

EXTENDED ABSTRACT

Predictive Modeling of Reservoir Inflow for Multi-Step Ahead to Improve Management of Alavian Dam Reservoir

Bagher Nikofar^a, Vahid Nourani^{b,*}, Aida Hosseini Baghanam^b

^a Ph.D. in Civil Engineering-Water Resources Management, Faculty of Civil Engineering, University of Tabriz, 5166616471, Iran

^b Faculty of Civil Engineering, University of Tabriz, Tabriz 5166616471, Iran

Received: 17 December 2023; **Reviewed:** 21 May 2024; **Accepted:** 06 July 2024

Keywords:

Inflow forecasting, Multi-step modeling, ANN, Weka software, Alavian dam reservoir.

1. Introduction

Accurate reservoir inflow forecasting is essential for efficient water resources management, particularly under increasing pressure on available water supplies. This study evaluated the performance of an Artificial Neural Network (ANN) and several data-mining models developed in the Weka environment for forecasting inflow to the Alavian Dam reservoir up to 12 months ahead. Monthly data from 1997 to 2022 were used to build the models. The input variables included reservoir inflow, temperature, evaporation, precipitation, basin snow cover, and drought indices. Model performance was assessed during the validation phase and compared across different forecasting horizons. The results showed that all models provided acceptable and reliable predictions. However, ANN, RF, and RT outperformed the other models in overall accuracy. Their average correlation coefficients were 0.85, 0.90, and 0.83, respectively. These models also showed strong agreement with observed data, especially in reproducing peak inflow values. The findings demonstrate the effectiveness of AI-based approaches for multi-step reservoir inflow prediction.

2. Methodology

2.1. Study area and data

The dataset utilized in this study spans 25 years (from 1997 to 2022), comprising information on the inflow to the Alaviyan dam reservoir and additional parameters such as precipitation, flow in hydrometric stations, temperature, evaporation, and snow cover within the dam's catchment area. To incorporate the impact of drought in the model, the Standardized Precipitation Index a drought parameter has been computed based on the previous months' rainfall and integrated into the model's input matrix.

2.2. Methodology

This study, in its initial phase, the collected time-series data were reconstructed, categorized, and the most effective input variables were selected using correlation analysis in WEKA. Due to the nonlinear and uncertain nature of runoff series, different input combinations were tested and the best-performing ones were retained.

Publisher: University of Tabriz

<https://doi.org/10.22034/CEEJ.2024.59627.230>

* Orcid Code Corresponding Author: 0000-0002-6931-7060

E-mail addresses: baghernik@yahoo.com (Bagher Nikofar), nourani@tabrizu.ac.ir (Vahid Nourani), hosseinibaghanam@tabrizu.ac.ir (Aida Hosseini Baghanam).



Online ISSN: 2717-4077

In the second step, GP, SM, RF, RT models and an Artificial Neural Network (ANN) were trained to predict monthly inflow to the reservoir. The dataset was divided into 75% for training and 25% for validation, with the first 18 years used for training and the last 7 years for testing. The backpropagation algorithm was used to estimate the network weights. The model was developed separately for short-term and long-term monthly forecasting, using lagged inputs from previous months and the corresponding month of the previous year. Snow cover was also included as a percentage of the total catchment area, making the model more comprehensive for inflow prediction.

3. Results and discussion

To evaluate model performance, ANN, GP, SM, RF, and RT models were compared against observed monthly inflow using correlation coefficients and RMSE. Analysis of the validation period (Figs. 1-2) indicates that ANN, RF, and RT models consistently outperform others, demonstrating high accuracy in predicting peak flows. Specifically, these models show strong correlation and lower error rates over the 12-month forecast horizon, particularly during the first three months of the water year. Overall, the integration of ANN and decision tree architectures significantly enhances the reliability of inflow forecasting for the Alaviyan dam reservoir.

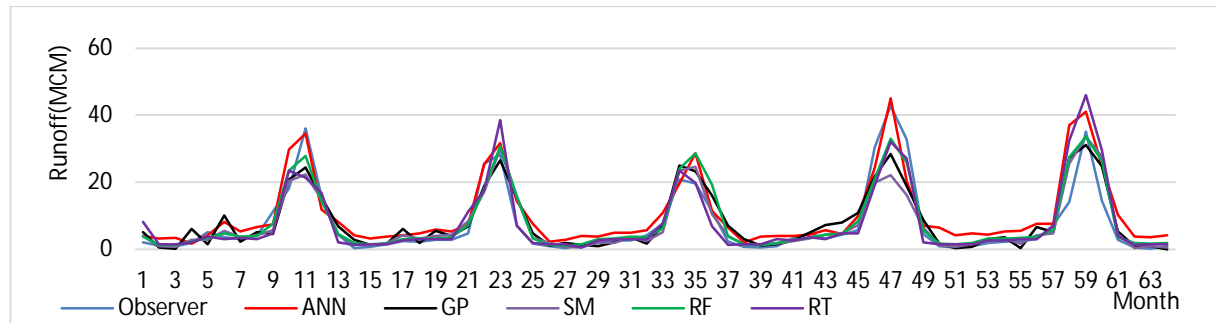


Fig. 1. Time series graph of observed runoff compared to ANN, GP, SM, RF, and RT calculations for validation data

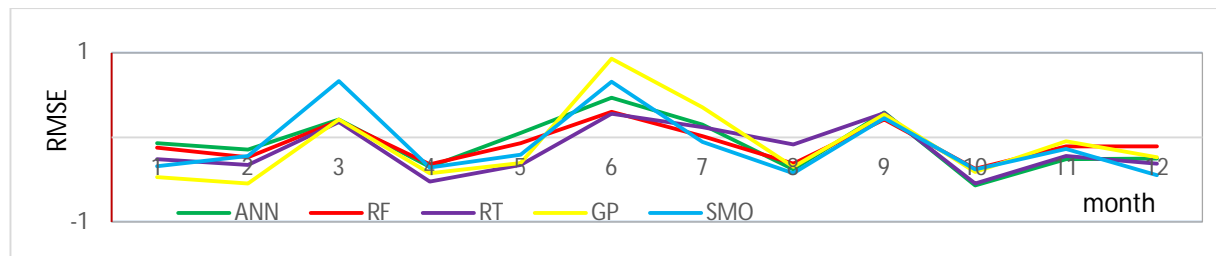


Fig. 2. Comparison of normalized root mean square error of monthly time series computational input modeling results for five models in the validation phase

4. Conclusions

This research focused on predicting monthly inflow to the Alaviyan dam over a 12-month period using various models. The Artificial Neural Network (ANN) model demonstrated strong predictive capabilities, achieving an average correlation of 85%. Decision tree models, specifically Random Forest (RF) and Random Tree (RT), showed superior performance with correlations of 90% and 83%, respectively, outperforming GP and SM models. The ANN's effectiveness stems from its nonlinear FFNN structure, which accurately captures the complex, uncertain dynamics of reservoir inflow. The RF model also proved highly adaptive, maintaining predictive accuracy even with data gaps. Both ANN and RF/RT models are suitable for forecasting Alaviyan dam inflow. While all models performed well for short-term predictions (1-3 months), forecast accuracy decreased for longer horizons (6-12 months). Temperature and precipitation were identified as critical input parameters.

مدل‌سازی پیش‌بینی جریان ورودی به مخزن سد در 12 ماه آتی برای بهبود مدیریت مخزن سد علویان

باقر نیکوفر¹، وحید نورانی^{2*}، آیدا حسینی بقانام³

¹ دانشجوی دکتری مهندسی عمران - مدیریت منابع آب، دانشگاه تبریز

² استاد گروه مهندسی آب، دانشکده مهندسی عمران، دانشگاه تبریز

³ استادیار گروه مهندسی آب، دانشکده مهندسی عمران، دانشگاه تبریز

دریافت: 1402/9/26، بازنگری: 1403/3/1، پذیرش: 1403/4/16، نشر آنلاین: 1403/4/16

چکیده

با توجه به نقش اساسی آب در زندگی بشر و محدودیت‌های منابع آب، برآورد دقیق میزان پتانسیل منابع آب و استفاده بهینه از آن، یک مسئله مهم و قابل توجه در محافل علمی و صنعت آب است. پیش‌بینی چند ماه آینده جریان ورودی به مخزن سدها موجب بهبود مدیریت بهره‌برداری از مخزن سدها می‌گردد؛ بنابراین پیش‌بینی جریان آب ورودی به مخازن سدها با دقت بالا، با استفاده از مدل‌های مبتنی بر هوش مصنوعی و روش‌های نوین روز، ضروری می‌باشد. در این مقاله از توانایی مدل شبکه عصبی مصنوعی (Artificial Neural Network (ANN)) با انتشار برگشتی خطا و از سایر مدل‌های مبتنی بر داده‌کاوی در بستر نرم‌افزار WEKA (Waikato Environment for Knowledge Analysis) در مدل‌سازی پیش‌بینی چند گام بعد (تا 12 ماه بعد) با به‌کارگیری ورودی‌های مشاهداتی استفاده شده است. در انتها، کارایی همه مدل‌ها با استفاده از معیارهای ارزیابی بررسی شده و مدل‌ها با یکدیگر مقایسه شده‌اند. هدف این مقاله، پیش‌بینی چند ماه آینده جریان ورودی به مخزن سد علویان می‌باشد. در این تحقیق، ورودی‌های مدل شامل رواناب ورودی به مخزن سد، دما، تبخیر، بارش و پوشش برف حوضه آبریز سد و شاخص‌های خشک‌سالی می‌باشد و این داده‌ها به‌صورت ماهانه و به‌مدت 25 سال (1376-1401) استفاده گردیده است. مقایسه نتایج مدل ANN با سایر مدل‌های RF، RT، GP و SM در مرحله صحت‌سنجی، نشان‌دهنده این است که در همه گام‌ها، نتایج پیش‌بینی برای همه مدل‌ها مناسب بوده است و از بین 5 مدل که نتایج آن‌ها باهم مقایسه گردیده است، نتایج متوسط ضریب همبستگی مدل‌های ANN و RF و RT به ترتیب 85%، 90% و 83% در مقایسه با سایر مدل‌ها در پیش‌بینی جریان ورودی به مخزن سد علویان در چند گام بعد (1 تا 12 ماه) مناسب و نزدیک به داده‌های مشاهداتی به‌ویژه در نقاط پیک بوده است.

کلیدواژه‌ها: پیش‌بینی جریان ورودی، مدل‌سازی چند گام، ANN، نرم‌افزار WEKA، مخزن سد علویان.

1- مقدمه

دمای حداقل و حداکثر، بارش و پوشش برف و ... و همچنین رهاسازی آب از مخزن سدها برای تأمین نیازهای پایین‌دست می‌باشد. در حل مسئله بهره‌برداری بهینه یک سد، هرگز نمی‌توان کاملاً از حجم آب ورودی به سد در ماه‌های آتی باخبر و مطمئن شد. بنابراین اولین پیچیدگی این مسئله، نامشخص بودن مقدار آورد جریان ورودی به مخزن سد در آینده است؛ به‌ویژه در سیاست‌های بهره‌برداری در سطح وزارت نیرو، برنامه‌ریزی برای یک سال آبی آینده مدنظر می‌باشد. لذا ضروری است در یک دوره 12 ماهه آینده (یک سال آبی) میزان جریان ورودی به مخزن

با توجه به نقش اساسی آب در زندگی بشر و محدودیت‌های منابع آب، برآورد دقیق میزان پتانسیل منابع آب و استفاده بهینه از آن یک مسئله مهم و قابل توجه در محافل علمی و صنعت آب است. بنابراین پیش‌بینی جریان آب ورودی به مخازن سدها با دقت بالا و بهینه‌سازی منابع آب مخازن سدها در مدیریت منابع آب یک راه‌کار اصولی و ضروری است. یکی از عوامل مهم تأثیرگذار در نحوه مدیریت و بهره‌برداری از مخزن سدها، مجموعه داده‌های هیدرولوژیکی مثل ورودی جریان آب به مخازن سدها، تبخیر،



در سال‌های اخیر، مطالعات زیادی در داخل و خارج کشور جهت پیش‌بینی جریان، به‌ویژه در جریان ورودی به مخازن سدها با بهره‌گیری از هوش مصنوعی صورت گرفته است که این موضوع نشانگر اهمیت موضوع پیش‌بینی جریان ورودی به مخازن سدها هست.

به‌عنوان مثال Sanikhani و همکاران (2015)، Ibrahimi Mohammadi و Seghale (2011)، Kisi (2004) و Lin و همکاران (2006) از هوش مصنوعی مثل شبکه‌های عصبی مصنوعی نسبت به پیش‌بینی جریان ورودی به‌صورت ماهانه استفاده کردند و نتایج حاکی از دقت بالاتر شبکه‌های عصبی مصنوعی نسبت به مدل‌های دیگری مثل رگرسیون و ... بوده است. Sattari و همکاران (2012) کارایی شبکه‌های عصبی بازگشتی با تأخیر زمانی و شبکه‌های عصبی با الگوریتم پس انتشار خطا را در پیش‌بینی جریان ورودی به مخزن سد علویان را مورد بررسی قرار دادند. نتایج به‌دست‌آمده نشان داد که با توجه به شاخص‌های آماری خطا، هر دو مدل توانمندی قابل‌قبولی در پیش‌بینی مقادیر واقعی جریان دارند.

با توجه به این‌که هدف این تحقیق پیش‌بینی جریان ورودی تا 12 گام بعد می‌باشد، در این ارتباط ظرفیت و توانایی بالای هوش مصنوعی در پیش‌بینی یک گام بعد در بسیاری از زمینه‌ها اثبات شده است و در مقالات متعددی از کاربرد موفق ANN در مدل‌سازی فرآیندهای هیدرولوژیکی و منابع آب برای یک گام بعد ارائه شده است. به‌عنوان مثال می‌توان به مقالات Kim و Valdés (2003)، Antar و همکاران (2006)، Dinu و همکاران (2017)، Sharghi و همکاران (2019)، Mohammadi و همکاران (2020) و Nourani و همکاران (2021) اشاره کرد. اما در زمینه پیش‌بینی جریان ورودی به مخازن سدها در چند گام بعد (12 ماه بعد) کمتر مورد توجه قرار گرفته است و Najafi و همکاران (2022) طی مقاله‌ای توانایی مدل ANN با انتشار برگشتی خطا و همچنین ترکیب مدل ANN و تبدیل موجک (WANN) در مدل‌سازی فصلانه فرآیندهای هیدروکلیماتولوژیکی برای پیش‌بینی چند گام بعد (یک تا دوازده گام) با به‌کارگیری کمترین ورودی بررسی نموده‌اند.

همچنین Mohammadi Klaleg و همکاران (2015) طی مطالعه‌ای با استفاده از مدل درختی M5 در مقایسه با روش شبکه‌های عصبی مصنوعی اقدام به پیش‌بینی جریان ورودی به سد سبلان در افق زمانی 12 ماه آتی نمودند، نتایج به‌دست آمده نشان داد، مدل درختی M5 ضمن داشتن نتایج نسبتاً دقیق و با ارائه روابط خطی ساده و قابل‌فهم‌تر جنبه کاربردی بیشتری دارد.

سدها با دقت بالا پیش‌بینی گردد. با علم بر این‌که پارامترهای هیدرولوژیکی و آب و هوایی منطقه در آینده از جمله بارش، تبخیر، دما، جریان آب در ایستگاه‌های هیدرومتری و جریان آب ورودی به مخزن سد و ... در سطح حوضه آبریز نامشخص است. بنابراین به‌دلیل عدم قطعیت پارامترهای مذکور، پیش‌بینی جریان‌های ورودی به مخزن سد و بهینه‌سازی را برای عوامل بهره‌بردار سدها دشوار می‌سازد. حال اگر علم هیدرولوژی و پیش‌بینی‌های هواشناسی به‌کمک بهره‌برداری سد آمده و با سطح مشخصی از اطمینان، مجموعه داده‌های سری زمانی آورد مخزن، بارش و داده‌های هواشناسی برای بهره‌بردار سد مشخص گردد، می‌توان با استفاده از ابزارها و تکنیک روز و نوین جهت پیش‌بینی جریان‌های ورودی به مخازن سدها در 12 گام بعدی استفاده نمود.

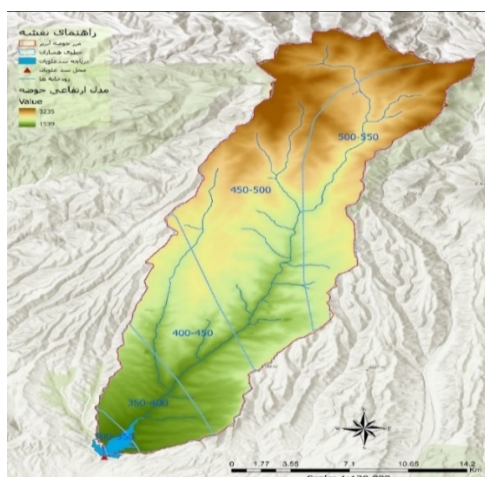
همان‌طوری‌که اشاره گردید، پیش‌بینی جریان ورودی به مخزن در بهینه‌سازی مدیریت مخازن آب بسیار مؤثر می‌باشند؛ اما فاکتورها و عوامل مختلفی بر روی این پدیده‌ها تأثیرگذارند که تحلیل آن‌ها را مشکل می‌سازد. مدل‌های آماری و رگرسیون از معمول‌ترین روش‌های تحلیلی می‌باشند که غالباً با توجه به حل خطی این پدیده‌ها، نتایجی همراه با خطا بوده و نمی‌توانند تغییرات زمانی پدیده موردنظر را با دقت قابل‌قبول مدل کنند. امروزه سیستم‌های "هوش مصنوعی"¹ و روش‌های داده‌کاوی با توجه به توانایی در حل پدیده‌های غیرخطی و پیچیده، کاربردهای فراوانی در مسائل مختلف مهندسی منابع آب از جمله مدیریت مخازن سدها پیدا کرده‌اند؛ بنابراین ارائه مدلی متأثر از تمام عوامل مؤثر در پدیده‌های هیدرولوژی حوضه‌های آبریز مخازن سدها به دلیل تعدد آن‌ها و همچنین وابستگی آن‌ها به شرایط مکانی و زمانی امکان‌پذیر نیست، به‌همین دلیل برحسب استفاده از روش‌های گوناگون برای ساده‌سازی، انواع مدل‌ها توسعه می‌یابند. در این راستا از مدل‌های جعبه سیاه که با دریافت ورودی و انجام یک سری عملیات ریاضی، خروجی موردنظر را می‌دهند، استفاده می‌شود.

در سال‌های اخیر روش‌های هوش مصنوعی توانایی قابل ملاحظه‌ای در مدل‌سازی و پیش‌بینی سری‌های زمانی هیدرولوژی از خود نشان داده‌اند. این روش‌ها چون مبتنی بر داده هستند، به دانش اولیه نسبت به پدیده مورد مطالعه نیاز ندارند. مدل‌های مبتنی بر هوش مصنوعی، ابزاری کارآمد برای کار کردن بر روی تعداد وسیعی از داده‌های دینامیکی و غیرخطی می‌باشند به‌خصوص اگر روابط بین عوامل فیزیکی دخیل در پدیده به‌صورت کامل مشخص نباشد. مدل‌های ANN از جمله روش‌های هوش مصنوعی می‌باشند که حالت خود همبسته دارند و به‌صورت غیرخطی توانایی کنترل جزء خود همبسته سری زمانی را دارند.

2- مواد و روش تحقیق

2-1- محدوده مورد مطالعه و داده‌ها

حوضه آبریز سد علویان واقع در حوضه آبریز دریاچه ارومیه با مساحت 275 کیلومترمربع در محدوده‌ای به عرض جغرافیایی 37 درجه و 11 دقیقه تا 38 درجه و 28 دقیقه شمالی و طول جغرافیایی 46 درجه تا 46 درجه و 25 دقیقه شرقی قرار دارد (شکل 1). رودخانه صوفی‌چای با طول 120 کیلومتر از دامنه‌های جنوب غربی کوه سهند واقع در استان آذربایجان شرقی سرچشمه می‌گیرد و جریان‌های سطحی این رودخانه از مجموع چشمه‌سارهایی که در دامنه سهند ظاهر می‌شوند و همچنین ذوب ذخیره و پوشش برفی این منطقه تأمین می‌شود. جهت رودخانه تقریباً به طرف جنوب بوده و در دره باریک و کم‌عرض جریان دارد. شاخه فرعی اسپیران‌چای به طول 13/5 کیلومتر در پایین‌دست، در محلی به نام کهلک‌دره‌سی (اسپیران) از طرف راست وارد رودخانه می‌شود. رودخانه صوفی‌چای بعد از طی مسافتی در حدود 4 کیلومتر از محل اتصال اسپیران، به تازه‌کند قاسم‌خان می‌رسد. در این محل، ایستگاه هیدرومتری تازه‌کند قرار دارد. شاخه افسستنج یکی دیگر از شاخه‌های فرعی رودخانه صوفی‌چای است که در فاصله 2/5 کیلومتری پایین‌دست تازه‌کند قاسم‌خان وارد آن می‌شود. کمی پایین‌تر از این محل نیز، سد علویان قرار دارد که رودخانه نهایتاً در پایین‌دست سد علویان به دریاچه ارومیه ختم می‌شود (Mahab Ghods Eng. Cons., 2015).



شکل 1- حوضه آبریز سد علویان، از زیر حوضه‌های آبریز دریاچه ارومیه

سد علویان در حوضه آبریز دریاچه ارومیه، بر روی رودخانه صوفی‌چای در 5/3 کیلومتری شمال غربی شهر مراغه در استان آذربایجان شرقی احداث شده است. سد علویان در سال 1374 به

بنابراین در این تحقیق به‌عنوان نوآوری از توانایی مدل شبکه عصبی مصنوعی (ANN) با انتشار برگشتی خطا و از سایر مدل‌های مبتنی بر داده‌کاوی در بستر نرم‌افزار Weka در مدل‌سازی پیش‌بینی جریان ورودی به مخزن سد علویان در چند گام بعد (تا 12 ماه بعد) با به‌کارگیری داده‌های مشاهداتی سد علویان از جمله پوشش برف و شاخص خشک‌سالی استفاده شده است. ابتدا در راستای پیش‌پردازش و آماده‌سازی داده‌ها جهت انجام فرآیند مدل‌سازی و همین‌طور در راستای همگنی، دقت و صحت داده‌های تاریخی حوضه آبریز سد علویان مورد بررسی قرار گرفته است. همچنین با بررسی مقادیر خودهمبستگی جریان ورودی به مخزن سد با بارش، تبخیر، دمای حداقل و حداکثر، پوشش برف و سایر پارامترها، ترکیب متغیرهای ورودی در فرآیند مدل‌سازی شناسایی گردید. براین اساس حجم آب ورودی ماهانه به مخزن سد علویان برای یک سال آبی در آینده به‌عنوان مقادیر هدف و خروجی مدل و سایر پارامترهای جریان آب، بارش، تبخیر، دمای حداقل و حداکثر، پوشش برف، شاخص خشک‌سالی به‌عنوان ورودی مدل منظور شده است. جهت بررسی دقیق‌تر از چندین ترکیب مختلف از پارامترهای ورودی گوناگون به کمک معیار Correlation خطی انتخاب شده‌اند، به‌گونه‌ای که ارتباط ورودی با خروجی مدل توسط معیار Correlation خطی بررسی شده و عملکرد و دقت آن‌ها در پیش‌بینی جریان ورودی برای یک سال آبی در آینده مورد بررسی قرار گرفته و بهترین ترکیب ورودی انتخاب شده است (Sharghi و همکاران، 2018).

همان‌طوری که اشاره گردید، هدف این مقاله پیش‌بینی جریان ورودی تا 12 گام بعد می‌باشد، بنابراین چالش‌های ذیل در این مطالعه مدنظر بوده است:

(الف) با توجه به این‌که هر مدل با یک ترکیب ورودی سروکار دارد، انتخاب ترکیب ورودی که بتواند برای خروجی مدل نتایج مناسبی داشته باشد، جزء چالش‌های اصلی می‌باشد.

(ب) بررسی تعداد گام‌های مناسب برای پیش‌بینی، از چالش‌های مهم این مقاله می‌باشد.

(ج) بررسی تأثیر میزان بارش، رواناب حوضه آبریز سد، دما، تبخیر و پوشش برف در پیش‌بینی ورودی به مخزن سد

همان‌طوری که اشاره شد در این مطالعه، مدل‌سازی پیش‌بینی جریان ورودی به مخزن سد علویان با استفاده از روش شبکه عصبی مصنوعی و ارزیابی نتایج آن با سایر مدل‌های مبتنی بر داده‌کاوی در بستر نرم‌افزار WEKA شامل GP²، SMO³، RF⁴ و RT⁵ در 12 گام بعدی بررسی شده است. در انتها، کارایی همه مدل‌ها با استفاده از معیارهای ارزیابی بررسی شده و مدل‌ها با یکدیگر مقایسه شده‌اند.

4. Random Forest
5. Random Tree

2. Gaussian Processes
3. Support Machine Optimizer regression

برای این که اثر خشک‌سالی در مدل اعمال گردد پارامتر خشک‌سالی تحت عنوان SPI⁶ از طریق بارش ماه‌های قبل محاسبه و در ماتریس ورودی‌های مدل لحاظ گردیده است.

2-2- نرم‌افزار WEKA

در سال 1993 در دانشگاه Waikato کشور نیوزیلند، طرح اولیه برای تولید نرم‌افزاری که بتواند عملیات "داده‌کاوی" را انجام دهد، مطرح شد. در سال 1997 نیز این برنامه از پایه با زبان برنامه‌نویسی JAVA نوشته شد و توسعه یافت. در سال 2005، نرم‌افزار WEKA به معنی «محیط وایکاتو برای تحلیل دانش» در قالب مدل‌های مختلف معرفی شده است.

2-3- داده‌کاوی⁷

برای پیش‌بینی جریان ورودی به مخزن در گام اول داده‌ها بازسازی و دسته‌بندی شده است، در گام دوم انتخاب داده‌ها و تشکیل ماتریس داده‌های ورودی به مدل صورت گرفته است، همان طوری که می‌دانیم به لحاظ ماهیت غیرخطی و عدم قطعیت سری زمانی داده‌های حوضه آبریز سدها، انتخاب داده‌ها یکی از چالش‌های مهم مدل‌سازی می‌باشند، برای انتخاب داده‌های مؤثر بین انواع داده‌های ورودی و داده هدف، ضریب همبستگی به کمک معیار Correlation محاسبه و انواع حالت‌های مختلف در مدل تست گردید و پس از سعی و خطا بهترین داده‌ها که نتایج مناسبی را می‌دادند، انتخاب شدند. داده‌های ورودی به مدل، شامل جریان آب ورودی به مخزن، بارش، دمای حداقل و حداکثر، درصد پوشش برف، شاخص‌های خشک‌سالی و شاخص عددی برای همراه در طول دوره آماری 25 سال می‌باشند.

قابل ذکر است که استفاده از SPI در کنار بارش به دلیل ارائه یک دید جامع‌تر از وضعیت آب و هوایی است. درحالی که بارش می‌تواند نشان‌دهنده مقدار بارش در یک دوره زمانی خاص باشد، SPI با ارائه اطلاعات در مورد تاریخچه بارش و تراکم آب زیرزمینی، از جامعیت بیشتری برخوردار است. این ارتباط نه تنها به تخمین جریان ورودی به مخزن سد کمک می‌کند، بلکه امکان بهبود مدیریت منابع آبی را فراهم می‌کند، زیرا می‌تواند الگوهای خشک‌سالی را در طول زمان شناسایی کند و از پایداری و عملکرد سدها حمایت کند.

یکی دیگر از داده‌های مهم و از نوع‌آوری‌های این تحقیق استفاده از داده پوشش برف به‌عنوان ورودی مدل پیش‌بینی می‌باشد که با توجه به مساحت حوضه آبریز سد علویان (272 کیلومتر مربع) مقادیر سطوح پوشش برف در هریک از ماه‌ها

بهره‌برداری رسیده است و نوع سد خاکی با هسته رسی دارای حجم کل اولیه مخزن 60 میلیون مترمکعب شامل 57 میلیون مترمکعب حجم مفید و 3 میلیون مترمکعب حجم مرده می‌باشد. از اهداف اصلی احداث سد تأمین آب مورد نیاز کشاورزی در محدوده شهرستان مراغه و بناب با وسعتی برابر با 12200 هکتار به‌عنوان مبنای طراحی و در حال حاضر 6500 هکتار از اراضی منطقه بهبود، تأمین بخشی از آب شرب شهرستان مراغه، تأمین آب موردنیاز صنایع منطقه و همچنین مهار سیلاب رودخانه صوفی‌چای و محیط‌زیست می‌باشد.

مطابق جدول (1)، در بالادست سد علویان، ایستگاه هیدرومتری تازه‌کند و ایستگاه‌های هیدرومتری اسفستانج و کهلیک‌دره‌سی وجود دارند، لازم به ذکر است که جریان عبوری از شاخه فرعی اسپیران‌چای (ایستگاه کهلیک‌دره‌سی) در بالادست ایستگاه هیدرومتری تازه‌کند بعد از پیوستن به شاخه اصلی وارد مخزن سد می‌شود (شکل (1)). میانگین جریان ورودی به مخزن سد 3 مترمکعب بر ثانیه بوده است (Mahab Ghods Eng. Cons., 2015).

براساس آمار و اطلاعات ایستگاه‌ها، متوسط بارندگی سالانه 350 میلی‌متر می‌باشد که حدود 34 درصد بارندگی در زمستان، 35 درصد آن در بهار، 28 درصد در پاییز و در حدود 3 درصد در تابستان می‌باشد. ارتفاع حوضه آبریز، در بالادست سد علویان، از 1547 تا 3440 متر متغیر بوده و ارتفاع متوسط حوضه (ارتفاع مربوط به 50 درصد مساحت حوضه)، 2270 متر نسبت به سطح دریا می‌باشد. با توجه به اختلاف شدید ارتفاعی درون حوضه، در حدود 89 درصد از بارش‌های این منطقه در بازه زمانی اوایل آبان تا اواخر اردیبهشت که مصادف با ورود جریان‌های مدیرانه‌ای به کشور است، اتفاق می‌افتد (Mahab Ghods Eng. Cons., 2015). داده‌های مورد استفاده در این بررسی شامل یک‌سری داده‌های 25 ساله (1376-1401) اعم از دبی ورودی به مخزن سد علویان و سایر اطلاعات مثل بارش، دبی جریان در ایستگاه‌های هیدرومتری، دما، تبخیر و پوشش برف حوضه آبریز سد می‌باشد.

جدول 1- مشخصات ایستگاه‌های هیدرومتری بالادست سد

| علویان | | | |
|------------------------------------|----------|----------|--------------|
| ایستگاه | تازه‌کند | اسفستانج | کهلیک‌دره‌سی |
| رودخانه | صوفی‌چای | اسفستانج | اسپیران‌چای |
| مساحت حوضه (کیلومتر) | 202 | 33 | 40 |
| متوسط آورد سالانه (میلیون مترمکعب) | 79 | 4/5 | 11/5 |

2-4- مدل سازی و آموزش شبکه

جهت انجام مدل سازی با استفاده از شبکه عصبی مصنوعی، ابتدا سری زمانی داده های موجود پس از آماده سازی به دو دسته به نسبت های 75% و 25% از کل تعداد داده های موجود به ترتیب به منظور انجام مراحل آموزش و صحت سنجی تقسیم و این دو مرحله برای هر یک از مدل ها و الگوهای موجود در نرم افزار WEKA، برای پیش بینی کوتاه مدت ماهانه و پیش بینی بلندمدت ماهانه انجام شده است. در مرحله آموزش پارامترهای مدل که شامل وزن های مرتبط بین نرون های مختلف لایه های ورودی-میانی و میانی-خروجی با استفاده از الگوریتم آموزش پس انتشار خطا، تخمین زده می شود (Sharghi و همکاران، 2019).

در این تحقیق، برای مدل سازی پیش بینی ورودی به مخزن سد علویان از سری داده های ماهانه ایستگاه های (تازه کند و ...) برای رواناب و سایر اطلاعات اعم از دما، تبخیر، بارش و پوشش برف در ایستگاه هایی که در حوضه آبریز سد علویان قرار دارند، استفاده شده است (شکل (1)).

در این تحقیق، از سری زمانی داده های 25 سال (1376-1401) دوره بهره برداری سد علویان استفاده شده است و با توجه به این که ترتیب زمانی در مدل سازی پیش بینی جریان ورودی به مخزن سد خیلی مهم است و دارای فرآیند اتورگرسیو⁸ و زنجیره مارکوف⁹ دارد هر زمان به زمان های قبلی ارتباط داده می شود و بایستی به صورت پیوسته داده ها انتخاب شوند، لذا داده های 18 سال اول برای آموزش و داده های 7 سال آخر دوره بهره برداری برای صحت سنجی در نظر گرفته شده است و همین طور ویژگی های آماری متغیرهای ورودی مدل پیش بینی جریان ورودی به مخزن سد علویان محاسبه و در جدول (3) قید شده است.

به صورت درصدی از کل مساحت حوضه آبریز محاسبه و در مدل استفاده شده است.

در گام سوم برای پیش بینی جریان آب ورودی به مخزن در هر یک از ماه ها که از مهرماه سال آبی شروع (t) تا شهریور سال بعد (t+11)، به عبارتی 12 گام بعد با مشخصات داده های مورد اشاره مطابق جدول (2) تشکیل می گردد (Najafi و همکاران، 2022).

در این جدول ترکیب ورودی های مدل برای پیش بینی خروجی، شامل داده های ورودی به مخزن، بارش، دمای حداقل و حداکثر، پوشش برف، شاخص خشک سالی و شاخص ماه بوده است، به عنوان مثال برای پیش بینی ورودی جریان آب به مخزن سد علویان در طول سال آبی به شکل زیر عمل می شود، برای پیش بینی جریان ورودی مهرماه، نسبت به یک گام قبل (شهریور ماه)، ضریب همبستگی 0/7 و برای 12 ماه قبل (مهرماه یک سال قبل)، ضریب همبستگی 0/89 به دست آمد در حالی که برای بقیه ماه ها این ضریب کمتر بود و همین طور برای پیش بینی آبان ماه، دو ماه قبل و یک سال قبل ارزیابی و ضرایب همبستگی مناسبی به دست داد، به همین ترتیب برای پیش بینی مردادماه از یازده گام زمانی قبل (شهریورماه) و داده های یک سال قبل یعنی مردادماه سال گذشته استفاده می شود و در نهایت برای پیش بینی شهریورماه از دوازده گام زمانی قبل (شهریورماه) که همان داده های یک سال قبل یعنی شهریورماه سال گذشته است، استفاده می شود. همچنین تعداد ماتریس ورودی برای هر یک از ماه های مهر لغایت مردادماه 16 مورد (ترکیب داده های ورودی مدل) و صرفاً برای شهریورماه 8 مورد می باشد. نتایج در جدول (2) برای تمامی ماه ها قید شده است.

جدول 2- گام های پیش بینی جریان ورودی

| log ₆ اسفند | | log ₅ بهمن | | log ₄ دی | | log ₃ آذر | | log ₂ آبان | | log ₁ مهر | |
|--------------------------|----------|-------------------------|----------|-----------------------|----------|------------------------|-----------|---------------------------|-----------|--------------------------|-----------|
| Target | Input | Target | Input | Target | Input | Target | Input | Target | Input | Target | Input |
| t+5 | t-1, t-7 | t+4 | t-1, t-8 | t+3 | t-1, t-9 | t+2 | t-1, t-10 | t+1 | t-1, t-11 | t | t-1, t-12 |
| log ₁₂ شهریور | | log ₁₁ مرداد | | log ₁₀ تیر | | log ₉ خرداد | | log ₈ اردیبهشت | | log ₇ فروردین | |
| Target | Input | Target | Input | Target | Input | Target | Input | Target | Input | Target | Input |
| t+11 | t-1 | t+10 | t-1, t-2 | t+9 | t-1, t-3 | t+8 | t-1, t-4 | t+7 | t-1, t-5 | t+6 | t-1, t-6 |

9. Markov

8. Autoregressive

جدول 3- ویژگی‌های آماری متغیرهای ورودی مدل پیش‌بینی جریان ورودی به سد علویان

| Title | Runoff time series (MCM/month) | | | | | Precipitation time series (mm) | | | | | Evaporation Time Series (MCM) | | | | |
|-----------------|--------------------------------|-----|------|------|-----|--------------------------------|-----|------|------|-----|-------------------------------|-----|------|------|-----|
| | Max | Min | Mean | St D | C.C | Max | Min | Mean | St D | C.C | Max | Min | Mean | St D | C.C |
| Training data | 52 | 0.4 | 8 | 9.8 | 0.9 | 147 | 0 | 26 | 29 | 0.6 | 0.9 | 0 | 0.2 | 0.2 | 0.6 |
| Validation data | 43 | 0.3 | 8 | 10 | 0.8 | 124 | 0 | 31 | 33 | 0.5 | 0.7 | 0 | 0.2 | 0.2 | 0.5 |

| Title | Minimum temperature time series (C) | | | | | Maximum temperature time series (C) | | | | | Snow cover (percentage) | | | | |
|-----------------|-------------------------------------|-----|------|------|-----|-------------------------------------|-----|------|------|-----|-------------------------|-----|------|------|-----|
| | Max | Min | Mean | St D | C.C | Max | Min | Mean | St D | C.C | Max | Min | Mean | St D | C.C |
| Training data | 22 | -9 | 6 | 8 | 0.2 | 36 | -2 | 10 | 11 | 0.3 | 99 | 0 | 28 | 33 | 0.2 |
| Validation data | 22 | -8 | 8 | 9 | 0.1 | 36 | 2 | 19 | 11 | 0.2 | 87 | 0 | 27 | 32 | 0.1 |

| Title | SPI1 | | | | | SPI6 | | | | | SPI12 | | | | |
|-----------------|------|-----|------|------|-----|------|-----|------|------|-----|-------|-----|------|------|-----|
| | Max | Min | Mean | St D | C.C | Max | Min | Mean | St D | C.C | Max | Min | Mean | St D | C.C |
| Training data | 3.8 | -3 | 0.1 | 0.1 | 0.1 | 2 | -3 | 0.1 | 0.2 | 0.3 | 2 | -2 | 0.1 | 0.3 | 0.2 |
| Validation data | 3.5 | -2 | 0.1 | 0.2 | 0.1 | 2 | -2 | 0.1 | 0.2 | 0.2 | 2 | -1 | 0.1 | 0.2 | 0.1 |

3- نتایج و بحث

برآورد پتانسیل منابع آب، به‌ویژه پیش‌بینی ورودی‌های جریان آب در 12 گام بعدی (یک تا 12 ماه بعد) یکی از گام‌های مهم در مدیریت منابع آب و بهره‌برداری از مخازن سدها محسوب می‌شود؛ بنابراین استفاده از روش‌هایی که بتواند عملکرد و کارایی مخازن سدها را در زمینه پیش‌بینی جریان‌های ورودی بهبود بخشد از اهمیت ویژه‌ای برخوردار است. از این‌رو، در این مطالعه، جهت بهبود مدیریت مخزن سد علویان، جریان ورودی به مخزن سد علویان در 12 گام بعدی با استفاده از روش‌های هوش مصنوعی و مدل‌های نرم‌افزار WEKA پیش‌بینی گردیده است.

جهت مدل‌سازی چندگامه از توانایی مدل ANN با انتشار برگشتی خطا و همچنین از سایر الگوریتم‌ها در بستر نرم‌افزار WEKA در مدل‌سازی برای پیش‌بینی جریان ورودی با به‌کارگیری داده‌های ورودی مشاهداتی استفاده شده است. به‌همین منظور ابتدا از مدل ANN و سپس از سایر مدل‌های نرم‌افزار WEKA برای پیش‌بینی ورودی جریان آب به مخزن سد علویان یک تا دوازده گام بعد (12 ماه بعد) استفاده شده است. در انتها کارایی همه مدل‌ها با استفاده از معیارهای ارزیابی بررسی شده و مدل‌ها با یکدیگر مقایسه شده‌اند. قبل از تحلیل و ارزیابی نتایج مدل‌ها، مزایا و کاربرد هر یک از مدل‌های هوش مصنوعی و داده‌کاوی توضیح داده شده است.

3-1- شبکه عصبی مصنوعی

شبکه‌های عصبی مصنوعی دسته‌ای از سیستم‌های دینامیکی هستند که با پردازش روی داده‌های تجربی دانش یا قانون نهفته در ورای داده‌ها را به ساختار شبکه منتقل می‌کند، به‌همین خاطر

2-5- معیارهای ارزیابی

صحت‌سنجی مدل‌های پیش‌بینی به‌منظور سنجش دقت مدل‌ها در انجام پیش‌بینی صورت می‌گیرد. برای این منظور معیارهای مختلفی تدوین شده و مورد استفاده قرار می‌گیرند. از پرکاربردترین آن‌ها در مدل‌سازی پیش‌بینی جریان‌های ورودی به مخازن سدها، معیارهای متداول ضریب همبستگی (CC)¹⁰ و جذر میانگین مربعات خطا (RMSE)¹¹ می‌باشند که با توجه به روابط (1) و (2) باید مقادیر آن‌ها در هر دو مرحله آموزش و صحت‌سنجی در محدوده مطلوب قرار گیرند:

$$CC = \frac{n(\sum_{i=1}^n P_i \cdot O_i) - (\sum_{i=1}^n P_i) \cdot (\sum_{i=1}^n O_i)}{\sqrt{(n \sum_{i=1}^n P_i^2 - (\sum_{i=1}^n P_i)^2) \cdot (n \sum_{i=1}^n O_i^2 - (\sum_{i=1}^n O_i)^2)}} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (2)$$

که در این روابط O_i بیانگر داده‌های مشاهداتی، P_i داده‌های محاسباتی و N تعداد کل داده‌های موجود می‌باشد. ضریب همبستگی معیاری آماری است که دارای مقدار بین -1 و 1 بوده و بیانگر میزان همبستگی خطی بین داده‌های مشاهداتی و محاسباتی می‌باشد که در بهترین حالت برابر 1 است. معیار RMSE نیز معیاری جهت بررسی خطای مدل به‌کار می‌رود که هم واحد داده‌ها بوده و هرچه به صفر میل کند، مدل موردنظر بهینه خواهد بود (Ghorbani و همکاران، 2018).

داده‌ها به دو بخش آموزش (75%) و صحت‌سنجی (25%) تقسیم‌بندی شدند. جهت مدل‌سازی چندگامه (یک تا دوازده ماه بعد) از توانایی مدل FFNNs با انتشار برگشتی خطا برای پیش‌بینی جریان ورودی به مخزن سد علویان با به‌کارگیری ورودی‌های مشاهداتی، شبکه آموزش داده شده و از تابع تانژانت سیگموئید به‌عنوان یک تابع محرک در لایه‌های میانی و خروجی استفاده گردید. تعداد تکرارها و نرون‌های لایه میانی با استفاده از روش آزمون‌وخطا به دست آمد. نتایج در جدول (4) ارائه شده است.



شکل 2- پیش‌بینی با مدل‌سازی شبکه‌های عصبی FFNN (Najafi و همکاران، 2022)

جدول 4- نتایج و ساختار مدل‌سازی رواناب ورودی به مخزن سد علویان

| براساس متغیرهای ورودی جدول (2) | Step | Target | ANN | | | | GP | | | |
|--------------------------------|------|----------|-------|---------|------------|---------|-------|---------|------------|---------|
| | | | CC | | RMSE (MCM) | | CC | | RMSE (MCM) | |
| | | | Train | Validat | Train | Validat | Train | Validat | Train | Validat |
| Inputs (t-1, t-12) | 1 | Q (t) | 0.99 | 0.89 | 2.5 | 3.2 | 0.89 | 0.88 | 4.5 | 4.7 |
| Inputs (t-1, t-11) | 2 | Q (t+1) | 0.97 | 0.87 | 2.8 | 4.5 | 0.88 | 0.86 | 4.7 | 5.2 |
| Inputs (t-1, t-10) | 3 | Q (t+2) | 0.95 | 0.88 | 2.5 | 3.7 | 0.88 | 0.88 | 4.8 | 4.9 |
| Inputs (t-1, t-9) | 4 | Q (t+3) | 0.94 | 0.88 | 2.4 | 3.5 | 0.87 | 0.87 | 4.9 | 5.1 |
| Inputs (t-1, t-8) | 5 | Q (t+4) | 0.95 | 0.87 | 2.9 | 3.8 | 0.85 | 0.84 | 5.1 | 5.5 |
| Inputs (t-1, t-7) | 6 | Q (t+5) | 0.95 | 0.85 | 3.7 | 4.1 | 0.88 | 0.85 | 4.9 | 5.2 |
| Inputs (t-1, t-6) | 7 | Q (t+6) | 0.96 | 0.86 | 2.6 | 3.2 | 0.87 | 0.85 | 4.9 | 5.1 |
| Inputs (t-1, t-5) | 8 | Q (t+7) | 0.96 | 0.85 | 2.9 | 3.1 | 0.86 | 0.84 | 4.9 | 5.2 |
| Inputs (t-1, t-4) | 9 | Q (t+8) | 0.95 | 0.85 | 3.2 | 4.5 | 0.87 | 0.83 | 4.9 | 5.2 |
| Inputs (t-1, t-3) | 10 | Q (t+9) | 0.94 | 0.86 | 3.1 | 4.6 | 0.87 | 0.86 | 4.8 | 5.2 |
| Inputs (t-1, t-2) | 11 | Q (t+10) | 0.93 | 0.84 | 2.3 | 3.5 | 0.85 | 0.83 | 5.1 | 5.4 |
| Inputs (t-1) | 12 | Q (t+11) | 0.93 | 0.82 | 2.1 | 2.8 | 0.85 | 0.82 | 5.2 | 5.7 |

| SMO | | | | RF | | | | RT | | | |
|-------|-------|------------|---------|-------|---------|------------|---------|-------|---------|------------|---------|
| CC | | RMSE (MCM) | | CC | | RMSE (MCM) | | CC | | RMSE (MCM) | |
| Train | Valid | Train | Validat | Train | Validat | Train | Validat | Train | Validat | Train | Validat |
| 0.88 | 0.87 | 4.7 | 5 | 0.99 | 0.92 | 1.6 | 3.9 | 0.99 | 0.88 | 0.17 | 5.6 |
| 0.86 | 0.84 | 5 | 5.5 | 0.99 | 0.92 | 1.6 | 4 | 0.99 | 0.88 | 0.17 | 5.9 |
| 0.86 | 0.86 | 4.9 | 5.2 | 0.98 | 0.92 | 1.8 | 4.1 | 0.99 | 0.88 | 0.17 | 6.3 |
| 0.85 | 0.83 | 5.2 | 3.7 | 0.99 | 0.88 | 1.7 | 4.9 | 0.99 | 0.88 | 0.17 | 6.3 |
| 0.85 | 0.82 | 5.2 | 5.8 | 0.98 | 0.9 | 1.8 | 4.5 | 0.99 | 0.86 | 0.18 | 5.9 |
| 0.85 | 0.84 | 5.2 | 5.6 | 0.98 | 0.92 | 5.2 | 5.5 | 0.99 | 0.86 | 0.18 | 7.9 |
| 0.85 | 0.84 | 5.3 | 5.5 | 0.98 | 0.91 | 1.8 | 4.3 | 0.99 | 0.86 | 0.15 | 5.5 |
| 0.96 | 0.87 | 2.9 | 6 | 0.98 | 0.9 | 1.9 | 4.4 | 0.99 | 0.85 | 0.16 | 6.5 |
| 0.86 | 0.83 | 5.1 | 5.7 | 0.98 | 0.89 | 1.9 | 4.8 | 0.99 | 0.85 | 0.18 | 6.6 |
| 0.86 | 0.84 | 5 | 5.6 | 0.98 | 0.9 | 1.9 | 4.5 | 0.99 | 0.84 | 0.17 | 6.9 |
| 0.84 | 0.83 | 5.3 | 5.7 | 0.98 | 0.9 | 2 | 4.4 | 0.99 | 0.85 | 0.18 | 8.3 |
| 0.84 | 0.82 | 5.3 | 5.8 | 0.98 | 0.88 | 1.9 | 4.8 | 0.99 | 0.84 | 0.17 | 4.9 |

3-2- الگوریتم GP

الگوریتم GP فرآیندهای گاوسی را برای رگرسیون بدون تنظیم هاپیر پارامتر پیاده‌سازی می‌کند (Schulz و همکاران، 2016). نتایج و ساختار مدل‌سازی رواناب ورودی به مخزن سد علویان با استفاده از الگوریتم GP در جدول (4) نشان داده شده است.

3-3- الگوریتم SMO

الگوریتم SMO ماشین بردار پشتیبانی را برای رگرسیون پیاده‌سازی می‌کند. پارامترها را می‌توان با استفاده از الگوریتم‌های مختلف یاد گرفت. الگوریتم با تنظیم RegOptimizer انتخاب می‌شود (Sharma و Chug، 2015). نتایج و ساختار مدل‌سازی رواناب ورودی به مخزن سد علویان با استفاده از الگوریتم SMOreg در جدول (4) نشان داده شده است.

3-4- الگوریتم RF

مدل جنگل تصادفی (RF) یکی از بهترین و متداول‌ترین مدل‌های هوشمند محسوب می‌شود و از بارزترین ویژگی‌های آن می‌توان به نداشتن مشکل پیش‌برازش داده‌ها اشاره کرد. یکی از بزرگترین مزایای جنگل تصادفی تطبیق‌پذیری و استفاده برای کارهای رگرسیون و طبقه‌بندی است و همچنین مشاهده اهمیت نسبی که به ویژگی‌های ورودی اختصاص می‌دهد. جنگل تصادفی اغلب نتیجه پیش‌بینی خوبی ایجاد می‌کند و درک بیش از حد پارامترها کاملاً ساده است. این الگوریتم بسیار انعطاف‌پذیر است زیرا جنگل تصادفی می‌تواند وظایف رگرسیون و طبقه‌بندی را با درجه دقت بالایی کنترل کند و به این دلیل یک روش محبوب در بین دانشمندان داده است. دسته‌بندی کردن ویژگی در طبقه‌بندی جنگل تصادفی را به ابزاری مؤثر برای تخمین مقادیر از دست‌رفته تبدیل می‌کند زیرا در هنگام از دست دادن بخشی از داده‌ها دقت را حفظ می‌کند.

این مدل در مطالعات متعددی جهت مدل‌سازی و پیش‌بینی پارامترهای هیدرولوژیکی مورد استفاده و ارزیابی قرار گرفته است. از این مطالعات، می‌توان به مدل‌سازی دبی جریان رودخانه و همچنین تخمین تراز سطح آب دریاچه اشاره کرد. مدل جنگل تصادفی مبتنی بر دسته‌ای از درخت‌های رگرسیونی تصادفی (درخت‌های تصمیم) می‌باشد که با تشکیل مجموعه‌ای از درخت‌های رگرسیونی و تجمیع نتایج آن‌ها، مدلی رگرسیونی جهت تخمین داده‌ها ایجاد می‌کند (Li و همکاران، 2019؛ Li و همکاران، 2016).

نتایج و ساختار مدل‌سازی رواناب ورودی به مخزن سد علویان با استفاده از الگوریتم Random Forest در جدول (4) نشان داده شده است.

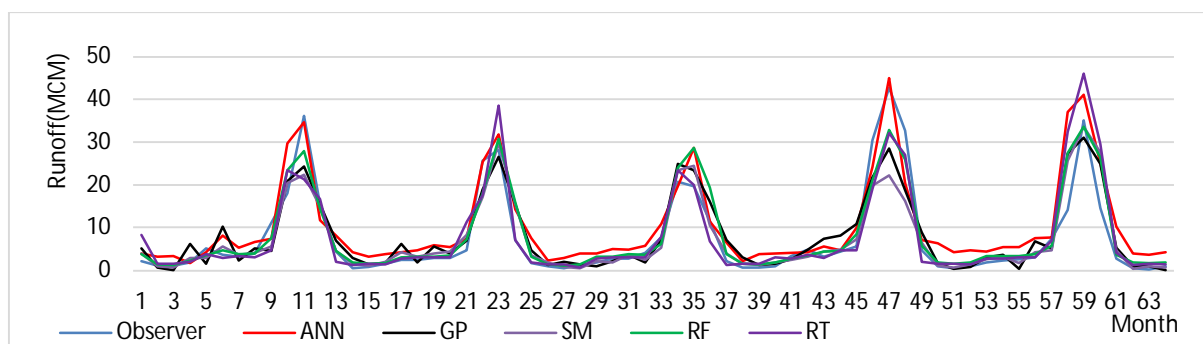
3-5- الگوریتم RT

کلاسی برای ساختن درختی که با ویژگی‌های K انتخاب تصادفی را در هر گره در نظر می‌گیرد هرس را انجام نمی‌دهد. همچنین دارای گزینه‌ای برای تخمین احتمالات کلاس (یا میانگین هدف در حالت رگرسیون) بر اساس یک مجموعه نگاه‌دارنده است (Kulkarni و همکاران، 2021).

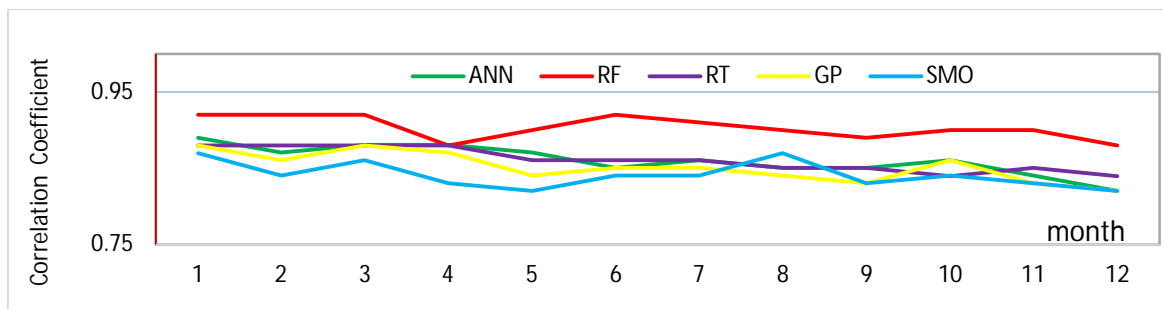
نتایج و ساختار مدل‌سازی رواناب ورودی به مخزن سد علویان با استفاده از الگوریتم RT در جدول (4) نشان داده شده است.

3-6- مقایسه مدل‌ها

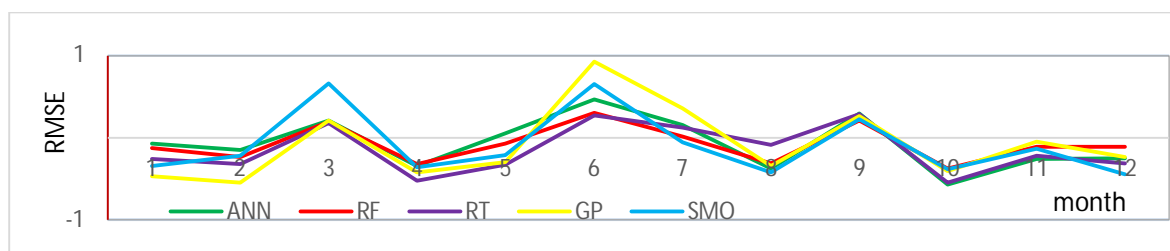
به منظور ارزیابی کیفیت مدل‌سازی و بررسی دقیق‌تر، نتایج مدل‌سازی‌های مختلف ناشی از مدل‌های GP، ANN، SM، RF و RT بر اساس معیارهای ضریب همبستگی و جذر میانگین مربعات خطا در مرحله صحت‌سنجی با سری زمانی ماهانه رواناب ورودی مشاهداتی مقایسه شده‌اند (شکل‌های (3) تا (7)).



شکل 3- نمودار سری زمانی رواناب مشاهداتی در مقایسه با محاسباتی ANN، GP، SM، RF و RT برای داده‌های صحت‌سنجی



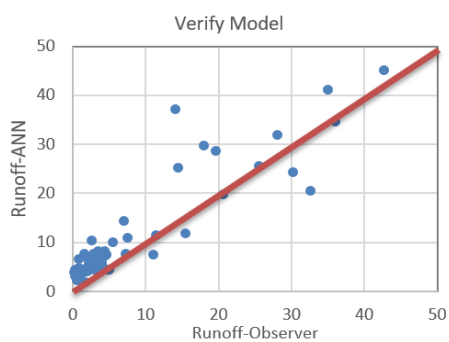
شکل 4- مقایسه ضریب همبستگی ماهانه نتایج مدل سازی ورودی محاسباتی برای پنج مدل در مرحله صحت‌سنجی



شکل 5- مقایسه نرمال شده جذر میانگین مربعات خطا سری زمانی ماهانه نتایج مدل سازی ورودی محاسباتی برای پنج مدل در مرحله صحت‌سنجی

نشان داد که در هر 12 گام بعد پیش‌بینی‌شده، مدل‌های ANN، RT و RF عملکرد بهتر و نزدیک به هم در مقایسه با سایر مدل‌ها داشته است. جهت مقایسه دقیق‌تر مقادیر پیش‌بینی با مقادیر مشاهداتی لازم است نمودارهای پراکنش ترسیم گردد تا از روی نحوه پراکنش نقاط، تفسیرهای نسبتاً دقیقی در مورد نتایج صورت پذیرد.

در شکل (6) ارتباط سری داده‌های مدل ANN با سری داده‌های مشاهداتی مورد بررسی قرار گرفت و نتایج نشان می‌دهد که همبستگی مناسبی بین سری داده‌های مشاهداتی با نتایج مدل ANN دارد و همچنین در شکل (7) ارتباط سری داده‌های مدل RT با سری داده‌های مشاهداتی مورد بررسی قرار گرفت و نتایج نشان داد که همبستگی مناسبی بین سری داده‌های مشاهداتی با نتایج مدل RT نیز دارد.



شکل 6- مقایسه سری زمانی مشاهداتی رواناب با نتایج مدل ANN

طبق شکل (3) نتایج خروجی مدل‌های پنج‌گانه برای داده‌های صحت‌سنجی باهم مورد بررسی قرار گرفت و با تحلیل روند نمودارها و مقایسه آن‌ها با سری داده‌های مشاهداتی نشان داد که نتایج سری داده‌های مدل ANN و RF و RT باهم و با داده‌های مشاهداتی به‌ویژه در پیک جریان‌های ورودی هم‌خوانی خوبی دارند و نتایج مدل‌سازی بیانگر این است که استفاده از شبکه‌های عصبی مصنوعی و مدل‌های درخت تصمیم به‌عنوان پیش‌پردازش داده‌ها موجب شده با داده‌های ورودی مشاهداتی به مخزن سد علویان، دقت نتایج مدل‌سازی در همه گام‌های پیش‌بینی به‌صورت مطلوبی افزایش یابد.

در شکل (4) نمودار همبستگی متقابل بین مقادیر جریان ورودی ماهانه مشاهداتی و جریان محاسباتی در 12 ماه آتی برای داده‌های صحت‌سنجی ارائه گردیده است. ارزیابی نتایج مدل‌سازی نشان داد که در هر 12 گام بعد پیش‌بینی شده، مدل‌های ANN، RT و RF عملکرد بهتر و نزدیک به هم در مقایسه با سایر مدل‌ها داشته است. با بررسی مقادیر ضریب همبستگی برای هر یک از مدل‌های پنج‌گانه به‌ویژه در سه ماه اول سال آبی، سه مدل مذکور در پیش‌بینی جریان آب ورودی، همبستگی خوبی با داده‌های مشاهداتی نشان می‌دهند و در بقیه ماه‌ها همبستگی مناسب تشخیص داده شده است. همچنین در شکل (5) نمودار نرمال‌شده جذر میانگین مربعات خطا سری زمانی ماهانه رواناب ورودی مشاهداتی با جریان محاسباتی در 12 ماه آتی برای داده‌های صحت‌سنجی مقایسه شده است. ارزیابی نتایج مدل‌سازی

افزایش ضریب همبستگی در پیش‌بینی جریان ورودی به مخزن سد علویان در 12 گام بعدی شده است و به‌همین ترتیب مدل جنگل تصادفی نیز در مقایسه با داده‌های ورودی مشاهداتی در همه گام‌های پیش‌بینی جریان ورودی به‌صورت مطلوب ارزیابی می‌گردد.

با توجه به این‌که هدف این مقاله پیش‌بینی جریان ورودی تا 12 گام بعد به‌صورت ماهانه به مخزن سد علویان بود، انتخاب ترکیب ورودی که بتواند برای همه گام‌های پیش‌بینی نتایج مناسبی را ارائه دهد، اهمیت ویژه‌ای داشت؛ بنابراین با توجه به اهداف این مقاله نکات مهم زیر نتیجه‌گیری می‌شود:

- به‌طور کلی، سیستم‌های هوشمند دبی ورودی به سد را با دقت نسبتاً بالا پیش‌بینی می‌کنند. دو مدل شبکه عصبی مصنوعی ANN و مدل تصمیم درخت (RT و RF) نشان دادند که می‌توانند دبی ورودی به سد را در 1 تا 3 ماه آینده، به‌خصوص در مرحله آموزش و همچنین مرحله صحت‌سنجی به‌خوبی پیش‌بینی کنند و در 3 تا 6 ماه آینده پیش‌بینی‌ها مناسب و در حد قابل‌قبول و در 6 تا 12 ماه آخر سال آبی در همه مدل‌ها پیش‌بینی‌ها به‌همدیگر نزدیک و همراه با همبستگی ضعیف با داده‌های مشاهداتی می‌باشند.

- نتایج به‌دست‌آمده نشان داد مدل درخت تصمیم در مقایسه با سایر روش‌ها به‌علت ارائه روابط خطی ساده و نسبتاً دقیق از دیدگاه کاربردی مناسب می‌باشد.

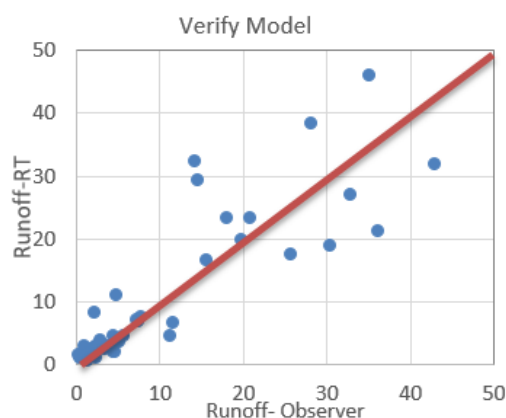
- شبکه‌های عصبی و مدل‌های درخت تصمیم مقدار دبی را بیشتر از مقدار واقعی پیش‌بینی می‌کنند.

- از ترکیب داده‌های ورودی به مدل، رواناب حاصل در ایستگاه‌های هیدرومتری به‌دلیل همبستگی ضعیف حذف شده است و به‌جای آن سری زمانی جریان ورودی به مخزن سد علویان که با سایر پارامترهای حوضه آبریز همبستگی مناسبی داشت، در نظر گرفته شده است. البته به لحاظ شرایط فیزیولوژیکی حوضه آبریز، تمامی جریان‌های گذرنده از ایستگاه‌های هیدرومتری در ورودی مخزن سد علویان تأثیر مستقیمی دارد.

- نتایج به‌دست‌آمده نشان داد که به‌علت برف‌گیر بودن حوضه آبریز سد علویان، پارامتر دما در کنار بارش به‌عنوان یکی از پارامترهای مهم در فرآیند مدل‌سازی بوده است.

- نتایج این مطالعه نشان می‌دهد که مدل شبکه عصبی (ANN) و مدل تصمیم درخت (RT و RF) به‌عنوان مدل‌های کارآمد و دقیق می‌توانند برای پیش‌بینی جریان ورودی مخزن سد علویان به‌کار گرفته شوند.

در نهایت برای مطالعات آینده، پیشنهاد می‌شود با توجه به وجود مدل‌های متنوع و جدید دیگر، در راستای افزایش دقت مدل‌سازی، از توانایی سایر روش‌های هوش مصنوعی به‌جای ANN استفاده گردد و تأثیر روش اعمالی جدید بر روی دقت مدل‌سازی‌ها



شکل 7- مقایسه سری زمانی مشاهداتی رواناب با نتایج مدل RT

4- نتیجه‌گیری

بر اساس بررسی نتایج خروجی مدل‌ها مشخص گردید که مدل‌های تصمیم درخت (RT و RF) و مدل‌های GP و SM در مقایسه با روش شبکه عصبی مصنوعی ANN روشی جدید و کمتر شناخته شده‌ای است. در بین دو روش مختلف شبکه‌های عصبی و داده‌کاوی، به‌منظور افزایش دقت داده‌های جریان ورودی محاسباتی به مخزن سد علویان، کارایی همه مدل‌ها با استفاده از معیارهای ارزیابی بررسی شده با یکدیگر مقایسه شده‌اند و با بررسی نتایج خروجی مدل‌ها، عملکرد متوسط ضریب همبستگی مدل شبکه عصبی 85% و مدل‌های درخت تصمیم شامل جنگل تصادفی 90% و درخت تصادفی 83% در مقایسه با سایر مدل‌های GP و SM کارا تر و مناسب‌تر در مرحله صحت‌سنجی و در پیش‌بینی جریان ورودی سد علویان در افق زمانی 12 ماه آینده به‌منظور برنامه‌ریزی و بهبود مدیریت مخزن سد علویان تشخیص داده شد و موفقیت مدل شبکه عصبی بر این اساس است که به علت استفاده از مدل FFNN که دارای ماهیت غیرخطی است به‌طوری‌که می‌تواند رفتار غیرخطی پدیده را دقیق‌تر از سایر مدل‌های مذکور مدل‌سازی کند. از این‌رو به‌کار بردن مدل غیرخطی با شبکه عصبی می‌تواند روش مناسبی برای مدل‌سازی پدیده‌های غیرخطی و دارای عدم قطعیت، همچون پیش‌بینی جریان ورودی به مخزن سد علویان در 12 گام بعدی باشد و در مقایسه با داده‌های ورودی مشاهداتی در همه گام‌های پیش‌بینی جریان ورودی به‌صورت مطلوب ارزیابی می‌گردد. همچنین مدل جنگل تصادفی با توجه به ماهیت تطبیق‌پذیری، انعطاف‌پذیری و با انتخاب تصادفی مشاهدات، توانسته وظایف رگرسیون و طبقه‌بندی را با درجه دقت بالایی کنترل کند و باعث شده در دسته‌بندی کردن ویژگی در طبقه‌بندی ابزاری مؤثر برای تخمین مقادیر در هنگام از دست دادن بخشی از داده‌ها دقت را حفظ کند و این ماهیت مدل جنگل تصادفی باعث

- <https://doi.org/10.1080/02626667.2019.1680846>
Lin JY, Cheng CT, Chau KW, "Using support vector machines for long-term discharge prediction", *Hydrological Sciences Journal*, 2006, 51 (4), 599-612. <https://doi.org/10.1623/hysj.51.4.599>
- Mahab Ghods Eng. Cons. "Standards for operation of Alavian Dam: Hydrologic Report", East Azarbaijan Water Organization, 2015.
- Mohammadi B, Ahmadi F, Mehdizadeh S, Guan Y, Pham Q-B, Linh N-T, Tri D-Q, "Developing novel robust models to improve the accuracy of daily streamflow modeling", *Water Resources Management*, 2020, 34, 3387-3409.
<https://doi.org/10.1007/s11269-020-02619-z>
- Mohammadi Klaleg Sh, Sattari MT, Rasouli A, Khyat Rostami B, "Prediction of monthly input discharge into Sabalan Dam reservoir via ANN and M5 for one water year", 1st Int. Water and Environment and Development, 2015.
- Najafi H, Nourani V, Sharghi E, Babaeian Amini A, "Multistep modeling of hydroclimatic phenomena using wavelet-neural network seasonal model", *Journal of Civil and Environmental Engineering*, 2022, 52 (107), 139-150.
<https://doi.org/10.22034/jcee.2020.16962.1418>
- Nourani V, Najafi H, Amini A-B, Tanaka H, "Using Hybrid Wavelet-Exponential Smoothing Approach for Streamflow Modeling", *Complexity*, 2021.
<https://doi.org/10.1155/2021/6611848>
- Sanikhani H, Dinpashoh Y, Ghorbani M. "Baranduz-chay river flow modeling using the k-nearest neighbor and intelligent methods", *Water and Soil Science*, 2015, 25 (1), 219-233.
- Sattari MT, Yurekli K, Pal M, "Performance evaluation of artificial neural network approaches in forecasting reservoir inflow", *Applied Mathematical Modeling*, 2012, 36, 2649-2657.
<https://doi.org/10.1016/j.apm.2011.09.048>
- Schulz E, Speekenbrink M, Krause A, "A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions", *Journal of Mathematical Psychology*, 2018, 85, 1-16.
<https://doi.org/10.1016/j.jmp.2018.03.001>
- Sharghi E, Nourani V, Molajou A, Najafi H, "Conjunction of emotional ANN (EANN) and wavelet transform for rainfall-runoff modeling", *Journal of Hydroinformatics*, 2019, 21.
<https://doi.org/10.2166/hydro.2018.054>
- Sharghi E, Nourani V, Najafi H, Molajou A, "Emotional ANN (EANN) and wavelet-ANN (WANN) approaches for Markovian and seasonal based modeling of rainfall-runoff process", *Water Resources Management*, 2018, 32, 3441-3456.
<https://doi.org/10.1007/s11269-018-2000-y>
- Sharma H, Chug A, "Dynamic metrics are superior than static metrics in maintainability prediction: An empirical case study", In 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions) (pp. 1-6). IEEE, 2015.
<https://doi.org/10.1109/ICRITO.2015.7359354>
- با مطالعه حاضر مقایسه گردد. توصیه می‌شود مدل ANN در ترکیب با تبدیل موجک (WANN) را در کنترل اجزاء اصلی سری زمانی و پیش‌بینی چند گام بعد مورد بررسی قرار داد. همچنین توصیه می‌شود برای ارزیابی بیشتر از توانایی مدل WANN در مدل‌سازی چندگامه استفاده شود و برای ارزیابی و مقایسه نتایج از رگرسیون خطی چندمتغیره یا از مدل سری زمانی آرما یا آرما استفاده گردد.
- ### 5- مراجع
- Antar MA, Elassiouti I, Allam MN, "Rainfall-runoff modelling using artificial neural networks technique: a Blue Nile catchment case study", *Hydrological Processes*, 2006, 20, 1201-1216.
<https://doi.org/10.1002/hyp.5932>
- Dinu C, Drobot R, Pricop C, Blidaru TV, "Flash-flood modelling with artificial neural networks using radar rainfall estimates", *Mathematical Modelling in Civil Engineering*, 2017, 13, 10-20.
<https://doi.org/10.1515/mmce-2017-0008>
- Ghorbani MA, Khatibi R, Karimi V, Yaseen ZM, Zounemat-Kermani M, "Learning from multiple models using artificial intelligence to improve model prediction accuracies: application to river flows", *Water Resources Management*, 2018, 32 (13), 4201-4215.
<https://doi.org/10.1007/s11269-018-2038-x>
- Ibrahimi Mohammadi Sh, Boushra Seghale M, "Modeling and forecasting of monthly discharge, case study: Ghara So", 4th Iran Water Resources Conference, Amirkabir University, Tehran, 2011.
- Kim TW, Valdés JB, "Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks", *Journal Hydrologic Engineering*, 2003, 8, 319-328.
[https://doi.org/10.1061/\(ASCE\)1084-0699\(2003\)8:6\(319\)](https://doi.org/10.1061/(ASCE)1084-0699(2003)8:6(319))
- Kisi O, "River flow modeling using artificial neural networks", *Journal of Hydrologic Engineering*, 2004, 9 (1), 60-63. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2004\)9:1\(60\)](https://doi.org/10.1061/(ASCE)1084-0699(2004)9:1(60))
- Kulkarni PS, Londhe S, Sainkar N, Rote S, "Estimation of discharge and total water level at yedgaon dam using data driven techniques", In IOP Conference Series: Materials Science and Engineering, 2021, 1197 (1), 012021. <https://doi.org/10.1088/1757-899X/1197/1/012021>
- Li B, Yang G, Wan R, Dai X, Zhang Y, "Comparison of random forests and other statistical methods for the prediction of lake water level: a case study of the Poyang Lake in China", *Hydrology Research*, 2016, 47 (1), 69-83.
<https://doi.org/10.2166/nh.2016.264>
- Li X, Sha J, Wang ZL, "Comparison of daily streamflow forecasts using extreme learning machines and the random forest method", *Hydrological Sciences Journal*, 2019, 64 (15), 1857-1866.