

Uncertainty analysis of artificial neural network (ANN) and support vector machine (SVM) models in predicting monthly river flow (Case study: Ghezelozan River)

Majid Mohammadi ¹, Pouya Allahverdipour ^{2*}

¹ Expert, Regional Water Company of Zanjan, Zanjan, Iran

² M.Sc. Student, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

Abstract

Introduction

River flow forecasting has been one of the important challenges of water resources management in recent decades, so many researchers have proposed different methods to improve the performance of forecasting models. In the last decade, artificial intelligence methods have been most widely used in the simulation of various processes, including hydrological processes, due to their flexibility and high accuracy in modeling. However, the results of these methods have always been associated with uncertainty due to several factors such as structure, algorithm, input data, and the type of method chosen for data calibration. One of the methods that can somewhat solve this problem is the uncertainty analysis of the predictions made by these models.

Materials and Methods

In this study, the uncertainty of the results of artificial neural network (ANN) and support vector machine (SVM) models in predicting the monthly flow of the river has been evaluated. In this research, the time series of the monthly flow of the Ghezelozan River using the data of the Bianlu-Yasaul Stream gauging station in 39 years from 1976 to 2014 was used, where 75% and 25% of the data was used for training and testing the models, respectively. In these models, to estimate the monthly flow of the Ghezelozan River, six different input combinations including the flow of one, two, and three months before and the number of months of the flow were used. Then, the accuracy and performance of the models were compared using the coefficient of determination (R) and root mean square of errors (RMSE). Finally, the uncertainty of the results of ANN and SVM models in predicting the monthly flow of the river was evaluated by the Monte-Carlo method using d-factor and 95PPU values.

Results and Discussion

The evaluation of the performance of the ANN model shows that the best performance is related to the combination where the flow of the previous two months and the number of the month of the flow are the inputs of the model so that R and RMSE indexes were obtained as 0.757 and 9.45, respectively. In the SVM model for the monthly river flow series, the best performance is related to the combination where the flow of one, two, and three months ago and the number of months of the flow were the inputs of the model, and the R and RMSE indexes for this input pattern were 0.729 and 10.946, respectively. After studying the uncertainty of the models, the results showed that the ANN model has more uncertainty in the output values compared to the SVM model, and this is while the d-factor of the ANN model, both in the training and test phase, it was more than the SVM model. The comparison of the uncertainty analysis of the results of the ANN and SVM models showed that the SVM model with d-factor and 95PPU values equal to 0.155 and 17.241, respectively, compared to the ANN model with d-factor and 95PPU values equal to 0.993 and 85.470, respectively, has less uncertainty in the output values. So the number of observation data placed in the 95% confidence range (95PPU) of the ANN model compared to the SVM model has increased significantly in both the training and testing phases. Also, the results showed that both models have more uncertainty in the months with a large volume of water, which can be due to the complexity of the process and the involvement of uncertain factors in these months, as well as the effect of factors that are not considered in the structure of predictive models.

Conclusion

The results of ANN and SVM models in predicting the monthly flow of the Ghezeloan River showed that although the ANN model with R-value equal to 0.757 and RMSE value equal to 9.45 has a good performance compared to the SVM model with R-value equal to 0.729 and RMSE value equal to 10.946 in predicting the river flow, the results of this model are associated with high uncertainty. The comparison of the uncertainty analysis of the results of ANN and SVM models by Monte-Carlo method showed that the SVM model with d-factor and 95PPU values equal to 0.155 and 17.241, respectively, compared to the ANN model with d-factor and 95PPU values equal to 0.993 and 85.470, respectively, has less uncertainty in predicting the monthly flow of the Ghezeloan River and it is better than ANN model. According to the results of this research, taking into account the fact that advanced artificial intelligence models also have uncertainty, it is necessary to apply these methods in the fields of risk management and future planning to obtain the best performance.

Keywords: Artificial Neural Network, Flow prediction, Ghezeloan River, Support Vector Machine, Uncertainty

Article Type: Research Article

Acknowledgement

The authors gratefully acknowledge the Regional Water Company of Zanjan for providing the necessary data.

Conflicts of interest

The authors of this article declared no conflict of interest regarding the authorship or publication of this article.

Data availability statement

All data generated during the manuscript analysis are included in the article. Further datasets are available upon a reasonable request to the corresponding author.

Authors' contribution

Majid Mohammadi: Performing software/statistical analysis, writing the first version of the article; **Pouya Allahverdipour:** Performing software/statistical analysis, writing the first version of the article, manuscript editing, conceptualization, final editing

*Corresponding Author, E-mail: a.pouya1400@ms.tabrizu.ac.ir

Citation: Mohammadi, M., & Allahverdipour, P. (2024). Uncertainty analysis of artificial neural network (ANN) and support vector machine (SVM) models in predicting monthly river flow (Case study: Ghezeloan River). *Water and Soil Management and Modeling*, 4(2), 311-326.

DOI: 10.22098/mmws.2023.12702.1267

Received: 13 April 2023, Received in revised form: 05 May 2023, Accepted: 06 May 2023, Published online: 06 May 2023

Water and Soil Management and Modeling, Year 2024, Vol. 4, No. 2, pp. 311-326

Publisher: University of Mohaghegh Ardabil

© Author(s)





تحلیل عدم قطعیت مدل های شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) در پیش بینی جریان ماهانه رودخانه (مطالعه موردی: رودخانه قزل اوزن)

مجید محمدی^۱، پویا اللهویردی پور^{۲*}

^۱ کارشناس، شرکت سهامی آب منطقه ای زنجان، زنجان، ایران

^۲ دانشجوی کارشناسی ارشد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

چکیده

در یک دهه اخیر، روش های هوش مصنوعی بیشترین کاربرد را در شبیه سازی فرآیندهای مختلف از جمله فرآیندهای هیدرولوژیکی داشته اند، اما نتایج این روش ها همواره با عدم قطعیت همراه بوده اند. یکی از راه حل هایی که می تواند تا حدودی این مشکل را حل نماید، تحلیل عدم قطعیت پیش بینی های صورت گرفته است. در مطالعه حاضر عدم قطعیت نتایج مدل های شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) در پیش بینی جریان ماهانه رودخانه با استفاده از شبیه سازی مونت-کارلو و مقادیر 95PPU و d-factor مورد ارزیابی قرار گرفته است. در این پژوهش از داده ها و سری زمانی جریان ماهانه رودخانه قزل اوزن در یک دوره ۳۹ ساله از سال ۱۳۵۵ تا ۱۳۹۳ برای ایستگاه آب سنجی بیابلو-یساول استفاده شده است که ۷۵ درصد داده ها برای آموزش و ۲۵ درصد برای آزمون مدل ها به کار رفته است. در این مدل ها به منظور تخمین جریان رودخانه، شش ترکیب مختلف ورودی شامل جریان یک، دو و سه ماه قبل و شماره ماه های جریان مورد استفاده قرار گرفت. برای ارزیابی مدل ها از معیارهای آماری ضریب همبستگی (R) و ریشه میانگین مربعات خطا (RMSE) استفاده شد. نتایج نشان داد که اگر چه مدل ANN با مقادیر R مساوی با ۰/۷۵۷ و RMSE مساوی با ۹/۴۵ دارای عملکرد خوبی نسبت به مدل SVM با مقادیر R مساوی با ۰/۷۲۹ و RMSE مساوی با ۱۰/۹۴۶ در پیش بینی جریان رودخانه است. اما نتایج این مدل با عدم قطعیت زیادی همراه است. مقایسه تحلیل عدم قطعیت نتایج مدل ها نشان داد که مدل SVM با مقادیر d-factor و 95PPU به ترتیب برابر با ۰/۱۵۵ و ۱۷/۲۴۱ نسبت به مدل ANN با مقادیر d-factor و 95PPU به ترتیب برابر با ۰/۹۹۳ و ۸۵/۴۷۰ از عدم قطعیت کمتری برخوردار است و از این لحاظ بر مدل ANN برتری دارد. مطابق نتایج این پژوهش باید با در نظر گرفتن این نکته که مدل های پیشرفته هوش مصنوعی نیز دارای عدم قطعیت هستند، نسبت به کاربرد این روش ها در زمینه های مدیریت ریسک و برنامه ریزی های آینده اقدام کرد تا بهترین عملکرد را به دست آورد.

واژه های کلیدی: پیش بینی جریان، رودخانه قزل اوزن، شبکه عصبی مصنوعی، عدم قطعیت، ماشین بردار پشتیبان

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: a.pouya1400@ms.tabrizu.ac.ir

استناد: محمدی، مجید، و اللهویردی پور، پویا (۱۴۰۳). تحلیل عدم قطعیت مدل های شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) در پیش بینی جریان ماهانه رودخانه (مطالعه موردی: رودخانه قزل اوزن). *مدل سازی و مدیریت آب و خاک*، ۴(۲)، ۳۱۱-۳۲۶.

DOI: 10.22098/mmws.2023.12702.1267

تاریخ دریافت: ۱۴۰۲/۰۱/۲۴، تاریخ بازنگری: ۱۴۰۲/۰۲/۱۵، تاریخ پذیرش: ۱۴۰۲/۰۲/۱۶، تاریخ انتشار: ۱۴۰۲/۰۲/۱۶

مدل سازی و مدیریت آب و خاک، سال ۱۴۰۳، دوره ۴، شماره ۳، صفحه ۳۱۱ تا ۳۲۶

© نویسندگان

ناشر: دانشگاه محقق اردبیلی



۱- مقدمه

هستند. اما از نظر گرافیکی و تخمین دبی‌های کمینه و بیشینه، مشخص است که هر دو مدل در پیش‌بینی دبی بیشینه خطای بیش‌تری دارند اما در پیش‌بینی دبی متوسط و کمینه عملکرد مدل SVM به مراتب بهتر از شبکه‌ی بیزین است.

تحلیل عدم قطعیت مدل‌های هوشمند با توجه به این‌که پیش‌بینی‌ها قطعی نیستند، می‌تواند یکی از جنبه‌های مهم پژوهشی در مدل‌های مذکور باشد. دلیل اصلی عدم قطعیت در مدل‌های هوشمند، ساختار و الگوریتم آن‌ها است. مدل‌های هوشمند معمولاً به‌عنوان مدل‌های داده‌محور شناخته شده و عملکرد آن‌ها به‌شدت وابسته به روش انتخابی برای واسنجی است. بنابراین، این مدل‌ها همیشه حساس به داده‌های ورودی بوده و نتایج با تغییر روش واسنجی انتخابی، تغییر می‌کند (Noori et al., 2015). بنابراین شناخت و ارائه‌ی تحلیل عدم قطعیت به همراه مقادیر پیش‌بینی شده می‌تواند نقش مهمی در تصمیم‌گیری‌های حساس داشته باشد. با بررسی پژوهش‌های انجام گرفته در رابطه با مدل‌های ANN و SVM کمبود چنین پژوهش‌هایی در زمینه‌ی مطالعات منابع آب به‌وضوح دیده می‌شود. هم‌چنین، بررسی مراجع نشان می‌دهد که روش‌های محدودی برای تعیین عدم قطعیت در این مدل‌ها ارائه شده است که با بررسی این منابع، روش پیشنهاد شده توسط Marcé et al. (2004) به‌دلیل جدید بودن و عملکرد مناسب گزارش شده از آن، در این پژوهش مورد استفاده قرار گرفت. این روش بر مبنای قرارگیری مدل‌ها در قالب یک پروسه‌ی نمونه‌گیری تصادفی مونت-کارلو^۲ است که اولین بار برای برآورد میزان خروج سالانه‌ی مواد معدنی مغذی محلول در جریان رودخانه از حوضه، با استفاده از مدل نروفازی به‌کار گرفته شده است (Marcé et al., 2004; Farokhnia and Morid, 2010).

در مطالعه‌ی Aqil et al. (2007) با استفاده از روش مونت-کارلو، عدم قطعیت مقادیر خروجی مدل نروفازی را در پیش‌بینی جریان هفتگی رودخانه مورد بررسی قرار داده و آن را برای مدل نروفازی، مناسب گزارش نمودند. در ادامه، Akbarzadeh et al. (2010) به‌منظور تحلیل عدم قطعیت مدل‌های ANN و نروفازی تطبیقی^۸ (ANFIS) در پیش‌بینی ضریب انتشار طولی^۹ (LDC) رودخانه‌های طبیعی، از روش مونت-کارلو استفاده کردند. نتایج پژوهش آن‌ها بیان‌گر این مطلب بود که اگر چه مدل ANN در پیش‌بینی ضریب انتشار طولی رودخانه دارای عملکرد خوبی است، اما نتایج این مدل با

نیاز روزافزون به آب سبب شده است که برنامه‌ریزی‌های مدیریتی به‌منظور کنترل مصرف آب در آینده از اهمیت بسیار بالایی برخوردار باشد. با پیش‌بینی نمودن جریان رودخانه‌ها علاوه‌بر مدیریت بهره‌برداری از منابع آب، می‌توان حوادث طبیعی نظیر سیل و خشکسالی را نیز پیش‌بینی و مهار نمود. پیش‌بینی دبی رودخانه و نیز تحلیل عدم قطعیت^۱، یک موضوع پژوهشی پویا برای برنامه‌ریزی و مدیریت حوضه‌ها و ارزیابی ریسک در مسائل مربوط به آب است. در سال‌های اخیر استفاده از روش‌های هوشمند مانند شبکه‌های عصبی مصنوعی^۲ (ANN) و ماشین بردار پشتیبان^۳ (SVM) در علوم مهندسی آب و به‌ویژه پیش‌بینی جریان، گسترش چشم‌گیری داشته است. در این راستا، Wu et al. (2005) برای پیش‌بینی رواناب یک حوضه‌ی کوچک در آمریکا از مدل ANN استفاده کردند. این مدل در پیش‌بینی چندساعتی رواناب با بازه‌های زمانی ۱۵ دقیقه‌ای نتایج قابل‌قبولی ارائه داد. نتایج آن‌ها نشان داد که می‌توان از مدل ANN برای هشدار سیل در حوضه‌های شهری به‌صورت مفید و مؤثر استفاده نمود. در پژوهشی (Asefa et al., 2006) از مدل SVM با تابع کرنل خطی برای پیش‌بینی جریان‌های فصلی و ساعتی رودخانه‌ی سوئیتر^۴ در ایالات متحده استفاده کردند. آن‌ها در این پژوهش با توجه به حجم جریان معادل برف و حجم جریان در دوره‌های پیشین، میزان جریان در دوره‌های شش ماهه و ۲۴ ساعته را پیش‌بینی نمودند که نتایج مطالعه آن‌ها از دقت مطلوبی برخوردار بود. در یک پژوهش دیگری، Adamowski (2013) برای شبیه‌سازی بارش-رواناب در حوضه‌ی کوهستانی یوتارانچال هندوستان از دو مدل SVM و ANN را همراه با داده‌های محدود استفاده کرد. نتایج بررسی ایشان نشان از توانایی مدل SVM با تابع پایه‌ی شعاعی^۵ در پیش‌بینی رواناب مستقیم، جریان پایه و جریان کل این منطقه کوهستانی و پر از تپه داشت. هم‌چنین، Ahmadi et al. (2016) جهت پیش‌بینی جریان روزانه‌ی رودخانه‌ی باراندوزچای از دو روش SVM و شبکه‌های بیزین^۶ استفاده نمودند. برای ایجاد مدل از داده‌های جریان روزانه‌ی این رودخانه در بین سال‌های ۱۳۸۵ تا ۱۳۸۹ استفاده کردند. نتایج نشان داد که عملکرد روش SVM و شبکه‌ی بیزین از نظر تحلیل‌های آماری نزدیک به هم

¹ Uncertainty analysis

² Artificial neural network

³ Support vector machine

⁴ Sevier river basin

⁵ Radial base function (RBF)

⁶ Bayesian networks

⁷ Monte-Carlo

⁸ Adaptive neuro fuzzy inference system

⁹ Longitudinal dispersion coefficient

Etminan et al. (2023) به ارزیابی عدم قطعیت پارامترهای هیدرولیکی مدل هایدروس پرداختند. نتایج نشان داد روند تغییرات مقدار شاخص‌های p-factor و d-factor برای لایه‌ی سطحی ۸۳/۶ و ۰/۱۳ و برای لایه‌ی زیرین ۱۰ و ۰/۱۴ به ترتیب بیان‌گر افزایش عدم قطعیت پارامترها و ساختار مدل HYDRUS در برآورد جریان آب در خاک تحت روش آبیاری ستیریوت در امتداد پروفیل خاک بود. در نهایت، Mehdi Zare Anari et al. (2023) برای بررسی عدم قطعیت مدل‌های هوشمند برنامه‌نویسی بیان ژن^۵ (GEP)، اسپیلین رگرسیون تطبیقی چندمتغیره^۶ (MARS) و درخت مدل^۷ (MT) در برآورد جریان رودخانه کارون مطالعه‌ای انجام و نشان دادند که مدل MT با داشتن شاخص $d\text{-factor}=1.67$ و $95\text{PPU}=55.5$ درصد از لحاظ عدم قطعیت بهترین عملکرد را نسبت به سایر مدل‌ها دارد.

با توجه به موارد فوق در این پژوهش سعی شده است که با کاربرد عدم قطعیت بر اساس مفاهیم روش شبیه‌سازی مونت-کارلو به بررسی میزان دقت و کارایی دو روش داده محور؛ مدل ANN و مدل SVM با استفاده از ترکیبات مختلف ورودی، برای پیش‌بینی جریان ماهانه رودخانه قزل‌اوزن استان زنجان پرداخته و میزان دقت و کارایی این دو روش در فرآیند شبیه‌سازی دبی رودخانه ارزیابی شود.

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه و داده‌های مورد استفاده

سیستم رودخانه‌ای قزل‌اوزن که در شمال غرب کشور ایران واقع شده است، یکی از سیستم‌های رودخانه‌ای خزر است که در دوره‌های اخیر زمین‌شناسی دچار تحولات زیادی شده است. سرچشمه رودخانه قزل‌اوزن از ارتفاعات چهل چشمه استان کردستان بوده و با طولی بالغ بر ۵۵۰ کیلومتر پس از عبور از استان‌های زنجان، آذربایجان شرقی و اردبیل، ضمن دریافت شاخه‌های متعدد، با رودخانه شاهرود تلاقی و وارد مخزن سفیدرود می‌شود و وسعت حوزه آبخیز آن نزدیک به ۴۹۴۰۰ کیلومتر مربع است (Rezaei Moghaddam et al., 2011). این رودخانه بین استان کردستان، زنجان، آذربایجان شرقی، اردبیل، همدان و بخش کوچکی از استان‌های قزوین، آذربایجان غربی و گیلان قرار گرفته است.

عدم قطعیت زیادی همراه است. با مقایسه نتایج به دست آمده از تحلیل عدم قطعیت دو مدل ANN و ANFIS مشخص شد که مدل ANFIS نسبت به مدل ANN از عدم قطعیت کمتری برخوردار بوده و از این لحاظ بر مدل ANN برتری دارد. همچنین، Riahi Madvar et al. (2010) به منظور تعیین هندسه رژیم کانال از دو مدل ANN و سیستم استنتاج فازی عصبی تطبیقی استفاده نمودند. نتایج ارزیابی عدم قطعیت بر اساس شبیه‌سازی مونت-کارلو نشان داد که مدل سیستم استنتاج فازی عصبی تطبیقی عدم قطعیت کمتری نسبت به مدل ANN دارد. در مطالعه Farokhnia and Morid (2010) برای تحلیل عدم قطعیت نتایج مدل‌های ANN و ANFIS از روش مونت-کارلو در پیش‌بینی یک تا سه ماه آینده جریان رودخانه صوفی چای در محل ایستگاه تازه‌کند استفاده شد. نتایج نشان داد که مدل ANFIS از دقت بالاتر در پیش‌بینی و عدم قطعیت کمتری برخوردار است، اما به طور مشخص برای پرابی‌ها در هر دو مدل عدم قطعیت زیاد بوده که خطرپذیری و ریسک کاربرد نتایج را افزایش می‌دهد.

از سایر پژوهش‌ها، Seefi et al. (2013) به منظور پیش‌بینی تبخیر-تعرق^۱ (ET) لایسیمتری ایستگاه سینوپتیک کرمان از مدل ماشین بردار پشتیبان حداقل مربعات^۲ (LS-SVM) با سه نوع تابع کرنلی استفاده نمودند. نتایج نشان داد که با مدل‌های ANN و ANFIS، روش پنمن-مانتیت فائو^۳ و داده‌های لایسیمتری مقایسه شدند. از روش مونت-کارلو جهت ارزیابی عدم قطعیت مدل‌ها استفاده شد و نتایج حاکی از آن بود که مدل LS-SVM دارای عدم قطعیت کمتری نسبت به مدل‌های ANN و ANFIS است. در پیش‌بینی جریان ماهانه رودخانه زربنده رود نیز Ghorbani et al. (2016) عدم قطعیت مدل‌های MLP، RBF و SVM را مورد ارزیابی قرار دادند و با توجه به تعیین مقادیر 95PPU و $d\text{-factor}$ نتایج حاکی از آن بود که مدل SVM دارای عدم قطعیت کمتری نسبت به دو مدل دیگر است. در پژوهشی، Eskandari et al. (2019) برای ارزیابی عدم قطعیت مدل‌های ANN و ANFIS در تخمین جریان ورودی به سد رئیس‌علی دلواری از دو شاخص $d\text{-factor}$ و 95PPU استفاده کردند. مطابق نتایج آن‌ها مدل ANFIS با مقدار $d\text{-factor}$ کمتر و مقدار 95PPU بزرگ‌تر، از عدم قطعیت کمتری نسبت به مدل ANN برخوردار بود. همچنین،

¹ Evapotranspiration

² Least square support vector machine

³ FAO penman-monteith

⁴ 95 Percent of predicted uncertainties

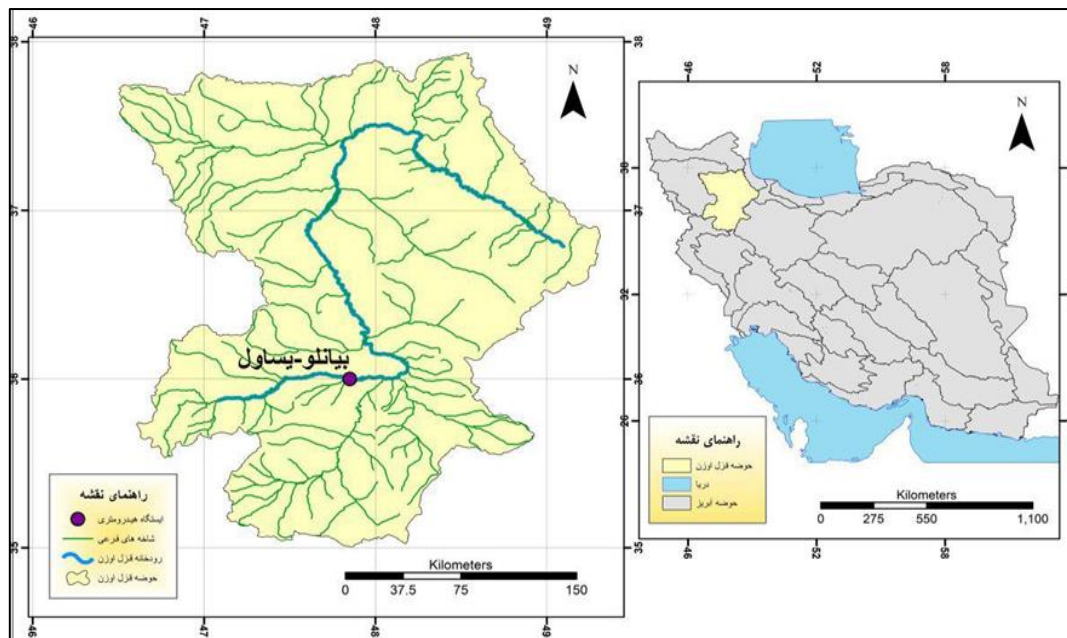
⁵ Gene expression programming

⁶ Multivariate adaptive regression splines

⁷ Model tree

شکل ۱ رودخانه قزل اوزن و موقعیت ایستگاه آب سنجی بیانلو-یساول واقع در استان زنجان را نشان می دهد. در جدول ۱ مشخصات جغرافیایی و آماری ایستگاه آب سنجی بیانلو-یساول ارائه شده است. شکل ۲ نیز نشان دهنده نمودار سری زمانی جریان ماهانه رودخانه قزل اوزن است.

در این مطالعه از داده های دبی جریان ماهانه ایستگاه آب سنجی بیانلو-یساول واقع در رودخانه قزل اوزن زنجان در دوره آماری ۱۳۹۳-۱۳۵۵ استفاده شد. ایستگاه آب سنجی بیانلو-یساول در عرض های جغرافیایی ۳۶ درجه و صفر دقیقه عرض شمالی و ۴۷ درجه و ۵۱ دقیقه طول شرقی قرار دارد.

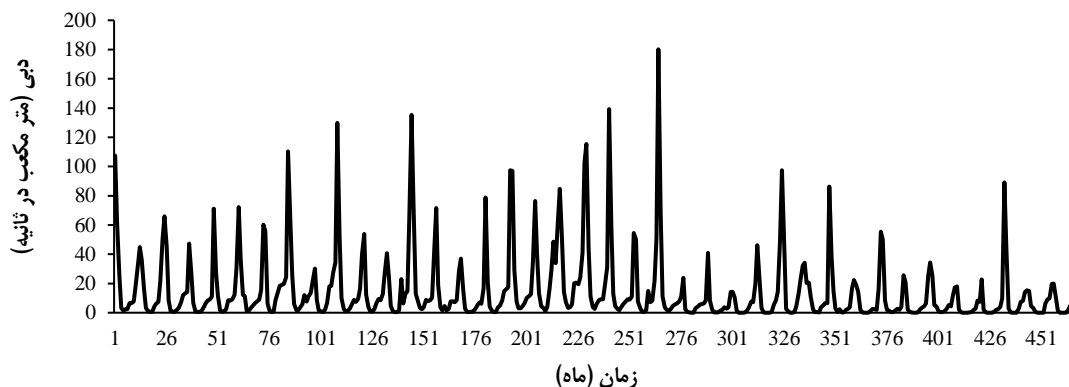


شکل ۱- موقعیت رودخانه قزل اوزن و ایستگاه آب سنجی بیانلو-یساول
Figure 1- location of Ghezelozan River and Bianlu-Yasaul Stream gauging station

جدول ۱- خلاصه مشخصات جغرافیایی ایستگاه آب سنجی بیانلو-یساول و ویژگی های آماری داده ها

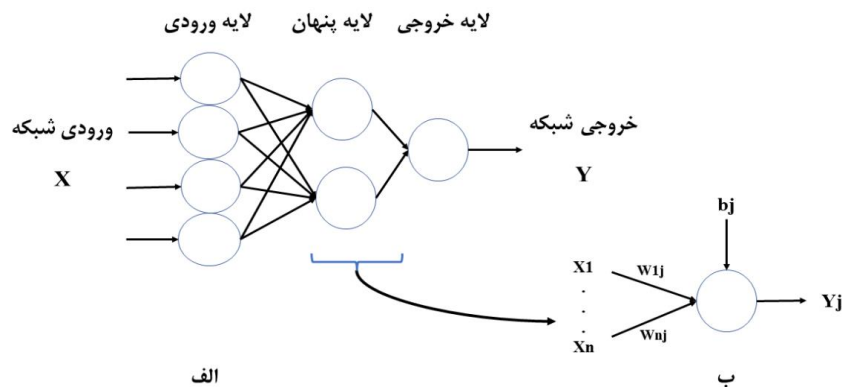
Table 1- Summary of geographic characteristics of the Bianlu-Yasaul water measuring station and statistical features of the data

ایستگاه هیدرومتری	دوره آماری	طول جغرافیایی	عرض جغرافیایی	تعداد داده ها	مشخصات آماری دبی ماهانه (متر مکعب بر ثانیه)
بیانلو-یساول	1355-1393	47° 51' 00"	36° 00' 00"	468	میانگین 14.832
					واریانس 558.137
					انحراف معیار 23.625
					بیشینه 180.13
					کمینه 0.00



شکل ۲- سری زمانی داده های جریان مشاهداتی رودخانه قزل اوزن در ایستگاه آب سنجی بیانلو-یساول
Figure 2- Time series of observational flow data of Ghezelozan River at Bianlu-Yasaul Stream gauging station

را نشان می‌دهد که در آن، $X(x_1, x_2, \dots, x_i, i=1, 2, \dots, n)$ بردار ورودی‌ها، W_{ij} وزن اتصالی از تأمین‌گره به تأمین‌گره در لایه بعد، b_j بایاس مربوط به گرّه j ام، y_i ها خروجی‌های هر لایه و Y خروجی نهایی ANN است. در این پژوهش از شبکه پرسپترون سه لایه، تابع تانژانت سیگموئیدی و محرک خطی برای نگاشت اطلاعات به ترتیب از لایه ورودی به لایه پنهان و از لایه پنهان به لایه خروجی و از الگوریتم پس انتشار لونیگ-مارکوارت استفاده شده است.



شکل ۳- شبکه‌های پیشروی سه لایه (الف) و شمای زامین گرّه (ب)
Figure 3- three-layer forward networks (a) and jth node scheme (b)

RBF به عنوان بهترین انتخاب از بین دیگر توابع کرنل گزارش شده است (Dibike et al., 2001). فرض می‌شود که تعداد N سری داده آموزشی بصورت $\{(x_i, d_i)\}_i^N$ موجود باشد. رابطه رگرسیون خطی که در فضای ویژگی با بعد بالا به داده‌ها برازش داده می‌شود، به صورت رابطه (۱) است:

$$y_i = f(x) = \langle W_i, \phi_i(x) \rangle + b \quad (1)$$

در آن، W_i و b ضرایب رابطه در فضای ویژگی در بعد بالا و $\phi_i(x)$ تابعی است که داده‌ها را از فضای ورودی به فضای ویژگی افراز می‌کند. مانند هر مدل دیگر جهت تخمین پارامترهای این مدل باید یک معیار خطا در نظر گرفته شود. در این رابطه توابع خطای مختلف مثل خطی، چندجمله‌ای، توانی و غیره استفاده می‌شود. از معیار $vapnik's \epsilon$ (رابطه ۲) برای بهینه‌سازی پارامترهای مدل فوق استفاده می‌شود.

$$L_\epsilon(v, g(u)) = \begin{cases} 0 & \text{for } |v - g(u)| \leq \epsilon \\ |d - g(u)| - \epsilon & \text{otherwise} \end{cases} \quad (2)$$

با در نظر گرفتن این معیار خطا، تخمین پارامترهای رابطه رگرسیونی به مسأله بهینه‌سازی تبدیل می‌شود که به صورت رابطه‌های (۳) و (۴) بیان می‌شود.

$$\min \frac{1}{2} \|W\|^T + C \sum_{i=1}^n \epsilon_i + \epsilon_i^* \quad (3)$$

۳-۲- ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان یکی از روش‌های یادگیری تحت نظارت است که هم برای طبقه‌بندی و هم رگرسیون قابل استفاده است. این روش توسط وپنیک در سال ۱۹۹۸ بر پایه تئوری یادگیری آماری بنا نهاده شده است. روش SVM روشی برای طبقه‌بندی دوتائی در فضای ویژگی‌های دلخواه است و از این‌رو، روشی مناسب برای مسائل پیش‌بینی به‌شمار می‌رود (Pai and Hong, 2007). مدل پیش‌بینی روشی نو مبتنی بر تئوری آماری بوده و از نگاشت غیرخطی برای آموزش داده‌ها استفاده می‌کند. در این حالت، الگوهای ورودی به فضایی با ابعاد بالا نگاشت می‌شوند. به طوری که می‌توان آن‌ها را به صورت خطی در فضای نگاشت شده پسرقت کرد. دقت رگرسیون بردار پشتیبان به کمینه کردن تابع خطا مربوط می‌شود. مدل SVM از توابع کرنل برای تغییر ابعاد مسأله برای حل مسائل غیرخطی استفاده می‌کند. انتخاب کرنل به مواردی نظیر اندازه داده‌های آموزشی و نیز ابعاد بردار ویژگی بستگی دارد. با توجه به این پارامترها باید تابع کرنلی استفاده شود که قابلیت آموزش ورودی‌های مسئله را داشته باشد (Allahverdipour and Sattari, 2023). معمولاً SVM با سه تابع کرنل پایه شعاعی (RBF)، چندجمله‌ای درجه d و خطی استفاده می‌شود (Kavzoglu and Colkesen, 2009) و تابع

در این روش ابتدا از اطلاعاتی که برای آموزش مدل در نظر گرفته شده‌اند به دفعات زیاد (معمولاً ۱۰۰۰ بار) بدون تکرار، نمونه‌گیری می‌شود. به نحوی که تعداد داده‌های آموزش و نظارت بر آموزش در هر نمونه یکسان باشد و برای هر نمونه، عملیات آموزش مدل انجام شود (Aqil et al., 2007). در نتیجه پارامترهای مدل به دفعات زیاد و هر بار بر اساس نمونه‌ای از وقایع اتفاق افتاده در گذشته به دست می‌آیند. در نتیجه با استفاده از این تکنیک محاسباتی حجیم به جای پیش‌بینی یک عدد مطلق، مطابق با آنچه در پژوهش‌های مشابه انجام گرفته، محدوده‌ای از خروجی که نتیجه عدم قطعیت مدل ساخته شده است، حاصل خواهد شد (Farokhnia and Morid, 2010).

برای بررسی عدم قطعیت، حدود اطمینان ۹۵ درصد (95PPU) برای پیش‌بینی‌ها در نظر گرفته می‌شود که طریقه به دست آوردن آن، تعیین حدود ۲/۵ درصد (XL) و ۹۷/۵ درصد (XU) توزیع تجمعی احتمال تجربی به دست آمده از ۱۰۰۰ پیش‌بینی انجام شده برای هر ماه است. حدود اطمینان مناسب حدودی هستند که بالای ۸۰ درصد از آبدهی مشاهده شده در آن حدود قرار گرفته باشند و دارای عرض متوسط قابل قبولی باشند (d-factor → 0). برای سنجش عرض متوسط باند اطمینان شاخص عامل عرض باند توسط Abbaspour et al. (2007) مطابق رابطه (۷) پیشنهاد شده است:

$$d - \text{factor} = \frac{\bar{d}_x}{\sigma_x}, \quad 0 \leq d - \text{factor} \leq \infty \quad (7)$$

در رابطه بالا، عامل σ_x انحراف استاندارد داده‌های مشاهده‌ای و \bar{d}_x عرض متوسط باند اطمینان است و به صورت رابطه (۸) تعریف می‌شود. درصد داده‌های قرار گرفته در باند اطمینان ۹۵ درصد نیز به صورت رابطه (۹) محاسبه می‌شود:

$$\bar{d}_x = \frac{1}{n} \sum_{i=1}^k (X_u - X_L)_i \quad (8)$$

Bracketed by 95PPU

$$= \frac{1}{k} \text{Count} (j | X_L^1 \leq X_{reg}^1 \leq X_U^1) \times 100 \quad (9)$$

در رابطه فوق، k تعداد داده‌های مشاهده‌ای، l به معنای ماه فعلی است که از یک تا k تغییر می‌کند. X_U^1 و X_L^1 حد پایین و بالای عدم قطعیت و X_{reg}^1 مقادیر ثبت شده در ماه فعلی l است. هر بار که متغیر ثبت شده برای ماه فعلی l محدوده عدم قطعیت را شامل می‌شود، یک واحد به z اضافه می‌شود و حداکثر مقدار z در $l=k$ است. اگر همه مقادیر ثبت شده در داخل باند عدم قطعیت قرار بگیرند، Bracketed by 95PPU=100 خواهد شد. به گزارش Abbaspour et al. (2007) اگر داده‌های اندازه‌گیری شده از کیفیت خوبی برخوردار باشند، باید ۱۰۰-۸۰ درصد آن‌ها سطح 95PPU را شامل شوند. اما در بعضی مناطق

$$\text{subject to} \begin{cases} V_i - (w \cdot \phi(u_i) + b) \leq \varepsilon + \Sigma_i \\ (w \cdot \phi(u_i) + b) - V_i \leq \varepsilon + \Sigma_i^* \\ \Sigma_i, \Sigma_i^* \geq 0, \quad i = 1, 2, \dots, n \end{cases} \quad (4)$$

در روابط مذکور Σ_i و Σ_i^* ضرایب کمبود (که حد بالا و پایین خطای آموزش را تعیین می‌کنند)، c پارامتر تنظیم‌کننده و ε حد بالای خطای خروجی هستند. با در نظر گرفتن ضرایب لاگرانژ و شرایط کان-تاکر، فرم دوگانه مسئله بهینه‌سازی بالا تبدیل به مسئله بهینه‌سازی برنامه‌ریزی چندجمله‌ای درجه دو می‌شود که با بهینه‌سازی آن ضرایب لاگرانژ a_i و \bar{a}_i بردار وزن‌های بهینه ابر صفحه رگرسیون محاسبه و در نهایت برای بردارهای پشتیبان، رابطه رگرسیون (رابطه ۵) بیان می‌شود.

$$g(x) = \sum_{i=1}^{nl} (a_i - \bar{a}_i) K < \phi(x_i), \phi(x_j) > + b \quad (5)$$

در آن، nl نشان‌دهنده تعداد بردار پشتیبان است. در رابطه (۵) ضرب داخلی در فضای ویژگی توسط تابع کرنل انجام می‌شود. یعنی اگر دو بردار $\phi(x_i)$ و $\phi(x_j)$ وجود داشته باشد، ضرب داخلی آن‌ها به وسیله تابع کرنل انجام می‌شود (Vapnik, 1995). در این پژوهش به منظور مدل‌سازی جریان ماهانه رودخانه قزل‌اوزن توسط مدل SVM تابع RBF مورد بررسی قرار گرفت و از طریق سعی و خطا، مقادیر بهینه پارامترهای مورد نیاز مدل ماشین بردار پشتیبان با این نوع تابع کرنل به دست آمد.

$$c = 2.227 \quad \varepsilon = 4.239 \quad (6)$$

۲-۴- برآورد عدم قطعیت در پیش‌بینی‌ها

یکی از مهم‌ترین اهداف پژوهش حاضر بررسی و تخمین عدم قطعیت در خروجی مدل‌های ANN و SVM است که بدین منظور روش پیشنهاد شده توسط Marcé et al. (2004) به دلیل جدید بودن و کاربرد زیاد آن، مینا قرار داده شده است. در این روش مدل‌های ANN و SVM در قالب یک نمونه‌گیری مونت-کارلو به کار گرفته می‌شوند. اساس شبیه‌سازی مونت-کارلو بر پایه استفاده از اعداد تصادفی و حصول تابع توزیع احتمالاتی خروجی یک مدل است. بدین صورت که متغیرها (ورودی به مدل) با استفاده از تابع توزیع احتمالاتی مربوط به طور تصادفی تولید شده و سپس خروجی معادل آن‌ها از مدل مورد نظر به دست می‌آید و این عمل بارها تکرار می‌شود. در مرحله بعد عدم قطعیت خروجی از طریق محاسبه شاخص‌های آماری و یا تعیین تابع توزیع احتمالاتی برآورد می‌شود (Eckhardt et al., 2003). تفاوت روش مورد بحث در این پژوهش با تعریف بالا، استفاده از نمونه‌های تصادفی گرفته شده به جای اعداد تصادفی ساخته شده است (Efron and Tibshirani, 1994).

توجه به این معیارها، مدلی خواهد بود که مقدار این معیارها برای آن، به ترتیب نزدیک به یک و صفر باشد.

$$R = \frac{\sum_{i=1}^N (Q_{xi} - \bar{Q}_x)(Q_{yi} - \bar{Q}_y)}{\sqrt{\sum_{i=1}^N (Q_{xi} - \bar{Q}_x)^2 \sum_{i=1}^N (Q_{yi} - \bar{Q}_y)^2}} \quad 0 \leq R \leq 1 \quad (10)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_{xi} - Q_{yi})^2} \quad 0 \leq RMSE \leq \infty \quad (11)$$

۳- مواد بحث و نتایج

۳-۱- نتایج مدل‌سازی جریان ماهانه رودخانه قزل‌اوزن با استفاده از مدل ANN

در مطالعه حاضر، برای مدل‌سازی جریان ماهانه رودخانه قزل‌اوزن با استفاده از مدل ANN، داده‌های ماهانه ۲۹ سال اول (۸۳-۱۳۵۵) برای آموزش و ۱۰ سال آخر (۱۳۸۴-۱۳۹۳) برای صحت‌سنجی مدل انتخاب شدند. به عبارت دیگر ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد برای صحت‌سنجی مدل در نظر گرفته شد. پارامترهای آماری مربوط به هر دو مرحله آموزش و آزمون در جدول ۲ برای ایستگاه بیانلو-یساول آمده است.

جدول ۲- مشخصات آماری سری جریان ماهانه رودخانه قزل‌اوزن در محل ایستگاه آب‌سنجی بیانلو-یساول

ایستگاه	مرحله	واحد	تعداد	حداکثر	حداقل	میانگین	انحراف معیار	ضریب تغییرات
بیانلو-یساول	آموزش	مترمکعب بر ثانیه	350	180.13	0.00	17.392	25.790	1.483
	تست	مترمکعب بر ثانیه	118	89.196	0.00	7.152	12.677	1.772

جدول ۴- الگوهای ورودی مورد استفاده در مقیاس زمانی ماهانه رودخانه قزل‌اوزن

Table 4- The input patterns used in the monthly time scale of Ghezelozan River

شماره ترکیب	تعداد ورودی	ترکیب ورودی	خروجی
1	1	Q(t-1)	Q(t)
2	2	Q(t-1), Month	Q(t)
3	2	Q(t-1), Q(t-2)	Q(t)
4	3	Q(t-1), Q(t-2), Month	Q(t)
5	4	Q(t-1), Q(t-2), Q(t-3), Month	Q(t)
6	3	Q(t-1), Q(t-2), Q(t-3)	Q(t)

جدول ۵ نشان‌دهنده نتایج ارزیابی الگوهای مختلف ورودی مدل ANN برای مدل‌سازی جریان ماهانه رودخانه قزل‌اوزن در مراحل آموزش و صحت‌سنجی است. با توجه به جدول می‌توان گفت، ترکیب شماره دو با $R=0.757$ و $RMSE=9/45$ دارای کم‌ترین خطا و بیش‌ترین دقت در مرحله صحت‌سنجی بوده و به‌عنوان بهترین الگو برای مدل ANN انتخاب می‌شود. شکل‌های ۴ و ۵ به ترتیب نشان‌دهنده مقادیر مشاهداتی و مقادیر

که داده‌ها از کیفیت مناسبی برخوردار نیستند باید فقط ۵۰ درصد داده‌ها سطح 95PPU را شامل شوند. به‌طورکلی باند 95PPU حد فاصل صفر تا ۱۰ درصد و d-factor نیز مابین صفر تا بی‌نهایت تغییر می‌کند. بدین ترتیب که مقادیر بی‌شتر معیار 95PPU به‌همراه مقادیر کمتر پارامتر d-factor نشان از عدم قطعیت کمتر مدل و عملکرد مناسب آن است (Mehdizadeh, Zare Anari et al., 2023).

۲-۵- برآورد معیارهای ارزیابی مدل‌ها

پس از تحلیل داده‌ها با انواع روش‌ها و مدل‌ها، لازم است عملکرد آن‌ها ارزیابی شود. برای این منظور در این پژوهش از جمله معیارهای پرکاربرد در مهندسی آب شامل ضریب همبستگی (R) و ریشه میانگین مربعات خطا (RMSE) استفاده و به ترتیب از طریق رابطه‌های (۱۰) و (۱۱) برآورد شدند. در این روابط، پارامترهای Q_{xi} و Q_{yi} به ترتیب مقادیر دبی مشاهداتی و محاسباتی در ماه t ام، N تعداد داده‌ها، \bar{Q}_x و \bar{Q}_y نیز میانگین مقادیر دبی مشاهداتی و محاسباتی می‌باشند. دقیق‌ترین مدل با

یکی از مهم‌ترین مشکلات در انتخاب متغیرهای ورودی، زمانی است که مدل‌های پیش‌بینی در حال توسعه هستند. بنابراین، تحلیل همبستگی متقاطع بین متغیرهای ورودی و خروجی به‌منظور تعیین تأخیر زمان مناسب و تعداد متغیرهای ورودی انجام می‌شود که در جدول ۳ ارائه شده است (Luk et al., 2000). جدول ۴ نیز نشان‌دهنده الگوهای مختلف ورودی است. با توجه به این جدول از نقش شماره ماه‌های جریان و حافظه ارتباط دبی با جریان ماه‌های ماقبل در مدل‌سازی جریان استفاده شده است. در این جداول پارامتر Month شماره ماه فعلی، $Q(t)$ جریان ماه فعلی، $Q(t-1)$ جریان یک ماه قبل، $Q(t-2)$ جریان دو ماه قبل و $Q(t-3)$ جریان سه ماه قبل است.

جدول ۳- همبستگی بین جریان ماهانه با جریان ماه‌های قبل
Table 3- Correlation between the monthly flow and the flow of previous months

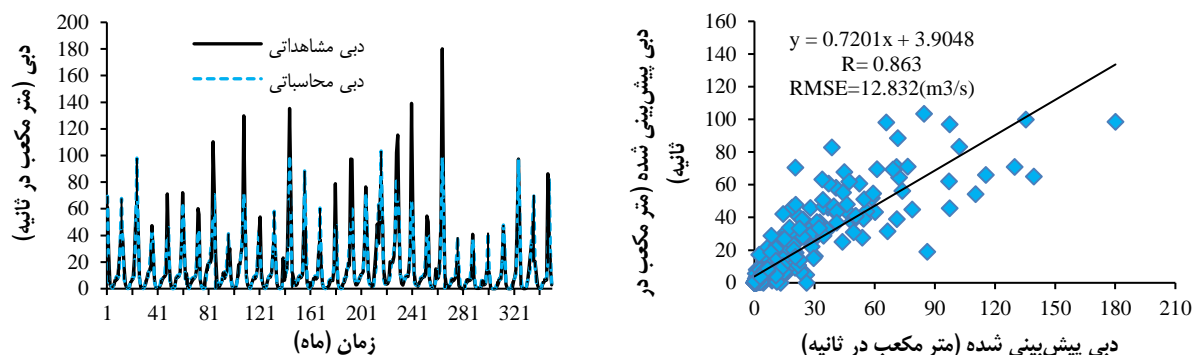
Month	Q(t-3)	Q(t-2)	Q(t-1)
Q(t)	-0.277	-0.060	0.138

پیش‌بینی شده برای الگوی بهینه ترکیب شماره دو سری ماهانه جریان رودخانه قزل‌اوزن در مرحله آموزش و صحت‌سنجی است.

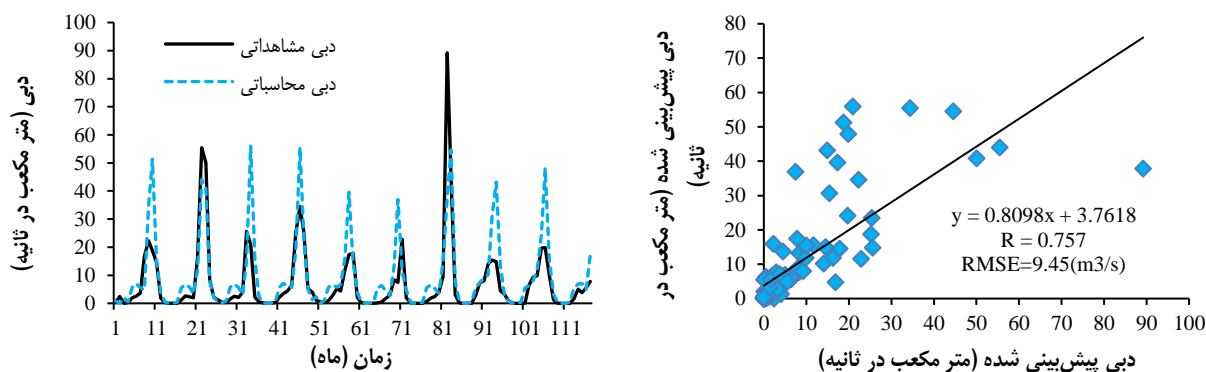
جدول ۵- نتایج مدل ANN در تخمین جریان ماهانه رودخانه قزل‌اوزن

Table 5- The results of the artificial neural network model in the estimation of the monthly flow of the Ghezeloan River

صحت‌سنجی				ساخار	شماره ترکیب	ایستگاه
R	RMSE (مترمکعب بر ثانیه)	R	RMSE (مترمکعب بر ثانیه)			
0.5734	10.824	0.612	20.08	1-5-1	1	بیانلو-یساول
0.757	9.45	0.863	12.832	2-15-1	2	
0.656	11.30	0.728	17.39	2-8-1	3	
0.765	9.87	0.84	14.32	3-8-1	4	
0.783	11.88	0.809	15.27	4-2-1	5	
0.673	9.844	0.712	18.035	3-5-1	6	



شکل ۴- نمودارهای مقادیر مشاهداتی و پیش‌بینی شده حاصل از بهترین مدل ANN در مرحله آموزش
Figure 4- Plots of observed and predicted values resulting from the best ANN model in the training phase



شکل ۵- نمودارهای مقادیر مشاهداتی و پیش‌بینی شده حاصل از بهترین مدل ANN در مرحله صحت‌سنجی
Figure 5- Plots of observed and predicted values resulting from the best ANN model in the test phase

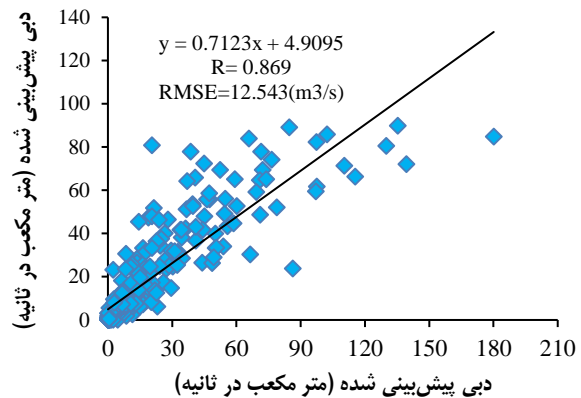
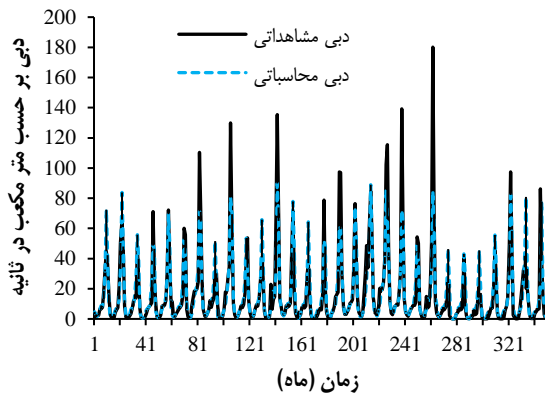
مراحل آموزش و صحت‌سنجی را نشان می‌دهد. با توجه به جدول می‌توان نتیجه گرفت که ترکیب شماره پنج با $R=0.729$ و $RMSE=10.946$ با کمترین خطا در مرحله صحت‌سنجی به عنوان بهترین الگو برای مدل SVM انتخاب می‌شود. در شکل های ۶ و ۷ نیز به ترتیب نمودار مقادیر مشاهداتی و مقادیر پیش‌بینی شده دوره‌های آموزش و صحت‌سنجی مدل SVM ارائه شده است.

۲-۳- نتایج مدل‌سازی جریان ماهانه رودخانه قزل‌اوزن با استفاده از مدل SVM

به منظور مدل‌سازی جریان ماهانه رودخانه قزل‌اوزن توسط مدل SVM در پژوهش حاضر، تابع کرنل پایه شعاعی مورد بررسی قرار گرفت. در این مدل مانند مدل ANN، ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد برای صحت‌سنجی در نظر گرفته شد. الگوهای ورودی مدل SVM نیز مانند مدل ANN انتخاب شد تا هر دو مدل به کار رفته از شرایط یکسانی در اطلاعات دریافتی برخوردار باشند. جدول ۶ نتایج ارزیابی الگوهای مختلف ورودی ماشین بردار پشتیبان برای مدل‌سازی جریان ماهانه رودخانه در

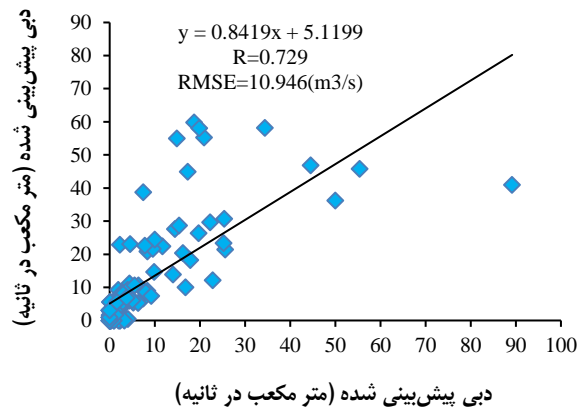
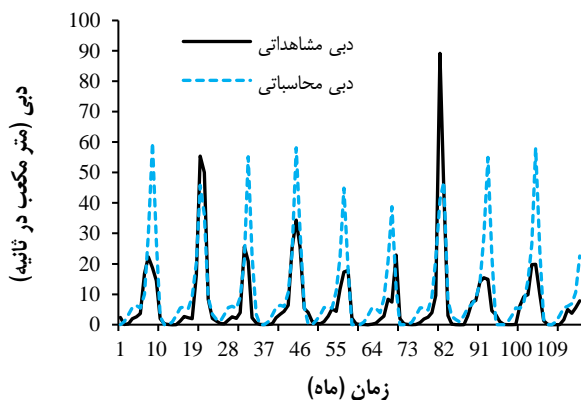
جدول ۶- نتایج مدل SVM در تخمین جریان ماهانه رودخانه قزل‌اوزن (RMSE بر حسب متر مکعب بر ثانیه)
 Table 6- The results of the SVM model in estimating the monthly flow of the Ghezeloan River (RMSE in $m^3 s^{-1}$)

صحت‌سنجی		آموزش		شماره ترکیب	ایستگاه
R	RMSE	R	RMSE		
0.56	11.414	0.621	19.851	1	بیانو یساوول
0.754	11.388	0.859	12.965	2	
0.628	11.406	0.737	17.096	3	
0.736	11.210	0.867	12.626	4	
0.729	10.946	0.869	12.543	5	
0.605	11.353	0.757	16.515	6	



شکل ۶- نمودارهای مقادیر مشاهداتی و پیش‌بینی شده حاصل از بهترین مدل SVM در مرحله آموزش

Figure 6- Plots of observed and predicted values resulting from the best Support Vector Machine (SVM) model in the training phase



شکل ۷- نمودارهای مقادیر مشاهداتی و پیش‌بینی شده حاصل از بهترین مدل SVM در مرحله صحت‌سنجی

Figure 7- Plots of observed and predicted values resulting from the best Support Vector Machine (SVM) model in the test phase

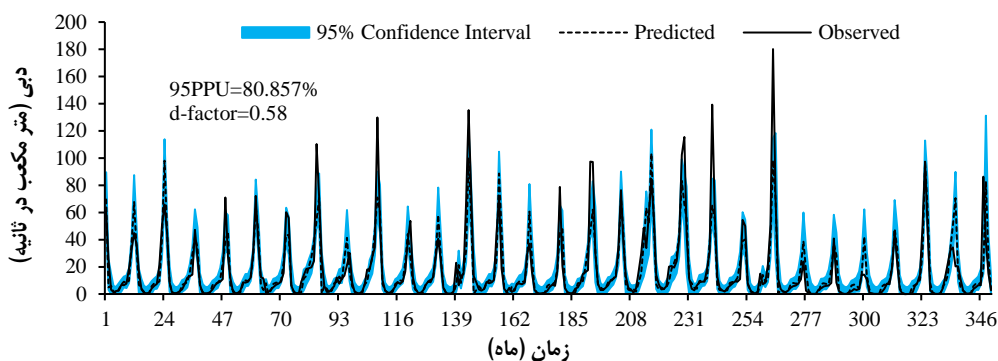
شماره دو است که جریان دو ماه قبل و شماره ماه جریان به عنوان ورودی‌های مدل هستند. به طوری که R و RMSE به ترتیب ۰/۷۵۷ و ۹/۴۵ به دست آمده است. در مدل SVM برای سری جریان ماهانه رودخانه، بهترین عملکرد مربوط به ترکیب شماره پنج است که جریان یک، دو و سه ماه قبل و شماره ماه های جریان به عنوان ورودی‌های مدل بودند و معیارهای R و RMSE برای این الگوی ورودی به ترتیب ۰/۷۲۹ و ۱۰/۹۴۶ محاسبه شد.

با توجه به نتایج به دست آمده مشاهده می‌شود که دو مدل SVM و ANN از نظر تحلیل‌های آماری عملکرد تقریباً یکسانی دارند اما در بررسی تخمین دبی‌های بیشینه و کمینه، مشخص است که هر دو مدل در پیش‌بینی دبی بیشینه خطای بیشتری دارند اما در پیش‌بینی دبی متوسط و کمینه، مدل SVM به مراتب عملکرد بهتری نسبت به مدل ANN داشته است. ارزیابی عملکرد مدل ANN بر مبنای دو معیار R و RMSE نشان می‌دهد که بهترین عملکرد مربوط به ترکیب

۳-۳- تحلیل عدم قطعیت مدل‌ها

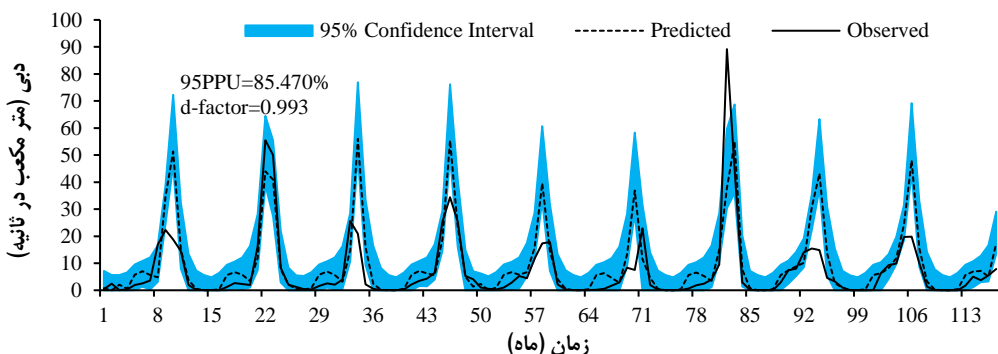
در این بخش ساختار بهینه هر یک از مدل‌ها که در قسمت قبل تعیین شده بودند، مورد استفاده قرار گرفته و تحلیل عدم قطعیت بر اساس شبیه‌سازی مونت-کارلو برای مدل‌های ANN و SVM در پیش‌بینی جریان ماهانه رودخانه انجام گرفت. پس از آموزش مدل‌ها و به‌دست آوردن ۱۰۰۰ مقدار برای پیش‌بینی هر ماه، حدود اطمینان ۹۵ درصد برای پیش‌بینی مدل‌های ANN و

SVM از طریق تعیین مقادیر ۲/۵ درصد و ۹۷/۵ درصد توزیع احتمالاتی تجربی به‌دست آمده برای هر ماه مشخص شد. حدود اطمینان ۹۵ درصد مدل‌های ANN و SVM در مراحل آموزش و صحت‌سنجی به‌ترتیب در شکل‌های ۸، ۹، ۱۰ و ۱۱ و شاخص‌های عدم قطعیت در جدول ۷ ارائه شده‌اند.



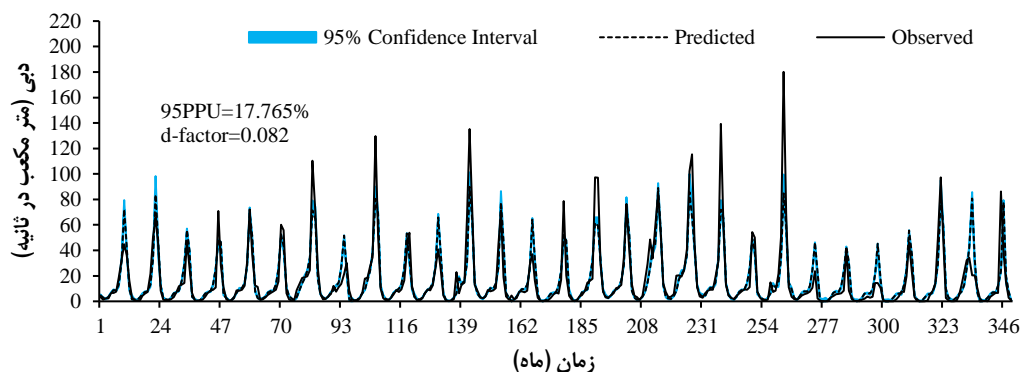
شکل ۸- مقادیر مشاهداتی و پیش‌بینی شده حاصل از مدل ANN و باند اطمینان ۹۵ درصد در مرحله آموزش

Figure 8- The observed and predicted values obtained from the ANN model and the 95% confidence interval in the training phase



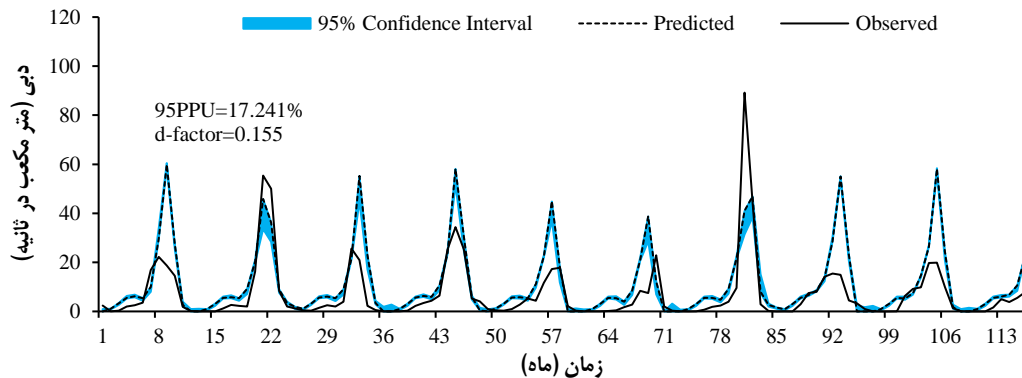
شکل ۹- مقادیر مشاهداتی و پیش‌بینی شده حاصل از مدل ANN و باند اطمینان ۹۵ درصد در مرحله صحت‌سنجی

Figure 9- The observed and predicted values obtained from the ANN model and the 95% confidence interval in the test phase



شکل ۱۰- مقادیر مشاهداتی و پیش‌بینی شده حاصل از مدل SVM و باند اطمینان ۹۵ درصد در مرحله آموزش

Figure 10- The observed and predicted values obtained from the SVM model and the 95% confidence interval in the training phase



شکل ۱۱- مقادیر مشاهداتی و پیش‌بینی شده حاصل از مدل SVM و باند اطمینان ۹۵ درصد در مرحله آزمایش

Figure 11- The observed and predicted values obtained from the SVM model and the 95% confidence interval in the test phase

جدول ۷- شاخص‌های عدم قطعیت مدل‌های ANN و SVM مربوط به مراحل آموزش و صحت‌سنجی

Table 7- Uncertainty indices of ANN and SVM models in the training and test phase

صحت‌سنجی	آموزش	شاخص خطا	مدل	ایستگاه
0.993	0.58	d-factor	ANN	بیانلو - یساول
85.470	80.85	95PPU		
0.155	0.082	d-factor	SVM	
17.241	17.765	95PPU		

با توجه به شکل‌های ۸، ۹، ۱۰ و ۱۱ و جدول ۷، مقادیر اندازه‌گیری شده با معیار 95PPU نشان می‌دهد که حدود ۸۵/۴۷۰ درصد مقادیر داده‌ها در مدل ANN و ۱۷/۲۴۱ درصد مقادیر داده‌ها در مدل SVM در مرحله صحت‌سنجی برای ایستگاه بیانلو-یساول در محدوده اطمینان قرار گرفته‌اند. این موضوع نشان می‌دهد که نتایج پیش‌بینی مدل ANN نسبت به مدل SVM از لحاظ فاصله اطمینان مناسب خروجی‌ها، از کیفیت مناسبی برخوردار است. همچنین، معیار ضریب d-factor دارای مقادیر ۰/۹۹۳ و ۰/۱۵۵ به ترتیب برای مدل ANN و SVM در مرحله صحت‌سنجی است. در حالت کلی مقادیر بسیار کوچک d-factor حاکی از عدم قطعیت بسیار پایین مدل SVM است. با توجه به مقادیر 95PPU و d-factor مربوط به مرحله صحت‌سنجی، می‌توان نتیجه گرفت که مدل SVM در مدل سازی جریان رودخانه دارای عدم قطعیت کم‌تری نسبت به مدل ANN است.

پس از بررسی عدم قطعیت مدل‌ها، نتایج نشان داد که مدل ANN در مقایسه با مدل SVM دارای عدم قطعیت بیشتری در مقادیر خروجی است و این امر در حالی است که عرض متوسط محدوده اطمینان (d-factor) در مدل ANN، هم در مرحله آموزش و هم در مرحله تست نسبت به مدل SVM بیش تر بوده است. به طوری که تعداد داده‌های مشاهداتی قرار گرفته در محدوده اطمینان ۹۵ درصد (95PPU) مدل ANN نسبت به مدل SVM در هر دو مرحله آموزش و آزمون افزایش قابل توجهی داشته است.

۴- نتیجه‌گیری

تعیین جریان رودخانه‌ها یکی از مؤلفه‌های مهم و تأثیرگذار در مدیریت منابع آب حوضه آبریز است. در این پژوهش دو مدل ANN و SVM برای پیش‌بینی جریان ماهانه رودخانه قزل‌اوزن زنجان در دوره آماری ۱۳۵۵ تا ۱۳۹۳ برای ایستگاه آب‌سنجی بیانلو-یساول همراه با لحاظ نمودن عدم قطعیت به روش مونت-کارلو در محاسبات مربوطه، مورد بررسی قرار گرفت. برای پیش‌بینی بهتر جریان ماهانه رودخانه شش ترکیب ورودی مختلف به عنوان ورودی مدل‌ها در نظر گرفته شد. نتایج حاصل بیانگر این بود که مدل‌های ANN و SVM از عملکرد قابل قبولی در پیش‌بینی جریان ماهانه رودخانه قزل‌اوزن برخوردار هستند، اما در حالت کلی، مدل ANN با بیش‌ترین دقت و کم‌ترین خطا از عملکرد نسبتاً بهتری برخوردار بود.

نتایج این پژوهش نشان داد که هر دو مدل در ماه‌های پر آب دارای عدم قطعیت بیش‌تری می‌باشند که این امر می‌تواند به علت پیچیدگی فرآیند و دخالت عوامل غیرقطعی (تصادفی) در این ماه‌ها و تأثیر عواملی باشد که به علت کمبود اطلاعات در ساختار مدل‌های پیش‌بینی در نظر گرفته نشده‌اند. به طور کلی می‌توان نتیجه گرفت که نتایج به دست آمده از پژوهش حاضر بیانگر قابلیت مناسب مدل‌های هوشمند در پیش‌بینی جریان رودخانه‌ها است. همچنین، اگر چه مدل ANN در پیش‌بینی جریان رودخانه دارای عملکرد خوبی است اما نتایج این مدل با عدم قطعیت زیادی همراه است. مقایسه تحلیل عدم قطعیت نتایج

اطمینان، سمانه، جلالی، وحیدرضا، محمودآبادی، مجید، خاشعی سیوکی، عباس، پوررضا بیلندی، محسن (۱۴۰۱). ارزیابی عدم قطعیت پارامترهای هیدرولیکی مدل هایدروس با استفاده از روش DREAM. *مدل‌سازی و مدیریت آب و خاک*، ۳(۴)، ۱-۱۵. doi:10.22098/mmws.2022.11659.1152

اسکندری، علی، نوری، روح‌اله، وصالی ناصح، محمدرضا، و سعیدی، فریماه (۱۳۹۸). ارزیابی عدم قطعیت مدل‌های ANN و ANFIS در تخمین جریان ورودی به سد رئیس‌علی دلواری. *علوم و تکنولوژی محیط‌زیست*، ۲۱(۷)، ۳۱-۴۷. doi:10.22034/JEST.2020.20068.2909

الهویردی پور، پویا، و ستاری، محمدتقی (۱۴۰۲). مقایسه عملکرد روش کلاسیک رگرسیون خطی چندگانه و روش‌های داده‌کاوی نوین در مدل‌سازی بارش سالانه (مطالعه موردی: شهر اهواز). *مدل‌سازی و مدیریت آب و خاک*، ۳(۲)، ۱۲۵-۱۴۲. doi:10.22098/mmws.2022.11337.1120.۱۴۲

رضایی مقدم، محمد حسین، ثروتی، محمد رضا، و اصغری سراسکانرود، صیاد (۱۳۹۰). بررسی مقایسه‌ای الگوی پیچانرود با استفاده از تحلیل هندسه فراکتالی و شاخص‌های زاویه مرکزی و ضریب خمیدگی (مطالعه موردی: رودخانه قزل‌اوزن). *پژوهشنامه مدیریت حوضه آبخیز*، ۲(۳)، ۱-۱۸.

http://jwmr.sanru.ac.ir/article-1-145-fa.html
سیفی، اکرم، میرلطیفی، مجید، و ریاحی، حسین (۱۳۹۲). معرفی و کاربرد ماشین بردار پشتیبان حداقل مربعات در برآورد تبخیر-تعرق مرجع و تحلیل عدم قطعیت نتایج؛ مطالعه موردی: شهر کرمان. *مهندسی آبیاری و آب ایران*، ۴(۱۳)، ۶۷-۷۹.

https://www.waterjournal.ir/article_70860.html
فرخ‌نیا، اشکان، و مرید، سعید (۱۳۸۸). تحلیل عدم قطعیت مدل‌های شبکه عصبی و نروفازی در پیش‌بینی جریان رودخانه. *تحقیقات منابع آب ایران*، ۵(۳)، ۱۴-۲۷.

https://www.iwrr.ir/article_15771.html
مهدی‌زاده زارع اناری، یاسر، نجف‌زاده، محمد، و انوری، صدیقه (۱۴۰۱). بررسی عدم قطعیت مدل‌های هوشمند در برآورد جریان رودخانه (مطالعه موردی: رودخانه کارون). *مهندسی عمران / امیرکبیر*، ۵۴(۱۱)، ۴۳۸۳-۴۳۹۶. doi:10.22060/ceej.2022.21208.7653

مدل‌ها نشان داد که مدل SVM نسبت به مدل ANN از عدم قطعیت کم‌تری برخوردار است و از این لحاظ بر مدل ANN برتری دارد. نتایج این پژوهش می‌تواند در زمینه‌های مدیریت ریسک و برنامه‌ریزی‌های آینده کمک‌کننده باشد. پیشنهاد می‌شود در پژوهش‌های آتی عملکرد سایر مدل‌ها نیز بررسی و عدم قطعیت آن‌ها در مقایسه با یک‌دیگر سنجیده شود.

سپاسگزاری

نویسندگان از شرکت آب منطقه‌ای استان زنجان جهت همکاری در ارائه داده‌های لازم، صمیمانه تشکر می‌کنند.

تضاد منافع نویسندگان

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تضاد منافی در خصوص نگارش و انتشار مطالب و نتایج این پژوهش وجود ندارند.

دسترسی به داده‌ها

داده‌ها و نتایج استفاده شده در این پژوهش از طریق مکاتبه با نویسنده مسئول در اختیار قرار خواهد گرفت.

مشارکت نویسندگان

مجید محمدی: انجام تحلیل‌های نرم‌افزاری/آماري، نگارش نسخه اولیه مقاله؛ پویا الهویردی پور: انجام تحلیل‌های نرم‌افزاری/آماري، مفهوم‌سازی، ویرایش و بازبینی مقاله، کنترل نتایج.

منابع

احمدی، فرشاد، رادمنش، فریدون، و میرعباسی نجف‌آبادی، رسول (۱۳۹۴). مقایسه عملکرد روش‌های ماشین بردار پشتیبان و شبکه‌های بیزین در پیش‌بینی جریان روزانه رودخانه (مطالعه موردی: رودخانه باراندوزچای). *پژوهش‌های حفاظت آب و خاک*، ۲۲(۶)، ۱۷۱-۱۸۶. dor: 20.1001.1.23222069.1394.22.6.10.2

اکبرزاده، عباس، نوری، روح‌اله، فرخ‌نیا، اشکان، خاکپور، امیر، و صباحی، محمد سلمان (۱۳۸۹). تحلیل دقت و عدم قطعیت مدل‌های هوشمند در پیش‌بینی ضریب انتشار طولی رودخانه‌ها. *آب و فاضلاب*، ۲۱(۳)، ۹۹-۱۰۷. https://www.wjjournal.ir/article_1288.html

Hydrology, 333(2-4), 413-430.

doi:10.1016/j.jhydrol.2006.09.014

Adamowski, J. (2013). Using support vector regression to predict direct runoff, base flow and total flow in a mountainous watershed with limited data in Uttaranchal, India. *Annals*

References

Abbaspour, K.C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., & Srinivasan, R. (2007). Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of*

- of Warsaw University of Life Sciences-SGGW. *Land Reclamation*, 45(1), 71-83. doi: 10.2478/ssgw-2013-0007
- Ahmadi, F., Radmanesh, F., & Mirabbasi Najafabadi, R. (2016). Comparing the performance of support vector machines and bayesian networks in predicting daily river flow (Case study: Baranduz Chai River). *Journal of Water and Soil Conservation*, 22(6), 171-186 doi: 20.1001.1.23222069.1394.22.6.10.2. [In Persian]
- Akbarzadeh, A., Noori, R., Farokhnia, A., Khakpour, A., & Sabahi, M.S. (2010). Accuracy and uncertainty analysis of intelligent techniques for predicting the longitudinal dispersion coefficient in rivers. *Journal of Water and Wastewater*, 21(3), 99-107. [In Persian]
- Allahverdipour, P., & Sattari, M.T. (2023). Comparing the performance of the multiple linear regression classic method and modern data mining methods in annual rainfall modeling (Case study: Ahvaz city). *Water and Soil Management and Modeling*, 3(2), 125-142 doi:10.22098/mmws.2022.11337.1120. [In Persian]
- Aqil, M., Kita, I., Yano, A., & Nishiyama, S. (2007). Analysis and prediction of flow from local source in a river basin using a Neuro-fuzzy modeling tool. *Journal of Environmental Management*, 85(1), 215-223. doi:10.1016/j.jenvman.2006.09.009
- Asefa, T., Kemblowski, M., McKee, M., & Khalil, A. (2006). Multi-time scale stream flow predictions: The support vector machines approach. *Journal of Hydrology*, 318(1-4), 7-16. doi:10.1016/j.jhydrol.2005.06.001
- Dibike, Y.B., Velickov, S., Solomatine, D., & Abbott, M.B. (2001). Model induction with support vector machines: introduction and applications. *Journal of Computing in Civil Engineering*, 15(3), 208-216. doi: 10.1061/(ASCE)0887-3801(2001)15:3(208)
- Eckhardt, K., Breuer, L., & Frede, H.G. (2003). Parameter uncertainty and the significance of simulated land use change effects. *Journal of Hydrology*, 273(1-4), 164-176. doi: 10.1016/S0022-1694(02)00395-5
- Eskandari, A., Noori, R., Vesali Naseh, M.R., & Saeedi, F. (2019). Uncertainty evaluation of ANN and ANFIS Models in Inflow forecasting into the Raees-Ali Delvari Dam. *Journal of Environmental Science and Technology*, 21(7), 31-47. doi:10.22034/jest.2020.20068.2909 [In Persian]
- Etmiman, S., Jalali, V.R., Mamodabadi, M., Khashei-siuki, A., Pourreza-bilondi, M. (2023). Assessing the hydraulic parameter's uncertainty of the HYDRUS model using DREAM method. *Water and Soil Management and Modeling*, 3(4), 1-15. doi:10.22098/mmws.2022.11659.1152 [In Persian]
- Farokhnia, A., & Morid, S. (2010). Uncertainty analysis of artificial neural networks and neuro-fuzzy models in river flow forecasting. *Iran-Water Resources Research*, 5(3), 14-27. [In Persian]
- Ghorbani, M.A., Zadeh, H.A., Isazadeh, M., & Terzi, O. (2016). A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction. *Environmental Earth Sciences*, 75(6), 1-14. doi: 10.1007/s12665-015-5096-x
- Jain, S.K., Das, A., & Srivastava, D.K. (1999). Application of ANN for reservoir inflow prediction and operation. *Journal of Water Resources Planning and Management*, 125(5), 263-271. doi:10.1061/(ASCE)0733-9496(1999)125:5(263)
- Kavzoglu, T., & Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11(5), 352-359. doi:10.1016/j.jag.2009.06.002
- Luk, K.C., Ball, J.E., & Sharma, A. (2000). A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. *Journal of Hydrology*, 227(1-4), 56-65. doi: 10.1016/S0022-1694(99)00165-1
- Marcé, R., Comerma, M., García, J.C., & Armengol, J. (2004). A neuro-fuzzy modeling tool to estimate fluvial nutrient loads in watersheds under time-varying human impact. *Limnology and Oceanography: Methods*, 2(11), 342-355. doi:10.4319/lom.2004.2.342
- Mehdizadeh Zare Anari, Y., Najafzadeh, M., & Anvari, S. (2023). The Uncertainty Analysis of Artificial Intelligence Models in Forecasting River Flow (Case Study: Karun River). *Amirkabir Journal of Civil Engineering*, 54(11), 4383-4396. doi:10.22060/ceej.2022.21208.7653 [In Persian]
- Noori, R., Yeh, H.D., Abbasi, M., Kachooangi, F. T., & Moazami, S. (2015). Uncertainty analysis of support vector machine for online prediction of five-day biochemical oxygen

- demand. *Journal of Hydrology*, 527, 833-843. doi:10.1016/j.jhydrol.2015.05.046
- Pai, P.F., & Hong, W.C. (2007). A recurrent support vector regression model in rainfall forecasting. *Hydrological Processes*, 21(6), 819-827. doi:10.1002/hyp.6323
- Rezaei Moghadam, M.H., Sarvati, M.R., & Asghari Sareskanroodi, S. (2011). Compared Investigation Meandering Pattern by Fractal Geometrical Analysis and Central Angles and Sinuosity Ratio Indices. *Journal of Watershed Management*, 2(3), 1-18. [In Persian]
- Riahi, H., & Ayyoubzadeh, S.A. (2010). Uncertainty analysis of ANN and ANFIS techniques in comparison with regime equations for determining regime channel geometry. In Fifth International Conference on Water Resources and Environment Research, Quebec City, Canada.
- Seefi, A., Mirlatifi, M., & Reahi, H. (2013). Introduction and application of least square support vector machine (LSSVM) for simulation of reference evapotranspiration and uncertainty analysis of results, A case study of the Kerman City. *Irrigation and Water Engineering*, 4(1), 67-79. [In Persian]
- Efron, B., & Tibshirani, R.J. (1994). *An Introduction to the Bootstrap*. 1st Edition: Chapman and Hall/CRC, 456 pages. doi:10.1201/9780429246593
- Vapnik, V.N. (1995). *The Nature of Statistical Learning Theory*. 1st Edition: Springer, New York, NY, 188 pages. doi:10.1007/978-1-4757-2440-0
- Wu, J.S., Han, J., Annambhotla, S., & Bryant, S. (2005). Artificial neural networks for forecasting watershed runoff and stream flows. *Journal of Hydrologic Engineering*, 10(3), 216-222. doi:10.1061/(ASCE)1084-0699(2005)10:3(216)