

تعیین کیفیت آب رودخانه بابلرود با استفاده از الگوریتم‌های فراابتکاری

فاطمه آخونی پورحسینی¹، کیومرث ابراهیمی^{2*}، محمد حسین امید³

دانشجوی دکتری مهندسی منابع آب، گروه مهندسی آبیاری و آبادانی، دانشگاه تهران، کرج، ایران-1
Email: Fateme.Pourhosini@ut.ac.ir

استاد گروه مهندسی آبیاری و آبادانی، دانشگاه تهران، کرج، ایران (نویسنده مسئول)-2
Email: EbrahimiK@ut.ac.ir

استاد گروه مهندسی آبیاری و آبادانی، دانشگاه تهران، کرج، ایران-3
Email: MOmid@ut.ac.ir

چکیده

آلودگی آب در دهه‌های اخیر تهدید مهمی برای انسان و اکوسیستم‌های طبیعی تلقی می‌شود. در این پژوهش جهت تعیین کیفیت آب رودخانه بابلرود از روش هیبریدی SVM-PSO استفاده شد. به منظور برآورد کیفیت آب رودخانه بابلرود، در ایستگاه قرآن تالار در بازه زمانی 1394-1345 داده‌های ماهیانه کلسیم، منیزیم، بی‌کربنات، سدیم، سولفات، pH، EC و TDS تجزیه و تحلیل شد. ارزیابی دو مدل SVM-PSO و SVM، توسط معیارهای ضریب تبیین (R^2)، جذر میانگین مربعات خطا (RMSE) و درصد میانگین مطلق خطا (MAPE) نشان داد که مدل SVM-PSO با بیشترین ضریب تبیین برابر با 98/0 و کمترین جذر میانگین مربعات خطای 86/5 و کمترین درصد میانگین مطلق خطا 02/0 در مقایسه با ماشین بردار پشتیبان از کارایی بهترین برخوردار است.

واژه‌های کلیدی: بهینه سازی، کیفیت منابع آب، مدل هوشمند، واسنجی

مقدمه

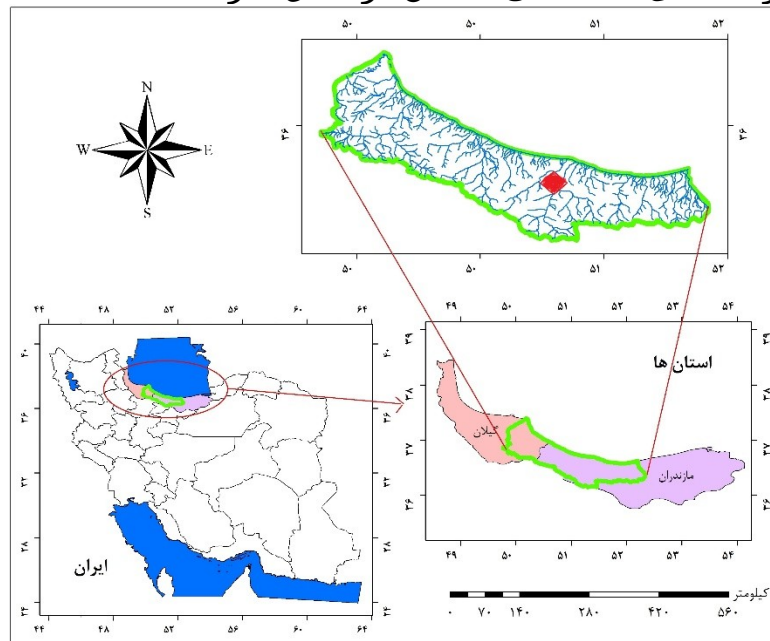
حفاظت و استفاده بهینه از منابع آب از اصول توسعه پایدار یک کشور می‌باشد. آب‌های سطحی و رودخانه‌ها از مهم‌ترین منابع آب هستند که نقش مهمی در تامین نیازهای مختلف دارند[1]. کل مواد جامد محلول یا TDS در آب برابر مجموع غلظت همه یون‌های موجود در آب می‌باشد. تخمین و ارزیابی کیفیت آب و مقدار TDS به منظور حفاظت منابع آب، همچنین بررسی و جلوگیری از آلودگی آن‌ها بسیار حائز اهمیت است[2]... بنابراین ارزیابی مستمر پارامترهای شیمیایی، فیزیکی و بیولوژیکی آب بخش ضروری برنامه‌های کنترل کیفیت آب می‌باشد. از زمانی که داده‌های کیفیت آب از طریق نمونه‌برداری و تجزیه و تحلیل جمع‌آوری شد، نیاز به بیان به شکل قابل فهم احساس می‌شود. یک راه توصیف کیفیت آب، فهرست کردن غلظت پارامتر عناصری است که در نمونه آب موجود باشند. چنین فهرستی فقط برای متخصصین آب قابل استفاده است. فرمول‌بندی و استفاده از شاخصها توسط سازمانهای مسئول و کنترل کیفیت آب بسیار توصیه شده است[3]. وگا و همکاران[4] در ارزیابی تغییرات فصلی و اثرات آلوده کننده کیفیت آب رودخانه اسپانیا با استفاده از آنالیزهای آماری، متغیرهای فیزیکی و شیمیایی آب در طول 5/2 سال از سه ایستگاه را بررسی کردند. در این مطالعه PCA نشان داد که مقدار مواد معدنی آلودگی انسانی و مقدار درجه حرارت کاهش یافته است. لفدانی و همکاران[5] جهت پیش‌بینی حجم رسوبات معلق رودخانه دوبرج، واقع در استان ایلام از روش آزمون گاما جهت تعیین متغیرهای ورودی به مدل ماشین بردار پشتیبان استفاده نمودند و نتیجه گرفتند که عملکرد مدل GT-SVM در مرحله آزمون بهتر از عملکرد مدل SVM می‌باشد. دیباک و همکاران[6] قابلیت‌های مدل SVM را در خصوص پیش‌بینی‌های هیدرولوژیکی نشان دادند. آن‌ها با استفاده از این روش در زمینه دسته‌بندی داده‌های سنجش از دور و مدل‌سازی بارش- رواناب و مقایسه آن با روش‌های شبکه عصبی مصنوعی به نتایج خوبی برای پیش‌بینی دست یافتند. محمدی و موذن زاده [7] برای پیش‌بینی جریان آبراه‌های رودخانه زرین رود از روش‌های هیبردی هوشمند استفاده کردند. نتایج نشان داد الگوریتم هیبریدی شبیه ساز تبرید در مقایسه با سایر مدل‌های هیبریدی عملکرد مطلوبی ارائه می‌دهد. مدل ماشین بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است[8]. اسکندری و همکاران[9] عملکرد دو روش ماشین بردار پشتیبان و شبکه عصبی مصنوعی را در پیش‌بینی اکسیژن خواهی بیوشیمیایی پنج‌روزه رودخانه سفیدرود مورد ارزیابی قرار دادند. نتایج مطالعات آن‌ها نشان داد از بین دو روش مذکور مدل ماشین بردار پشتیبان با دقت بیشتری شاخص‌های کیفیت آب رودخانه را شبیه‌سازی کرده است. دهقانی و همکاران[10] برای تعیین دبی روزانه رودخانه‌های حوضه دز از روش هیبریدی ماشین بردار پشتیبان و گیاهان مصنوعی استفاده کردند. نتایج نشان داد مدل هیبریدی مورد نظر در تخمین دبی روزانه عملکرد مطلوبی ارائه می‌دهد. در پژوهش حاضر به بررسی کارایی مدل هیبریدی ماشین بردار پشتیبان - ازدحام ذرات جهت تعیین کیفیت آب در رودخانه بابلرود ایستگاه قرآن تالار پرداختند.

مواد و روش

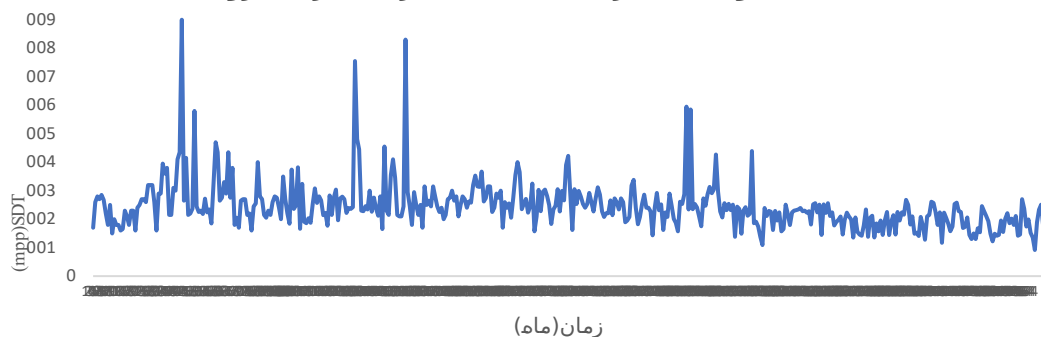
منطقه مورد مطالعه و داده‌های مورد استفاده

بابلرود یکی از رودخانه‌های پر آب شمال کشور می‌باشد که در استان مازندران قرار دارد. این رودخانه از کوه‌های البرز در شهرستان سوادکوه با حداکثر ارتفاع 3710 متر سرچشمه گرفته است. طول بابلرود 78 کیلومتر می‌باشد و پس از گذر از جنگل‌های سوادکوه، شالیزارها و شهرهای بابل و بابلسر در ارتفاع 10- متری به دریای خزر می‌ریزد. این رودخانه حوضه‌ای به مساحت 42/1746 کیلومترمربع را زهکشی می‌کند. برای دستیابی به اطلاعات لازم در مورد

ارزیابی کیفیت آب بابلرود از داده‌های اندازه‌گیری‌شده توسط شرکت مدیریت منابع آب ایران از 1394-1345 در ایستگاه هیدرومتری قرآنتالار استفاده شد. جانمایی ایستگاه هیدرومتری مذکور در شکل 1 و نمودار داده‌های مشاهداتی TDS آن در شکل 2 ارائه شده است.



شکل 1. موقعیت جغرافیایی ایستگاه قرآن تالار- بابلرود



شکل 2. نمودار مقادیر مشاهداتی TDS در ایستگاه هیدرومتری قرآنتالار از بابلرود

مدل ماشین بردار پشتیبان¹

مدل ماشین بردار پشتیبان یک سامانه‌ی یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که از اصل استقرای کمینه‌سازی خطای ساختاری استفاده کرده و منجر به یک جواب بهینه کلی میشود [11]. مشابه سایر مسائل رگرسیونی فرض می‌شود که رابطه میان متغیرهای مستقل و وابسته با تابع جبری مانند $f(x)$ به علاوه مقداری اغتشاش (خطای مجاز (ε)) مشخص شود.

(1)

¹ Support Vector Machine

$$f(x) = W^T \cdot \phi(x) + b$$

(2)

$$y = f(x) + \text{noise}$$

چنانچه بردار ضرایب W و مقدار ثابت b مشخصه‌های تابع رگرسیونی و نیز تابع کرنل باشد، آنگاه هدف پیدا کردن فرم تابعی برای $f(x)$ است. این مهم با آموزش مدل SVM توسط مجموعه‌ای از نمونه‌ها (مجموعه آموزش) محقق می‌شود. انواع رایج توابع کرنل قابل استفاده در مدل SVM رگرسیونی عبارت‌اند از: کرنل خطی، کرنل چندجمله‌ای و کرنل توابع پایه شعاعی (RBF) که میزان کارایی ماشین بردار پشتیبان با توجه به نوع تابع کرنل متفاوت می‌باشد. تابع پایه شعاعی به عنوان بهترین انتخاب از بین دیگر توابع کرنل گزارش شده است [8].

جدول 1. توابع کرنل [11]

توابع کرنل	رابطه ریاضی
تابع کرنل خطی	$K(x_i, x_j) = x_i^T \cdot x_j$
تابع کرنل چندجمله‌ای	$K(x_i, x_j) = c^d$
تابع کرنل با پایه شعاعی	$K(x, x_i) = \exp\left(\frac{- x_i - x_j ^2}{2\sigma^2}\right)$

الگوریتم ازدحام ذرات PSO

یک روش سراسری کمینه سازی است که با استفاده از آن می توان با مسائلی که جواب آن ها یک نقطه یا سطح در فضای N بعدی است، برخورد کرد. این الگوریتم براساس به اشتراک گذاشتن اطلاعات بین ذرات بار گذاری شده است. ابتدا برای هر متغیر در فضای جستجو به صورت تصادفی مقادیر تخصیص پیدا می کند. این عمل براساس جابه جایی و موقعیت ذرات دیگر و خود ذرات در تجربیات قبلی انجام می شود [13].

(3)

$$vit+1 = \omega vit + c1 r1 (pit - xit) + c2 r2 (pit - xit)$$

(4)

$$x_i^{t+1} = x_i^t + v_i^{t+1}$$

در معادله x_i^t موقعیت یا مقدار هر ذره i در فضای جستجو و t شماره تکرار هر ذره p_i^t موقعیت بهترین ذره در فضای جستجو که بسته به نوع جستجو متفاوت است. $r1$ و $r2$ عددهای تصادفی بین صفر و یک براساس تابع توزیع یکنواخت، $c1$ و $c2$ ضریب های شناختی و اجتماعی، ω ضریب اینرسی می باشد و x ضریب انقباض می باشد [14].

معیارهای ارزیابی

برای ارزیابی مدل‌های مورد استفاده از سه شاخص آماری معتبر و رایج شامل ضریب تبیین R^2 ، جذر میانگین مربعات خطا RMSE و درصد میانگین مطلق خطا MAPE استفاده شد. روابط شاخص‌های مذکور در زیر ارائه شده است. بر اساس این شاخص‌های کمی، مدل برتر دارای مقدار حداکثر ضریب تبیین و حداقل MAPE و RMSE خواهد بود.

$$R^2 = \left[\frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \cdot \sum_{i=1}^N (y_i - \bar{y})^2}} \right]^2 \quad (5)$$

(6)

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{x_i} \right| \quad (7)$$

که در آن‌ها x_i داده‌های مشاهداتی، y_i داده‌های محاسباتی و $\bar{x}$ مقدار متوسط مشاهداتی می‌باشد.

نتایج و بحث

نتایج مدل ماشین بردار پشتیبان

به منظور تخمین مواد جامد محلول در رودخانه بابلرود توسط مدل SVM توابع مختلف کرنل خطی، چندجمله‌ای و تابع پایه شعاعی که از انواع رایج مورداستفاده در هیدرولوژی می‌باشند را مورد بررسی و بهینه‌ترین انتخاب شد. نتایج حاصل از بررسی توابع کرنل در جدول 2 ارائه شده است. با توجه به جدول 2 تابع پایه شعاعی در رودخانه بابلرود ایستگاه قرآن تالار دارای بیشترین ضریب تبیین مساوی 91/0 و کمترین ریشه میانگین مربعات خطا 01/12 میلی‌گرم بر لیتر و کمترین درصد میانگین مطلق خطا 06/0 در مرحله صحت سنجی جواب بهینه‌تری نسبت به سایر مدل‌ها به همراه داشت.

جدول 2. نتایج مدل ماشین بردار پشتیبان در مراحل آموزش و صحت سنجی در بابلرود ایستگاه قرآن

تالار

مدل	آموزش			صحت سنجی		
	MAPE	RMSE	R^2	MAPE	RMSE	R^2
RBF	06/0	01/12	91/0	031/0	06/13	88/0
Poly	08/0	02/13	88/0	04/0	09/14	86/0

79/0	08/14	05/0	87/0	08/13	09/0	Lin
------	-------	------	------	-------	------	-----

نتایج مدل SVM-PSO

نتایج حاصل از مدل SVM-PSO به منظور تخمین مواد جامد محلول در رودخانه بابلرود در جدول 3 ارائه شده است. با توجه به جدول 3 ترکیب تابع پایه شعاعی در رودخانه بابلرود ایستگاه قرآن تالار دارای بیشترین ضریب تبیین مساوی 99/0 و کمترین ریشه میانگین مربعات خطا 04/11 میلی گرم بر لیتر و کمترین درصد میانگین مطلق خطا 03/0 و در مرحله صحت سنجی جواب بهینه تری به همراه داشت. جدول 3. نتایج مدل SVM-PSO در مراحل آموزش و صحت سنجی در بابلرود ایستگاه قرآن تالار

صحت سنجی			آموزش			مدل
R ²	RMSE	MAPE	R ²	RMSE	MAPE	
98/0	86/5	02/0	99/0	04/11	03/0	SVM-PSO

مقایسه عملکرد مدل‌ها

همان‌طور که از جدول 4 مشاهده می‌شود با انتخاب جواب بهینه هر کدام از مدل‌ها و مقایسه آن‌ها با یکدیگر مشخص شد در رودخانه بابلرود ایستگاه هیدرومتری قرآن تالار از بین دو مدل به کار رفته، مدل ماشین بردار پشتیبان-ازدحام ذرات با بیشترین دقت 98/0 و کمترین ریشه میانگین مربعات خطا 86/5 و کمترین درصد میانگین مطلق خطا 02/0 در مرحله صحت سنجی را دارا می‌باشد.

جدول 4. نتایج آموزش و صحت سنجی مدل‌های ماشین بردار پشتیبان، مدل هیبریدی SVM-PSO رودخانه بابلرود ایستگاه قرآن تالار

صحت سنجی			آموزش			مدل	
R ²	RMSE	MAPE	R ²	RMSE	MAPE		
88/0	06/13	031/0	91/0	01/12	06/0	SVM-RBF	TDS
98/0	86/5	02/0	99/0	04/11	03/0	SVM-PSO	

نتیجه گیری

در این پژوهش، به مقایسه دو روش SVM و SVM-PSO برای تخمین پارامتر کیفیت آب TDS در رودخانه‌های بابلرود پرداخته شد. هر دو روش مذکور در شبیه‌سازی عملکرد مناسبی دارند، ولی با توجه به سه معیار ارزیابی ضریب تبیین، جذر میانگین مربعات خطا و درصد میانگین مطلق خطا دقت شبیه‌سازی ماشین بردار پشتیبان-ازدحام ذرات نسبت به ماشین بردار پشتیبان بیشتر بوده و از خطای کمتری برخوردار می‌باشد. در مجموع نتایج تحقیق نشان داد که SVM-PSO با خطای کمتر و دقت بیشتر جهت شبیه‌سازی TDS در رودخانه‌های بابلرود ایستگاه قرآن تالار مناسب می‌باشد.

قدردانی

بدین وسیله از دانشگاه تهران و شرکت مدیریت منابع آب ایران به دلیل تأمین امکانات و داده‌های لازم جهت انجام این تحقیق و تهیه مقالات مربوطه تشکر و قدردانی می‌شود.

منابع

- [1]. Jabari, E. and N. Nader. 2006. Instruction for surface water quality monitoring. Vice presidency for strategic planning and supervision. Ministry of Energy publishing. 222.
- [2]. Faryadi, S., K. Shahedi, and M. Nabatpoor, Investigation of Water Quality Parameters in Tadjan River using Multivariate Statistical Techniques. Journal of Watershed Management 6(3):75-92.(In Persian).
- [3]. Abbasi,T., & Abbasi, S.A. (2012). Water quality indices. Environmental Earth Science, 71(1), 4625-4628. .(In Persian).
- [4]. Vega, M., R. Pardo, E. Barrado and L. Deban. 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Pergamon. 32: 3581-3592.
- [5]. Lafadani, E. Moghadamnia, A. Ahmadi,H. Ebrahimi, 2014. Evaluate the effect of processing the input variables to the model of support vector machine to predict the volume of sediment gamma test. Iranian Journal of Natural Resources .67(2):289-303. (In Persian)
- [6]. Dibike, Y.B. Velickov, S. Solomatine, D.P. Abbott, M.B. 2001. Model Induction with Support Vector Machines: Introduction and Application, ASCE. Journal of Computing in Civil Engineering, 15 (3): 208-216.
- [7].Mohammadi, B., Moazenzadeh, R. 2017. Prediction of Stream Flow Using Intelligent Hybrid Models in Monthly Scale (Case study: Zarrin roud River). Journal of Environmental Science and Technology,21(9):73-80. .(In Persian).
- [8]. Cristianini, N., and J. Shawe-Taylor, 2000. An Introduction to Support Vector Machines, Cambridge University Press, New York.
- [9]. Eskandari, A., R. Nouri, H. Meragi, and A.Kiaghaderi, 2012. Developing an appropriate model based on artificial neural network and support vector machine for prediction of timely 5-day biochemical oxygen demand. Journal of Ecology 38(1):71-82. .(In Persian).
- [10]. Dehghani R., Torabi Poudeh H., Younesi, H., Shahinejad B. 2020.Application of the Hybrid Model of Support Vector Machine-Algorithm Artificial Flora in Estimating the Daily Flow of Rivers (Case Study: Dez Basin), Iran-Water Resources Research,16(2):132-149. .(In Persian).
- [11].Jones, A.J. 2004. New tools in non-linear modeling and prediction. Computational Management Science 1: 109-149.
- [12]. Seifi,A, Myrloftfy,S, Riahi,H. 2012. Assessment and monitoring of weather station network using principal components analysis and factor analysis Case study: Kerman. Journal of Irrigation and Drainage.Number1, Volume5,30-42.(In Persian).
- [13].Kavosi, M., Khasheii, A. Pourreza, M. Najafi, M.H. 2018. Application of a new vector simulation algorithm LSSVM-PSO in the design of an optimal network of groundwater level monitoring. Journal of Echo Hydrology,5(4).1309-1319. .(In Persian).

Evaluation Water Quality of Babolrood River Using Meta-Heuristic Algorithms

Fatemeh Akhoni Pourhosseini¹, Kumars Ebrahimi^{2*}, Mohammad Hossein Omid³

¹ Ph.D student of Water Resources Engineering, University of Tehran, Karaj, IRAN

Email: Fateme.Pourhosini@ut.ac.ir

^{2} Professor, Department of Irrigation and Reclamation Engineering, University of Tehran, Karaj, IRAN*

Email: EbrahimiK@ut.ac.ir

³ Professor, Department of Irrigation and Reclamation Engineering, University of Tehran, Karaj, IRAN

Email: MOmid@ut.ac.ir

Abstract

Water pollution is one of the crucial issues which affect human and environmental ecosystems in the world. In this study, SVM-PSO hybrid method was used to determine the water quality of Babolrood river. Monthly hydrometric data including; Ca, Mg, HCO₃, Na, SO₄, EC, pH, TDS

were analyzed, from 1966-2015, which were measured and recorded at Quran-Talar hydrometric stations of Babolrood River. The criteria of the correlation coefficient (R^2), root mean square error (RMSE) and Mean absolute percent error (MAPE) were used to evaluate and performance compare of models. The results showed that however the models could be used to estimate with reasonable accuracy the amount of dissolved solids in water deal, but regarding to accuracy, SVM model in Babolrood Goran Talar hydrometric station with the highest correlation (0.98), minimum root mean square error (5.86) and Mean absolute percent error (0.02) were put in the verification phase. The results showed that the SVM-PSO model to estimate high minimum and maximum values of water quality.

Keywords: Calibration, Intelligent model, Optimization, Water Quality