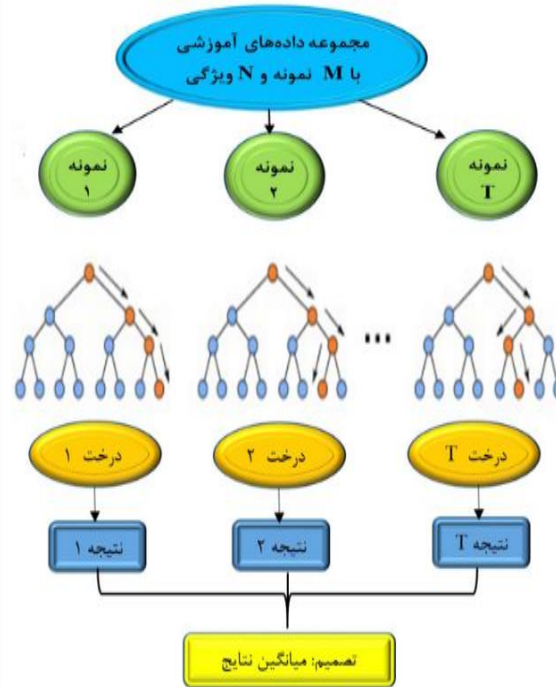
- تولید یک جمعیت اولیه از فرمول‌ها که این فرمول‌ها از ترکیب تصادفی مجموعه توابع (عملگرهای ریاضی مورد استفاده در فرمول‌ها) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) ایجاد می‌شوند.
- هر یک از افراد مذکور با استفاده از توابع برازش مورد ارزیابی قرار می‌گیرند.
- تولید یک جمعیت جدید از فرمول‌ها
- گام سوم تا نیل به حداکثر تعداد تولید، تکرار خواهد شد.

لازم به ذکر می‌باشد که در تحقیق حاضر، برای به کارگیری روش برنامه‌ریزی بیان ژن از نرم‌افزار GenXpro Tools 4.0 استفاده شد (دهقانی و همکاران، ۱۳۹۴).

جنگل تصادفی

جنگل تصادفی (RF) یک الگوریتم یادگیری ماشین با کاربرد آسان و منعطف است که در بیشتر مواقع حتی بدون تنظیم بیش از حد پارامترها نتیجه بسیار خوبی ایجاد می‌کند. همچنین به دلیل سادگی و تنوع، یکی از پرکاربردترین الگوریتم‌ها است. الگوریتم RF می‌تواند برای کارهای طبقه‌بندی و رگرسیونی استفاده شود. در الگوریتم RF برای

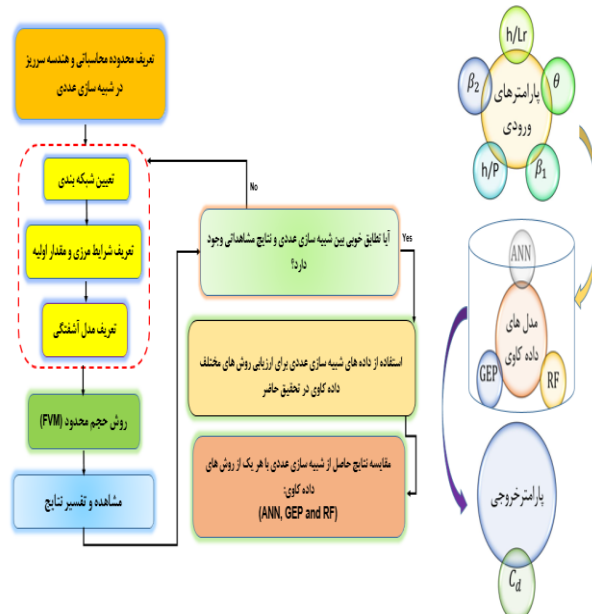
¹ Regression and Classification Tree



شکل (۴): ساختار کلی الگوریتم جنگل تصادفی (RF)

مدل‌ها ۷۵ درصد کل داده‌های موجود (۱۲۳ عدد) برای مرحله آموزش و ۲۵ درصد آن‌ها (۴۱ عدد) نیز برای مرحله آزمون مورد استفاده قرار گرفته است. لازم به ذکر است که عموماً ۷۰ الی ۸۰ درصد کل داده‌ها برای آموزش و ۲۰ الی ۳۰ درصد نیز برای آزمون در نظر گرفته شود که این مورد در تحقیق حاضر نیز رعایت شده است (Bowden et al., 2005). در شکل (۵) فرایند کلی مراحل طی شده برای تخمین ضریب دبی جریان در سرریزهای لبه‌پهن در تحقیق حاضر به صورت کامل آورده شده است.

در تحقیق حاضر از سه مدل هوشمند مدل ANN(MLP)، GEP و RF در تخمین مقدار ضریب دبی (C_d) با استفاده از ۵ پارامتر $\beta_1, \beta_2, h/Lr, h/P$ و θ استفاده شده است. برای این منظور از داده‌های شبیه‌سازی شده توسط مدل عددی در تحقیق حاضر استفاده گردید. برای رسیدن به برنامه و ساختاری که بتواند روابط بین ورودی‌ها و خروجی مدل را تعمیم دهد، لازم است برنامه آموزش داده شود. بدین منظور پارامترهای ورودی و خروجی به دو دسته تقسیم‌بندی گردید. در کل مدل‌های هوشمند پارامترهای $\beta_1, \beta_2, h/Lr, h/P$ و θ به عنوان ورودی و پارامتر C_d به عنوان خروجی معرفی شد و همچنین در همه



شکل (۵): مراحل انجام گرفته جهت برآورد و تخمین ضریب دبی جریان با مدل عددی و هوشمند

معیارهای ارزیابی دقت

جهت تخمین در تحقیق حاضر شامل R^2 , $RMSE$, $RE\%$ و KGE (Kling et al. 2012) می باشند که هر کدام از آن ها به ترتیب از روابط (۱۰)، (۱۱)، (۱۲) و (۱۳) مندرج در جدول (۱) محاسبه می گردند.

برای ارزیابی، توانایی و دقت عملکرد مدل های هوشمند به کار گرفته شده در تخمین ضریب دبی جریان (C_d)، از شاخص های آماری استفاده شده است. معیارهای ارزیابی

جدول (۱): شاخص های آماری مورد استفاده در تحقیق حاضر جهت برآورد دقت نتایج مدل های هوشمند

معیار ارزیابی	محدوده	کیفیت	رابطه
			(۱۰)
			(۱۱)

$$RE = \frac{\sum_{i=1}^n |O_i - P_i|}{\sum_{i=1}^n |O_i|} \times 100 \quad (12)$$

$$KGE = 1 - \sqrt{(R-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (13)$$

$0.7 < KGE < 1.00$	خیلی خوب
$0.6 < KGE < 0.7$	خوب
$0.5 < KGE \leq 0.6$	رضایت بخش
$0.4 < KGE \leq 0.5$	قابل قبول
$KGE \leq 0.4$	نا مطلوب

$$R = \sqrt{R^2}$$

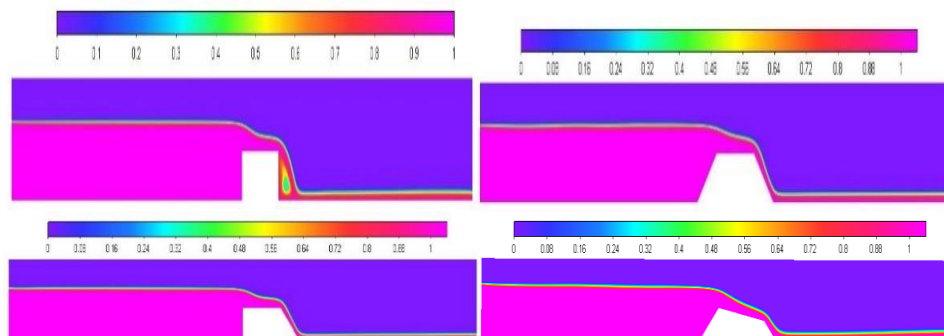
$$\beta = \frac{\bar{P}}{\bar{O}} \quad \gamma = \frac{CV_P}{CV_O} = \frac{\frac{\sigma_P}{\bar{P}}}{\frac{\sigma_O}{\bar{O}}}$$

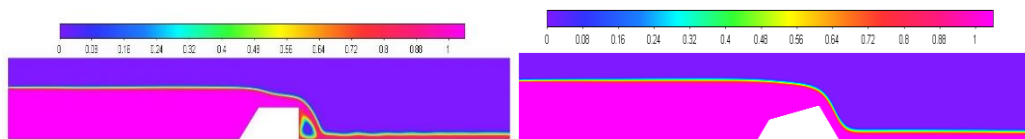
کارگیری نرم افزار ANSYS FLUENT که به روش حجم محدود به حل رابطه حاکم بر جریان می پردازد، پرداخته شد. در شکل (۶) نتایج شبیه سازی عددی پروفیل جریان آب بر روی سرریزهای مورد مطالعه در تحقیق حاضر برای مقادیر مختلفی از زاویه شیب وجه بالادست (۴۵ درجه)، زاویه شیب وجه پایین دست (۴۵ درجه) و تاج با شیب مثبت (زاویه ۴/۷۶ درجه)، تاج با شیب منفی (زاویه ۴/۷۶- درجه) و تاج افقی (زاویه صفر) در شرایطی که $h/Lr = 0/3$ و $h/P = 0/6$ نشان داده شده است.

در این روابط O_i مقادیر حاصل از مقادیر مشاهداتی (استخراجی از مدل عددی به کار گرفته شده در تحقیق حاضر)، P_i مقادیر حاصل از مدل های تخمین کننده (مدل های هوشمند)، \bar{O} و \bar{P} به ترتیب میانگین مقادیر مشاهداتی و تخمینی، CV_O و CV_P به ترتیب ضریب تغییرات مقادیر مشاهداتی و تخمینی، σ_O و σ_P به ترتیب انحراف معیار مقادیر مشاهداتی و تخمینی و n تعداد داده ها می باشد.

نتایج و بحث

همان طور که بیان گردید هدف از تحقیق حاضر برآورد ضریب دبی جریان در سرریزهای لبه پهن با استفاده از سه مدل (ANN(MLP)، GEP و RF می باشد. بدین منظور در تحقیق حاضر ابتدا به شبیه سازی شرایط هیدرولیکی جریان بر روی سرریزهای لبه پهن به روش عددی و با به





شکل (۶): پروفیل سطح جریان آب بر روی سرریزهای مورد مطالعه در تحقیق
به ازای شرایط $h/P = 0/6$ و $h/L_r = 0/3$.

رسیدن به این هدف در نرم‌افزارهای به کارگیری شده یعنی STATISTICA12 و Gene Xpro Tools 4.0 از سه چهارم کل داده‌ها برای آموزش مدل‌ها و یک چهارم کل داده‌ها برای آزمون مدل‌ها استفاده گردید. در جدول (۲) پارامترهای آماری مجموعه کل داده‌ها برای مرحله آموزش و آزمون نشان داده شده است. لازم به ذکر است که واحد زاویه‌ها بر حسب رادیان است.

در تحقیق حاضر در حالت کلی ۱۶۴ مدل عددی شبیه‌سازی گردید و سپس ضریب دبی جریان در هر یک از مدل‌های شبیه‌سازی شده استخراج گردید و با بهره‌گیری از نتایج به دست آمده در مدل‌سازی عددی، به تخمین ضریب دبی جریان با مدل‌های هوشمند و ارزیابی عملکرد هر یک از مدل‌های هوشمند فوق‌الذکر در برآورد ضریب دبی جریان با استفاده از پارامترهای موثر پرداخته شد. برای

جدول (۲): محدوده پارامترها برای داده‌های ورودی و خروجی در مرحله آزمون و آموزش

	داده‌های آموزش						داده‌های آزمون					
	$\frac{h}{L}$	$\frac{h}{p}$	β_1	β_2	θ	C_d	$\frac{h}{L}$	$\frac{h}{p}$	β_1	β_2	θ	C_d
حداقل	۰/۱۴	۰/۳	۰/۴۶	۰/۴۶	-۰/۱۶	۰/۳	۰/۱۵	۰/۳	۰/۴۶	۰/۴۶	-۰/۱۶	۰/۳۱
حداکثر	۰/۳۷	۱/۱۳	۱/۵۷	۱/۵۷	۰/۱۶	۰/۵	۰/۳۷	۱/۱۳	۱/۵۷	۱/۵۷	۰/۱۶	۰/۵۲
میانگین	۰/۲۲	۰/۸۳	۱/۱۰	۱/۱۰	۰/۳۳	۰/۱۹	۰/۲۶	۰/۵۵	۱/۰۳	۰/۹۷	-۰/۰۱	۰/۳۹
انحراف معیار	۰/۰۸	۰/۱۹	۰/۳۷	۰/۳۷	۰/۰۸	۰/۳۸	۰/۰۴	۰/۱۹	۰/۳۷	۰/۳۵	۰/۰۸	۱/۰۵۰
ضریب تغییرات	۰/۳۶	۰/۲۲	۰/۳۳	۰/۳۳	۰/۲۴	۲	۰/۱۵	۰/۳۴	۰/۳۶	۰/۳۶	-۸	۰/۱۳

برای مرحله آزمون و آموزش در جدول (۳) آورده شده است. نرون‌ها از توابع محرک متفاوتی جهت تولید خروجی استفاده می‌کنند، در این تحقیق برای مدل ANN(MLP) از توابع محرک Logistic، Tanh و Expontial یکسان در لایه ورودی و خروجی استفاده گردید که هر کدام از آن‌ها به ازای سه ساختار ۱-۳-۵، ۱-۶-۵ و ۱-۹-۵ به کار

نتایج مدل شبکه عصبی مصنوعی

برای تخمین ضریب دبی جریان (C_d) در سرریزهای لبه پهن با مدل ANN(MLP) از نوع تابع محرک^۱ برای لایه ورودی و خروجی با ساختارهایی مختلفی از تعداد نرون‌ها در لایه پنهان بهره برده شد که نتایج هر یک از این حالت‌ها

^۱ Active Function

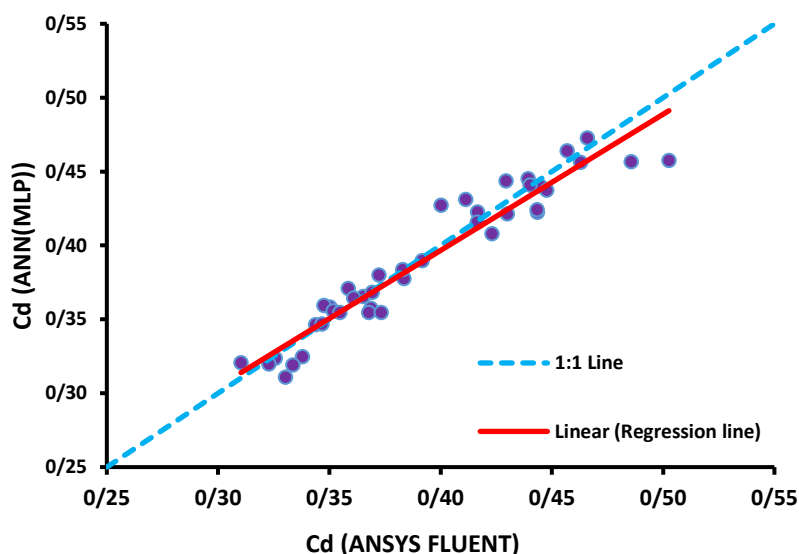


در تعداد نرون ۶ شبکه به آموزش خوبی می‌رسد. در شکل ۷ پراکندگی داده‌ها در مرحله آزمون برای معماری بهینه شبکه در مدل ANN(MLP) نشان داده شده است که تحت این شرایط مقادیر R^2 و RMSE به ترتیب ۰/۹۵۴ و ۰/۰۱۶ می‌باشند. با توجه به شکل ۶ مشاهده می‌گردد که بیشتر نقاط بر روی یا نزدیکی خط نیمساز قرار گرفته‌اند که این مقبولیت این مدل ANN (MLP) در تخمین مقدار ضریب دبی جریان (C_d) در سرریزهای لبه‌پهن را بیان می‌نماید.

گرفته شده‌اند. همان‌طور که مشاهده می‌گردد، مدل ANN(MLP) با تابع محرک Tanh برای لایه ورودی و خروجی در صورتی که معماری شبکه به صورت ۱-۶-۵ باشد، نسبت به سایر حالت‌های مورد بررسی در تحقیق حاضر به دلیل داشتن RMSE کمتر از دقت بالاتری برخوردار می‌باشد. در تعیین تعداد لایه پنهان، ابتدا تعداد نرون‌ها در لایه مخفی کمتر از تعداد ورودی‌ها انتخاب گردید سپس تعداد نرون‌های لایه پنهان افزایش داده شد و مشاهده گردید که در تعداد نرون‌های بیشتر شبکه همگرا و نتایج مطلوب‌تری را ارائه می‌دهد و در نهایت مشخص گردید

جدول (۳): تخمین ضریب دبی جریان با مدل ANN(MLP) با تابع انتقال و ساختارهای مختلف

نوع تابع انتقال	تعداد نرون‌ها در لایه پنهان	مرحله آموزش		مرحله آزمون	
		R^2	RMSE	R^2	RMSE
Tanh	۳	۰/۸۱۱	۰/۰۱۹	۰/۸۷۴	۰/۰۲۱
Tanh	۶	۰/۹۱۰	۰/۰۱۳	۰/۹۰۶	۰/۰۱۶
Tanh	۹	۰/۸۸۳	۰/۰۱۵	۰/۸۹۸	۰/۰۲۱
Logistic	۳	۰/۸۸۷	۰/۰۱۸	۰/۸۱۲	۰/۰۲۳
Logistic	۶	۰/۹۱۹	۰/۰۱۲	۰/۹۱۶	۰/۰۲۱
Logistic	۹	۰/۹۱۲	۰/۰۱۲	۰/۸۷۶	۰/۰۲۲
Exponential	۳	۰/۷۴۰	۰/۰۲۳	۰/۸۶۹	۰/۰۲۰
Exponential	۶	۰/۸۸۵	۰/۰۱۴	۰/۸۵۲	۰/۰۲۱
Exponential	۹	۰/۸۷۶	۰/۰۱۵	۰/۸۲۴	۰/۰۲۳



شکل (۷): نمودار پراکندگی داده‌ها در مرحله آزمون برای مدل ANN(MLP)

(y)، ضرب ($x \times y$)، تقسیم (x/y)، جذر (\sqrt{x})، نمایی ($\exp(x)$)، لگاریتم طبیعی ($\ln(x)$) و درجه دو (x^2) انتخاب گردید و سپس با انتخاب بهترین برازش^۸ نتایج مدل GEP استخراج گردید. مقادیر R^2 و RMSE برای این حالت در مرحله آموزش به ترتیب $0/834$ و $0/017$ در مرحله آزمون به ترتیب $0/789$ و $0/025$ می‌باشد. شکل (۸) پراکندگی داده‌ها در مرحله آزمون برای مدل GEP را نشان می‌دهد.

نتایج مدل برنامه‌ریزی بیان ژن

در مدل برنامه‌ریزی بیان ژن پارامترهای مختلفی وجود دارد که در دقت تخمین موثر واقع می‌گردند. در تحقیق حاضر اندازه سر^۱ برابر ۸، تعداد کروموزمها^۲ برابر با ۳۰، تعداد ژن‌ها^۳ در هر کروموزم برابر با ۳، نرخ جهش^۴ برابر با $0/044$ ، معیار خطای برازش^۵ RMSE و تابع پیوند^۶ جمع^۷ و توابع انتخابی از نوع جمع ($x+y$)، تفریق ($x-$)

⁵ Fitness Function Error Type

⁶ Linking Function

⁷ Addition

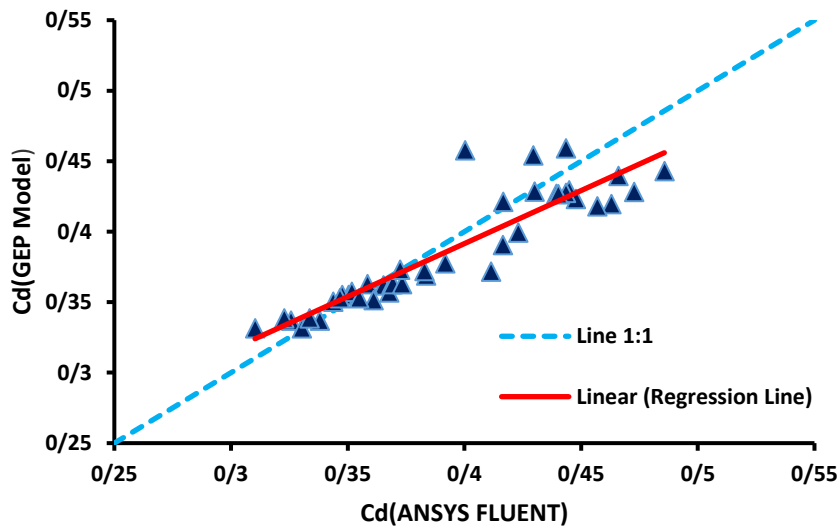
⁸ Maximum Fitness

¹ Head Size

² Chromosomes

³ Genes

⁴ Mutation Rate



شکل (۸): نمودار پراکندگی داده‌ها در مرحله آزمون برای مدل GEP

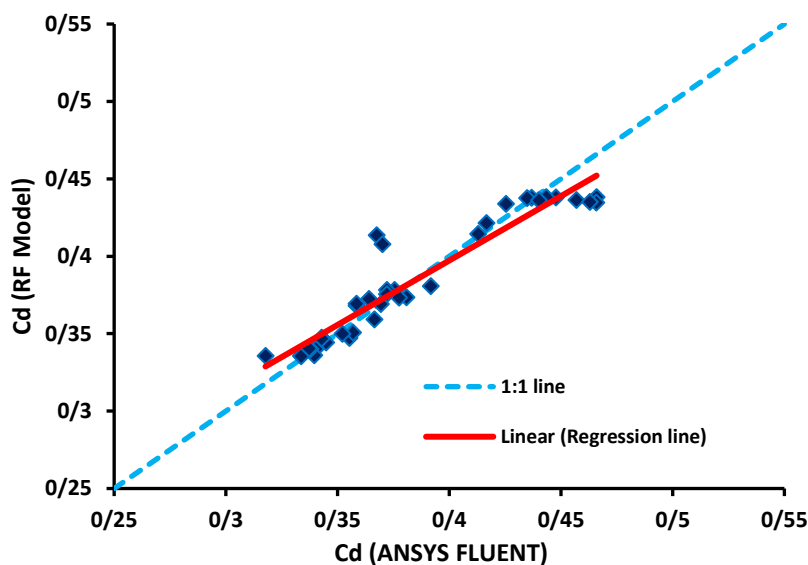
افزایش تعداد درخت هیچ تاثیری در دقت نتایج نداشته است. بنابراین مدل RF با تعداد درخت ۵۰۰ با داشتن مقادیر R^2 و RMSE به ترتیب ۰/۸۹۸ و ۰/۰۱۳ از دقت نسبتاً خوبی در مقایسه با سایر مدل‌های مورد بررسی با تعداد درخت کمتر را دارا می‌باشد. شکل (۹) پراکندگی داده‌ها در مرحله آزمون برای بهترین حالت مدل RF را نشان می‌دهد.

نتایج مدل جنگل تصادفی

برای تخمین ضریب دبی جریان (C_d) در سرریزهای لبه پهن با مدل جنگل تصادفی (RF) چندین حالت با تعداد درخت‌های مختلف مورد بررسی قرار گرفت که نتایج هر یک از این حالت‌ها برای مرحله آموزش و آزمون در جدول (۴) آورده شده است. همان‌طور که مشاهده می‌شود، دقت مدل RF از تعداد درخت ۵۰۰ به بالا تقریباً ثابت می‌ماند و

جدول (۴): تخمین ضریب دبی جریان با مدل RF با تعداد درخت‌های مختلف

تعداد درخت‌ها	مرحله آموزش		مرحله آزمون	
	R^2	RMSE	R^2	RMSE
۱۰۰	۰/۸۰۲	۰/۰۲۱	۰/۸۹۳	۰/۰۱۴
۳۰۰	۰/۸۰۸	۰/۰۲۰	۰/۸۹۵	۰/۰۱۳۵
۵۰۰	۰/۸۰۹	۰/۰۲۰	۰/۸۹۸	۰/۰۱۳
۷۵۰	۰/۸۰۹	۰/۰۲۰	۰/۸۹۸	۰/۰۱۳
۱۰۰۰	۰/۸۰۸	۰/۰۲۰	۰/۸۹۸	۰/۰۱۳

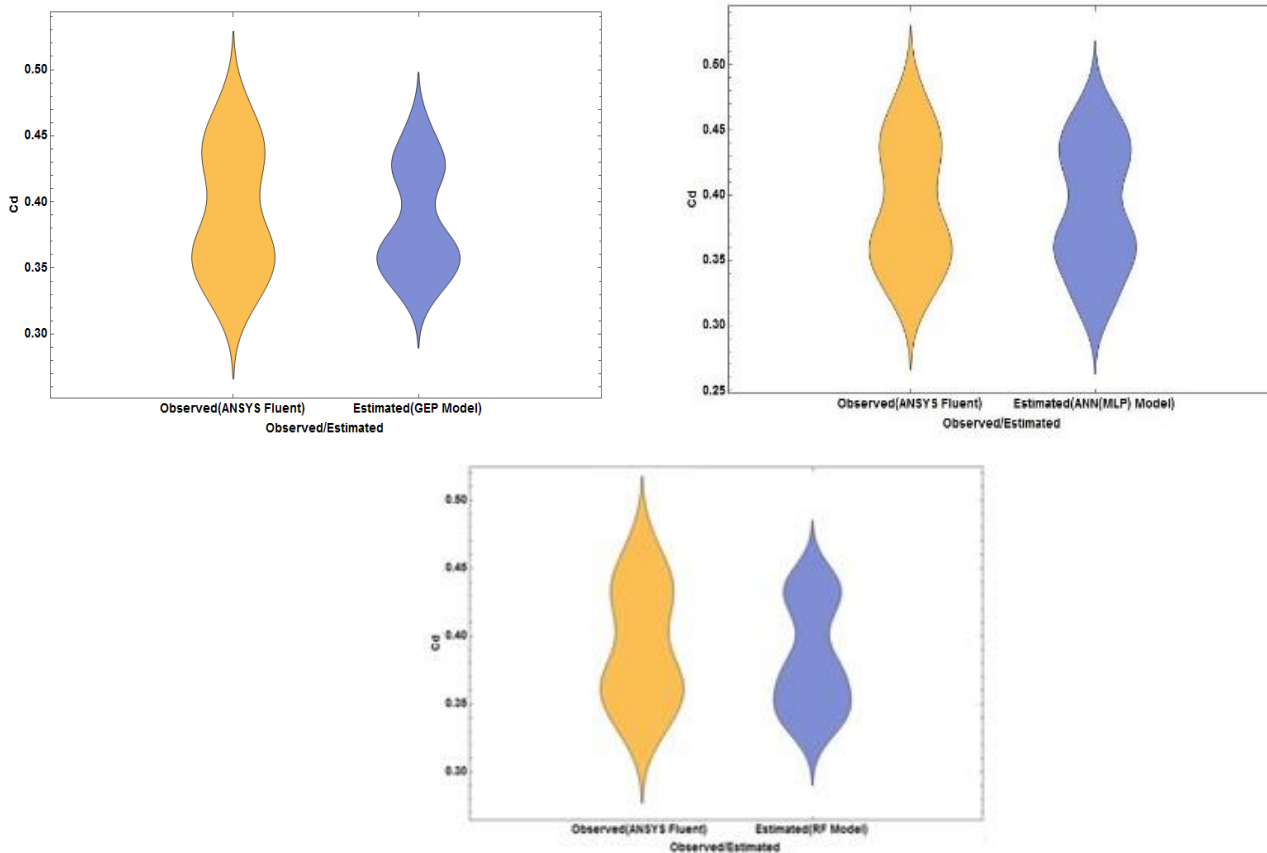


شکل (۹): نمودار پراکندگی داده‌ها در مرحله آزمون برای مدل RF

(۷۵٪)، میانه (۰/۵۰)، چارک اول (۰/۲۵) و حداقل در جدول نشان داده شده است. با مقایسه نمودارهای ویولنی برای هر سه مدل هوشمند و نتایج آن در جدول (۵) مشاهده می‌گردد که نمودار ویولنی مدل ANN(MLP) تطابق بسیار نزدیکی با داده‌های مدل عددی (ANSYS FLUENT) را دارد.

در شکل ۱۰ نمودار ویولنی^۱ داده‌ها در مرحله آزمون برای داده‌های شبیه‌سازی شده با مدل عددی و تخمین زده شده با مدل‌های هوشمند ANN(MLP)، GEP و RF نشان داده شده است. در جدول برای داده‌های شبیه‌سازی شده با مدل عددی و تخمینی با روش‌های هوشمند فوق‌الذکر مقادیر ضریب دبی جریان حداکثر، چارک سوم

¹ Violin plot



شکل (۱۰): نمودار ویولنی در مرحله آزمون برای داده‌های مشاهداتی و تخمینی با استفاده از مدل‌های ANN(MLP)، GEP و RF

جدول (۵): مقادیر شاخص‌های نمودار ویولنی داده‌ها در مرحله آزمون

	ANSYS Fluent	ANN(MLP)	RF	GEP
حداکثر	۰/۴۸۵	۰/۴۷۴	۰/۴۳۳	۰/۴۵۸
۷۵%	۰/۴۳۹	۰/۴۳۱	۰/۴۳۸	۰/۴۲۶
متوسط	۰/۳۸۳	۰/۳۷۹	۰/۳۷۳	۰/۳۷۱
۲۵%	۰/۳۵۱	۰/۳۵۴	۰/۳۴۷	۰/۳۵۵
حداقل	۰/۳۱۰	۰/۳۱۱	۰/۳۳۵	۰/۳۳۱

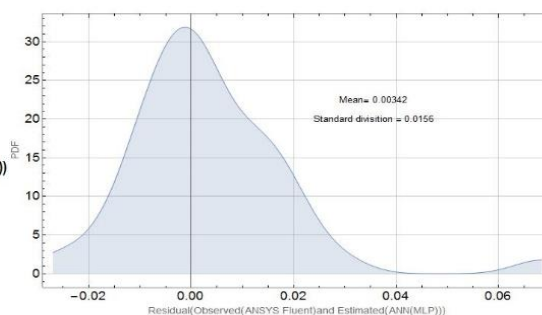
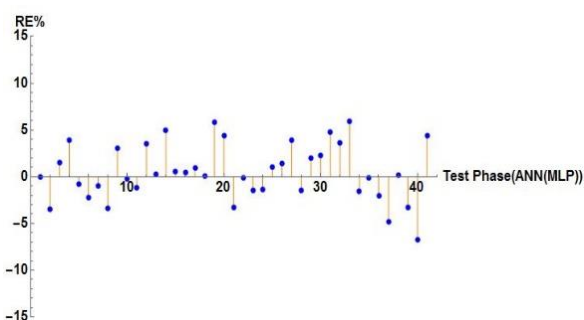
این نشان می‌دهد که ANN(MLP) با خطای کمتر یا مساوی ۵٪ می‌تواند ضریب دبی جریان را برآورد نماید. لازم به ذکر است که علامت و منفی و مثبت در مقادیر درصد خطای نسبی به معنای تخمین بیشتر و کمتر هر یک از مدل‌های هوشمند نسبت به مقادیر مشاهداتی را نشان می‌دهد. همچنین مشاهده می‌گردد که مدل GEP نسبت به مدل RF در اکثر داده‌های مرحله آزمون خطای نسبی بیش

در شکل ۱۱ نمودار تغییرات درصد خطای نسبی (RE%) و تابع توزیع احتمال باقی مانده‌ها (PDF)^۱ در مرحله آزمون برای هر یکی از مدل‌های هوشمند ANN(MLP)، GEP و RF نشان داده شده است. همان‌طور که مشاهده می‌گردد محدوده RE% در مدل ANN(MLP) $-5 \leq RE\% \leq +5$ در حالی که در مدل GEP و RF در خارج از این محدوده می‌باشد و

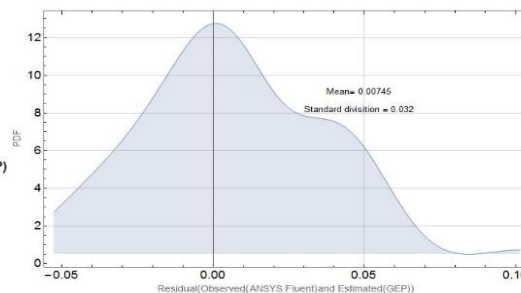
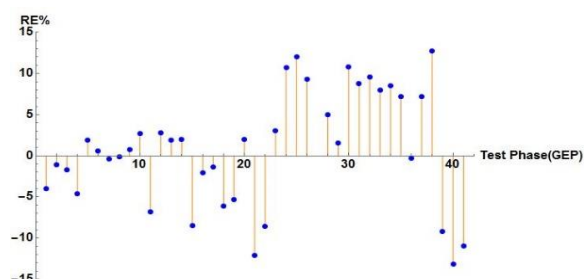
¹ Probability Distribution Function

ANN(MLP) مقدار میانگین و انحراف معیار به ترتیب ۰/۰۳۴ و ۰/۰۱۵۶ می‌باشد. در این مدل بر خلاف سایر مدل‌ها تابع توزیع احتمال به منحنی توزیع نرمال نزدیکتر است و توزیع مقادیر باقی‌مانده تقریباً به صورت متقارن می‌باشد.

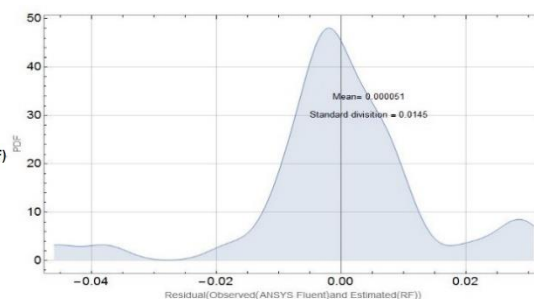
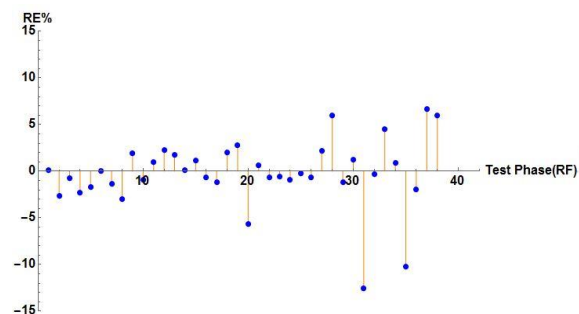
از $\pm 5\%$ درصد را دارد، بنابراین این مدل نسبت به مدل GEP از دقت بالاتری در تخمین ضریب دبی جریان را دارد. لازم به ذکر است که برای هر یک از مدل‌های فوق‌الذکر نمودار تابع توزیع احتمال باقی‌مانده‌ها نیز نشان داده شده است. همان‌طور که مشاهده می‌گردد در مدل



الف) برای مدل ANN(MLP)



ب) برای مدل GEP



ج) برای مدل RF

شکل (۱۱): نمودار تغییرات درصد خطای نسبی و تابع توزیع احتمال باقی‌مانده‌ها در مرحله آزمون

بنابراین می‌توان گفت که انتخاب یک ساختار بهینه با تابع محرک مناسب باعث کاهش خطا و افزایش دقت مدل تخمین‌کننده فوق‌الذکر شده است و می‌توان گفت شبکه عصبی مصنوعی نتایج بسیار مطلوبی نسبت به سایر مدل‌ها در تخمین ضریب دبی جریان را دارد و نتایج آن به نتایج مشاهداتی (استخراجی از مدل عددی در تحقیق حاضر) نزدیکتر می‌باشد.

در جدول (۶) معیارهای آماری برای هر یک از مدل‌های به کار رفته در تحقیق حاضر برای مرحله آزمون به صورت یکجا آورده شده است. همان‌طور که مشاهده می‌گردد مقدار KGE برای هر سه مدل هوشمند در محدوده خیلی خوب یعنی $(KGE \leq 0.17)$ قرار دارد، بنابراین از کیفیت خوبی در تخمین پارامتر مورد نظر یعنی ضریب دبی جریان برخوردار می‌باشند، اما مدل ANN(MLP) نسبت به مدل‌های GEP و RF از دقت نسبتاً بالایی برخوردار می‌باشد.

جدول (۶): مقایسه نتایج مدل ANN، GEP و RF در مرحله آزمون

مدل	R ²	RMSE	KGE
ANN	۰/۹۰۶	۰/۱۶	۰/۹۲۷
GEP	۰/۷۹۰	۰/۰۲۵	۰/۷۸۰
RF	۰/۸۹۸	۰/۰۱۳	۰/۸۴۱

نتیجه گیری

ارزیابی نتایج به دست آمده با استفاده از معیارهای آماری R^2 ، RMSE، و KGE و همچنین دیاگرام‌های گرافیکی انجام گردید. این مقادیر برای مدل ANN(MLP) ($R^2=0.906$ ، $RMSE=0.16$ و $KGE=0.927$)، برای مدل GEP ($R^2=0.790$ ، $RMSE=0.025$ و $KGE=0.780$) و برای مدل RF ($R^2=0.898$ ، $RMSE=0.013$ و $KGE=0.841$) به دست آمد. نتایج معیارهای آماری، برتری ANN(MLP) نسبت به سایر مدل‌ها را نشان می‌دهد، محدوده درصد خطای نسبی (RE%) برای مدل برتر کمتر از حدود $\pm 5\%$ حاصل گردید. دیاگرام ویولنی نشان داد که توزیع احتمال داده‌های تخمینی با مدل ANN(MLP) با داده‌های حاصل از روش حجم محدود (FVM) تطابق بسیار نزدیک و مشابهی دارد. همچنین نمودار تابع توزیع احتمال باقی‌مانده‌ها در مدل ANN(MLP) نسبت به سایر مدل‌های مورد بررسی متقارن‌تر و به توزیع نرمال نزدیکتر می‌باشد.

در این تحقیق به عملکرد مدل‌های هوشمند ANN و GEP و RF در تخمین ضریب دبی جریان (C_d) در سرریزهای لبه پهن با استفاده از پارامترهای h/P ، h/L_r ، β_1 ، β_2 و θ پرداخته شد. برای این منظور از داده‌های شبیه‌سازی شده به روش حجم محدود با به کارگیری نرم‌افزار ANSYS FLUENT استفاده گردید. برای رسیدن به برنامه و ساختاری که بتواند روابط بین ورودی‌ها و خروجی مدل را تعمیم دهد، لازم است برنامه آموزش داده شود. بدین منظور پارامترهای ورودی و خروجی به دو دسته تقسیم‌بندی گردید. در کل مدل‌های هوشمند پارامترهای h/P ، h/L_r ، β_1 ، β_2 و θ به عنوان ورودی و پارامتر C_d یعنی به عنوان خروجی معرفی گردید و همچنین در همه مدل‌ها ۷۵ درصد کل داده‌های موجود (۱۲۳ عدد) برای مرحله آموزش و ۲۵ درصد آن‌ها (۴۱ عدد) نیز برای مرحله آزمون مورد انتخاب شد.

منابع

دهقانی، ر.، قربانی، م. ع.، تشنه‌لب، م.، ریخته‌گر غیاثی، ا.، و اسدی، ا. ۱۳۹۴. مقایسه و ارزیابی مدل‌های شبکه عصبی بیزین، برنامه‌ریزی بیان ژن، ماشین بردار پشتیبان و رگرسیون خطی در تخمین بده جریان؛ مطالعه موردی حوضه صوفی چای. نشریه آبیاری و آب ایران. مرحله ۵. شماره ۴، ۸۵-۶۵.

عباسپور، ا.، عبدالله‌پور، م.، و سلماسی، ف. ۱۳۹۲. شبیه‌سازی عددی جریان بر روی سرریزهای لبه پهن مستطیلی با وجوه شیب‌دار بالادست و پایین‌دست بوسیله نرم‌افزار فلوئنت. دانش آب و خاک. جلد ۲۳. شماره ۴، ۲۶۵-۲۷۶.

فروودی، ج.، گودرزی، ا.، و شکری، ن. ۱۳۸۶. بررسی تاثیر شیب‌دار کردن وجه بالادست سرریزهای لبه‌پهن مستطیلی در ضریب تخلیه و مشخصات جریان، ششمین کنفرانس هیدرولیک ایران. شهرکرد. انجمن هیدرولیک ایران. دانشگاه شهرکرد.

فولادی‌پناه، م.، ماجدی اصل، م.، و حق‌گوپی، ا. ۱۳۹۹. کاربرد الگوریتم‌های هوشمند برای مدل‌سازی رابطه دبی-اشل در شرایط استغراق سرریزهای کنگره‌ای و خطی، نشریه علمی پژوهشی هیدرولیک. دوره ۱۵. شماره ۲، ۱۴۹-۱۶۴.

نورانی، ب.، نوروزی، ر.، رضایی، ف.، و سلماسی، ف. ۱۳۹۸. ارزیابی عملکرد مدل برنامه‌ریزی بیان ژن و شبکه عصبی مصنوعی در برآورد پارامترهای هیدرولیکی سرریز پلان مثلثی. نشریه آبیاری و آب ایران. دوره ۱۲. شماره ۱، ۱-۱۷.

ورجاوند، پ.، فرسادی زاده، د.، و خسروی‌نیا، پ.، رفیعی، ز. ۱۳۸۹. شبیه‌سازی جریان بر روی سرریز استوانه‌ای با استفاده از مدل فلوئنت و مقایسه نتایج با مدل فیزیکی. دانش آب و خاک، دوره ۲۰، شماره ۲، ۵۹-۶۹.

Aksoy, A., and Doğan, M. 2016. Experimental investigation of the approach angle effects on the discharge efficiency for broad crested weirs. *Anadolu University Journal of Science and Technology Applied Sciences and Engineering*, 17 (2): 279 – 286.

Azimi, A H., Rajaratnam, N., and David Z Zhu, M. 2013. Discharge Characteristics of Weirs of Finite Crest Length with Upstream and Downstream Ramps. *J. Irrig. Drain Eng.*, 139(1): 75-83.

Azimi, A. H., and Rajaratnam, N. 2009. Discharge characteristics of weirs of finite crest length. *J. Hydraul. Eng.*, 135(12): 1081– 1085.

Bayon A and Valero D. 2016. Performance assessment of OpenFOAM and FLOW-3D in the numerical modeling of a low Reynolds number hydraulic jump. *Environ. Model. Softw.*, 80, 322–335.

Bos, M G., 1989. Discharge Measurements structures. *J. Irrig. Drain Eng.*, 120(1): 13-26.

Breiman, L. 2001. Random Forests. *Machine Learning*, 45(1): 5–32.

Bowden, G.J., Dandy, G.C., and Maier, H.R. 2005. Input determination for neural network models in water resources applications. Part 1 – background and methodology. *Journal of Hydrology*, 301(1): 75–92.

Goguse, M., Defne, Z., and Ozkandemir, V. 2006. Discharge Broad crested weirs with rectangular compound cross sections. *J. Irrig. Drain Eng.*, 132(3): 272-280.

Hager, H., and Schwalt. M. 1989. Broad crested weir. Intl. Institute for Land Reclamation and Improvement, Wageningen NL.

Hager, H., and Fritz, M. 1998. Hydraulics of Embankment weirs. *J. Hydra Eng. ASCE*, 124(9).

Horton, R.E., and Murphy, E.C. 1906. Weir experiments, coefficients, and formulas; US government printing Office: Washington, DC, USA.

Hirt, C.W., and Nichols B.D. 1981. Volume of fluid (VOF) method for the dynamics of free boundaries. *J. Comput. Phys*, 39: 201–225.

Kling, H., Fuchs, M., and Paulin, M. 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.*, 424-425: 264–277.

Lei, J., Mingjun D., Haomiao, S., Yu, R. 2018. Numerical modeling of flow over a rectangular broad-crested weir with a sloped upstream face, *Water*, 10 (11), 1663; doi: 10.3390/w10111663.



- Madadi, M.R., Farsadizadeh, D., and Hosseinzadeh Dalir, A. 2014. Investigation of flow characteristics above trapezoidal broad-crested weirs, *Flow Measurement and Instrumentation* 38: 139 – 148.
- Norouzi, R., Daneshfaraz, R., and Ghaderi, A. 2019. Investigation of discharge coefficient of trapezoidal labyrinth weirs using artificial neural networks and support vector machines. *Applied Water Science*, 9: 148.
- Roushangar, K., Alami, M.T., Shiri, J., and Majedi Asl, M. 2018. Determining discharge coefficient of labyrinth and arced labyrinth weirs using support vector machine. *Hydrology Research*, 49 (3): 924–938.
- Roushangar, K., Alami, M.T Majedi Asl, M., and Shiri, J. 2017. Modeling discharge coefficient of normal and inverted orientation labyrinth weirs using machine learning techniques, *ISH Journal of Hydraulic Engineering*, 23(3): 331-340.
- Sarker, M.A., and Rhodes, D G. 2004. Calculation of free surface profile over a rectangular broad-crested weirs, *Flow Measurement and Instrumentation*, 15: 215 – 219.
- Sargison, J.E., and Percy, A. 2009. Hydraulics of broad-crested weirs with varying side slopes, *Journal of irrigation and drainage engineering*, 135(1): 115-118.
- Simsek, O., Sami Akoz, M., and Goksu Soydan, N. 2016. Numerical validation of open channel flow over a curvilinear broad-crested weir, *Progress in Computational Fluid Dynamics*, 16(6): 364-378.
- Toro, J.P., Bombardelli, F.A., and Paik, J. 2017. Detached Eddy Simulation of the No narrated Skim Flow over a Stepped Spillway. *J. Hydraul. Eng.*, 143(9): 04017032.



Investigation and Comparison Performance Three Data Mining Techniques ANN, GEP and RF in Estimating Flow Discharge Coefficients of Trapezoidal Broad-Crested Weirs

Bahram Nourani¹, Hadi Arvanaghi^{2*}, Farzin Salmasi³

Abstract

Accurate determination of the discharge coefficient plays an important role in estimating the discharge of the weirs. As a result, it is significant to estimate the discharge coefficient correctly. The purpose of current study is simulation and estimation the discharge coefficients of the broad-crested weirs. Hence, numerical simulation of hydraulic characteristics of these weirs were performed by ANSYS FLUENT software and results were obtained. Then, three intelligent models of ANN (MLP), GEP and RF were used to determine the discharge coefficients and the results of these models were compared. Assessment of the results were carried out using the statistical indices as R^2 , RMSE, KGE, and graphical diagrams. The amounts of R^2 , RMSE and KGE for ANN (MLP) 0.906, 0.016, 0.927, for GEP 0.790, 0.025, and 0.780, and for RF 0.898, 0.013, and 0.841 were obtained, respectively. Due to the statistical criteria results and the range of RE% ($\mp 5\%$), the ANN (MLP) was selected as the superior model. Violin plot illustrates a great and close agreement between the probability distribution of estimated data with ANN (MLP) and the obtained data from the numerical simulation using the finite volume method. Furthermore, the probability density function plot of residuals in ANN (MLP) model was closer to the normal distribution function, compared with the other models in this study.

Keywords: Broad-crested weirs, Intelligence models, Discharge coefficient, Finite volume method, ANSYS FLUENT.

¹ Ph. D Candidate of Hydraulic Structures, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran, Email: nourani.bahram@tabrizu.ac.ir, Tel: 09147822619

² Associated Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran, Email: arvanaghi.hadi@gmail.com, Tel: 09144195085 (*Corresponding Author)

³ Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran, Email: salmasi@tabrizu.ac.ir, Tel: 09143178332