

ارزیابی قابلیت روش تجزیه متعامد سره جهت تعیین ورودی به مدل شبکه عصبی برای پیش‌بینی جریان ماهانه ورودی به سد علویان

صابر معظمی^۱

محمد رضا وصالی ناصح^۲

m-vesalinaseh@araku.ac.ir

عباس اکبرزاده^۳

روح‌اله نوری^۴

تاریخ پذیرش: ۹۶/۲/۲۷

تاریخ دریافت: ۹۵/۸/۹

چکیده

زمینه و هدف: سدها به عنوان یکی از اساسی‌ترین منابع تامین آب شرب، کشاورزی، برق‌آبی و صنعتی از نقش مهمی در توسعه جوامع انسانی و محیط‌زیست اطراف خود برخوردارند. بنابراین با توجه به نقش اساسی سدها در پویایی محیط اطراف خود، برآورد جریان ورودی به آن‌ها از اهمیت ویژه‌ای برخوردار بوده و از ابزارهای مهم و مؤثر در مدیریت بهینه کمی و کیفی منابع آب است.

روش بررسی: در این تحقیق سعی شده تا با استفاده از مدل هوش مصنوعی شبکه عصبی مصنوعی (ANN) اقدام به مدل‌سازی جریان ماهانه رودخانه صوفی‌چای، ورودی به سد علویان، گردد. همچنین به منظور بهبود عملکرد مدل مذکور و با توجه به اطلاعات زیاد ورودی به این مدل، از روش تجزیه متعامد سره (POD) به منظور تعیین بهترین الگوی ورودی به مدل ANN استفاده گردید. در نهایت نیز عملکرد دو مدل ANN و مدل ترکیبی POD-ANN بر پایه آماره‌های ضریب تعیین (R^2)، میانگین خطای مطلق (MAE) و میانگین قدرمطلق خطای نسبی (AARE) مورد ارزیابی قرار گرفت.

یافته‌ها: نتایج این تحقیق مشخص نمود که اگرچه مقادیر پیش‌بینی شده دبی ورودی به مخزن سد توسط مدل ANN نزدیک به مقادیر مشاهده‌ای هستند اما عملکرد آن در نقاط با دبی بالا با خطای قابل توجهی همراه است. همچنین یافته‌های این تحقیق حاکی از عملکرد بهتر مدل POD-ANN نسبت به مدل ANN در نقاط با دبی بالا بود. در حالت کلی نتایج به دست آمده از مدل POD-ANN اجرا شده مشخص نمود که مقدار آماره‌های R^2 ، MAE و AARE مدل در هر دو مرحله واسنجی و صحت‌سنجی بهبود قابل توجهی نسبت

۱- استادیار مهندسی عمران، مرکز تحقیقات علوم زیست‌محیطی، واحد اسلامشهر، دانشگاه آزاد، تهران، ایران

۲- استادیار مهندسی محیط‌زیست، گروه مهندسی عمران، دانشگاه اراک، اراک، ایران. * (مسئول مکاتبات)

۳- استادیار محیط‌زیست، موسسه تحقیقات آب وزارت نیرو، تهران، ایران

۴- استادیار مهندسی محیط‌زیست، دانشکده تحصیلات تکمیلی محیط‌زیست، دانشگاه تهران، تهران، ایران

به مقادیر مشابه در مدل ANN داشته‌اند. مقدار آماره‌های R^2 ، MAE و AARE در مرحله صحت‌سنجی POD-ANN به ترتیب معادل ۰/۹۳، ۰/۷۹ و ۰/۵۴ بود.

بحث و نتیجه‌گیری: عملکرد بهتر مدل POD-ANN در دبی با مقادیر بالا نسبت به مدل ANN می‌تواند به دلیل عمل پیش‌پردازش بر روی متغیرهای ورودی و کاهش تعداد آن‌ها در مدل POD-ANN در مقایسه با مدل ANN باشد. بنابراین می‌توان نتیجه‌گیری نمود که عمل پیش‌پردازش بر روی متغیرهای ورودی به مدل ANN و کاهش تعداد متغیرهای ورودی به این مدل همراه با بهبود عملکرد آن بوده است.

واژه‌های کلیدی: شبکه عصبی مصنوعی، دبی ماهانه، سد علویان، تجزیه متعامد سره.

Proper Orthogonal Decomposition Performance to Determine the Inputs to the Artificial Neural Network for Prediction of Inflow into Alavian Dam

Saber Moazami¹

Mohammad Reza Vesali Naseh^{2*}

m-vesalinaseh@araku.ac.ir

Abbas Akbarzadeh³

Roohollah Noori⁴

Admission Date: May 17, 2017

Date Received: October 30, 2016

Abstract

Background and Objective: Dams play an important role in development of countries by drinking and agricultural water supply, flood control, hydropower energy supply and recreational purposes. Constructing a dam and making an artificial lake has an important effect on surrounding environment, so being able to forecast the inflow to the dam is an important issue for water resource management.

Method: In this study artificial neural network (ANN) was applied to forecast the monthly inflow from Soofichai River to Alavian Dam. Regarding the huge amount of input data to ANN model and for optimizing its application, proper orthogonal decomposition (POD) was used in order to determine the best inputs for ANN model. Finally, the application of ANN and POD-ANN models was evaluated by determination coefficient (R^2), mean absolute error (MAE) and average of absolute relative error (AARE).

Findings: Results of ANN and POD-ANN models indicated that although ANN output is close to the observed values of inflow to the dam, but it has significant errors. POD-ANN model showed better results than ANN model for high values of inflow. In general, comparing R^2 , MAE and AARE values of two models revealed that POD-ANN model had better performance in both calibration and verification steps in comparison with ANN model. R^2 , MAE and AARE in verification step of POD-ANN model were 0.93, 0.79, and 0.54, respectively.

Discussion and Conclusion: Preprocessing data contributes to better performance of POD-ANN than ANN model, especially in high values of inflow. Therefore, it can be concluded that applying data preprocessing and reducing inputs to ANN model enhances its performance.

Key words: Artificial Neural Network, Monthly Inflow, Alavian Dam, Proper Orthogonal Decomposition.

1- Assistant Professor, Civil Engineering, Environmental Science Research Center, Eslamshahr Branch, Islamic Azad University, Tehran, Iran.

2- Assistant Professor of Environmental Engineering, Department of Civil Engineering, Arak University, Arak, Iran. *(Corresponding Author)

3- Assistant Professor of Environment, Water Research Institute, Ministry of Energy, Tehran, Iran.

4- Assistant Professor of Environmental Engineering, Graduate Faculty of Environment, University of Tehran, Tehran, Iran.

مقدمه

کاهش اغتشاش^۷ در داده‌های ورودی به مدل بوده (مانند تبدیل موجک)، یا این که در راستای کاهش ابعاد بردارهای ورودی به مدل و در نتیجه دستیابی به مدلی با ساختار ساده‌تر و عملکرد بهتر بوده است (مانند الگوریتم ژنتیک).

به هر حال در غالب تحقیقات انجام شده برای پیش‌بینی جریان، زمانی که هدف کاهش ابعاد داده‌های ورودی به مدل‌های هوشمند بوده، از روش‌های بهینه‌سازی استفاده شده است. روش‌های بهینه‌سازی اگرچه در دستیابی به بهترین الگوی ورودی به مدل‌های هوشمند و حذف متغیرهای با اهمیت کم‌تر بسیار موفق عمل می‌نمایند، اما حذف مستقیم متغیرها در این روش‌ها می‌تواند با حذف مقداری از اطلاعات همراه باشد. این خاصیت روش‌های بهینه‌سازی در مواردی مانند پیش‌بینی جریان که در اکثر موارد تنها دستیابی به برخی از اطلاعات تاثیرگذار بر آن مانند اطلاعات جریان و متغیرهای اقلیمی امکان‌پذیر است، می‌تواند عملکرد مدل‌های هوشمند را تضعیف نماید. بنابراین در صورتی که بتوان با حفظ تمام متغیرها، ابعاد داده‌های ورودی به مدل را کاهش داد، می‌توان از حذف مستقیم متغیرها و به تبع آن برخی اطلاعات تاثیرگذار بر جریان جلوگیری نمود. در این راستا استفاده از روش تجزیه متعامد سره (POD)^۸ می‌تواند سودمند واقع گردد. POD عملاً با ایجاد مدهای غالب حاکم بر پارامتر هدف، که در واقع ترکیب خطی از متغیرهای ورودی هستند، قادر به کاهش معنادار داده‌های ورودی به مدل‌های هوشمند است. بنابراین در این تحقیق با توجه با موارد مذکور، هدف اصلی ارزیابی قابلیت روش POD در پیش‌پردازش متغیرهای ورودی به مدل ANN برای پیش‌بینی جریان است و سعی خواهد شد تا روش‌شناسی استفاده از این دو مدل در ترکیب با یکدیگر به طور شفاف ارائه گردد.

تخمین آورد رودخانه‌های ورودی به سدها، حجم آب مورد استفاده برای استفاده در پایین دست را مشخص نموده و امکان تخصیص آب برای کاربری‌های مختلف مانند شرب، کشاورزی و صنعت را امکان‌پذیر می‌نماید. اگرچه روش‌های متعددی برای تخمین جریان توسط محققین مختلف ارائه شده، اما استفاده از روش‌های هوش مصنوعی که بیش‌ترین انعطاف را نسبت به اطلاعات ورودی دارا می‌باشند، در مناطق با کمبود اطلاعات هیدرولوژیکی و اقلیمی، می‌تواند بسیار موثر واقع گردد (۱). در این راستا رابطه بین جریان ماهانه و دو متغیر بارش و درجه حرارت ماهانه در رودخانه‌ای واقع در کانزاس آمریکا توسط مدل شبکه عصبی مصنوعی (ANN)^۱ شبیه‌سازی گردید (۲). همچنین نایاک^۲ و همکاران (۳) با استفاده از سیستم استنتاج نروفازی تطبیقی (ANFIS)^۳ اقدام به مدل‌سازی جریان رودخانه بایتارانی واقع در هند نمودند. لین^۴ و همکاران (۴) برای پیش‌بینی جریان ماهانه رودخانه‌ای در کشور چین از مدل‌های ANN و ماشین بردار پشتیبان (SVM)^۵ استفاده کردند. نتایج این تحقیق بیانگر دقت بهتر مدل SVM در مقایسه با مدل ANN بود. هی^۶ و همکاران (۵) در پژوهشی به مقایسه عملکرد سه مدل هوش مصنوعی ANN، ANFIS و SVM در پیش‌بینی جریان رودخانه‌ای در ناحیه‌ای نیمه‌خشک کوهستانی واقع در کشور چین پرداختند.

با توجه به مطالب مذکور در مورد استفاده از مدل‌های هوشمند برای پیش‌بینی جریان مشخص است که اگرچه استفاده از این مدل‌ها کراراً مورد توجه محققین بوده، اما قابلیت استفاده از آن‌ها ارتباط مستقیمی به اطلاعات ورودی به مدل دارد. در همین راستا بیش‌تر مطالعات انجام شده نیز معمولاً از روش‌های پیش‌پردازش متغیرهای ورودی به منظور بهبود عملکرد مدل‌های هوشمند استفاده نموده‌اند. این روش‌ها یا در راستای

1- Artificial Neural Network (ANN)

2- Nayak

3-Adaptive Neuro-Fuzzy Inference System (ANFIS)

4- Lin

5- Support Vector Machine (SVM)

6- He

7- Noise

8- Proper Orthogonal Decomposition (POD)

مواد و روش‌ها

محدوده مورد مطالعه

محدوده مورد مطالعه در این تحقیق رودخانه صوفی‌چای و سد علویان واقع در استان آذربایجان شرقی است. این رودخانه از دامنه‌های جنوبی کوه‌های سهند سرچشمه گرفته و پس از عبور از قسمت غربی شهر مراغه و جنوب شهر بناب به دریاچه ارومیه می‌ریزد. در این تحقیق برای پیش‌بینی جریان یک ماه بعد، اطلاعات بارندگی (R)، تابش خورشیدی (Rad) و دمای هوا (T) به صورت کمینه، بیشینه و میانگین در مقیاس ماهانه، به مدت ۱۸ سال (از اواسط ۱۳۶۲ تا اواخر ۱۳۸۳) از ایستگاه سینوپتیک مراغه اخذ گردید. همچنین متوسط ماهانه دبی (Q) رودخانه صوفی‌چای نیز در محل ایستگاه هیدرومتری تازکند واقع در بالادست سد علویان تهیه گردید. با توجه به برف‌گیر بودن حوضه این رودخانه، استفاده از کلیه اطلاعات موجود مدنظر قرار گرفت. در گام بعد با توجه به اثرپذیری جریان رودخانه از سری زمانی اطلاعات مذکور، پارامترهای معرفی شده هر کدام با سه تاخیر زمانی به عنوان ورودی به مدل برای پیش‌بینی جریان ماه آینده مورد استفاده قرار گرفت. در نتیجه فضای بردار ورودی به مدل برای پیش‌بینی دبی ماه بعد شامل ۱۸ متغیر می‌باشند که به دلیل تعداد زیاد آن‌ها و همچنین هم‌پوشانی برخی از این اطلاعات، استفاده از برنامه‌های تعیین بهترین الگوی ورودی می‌تواند سودمند واقع گردد.

مدل POD

از POD به کرات در مدل‌سازی عددی سیالات جهت دستیابی به مدلی با ساختار ساده‌تر و زمان اجرای کم‌تر استفاده شده است (۶-۸). عملاً هدف از به کار بردن POD در ترکیب با مدل‌های هوش مصنوعی می‌تواند کاهش ابعاد بردارهای ورودی به این مدل‌ها جهت دستیابی به ساختاری ساده‌تر و به تبع آن بهینه نمودن ساختار ورودی به مدل باشد.

برای این منظور در این روش هدف می‌تواند یافتن یک تابع Φ برای بردارهای ورودی به مدل‌های هوشمند ($U^{(i)}$) می‌باشد که عملاً مشابه $U^{(i)}$ و دربرگیرنده اطلاعات آن است. یکی از روش‌های حل این مساله تصویر $U^{(i)}$ روی Φ است. برای این منظور باید ضرب داخلی ($\Phi, U^{(i)}$) را به دست آورد، به طوری که تا حد امکان این تابع با دنباله $U^{(i)}$ موازی باشد. ضرب داخلی برای دو تابع دلخواه برداری f و g به صورت رابطه (۱) تعریف می‌شود (۹).

$$(f, g) = \int_{\Omega} f(x)g(x)dx \quad (1)$$

بنابراین باید ضرب داخلی ($\Phi, U^{(i)}$) بیشینه و در شروع این ضرب داخلی بهنجار گردد. برای بهنجار کردن با توجه به فضای توابع Φ ضرب داخلی (Φ, Φ) موجود می‌باشد. به بیان ریاضی می‌بایست رابطه (۲) بیشینه شود (۱۰).

$$(\Phi, U^{(i)}) / (\Phi, \Phi) \quad (2)$$

به‌علاوه آن چه حائز اهمیت است بزرگی این ضرب داخلی است نه علامت آن. پس برای به حساب آوردن همه $U^{(i)}$ ‌ها از عبارت مربع میانگین استفاده می‌گردد. مجموعه $U^{(i)}$ ‌ها به صورت رابطه (۳) درنظر گرفته می‌شود (۱۱).

$$\Phi = \{U^{(i)}: 1 \leq i \leq N\} \quad (3)$$

بنابراین لازم است تابعی به صورت Φ که بهترین بیان از مجموعه Φ باشد، محاسبه شود. به بیان دیگر این تابع باید رابطه (۴) را بهینه نماید.

$$\frac{1}{N} \sum_{i=1}^N \frac{|(\Phi, U^{(i)})|^2}{(\Phi, \Phi)} \quad (4)$$

در حقیقت باید در جستجوی تابعی بود که بزرگ‌ترین تصویر مربع میانگین را روی مجموعه Φ داشته باشد. برای بدست آوردن این مهم رابطه (۵) درنظر گرفته می‌شود (۱۲).

$$K\Phi = \frac{1}{N} \sum_{i=1}^N \int_{\Omega} U^{(i)}(x)U^{(i)}(x')\Phi(x')dx' \quad (5)$$

بنابراین رابطه (۶) را داریم:

در رابطه (۷) λ در واقع ضریب لاگرانژ در مساله بهینه‌سازی است و ثابت می‌شود که λ همان مقدار ویژه سیستم می‌باشد. بنابراین فرض می‌شود که Φ^* تابعی باشد که λ را بیشینه کند. پس هر تابع دیگری را می‌توان به صورت $\Phi^* + \xi \Phi'$ نوشت. بنابراین برای λ رابطه (۸) بدست می‌آید (۱۳).

$$F(\xi) = \frac{(K\Phi^*, \Phi) + \xi(K\Phi^*, \Phi') + \xi(K\Phi', \Phi^*) + \xi^2(K\Phi', \Phi')}{(\Phi^*, \Phi^*) + \xi(\Phi^*, \Phi') + \xi(\Phi', \Phi^*) + \xi^2(\Phi', \Phi')} = \lambda \quad (۸)$$

که در آن ω_i ها طوری تعریف می‌شوند که Φ رابطه (۴) را بیشینه نماید. حال اگر روابط (۵) و (۱۱) در رابطه (۱۰) قرار داده شوند، پس از ساده‌سازی رابطه (۱۲) حاصل می‌شود (۱۴).

$$cw = \Lambda w \quad (۱۲)$$

که در آن داریم:

$$c_{i,j} = \frac{1}{N} \int_{\Omega} U^{(i)}(x)U^{(j)}(x) dx \quad \text{and} \quad (۱۳)$$

$$w = [w_1, w_2, \dots, w_n]$$

در رابطه (۱۳) C یک ماتریس غیرمنفی بوده و w ماتریس بردارهای ویژه است. همچنین Λ ماتریس قطری مقادیر ویژه بوده و به صورت $\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_N]$ نمایش داده می‌شود. بردارهای ویژه بدست آمده، متناظر با مقادیر ویژه $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ می‌باشند. اکنون می‌توان حل رابطه (۴) را به صورت رابطه (۱۴) نوشت (۱۴).

$$\Phi_1 = \sum_{i=1}^N \omega_1^1 U^{(i)}, \quad \Phi_2 = \sum_{i=1}^N \omega_1^2 U^{(i)}, \dots, \quad \Phi_N = \sum_{i=1}^N \omega_1^N U^{(i)} \quad (۱۴)$$

مدل ANN

شده است (۱۶). ساختار FFNN عملاً متشکل از چهار مولفه بوده که عبارتند از: الگوریتم آموزش، لایه‌های پنهان، تعداد نرون‌ها در لایه‌های پنهان، و توابع محرک. در این تحقیق از FFNN با یک لایه پنهان و الگوریتم لئونبرگ-ماردکویت (LM)^۲ برای بهینه‌سازی پارامترهای شبکه استفاده شده است. همچنین توابع محرک در لایه پنهان و خروجی به ترتیب تابع سیگموئید تانژانتی^۳ و تابع خطی انتخاب شدند. همچنین برای جلوگیری از بروز مشکل فوق‌برازشی و کم‌تخمینی FFNN

$$\begin{aligned} K\Phi, \Phi &= \frac{1}{N} \sum_{i=1}^N \iint U^{(i)}(x)\Phi(x)U^{(i)}(x')\Phi(x')dx dx' \\ &= \frac{1}{N} \sum_{i=1}^N |U^{(i)}, \Phi| \end{aligned} \quad (۶)$$

بعلاوه:

$$\frac{(K\Phi, \Phi)}{(\Phi, \Phi)} = \frac{\frac{1}{N} \sum_{i=1}^N |(\Phi, U^{(i)})|^2}{(\Phi, \Phi)} = \lambda \quad (۷)$$

با توجه به تعریف Φ^* مشخص است هنگامی که $\xi = 0$ باشد مقدار بیشینه λ اتفاق می‌افتد و در نتیجه $\frac{dF(\xi)}{d\xi} = 0$. این مساله منجر به رابطه (۹) می‌شود.

$$(K\Phi^*, \Phi') = \lambda(\Phi^*, \Phi') \quad (۹)$$

اکنون کاملاً مشخص است که مساله بیشینه رابطه (۴) همانند پیدا کردن مقادیر ویژه مساله مقدار ویژه رابطه (۱۰) می‌باشد (۷).

$$K\Phi^* = \lambda\Phi^* \quad (۱۰)$$

همچنین فرض می‌شود که Φ به صورت خاصی بر حسب جملات داده‌های اصلی $U^{(i)}$ تعریف شود.

$$\Phi = \sum_{i=1}^N \omega_i U^{(i)} \quad (۱۱)$$

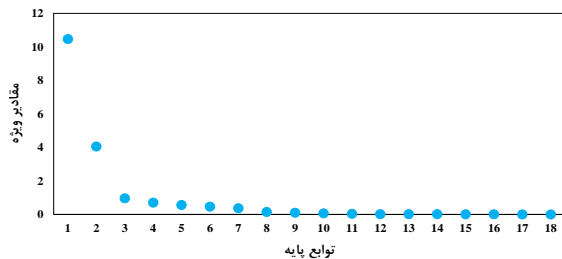
در این بخش با توجه به منابع فراوان در زمینه مدل ANN (۱۵) و همچنین جهت رعایت اختصار در مطالب، به معرفی مولفه‌های اصلی ANN استفاده شده در این تحقیق پرداخته شده است. ANN روشی هوشمند است که الگویی مناسب برای مدل‌سازی پدیده‌های پیچیده و غیرخطی را فراهم می‌نماید. در میان انواع مختلف ANN معرفی شده تا به حال، شبکه عصبی پیش‌خور (FFNN)^۱ با الگوریتم آموزش پس‌انتشار با عملکرد مناسب به عنوان تقریب‌زن جهانی شناخته

2-Levenberg-Marquardt (LM)

3-Tangent Sigmoid

1-Feed-Forward Neural Network (FFNN)

تابع پایه حدود ۱۰/۵ می‌باشد که بزرگ‌ترین مقدار در بین مقادیر ویژه محاسبه شده است. این موضوع بدین معناست که تابع پایه اول از اهمیت بیش‌تری نسبت به دیگر توابع پایه برخوردار است و بخش قابل توجهی از انرژی موجود در اطلاعات اولیه را مطابق شکل ۲ دربر می‌گیرد (حدود ۵۸ درصد). همچنین مقدار ویژه به دست آمده برای تابع پایه دوم نیز حدود ۴ بدست آمده که عددی بزرگ‌تر از مقادیر ویژه به دست آمده برای توابع بعد از خود است. بنابراین تابع پایه دوم از دیدگاه اهمیت در جایگاه دوم اهمیت قرار دارد به نحوی که معادل ۲۲ درصد انرژی داده‌های اصلی را شامل می‌شود. در نهایت می‌توان مشاهده نمود که آخرین مقدار ویژه که بیان‌گر اهمیت آخرین تابع ویژه است عملاً نزدیک به صفر می‌باشد. بنابراین با توجه به مطالب ذکر شده مشخص است که تنها تعداد اندکی از توابع پایه به دست آمده عملاً درصد بسیار بالایی از انرژی داده‌های اصلی را شامل می‌شوند. بنابراین در ادامه این پژوهش تنها از هفت تابع پایه متعامد اول به دست آمده که حدود ۹۸ درصد انرژی سیستم را پایسته نگه می‌دارند به عنوان بهترین مجموعه ورودی به مدل ANN استفاده خواهد شد. در این صورت با کاهش بردارهای ورودی به مدل از ۱۸ به ۷ مولفه، فضای بردارهای ورودی به مدل ANN به طور قابل توجهی کاهش می‌یابد. همچنین با حذف توابع پایه ۸ تا ۱۸، عملاً ۲ درصد از انرژی سیستم از دست خواهد رفت که این مساله با توجه به درصد اندک و ناچیز ذکر شده تاثیر مهمی بر عملکرد مدل شبکه عصبی نخواهد داشت.



شکل ۱- مقادیر ویژه محاسبه شده
Figure 1. Calculated eigenvalues

نیز از روش الگوریتم توقف آموزش (STA)^۱ پیشنهاد شده توسط کولیبالی و همکاران (۱۷) استفاده شد. در این روش عملاً متغیرهای ورودی به FFNN به سه دسته آموزش، اعتبارسنجی (مجموع آموزش و اعتبارسنجی تحت عنوان واسنجی شناخته می‌شوند) و صحت‌سنجی تقسیم می‌شوند. اطلاعات دسته آموزش برای بهینه نمودن پارامترهای شبکه، صحت‌سنجی برای ارزیابی عملکرد FFNN آموزش داده شده و اعتبارسنجی نیز برای توقف فرآیند آموزش FFNN در شرایطی که روند افزایشی خطا در مرحله آموزش شبکه زیاد گردد، استفاده می‌شود. STA، به دلیل توقف روند آموزش به محض برآورده شدن معیار تعریف شده غیرصفر، امکان استفاده از ساختارهای پیچیده مدل را فراهم نموده و در عین حال از مشکل فوق برآزشی و کم تخمینی در مدل نیز جلوگیری می‌کند.

نتایج و بحث

نتایج POD

همان‌طور که ذکر گردید در این تحقیق جمعاً ۱۸ متغیر به عنوان بردارهای ورودی به مدل ANN برای تخمین دبی یک ماه بعد استفاده شد. با توجه به تعداد زیاد ورودی‌ها به مدل ANN، در این تحقیق استفاده از روش POD برای کاهش فضای بردارهای ورودی استفاده شد. برای اجرای برنامه POD در گام اول ماتریس C بین ۱۸ متغیر ورودی مطابق رابطه (۱۳) تشکیل گردید. این ماتریس، ماتریسی مربعی و متقارن با ابعاد معادل متغیرهای ورودی (۱۸ بعد) می‌باشد. در گام بعد برای محاسبه توابع پایه متعامد (Φ)، لازم است مقادیر ویژه و بردارهای ویژه ماتریس C بین متغیرهای ورودی محاسبه گردد. با معلوم شدن مقادیر ویژه و بردارهای ویژه متناظر با آن‌ها، اطلاعات لازم برای تشکیل توابع پایه مطابق رابطه (۱۱) فراهم می‌گردد. با اجرای روند ذکر شده مقادیر ویژه ماتریس C محاسبه گردید که نتایج آن در شکل ۱ آمده است. مطابق این شکل مشخص است که مقدار ویژه به دست آمده برای اولین

بنابراین بر مبنای نتایج جدول ۱ می‌توان دقت مدل پیشنهادی ANN برای مدل سازی جریان یک ماه بعد ورودی به مخزن سد علویان را مطلوب ارزیابی نمود.

شکل ۳ به خوبی عملکرد مدل ANN را نشان می‌دهد. مطابق این شکل، مقادیر پیش‌بینی شده توسط مدل در مقابل مقادیر مشاهداتی در هر دو مرحله واسنجی و صحت‌سنجی نشان داده شده است. از این شکل مشخص است که خوش‌بختانه مدل ANN به خوبی آموزش داده شده به نحوی که مشکل کم‌تخمینی و بیش‌تخمینی در آن مشاهده نمی‌شود.

جدول ۱- نتایج مراحل واسنجی و صحت‌سنجی مدل‌های

ANN و POD-ANN

Table 1. Results of ANN and POD-ANN models during the calibration and verification steps

مرحله صحت‌سنجی	مرحله واسنجی	آماره	مدل
۰/۸۲	۰/۷۹	R^2	ANN
۱/۳۸	۱/۵۰	MAE	
۱/۱۰	۰/۹۵	AARE	
۰/۹۳	۰/۹۲	R^2	POD-ANN
۰/۷۹	۰/۹۴	MAE	
۰/۵۴	۰/۵۰	AARE	

در نهایت نتایج روند تغییرات دبی مشاهده‌ای در مقابل دبی پیش‌بینی شده توسط مدل ANN برای هر دو مرحله واسنجی و صحت به ترتیب در شکل‌های ۴ و ۵ نشان داده شده است. مطابق این شکل‌ها می‌توان مشاهده نمود که اگرچه مقادیر پیش‌بینی شده توسط مدل ANN نزدیک به مقادیر مشاهده‌ای هستند اما عملکرد مدل در برخی نقاط با دبی بالا در هر دو مرحله واسنجی و صحت‌سنجی با خطای قابل توجهی همراه است.

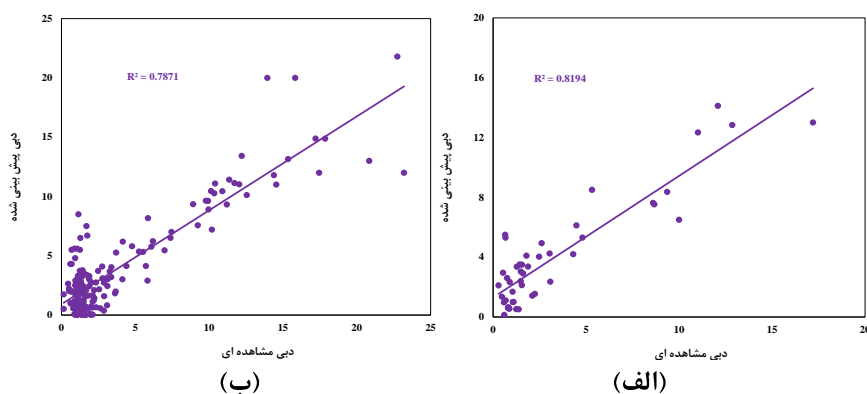


شکل ۲- درصد انرژی ذخیره شده توسط هر تابع پایه

Figure 2. Conserved percentage of energy by each basis function

نتایج ANN

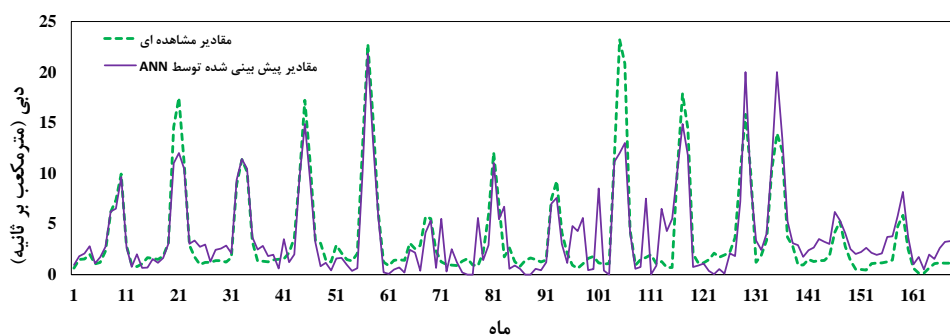
در این تحقیق حدود ۶۰ درصد اطلاعات (۱۳۳ ماه) برای آموزش، ۲۳ درصد برای اعتبارسنجی (۴۶ ماه) و ۱۷ درصد باقی‌مانده (۳۶ ماه) نیز برای مرحله صحت‌سنجی مدل ANN مورد استفاده قرار گرفت. در مرحله بعد با اجرای شبکه برای دفعات زیاد بهترین مدل ANN اجرا شده که مطابق آماره‌های مورد بررسی بالاترین دقت را داشت، به عنوان بهترین مدل پیش‌بینی جریان یک ماه بعد ورودی به مخزن سد علویان پیشنهاد گردید. در این راستا مدل ANN اجرا شده با ترکیب ۱۴ نرون در لایه پنهان به عنوان بهترین مدل در این مرحله انتخاب گردید که نتایج اجرای این مدل در مراحل واسنجی و صحت‌سنجی در جدول ۱ و همچنین شکل‌های ۳ تا ۵ نشان داده شده است. مطابق نتایج به دست آمده از مدل ANN بهینه اجرا شده مندرج در جدول ۱ می‌توان مشاهده نمود که مقدار آماره R^2 این مدل در مرحله واسنجی ۰/۷۹ و در مرحله صحت‌سنجی مدل نیز برابر ۰/۸۲ است. این واقعیت بیان‌گر دقت قابل قبول مدل است. همچنین آماره‌های MAE و AARE نیز برای هر دو مرحله واسنجی و صحت‌سنجی مدل بهینه اجرا شده مقادیر مناسبی هستند و به طور کلی هرچه این مقادیر نزدیک به صفر باشند بیان‌گر خطای کم‌تر مدل است.



شکل ۳- مقادیر پیش‌بینی شده دبی توسط ANN در مقابل مقادیر مشاهده‌ای آن در مراحل (الف) واسنجی و (ب)

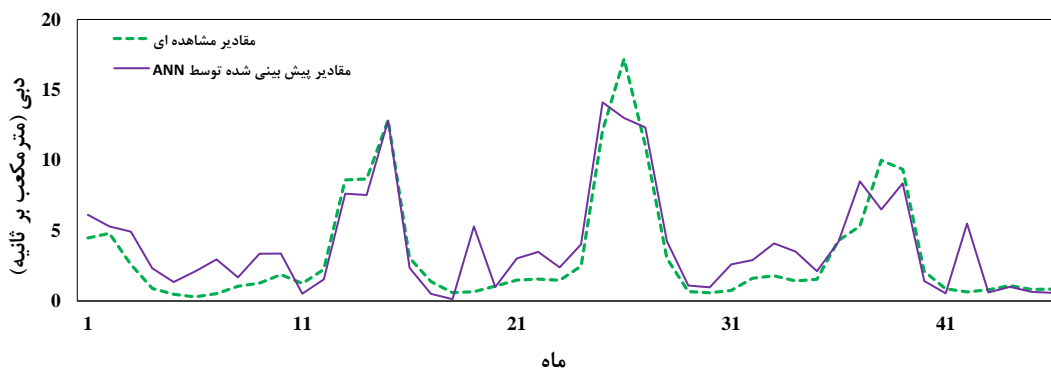
صحت‌سنجی

Figure 3. Predicted discharge by ANN vs observed values in (a) calibration and (b) verification steps



شکل ۴- مقادیر پیش‌بینی شده دبی توسط ANN در مقابل مقادیر مشاهده‌ای آن در مرحله واسنجی

Figure 4. Predicted discharge by ANN vs observed values in calibration step



شکل ۵- مقادیر پیش‌بینی شده دبی توسط ANN در مقابل مقادیر مشاهده‌ای آن در مرحله صحت‌سنجی

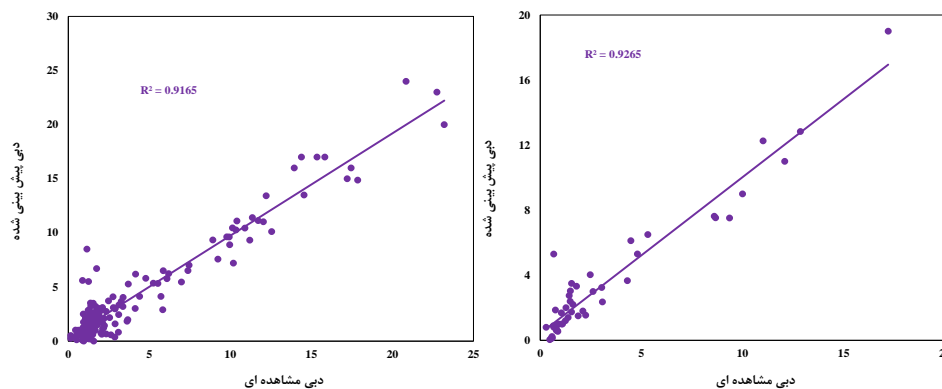
Figure 5. Predicted discharge by ANN vs observed values in verification step

نتایج ANN با فضای ورودی بهینه شده

هواشناسی مورد استفاده در مرحله قبل باشند، شامل ۷ تابع متعامد سره هستند که در اصل ترکیب خطی از همان متغیرهای اولیه هیدرولوژیکی-هواشناسی می‌باشند. در گام بعد

در این مرحله تعداد بردارهای ورودی به مدل ANN از ۱۸ متغیر به ۷ تابع متعامد کاهش داده شد. در واقع ورودی‌های مدل در این مرحله به جای این که متغیرهای هیدرولوژیکی-

صحت‌سنجی نیز برابر 0.93 است. این واقعیت بیان‌گر دقت مناسب مدل در هر دو مرحله است. در مقایسه با مدل ANN می‌توان مشاهده نمود که مقدار آماره R^2 در مدل POD-ANN در هر دو مرحله واسنجی و صحت‌سنجی بهبود قابل توجهی داشته است. همچنین آماره‌های MAE و AARE نیز از بهبود قابل توجهی در مقایسه با نتایج مشابه به دست آمده برای مدل ANN برخوردارند. بنابراین می‌توان دقت مدل پیشنهادی POD-ANN را نسبت به مدل ANN برای مدل‌سازی جریان یک ماه بعد ورودی به مخزن سد علویان بهتر از عملکرد مدل ANN ارزیابی نمود. مطابق شکل ۷ مشکل کم‌تخمینی و بیش‌تخمینی در مدل POD-ANN، همانند مدل ANN، وجود ندارد. همچنین دقت مدل نیز نسبت به مدل ANN در هر دو مرحله واسنجی و صحت‌سنجی مطلوب‌تر است. این واقعیت به خوبی از مقادیر نزدیک دبی پیش‌بینی شده توسط مدل به دبی مشاهده‌ای قابل‌رویت است.



شکل ۶- مقادیر پیش‌بینی شده دبی توسط POD-ANN در مقابل مقادیر مشاهده‌ای آن در مراحل (الف) واسنجی و (ب) صحت‌سنجی

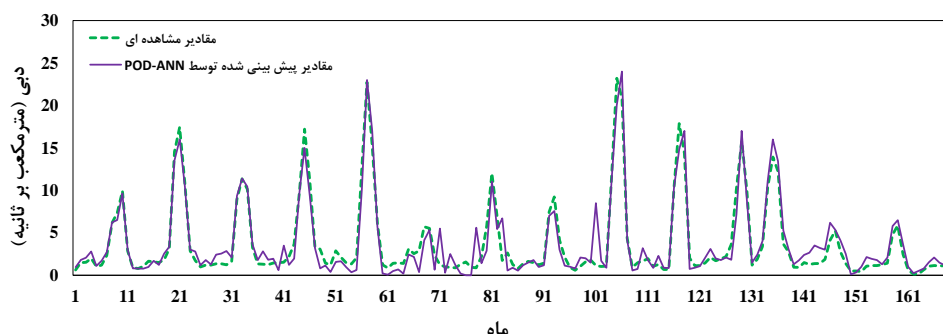
Figure 6. Predicted discharge by POD-ANN vs observed values in (a) calibration and (b) verification steps

است، ضعف مدل ANN در برآورد دبی‌های با مقادیر بالا را نیز به خوبی پوشش می‌دهد. این مهم به می‌تواند به دلیل عمل پیش‌پردازش بر روی متغیرهای ورودی به مدل ANN و کاهش تعداد ورودی‌ها در مدل POD-ANN در مقایسه با مدل ANN باشد. بنابراین می‌توان مشاهده نمود که عمل پیش‌پردازش بر روی متغیرهای ورودی به مدل ANN و کاهش تعداد متغیرهای ورودی به این مدل همراه با بهبود عملکرد آن بوده است.

مانند مرحله قبل حدود ۶۰ درصد اطلاعات (۱۳۳ ماه) برای آموزش، ۲۳ درصد برای اعتبارسنجی (۴۶ ماه) و ۱۷ درصد باقی‌مانده (۳۶ ماه) نیز برای مرحله صحت‌سنجی مدل ANN مورد استفاده قرار گرفتند. در این مرحله نیز با اجرای شبکه برای دفعات زیاد بهترین مدل POD-ANN اجرا شده که مطابق آماره‌های مورد بررسی بالاترین دقت را داشت به عنوان بهترین مدل پیش‌بینی جریان یک ماه بعد ورودی به مخزن سد علویان پیشنهاد گردید. در این راستا مدل POD-ANN اجرا شده با ترکیب ۴ نرون در لایه پنهان به عنوان بهترین مدل در این مرحله انتخاب گردید که نتایج اجرای آن در مراحل واسنجی و صحت‌سنجی در جدول ۱ و همچنین شکل‌های ۶ تا ۸ نشان داده شده است.

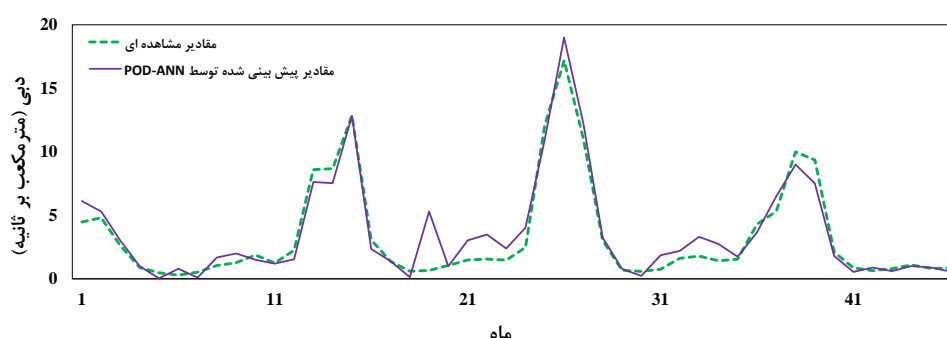
مطابق نتایج به دست آمده از مدل ترکیبی POD-ANN بهینه اجرا شده مندرج در جدول ۱ می‌توان مشاهده نمود که مقدار R^2 این مدل در مرحله واسنجی 0.92 و در مرحله

در نهایت با توجه به شکل‌های ۸ و ۹ می‌توان دریافت که دبی پیش‌بینی شده توسط مدل ترکیبی POD-ANN به خوبی روند دبی مشاهده‌ای را دنبال می‌کند و در بسیاری از نقاط عملاً تفاوت بسیار ناچیزی بین آن‌ها وجود دارد. همچنین مدل در هر دو مرحله واسنجی و صحت‌سنجی در دبی با مقادیر بالا از عملکرد مطلوبی برخوردار است. بنابراین مدل ترکیبی POD-ANN علاوه بر این که از دقت بالاتری نسبت به مدل ANN در پیش‌بینی جریان ماهانه ورودی به سد علویان برخوردار



شکل ۷- مقادیر پیش بینی شده دبی توسط POD-ANN در مقابل مقادیر مشاهده‌ای آن در مرحله واسنجی

Figure 7. Predicted discharge by POD-ANN vs observed values in calibration step



شکل ۸- مقادیر پیش بینی شده دبی توسط POD-ANN در مقابل مقادیر مشاهده‌ای آن در مرحله صحت‌سنجی

Figure 8. Predicted discharge by POD-ANN vs observed values in verification step

جمع بندی و نتیجه‌گیری

قابل توجهی از انرژی موجود در اطلاعات اولیه را دربر می‌گیرد (حدود ۵۸ درصد). همچنین مقدار ویژه به دست آمده برای تابع پایه دوم نیز حدود ۴ بدست آمد که عددی بزرگ‌تر از مقادیر ویژه به دست آمده برای توابع بعد از خود بود. بنابراین تابع پایه دوم از دیدگاه اهمیت در جایگاه دوم اهمیت قرار دارد به نحوی که معادل ۲۲ درصد انرژی سیستم را شامل می‌شد.

با توجه به نتایج به دست آمده از POD مشخص گردید • که تنها تعداد اندکی از توابع پایه به دست آمده عملاً درصد بسیار بالایی از انرژی داده‌های اصلی را شامل می‌شوند. بنابراین در این پژوهش تنها از هفت تابع پایه متعامد اول به دست آمده که حدود ۹۸ درصد انرژی سیستم را پایسته نگه می‌داشت به عنوان بهترین مجموعه ورودی به مدل ANN استفاده گردید.

با توجه به اهمیت سدها به عنوان یکی از اساسی‌ترین منابع تأمین آب شرب، کشاورزی، برق‌آبی و صنعتی و همچنین برآورد دبی ورودی به این سازه‌های عظیم، هدف اصلی این تحقیق نیز ارزیابی مدلی با استفاده از مدل ANN برای تخمین جریان ماهله رودخانه صوفی‌چای، ورودی به سد علویان، قرار داده شد. همچنین به منظور بهبود عملکرد مدل مذکور و با توجه به اطلاعات زیاد ورودی به این مدل، از روش POD به منظور تعیین بهترین ورودی به مدل ANN استفاده گردید. می‌توان مهم‌ترین رئوس نتایج به دست آمده از این تحقیق را در بندهای ذیل خلاصه نمود:

• مطابق نتایج روش POD مشخص گردید که مقدار ویژه به دست آمده برای اولین تابع پایه حدود ۱۰/۵ می‌باشد که بزرگ‌ترین مقدار در بین مقادیر ویژه محاسبه شده بود. این موضوع بدین معناست که تابع پایه اول از اهمیت بیش‌تری نسبت به دیگر توابع پایه برخوردار است و بخش

Reference

1. Noori, R., Karbassi, A.R., Moghaddamnia, A., Han, D., Zokaei-Ashtian, M.H., Farokhnia, A. (2011). Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology* 401(3): 177-189.
 2. Zhang, B., Govindaraja, R.S. (2000). Prediction of watershed runoff using Bayesian concepts and modular neural networks. *Water Resources Research* 36(3): 753-762.
 3. Nayak, P.C., Sudheer, K.P., Rangan, D.M., Ramasastri, K.S. (2004). A neuro-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology* 291: 52-66.
 4. Lin, J.Y., Cheng, C.T., Chau, K.W. (2006). Using support vector machines for long-term discharge prediction. *Hydrological Sciences Journal* 51(4): 599-612.
 5. He, Z., Wen, X., Liu, H., Du, J. (2014). A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *Journal of Hydrology* 509: 379-386.
 6. Noori, R., Asadi, N., Deng, Z. (2018). A simple model for simulation of reservoir stratification. *Journal of Hydraulic Research* 57(4):561-572. DOI:10.1080/00221686.2018.1499052
 7. Ravindran, S.S. (2000). A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. *Int. J. Numer. Methods Fluids* 34: 425-448.
- ارزیابی روند تغییرات دبی مشاهده‌ای در مقابل دبی پیش‌بینی شده توسط مدل ANN برای هر دو مرحله واسنجی و صحت‌سنجی مشخص نمود که اگرچه مقادیر پیش‌بینی شده توسط مدل نزدیک به مقادیر مشاهده‌ای بودند، اما عملکرد آن در نقاط با دبی بالا با خطای قابل توجهی همراه بود.
 - نتایج به دست آمده از مدل ترکیبی POD-ANN بهینه اجرا شده مشخص نمود که مقدار آماره‌های R^2 ، MAE و AARE در هر دو مرحله واسنجی و صحت‌سنجی بهبود قابل توجهی نسبت به مقادیر مشابه در مدل ANN داشتند.
 - نتایج روند تغییرات دبی مشاهده‌ای در مقابل دبی پیش‌بینی شده توسط مدل ترکیبی POD-ANN برای هر دو مرحله واسنجی و صحت‌سنجی نشان داد که این مدل از عملکرد بهتری نسبت به مدل ANN در نقاط با دبی بالا و همچنین در نقاط با دبی کم برخوردار است.
 - عملکرد مدل POD-ANN در دبی با مقادیر بالا نسبت به مدل ANN می‌تواند به دلیل عمل پیش‌پردازش بر روی متغیرهای ورودی و کاهش تعداد آن‌ها در مدل POD-ANN در مقایسه با مدل ANN باشد. بنابراین می‌توان نتیجه‌گیری نمود که عمل پیش‌پردازش بر روی متغیرهای ورودی به مدل ANN و کاهش تعداد متغیرهای ورودی به این مدل همراه با بهبود عملکرد آن بوده است.
- تشکر و قدردانی**
- تحقیق حاضر برگرفته از طرح پژوهشی مصوب دانشگاه آزاد اسلامی واحد اسلامشهر تحت عنوان "ارزیابی قابلیت روش تجزیه متعامد سره جهت تعیین ورودی به مدل شبکه عصبی برای پیش‌بینی جریان ماهانه ورودی به سد علویان" می‌باشد. بنابراین نگارندگان مراتب تشکر و قدردانی خود را از معاونت پژوهشی دانشگاه آزاد اسلامی واحد اسلامشهر جهت تأمین مالی این تحقیق ابراز می‌دارند.

13. Noori, R., Karbassi, A., Ashrafi, K., Ardestani, M., Mehrdadi, N. and Bidhendi, G.R.N. (2012). Active and online prediction of BOD5 in river systems using reduced-order support vector machine. *Environmental Earth Sciences* 67: 141-149.
14. Modabberi, A., Noori, R., Madani, K., Ehsani, A.H., Mehr, A.D., Hooshyaripor, F. and Kløve, B. (2020). Caspian Sea is eutrophying: the alarming message of satellite data. *Environmental Research Letters* 15(12):p.124047. DOI:10.1088/1748-9326/abc6d3.
15. Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*. Prentice Hall, New Jersey.
16. Noori, R., Khakpour, A., Omidvar, B. and Farokhnia, A. (2010). Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Systems with Applications* 37: 5856-5862.
17. Coulibaly, P., Ancti, F., Bobee, B. (2000). Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology* 230: 244-257.
8. Noori, R., Tian, F., Berndtsson, R., Abbasi, M.R., Naseh, M.V., Modabberi, A., Soltani, A., Kløve, B. (2019). Recent and future trends in sea surface temperature across the Persian Gulf and Gulf of Oman. *PloS One* 14(2):p.e0212790. DOI:10.1371/journal.pone.0212790.
9. Noori, R., Yeh, H.D., Ashrafi, K., Rezazadeh, N., Bateni, S.M., Karbassi, A., Kachoosangi, F.T., Moazami, S. (2015). A reduced-order based CE-QUAL-W2 model for simulation of nitrate concentration in dam reservoirs. *Journal of Hydrology* 530: 645-656.
10. Noori, R., Tian, F., Ni, G., Bhattarai, R., Hooshyaripor, F., Kløve, B. (2019). ThSSim: a novel tool for simulation of reservoir thermal stratification. *Scientific Reports* 9(1): 18524. DOI: 10.1038/s41598-019-54433-2.
11. Noori, R., Abbasi, M.R., Adamowski, J.F. and Dehghani, M. (2017). A simple mathematical model to predict sea surface temperature over the northwest Indian Ocean. *Estuarine, Coastal and Shelf Science* 197: 236-243.
12. Noori, R., Safavi, S. and Shahrokni, S.A.N. (2013). A reduced-order adaptive neuro-fuzzy inference system model as a software sensor for rapid estimation of five-day biochemical oxygen demand. *Journal of Hydrology* 495: 175-185.