



Environment and Water Engineering

Homepage: www.jewe.ir



ISSN: 2476-3683

Case Study

Application of Support Vector Machine and Boosted Tree Algorithm for Rainfall-Runoff Modeling (Case Study: Tabriz Plain)

Zeinab Bigdeli¹, Abolfazl Majnooni-Heris^{2*}, Reza Delirhasannia² and Sepideh Karimi³

¹Ph.D. Scholar, Department of Water Science Engineering, Faculty of Agriculture Engineering, University of Tabriz, Tabriz, Iran

²Assoc. Professor., Department of Water Science Engineering, Faculty of Agriculture Engineering, University of Tabriz, Tabriz, Iran

³Postdoc Researcher, Department of Water Science Engineering, Faculty of Agriculture Engineering, University of Tabriz, Tabriz, Iran

Article information

Received: October 23, 2022

Revised: April 27, 2023

Accepted: May 06, 2023

Keywords:

Aji Chai

Boosted Tree

Tabriz Plain

Support Vector Machine

Rain-Runoff Modeling

*Corresponding author:

majnooni@tabrizu.ac.ir



Abstract

This research focused on the application of Support Vector Machine (SVM) and Boosted Trees (BT) algorithms for simulating precipitation and runoff in two stations, Akhula and Pole Senikh, in the Tabriz Plain, Iran. Meteorological and hydrometric data were collected from 24 stations in the Tabriz watershed, obtained from the Regional Water Company and East Azerbaijan Meteorological Organization. Precipitation and runoff values were used as input to the model with a one-day time lag, and monthly runoff values were estimated and compared with monthly observations using evaluation criteria. The results showed that for both study periods, SVM model performed better than BT model for Akhula station, while BT model performed better than SVM model for Pole Senikh station. Additionally, the cross-correlation coefficient for the two study periods was found to be 0.83 and 0.82 for Akhula station, and 0.83 and 0.77 for Pole Senikh station, respectively. In the time series results, there was no clear trend in precipitation over the observation period. However, river flows at the Ahvaz and Pole Senikh stations, particularly after 1995, showed a significant decline, mainly due to factors such as runoff, agricultural expansion, and industrial development.

© Authors, Published by **Environment and Water Engineering** journal. This is an open-access article distributed under the CC BY (license <http://creativecommons.org/licenses/by/4.0/>).



Introduction

The statement emphasizes the importance of estimating runoff resulting from rainfall in order to effectively plan water resources, especially in regions with limited hydrometric station information like East Azarbaijan province in Iran. To overcome the lack of data, this research

aims to develop runoff estimation methods by utilizing longer-term precipitation data and statistical modeling techniques. The specific goal is to simulate the rainfall-runoff process in the terminal hydrometric stations of Tabriz plain, Iran using STATISTICA software and create a



rainfall-runoff model based on Support Vector Machine and Boosted Tree models.

Material and Methods

The study area is located in the Tabriz plain, East Azarbaijan province, in northwest Iran. The Tabriz sub-basin has an area of 5428 km², with an average annual rainfall of 230.7 mm and average daily temperature of 12.1°C. The research utilized monthly rainfall, average monthly temperature, and outlet monthly flow data from two stations in the area, collected from 1976 to 2019. The data was divided into two periods: 1976-1994 and 1995-2019. The average rainfall of the Tabriz plain was calculated using a polygonal method and combined with the discharge from Akhula and Pole Senikh stations with a one-day time lag. Two models, Support Vector Machine (SVM) and Boosted Trees (BT), were used to simulate rainfall-runoff, with the data split into 70% for training and 30% for validation. Different combinations of runoff and rainfall inputs were evaluated, and the best-performing combinations were selected. The models were evaluated using statistical metrics including RMSE, correlation coefficient (r),

coefficient of determination (R²), and Nash-Sutcliffe Efficiency Coefficient (NSE). Thiessen polygons were used to define rainfall polygons, weighting rainfall based on the covered surface area. Support Vector Machines are data mining algorithms used for classification and regression tasks. The mentioned statistical metrics were employed to evaluate and compare the efficiency of the models in this research.

Results

Table 1 provides rainfall and runoff statistics (average, minimum, maximum) for Akhula and Pole Senikh stations in two study periods: the first period (1976-1994) and the second period (1995-2019). The analysis of the time series data indicates that there is no clear trend in rainfall for both stations, with some years showing an increasing trend and others showing a decreasing trend. However, in the second study period, a downward trend in outflow from both stations is observed, indicating a significant decrease in recent years. This decline in outflow raises concerns about the declining water level in Lake Urmia.

Table 1 Statistical parameters of the rainfall and runoff data during 1976-1994 and (1995-2019)

Station	Parameter	Entire data 1976-1994 (1995-2019)		Training set		Test set	
		P (mm/month)	Q (m ³ /s)	BT	SVM	BT	SVM
Akhula	Mean Minimum Maximum	21.97 (22.40)	11.74 (5.07)	11.32 (4.53)	10.46 (5.73)	11.13 (5.97)	11.90 (3.38)
		0.00	0.00	0.00	0.00	0.00	0.00
		(0.81)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
		74.31	92.50	76.85	76.85	53.85	75.48
		(77.66)	(47.47)	(47.47)	(47.47)	(41.65)	(36.15)
Pole senikh	Mean Minimum Maximum	18.97 (20.39)	0.86 (0.52)	0.85 (4.53)	0.86 (0.54)	0.93 (0.51)	0.91 (0.41)
		0.00	0.00	0.00	0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
		61.38	3.63	3.63	3.86	3.86	3.70
		(81.25)	(3.55)	(3.55)	(2.73)	(2.21)	(2.23)

Based on the table, it is evident that in both sections, the Akhula station has the highest average rainfall and runoff values, while the Pole Senikh station has the lowest average rainfall and runoff values. Since the selected data for the two training sets differ for each model, it is expected that the statistical parameters of these two models would also differ from each other.

Rainfall-runoff modeling using SVM and BT models

As mentioned, in order to model rainfall-runoff data using SVM and BT models, part of the data was used for training and another part was used for validation. Table 2 shows the values of the calculated statistical indicators related to the results of the training and validation section in both SVM and BT models.

Table 2 Performance of SVM and BT models in the training and testing sets during 1994-2019

Station	Entire data (1976-2019)	Cross Correlation		R ²		NSE		RMSE	
		BT	SVM	BT	SVM	BT	SVM	BT	SVM
1976-1994									
Akhula	Train	0.90	0.90	0.82	0.81	0.81	0.81	6.86	6.59
	Test	0.81	0.89	0.65	0.80	0.64	0.72	8.13	8.09
Pole senikh	Train	0.84	0.83	0.70	0.69	0.70	0.69	0.45	0.46
	Test	0.83	0.82	0.69	0.67	0.69	0.67	0.51	0.54
1995-2019									
Akhula	Train	0.79	0.83	0.62	0.68	0.62	0.67	4.23	4.53
	Test	0.69	0.75	0.48	0.57	0.47	0.55	6.00	3.65
Pole senikh	Train	0.80	0.80	0.65	0.64	0.64	0.64	0.31	0.29
	Test	0.77	0.76	0.59	0.58	0.58	0.57	0.35	0.32

The results from Tables 2 indicate that the SVM model outperformed the BT model in simulating flow rate and monthly rainfall in Akhula station during both study periods. Conversely, the BT model exhibited better performance than the SVM model in the case of the Pole Senikh station for the same periods. The test set results showed that the Cross Correlation values for Akhula station were 0.89 and 0.75 for the first and second study periods, respectively. Similarly, for the Pole Senikh station, the Cross Correlation values were 0.83 and 0.77.

Conclusions

This research used intelligent artificial neural network models, specifically Support Vector Machine (SVM) with kernel function and Boosted Tree (BT), to estimate monthly runoff in Akhula and Pole Senikh stations. The models were trained using rainfall and runoff data with a one-day time lag as inputs, and the observed monthly runoff values were used for evaluation. The results indicated that the SVM model with kernel function performed better than the BT model for the Akhula station, while the BT model performed better for the Pole Senikh station, in both study periods. The mutual correlation values for the test sets were 0.89 and 0.75 for Akhula station, and 0.83 and 0.77 for Pole Senikh station, in the first and second study periods, respectively. The time series analysis

revealed that there was no significant trend in rainfall for both stations. However, a declining trend in water flow was observed in Ajichai and Pole Senikh rivers over a 43-year period, possibly attributed to water usage in agriculture and industry. The successful performance of the machine learning models in simulating rainfall-runoff suggests their potential applicability in other stations within the Lake Urmia region.

Acknowledgment

The authors would like to thank the Iran Meteorological Organization and Regional Water Company of East Azerbaijan in Iran for providing access to weather, hydrological, and agriculture databases of the study area.

Data Availability

The data of this research was obtained from (Regional Water Company and Meteorological Organization of East Azarbaijan Province), which can be presented after obtaining permission from (Regional Water Company and Meteorological Organization of East Azarbaijan Province).

Conflicts of interest

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



ISSN: 2476-3683

محیط زیست و مهندسی آب

Homepage: www.jewe.ir

مطالعه موردی

کاربرد الگوریتم ماشین بردار پشتیبان و درخت تقویت شده برای مدل سازی بارش - رواناب (مطالعه موردی: دشت تبریز)

زینب بیگدلی^۱، ابوالفضل مجنونی هریس^{۲*}، رضا دلیر حسن نیا^۲ و سپیده کریمی^۳

^۱ دانشجوی دکتری، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

^۲ دانشیار، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

^۳ محقق پسادکتری، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

چکیده	اطلاعات مقاله
این پژوهش به استفاده از روش های یادگیری ماشین، به ویژه الگوریتم ماشین بردار پشتیبان (SVM) و درخت تقویت شده (BT)، برای شبیه سازی بارش و رواناب در دو ایستگاه آخولا و پل سنخ پرداخته است. داده های هواشناسی و هیدرومتری از ۲۴ ایستگاه زیرحوضه تبریز از شرکت آب منطقه ای و سازمان هواشناسی استان آذربایجان شرقی جمع آوری شدند. مقادیر بارش و رواناب به عنوان ورودی به مدل با تأخیر زمانی یک روزه وارد شده و سپس مقادیر رواناب ماهانه با استفاده از معیارهای ارزیابی با مشاهدات ماهانه تخمین زده شده مقایسه شدند. نتایج نشان داد که در هر دو دوره مطالعه، برای ایستگاه آخولا، مدل SVM عملکرد بهتری نسبت به مدل BT داشت و در ایستگاه پل سنخ، مدل BT عملکرد بهتری نسبت به مدل SVM ارائه کرد. همچنین، مقدار همبستگی متقابل برای دو دوره مطالعاتی در ایستگاه آخولا به ترتیب برابر با ۰/۸۳ و ۰/۸۲ و ایستگاه پل سنخ ۰/۸۳ و ۰/۷۷ به دست آمد. در نتایج سری زمانی نیز، روند مشخصی برای بارش در طول دوره مشاهده نشد، اما دبی رودخانه ها در ایستگاه آخولا و پل سنخ، به ویژه بعد از سال ۱۳۷۴، روند کاملاً نزولی داشته است. از جمله دلایل اصلی این کاهش می تواند رواناب، توسعه کشاورزی و صنعت باشد.	تاریخ دریافت: [۱۴۰۱/۰۸/۰۱]
	تاریخ بازنگری: [۱۴۰۲/۰۲/۰۷]
	تاریخ پذیرش: [۱۴۰۲/۰۲/۱۶]
	واژه های کلیدی: آبی چای آخولا دشت تبریز شبیه سازی بارش-رواناب یادگیری ماشین
	*نویسنده مسئول: majnooni@tabrizu.ac.ir



۱- مقدمه

داده های جریان رواناب، تخصیص منابع آب به این مصارف را تعیین می کنند (Worland et al. 2018). از طرفی دیگر سیلاب ها عموماً در دره های خشک که داده های اندازه گیری شده ندارند، رخ می دهند. در استان آذربایجان شرقی برای

به طور کلی بخش قابل توجهی از نزولات آسمانی به رواناب تبدیل می شود. رواناب حاصل از بارش و ذوب برف یکی از منابع اصلی برای تحقق نیازهای کشاورزی، صنعتی و مصارف خانگی است؛ به طوری که برنامه ریزان منابع آب با تکیه بر



کشورهای دیگر با استفاده از مدل‌های هوشمند جهت پیش‌بینی جریان بارش- رواناب، سیلاب انجام شده است که در ادامه به برخی از آنها اشاره می‌شود.

Ghorbani et al. (2015) به بررسی فرآیند بارش رواناب روانه حوزه آبخیز بالیه لوچای با استفاده از ANN، SVM، هیبرید موجک ماشین بردار پشتیبان^۵ (WSVM) و هیبرید موجک شبکه عصبی^۶ (WANN) پرداخته‌اند. همه روش‌ها دارای دقت قابل قبولی بودند. Poursalehi et al. (2022) عملکرد الگوریتم جنگل تصادفی در پیش‌بینی نوسانات سطح ایستابی در مقایسه با دو مدل درخت تقویت‌شده و شبکه عصبی مصنوعی (مطالعه موردی: آبخوان آزاد دشت بیرجند) را مطالعه کردند. نتایج حاصل از شبیه‌سازی با استفاده از الگوریتم جنگل تصادفی نشان می‌دهد این الگوریتم توانایی نسبتاً زیادی در شبیه‌سازی تراز سطح ایستابی آبخوان دارد. در جدیدترین مطالعه انجام گرفته توسط Zarei et al. (2022) پتانسیل وقوع سیل با استفاده از مدل‌های داده‌کاوی ماشین بردار پشتیبان، چاید و جنگل تصادفی در حوزه آبخیز فریزی را ارزیابی نمودند. بر اساس نتایج، هم در داده‌های تعلیمی و هم در داده‌های آموزشی در الگوریتم (SVM) و مدل الگوریتم نوین جنگل تصادفی دارای نتایج دقیق‌تری می‌باشد. Bigdeli et al. (2023) در جدیدترین مطالعات خود مدل‌سازی بارش-رواناب حوزه آجی چای با استفاده از مدل‌های جنگل تصادفی و شبکه عصبی مصنوعی را بررسی کردند و به این نتیجه رسیدند که برای حوزه آجی چای مدل ANN کارایی بالاتری نسبت به RF دارد. Lee et al. (2017) از پیش‌بینی مکانی حساسیت به سیل با استفاده از مدل‌های جنگل تصادفی و درختان تقویت‌شده در کلان شهر سئول، کره جنوبی استفاده کردند. در نتیجه، مدل جنگل تصادفی برای الگوریتم‌های رگرسیون و طبقه‌بندی به ترتیب ۷۸/۷۸ و ۱۸/۷۹٪ و مدل درخت تقویت‌شده دقت اعتبارسنجی ۷۷/۵۵ و ۷۷/۲۶٪ را برای الگوریتم‌های رگرسیون و طبقه‌بندی نشان داد. Hussain et al. (2020) جهت پیش‌بینی جریان رودخانه هنزه در پاکستان از مدل‌های شبکه عصبی مصنوعی پرسپترون چندلایه، (MLP) رگرسیون بردار پشتیبان (SVR) و جنگل تصادفی (RF) استفاده نموده‌اند. مجموعه‌ای از پنج پارامتر به‌اضافه سری زمانی جریان رودخانه با تأخیرهای مختلف به‌عنوان

حوضه‌ی کوچکتر از ۴۰ km² اطلاعات ایستگاه‌های هیدرومتری موجود نمی‌باشد. بنابراین توسعه روش‌های تخمین رواناب برای این حوضه‌ها ضروری است. از آنجایی که اغلب داده‌های بارش در مدت‌های زمانی طولانی‌تری نسبت به داده‌های رواناب موجود می‌باشند، می‌توان از این موجودیت برای پر کردن خلأهای آماری داده‌های رواناب استفاده نمود. در این راستا، در دسترس بودن اطلاعات پیوسته بارش و سایر متغیرهای اقلیمی که می‌توانند برای مدل‌سازی مقدار رواناب مورد استفاده قرار گیرد، بسیار ضروری می‌باشد (Lorrai and Sechi 1995). مدل‌های زیادی از اوایل دهه ۱۹۶۰ توسعه داده شده است تا فرایند بارش-رواناب را شبیه‌سازی کنند. مدل ابزاری است که می‌تواند بخشی از واقعیت را شبیه‌سازی کند. هدف از مدل‌سازی، نشان دادن عملکرد و رفتار یک پدیده طبیعی تحت شرایط معین می‌باشد (Phomcha et al. 2011). به‌طور کلی، مدل‌های ریاضی شناخته شده بارش-رواناب را می‌توان در دو گروه مدل‌های دارای مبنای فیزیکی (همچون مدل‌های هیدرولوژیکی مفهومی) و مدل‌های داده محور (همچون شبکه عصبی مصنوعی) تقسیم‌بندی نمود. یک سری جدید از مدل‌های رگرسیون مبتنی بر رویکردهای یادگیری ماشین برای حل مشکلات مدل‌های سری زمانی کلاسیک ایجاد شده است. با استفاده از مدل‌های هوشمند می‌توان روابط ذاتی میان داده‌ها را که ریاضیات کلاسیک قادر به حل آنها نمی‌باشد، مدل‌سازی نمود. یادگیری ماشین و رویکردهای مدرن به‌طور گسترده‌ای در مطالعه منابع آب در سال‌های اخیر استفاده شده است. از جمله مدل‌های هوشمند می‌توان به مدل‌های شبکه‌های عصبی مصنوعی^۱ (ANNs)، برنامه‌ریزی بیان ژن^۲ (GEP) و مدل‌هایی همچون ماشین بردار پشتیبان^۳ (SVM) و درخت تقویت‌شده^۴ (BT) اشاره کرد. الگوریتم‌های ماشین بردار پشتیبان و درخت تقویت‌شده دو نوع یادگیری ماشینی هستند که از الگوریتم‌های ضروری برای پیش‌بینی‌های چندگانه مکرر استفاده می‌کنند (Kisi and Parmar 2016). از اولین کاربردهای روش SVM در مسائل آب می‌توان به Dibikey et al. (2001) برای شبیه‌سازی بارش رواناب اشاره کرد. تاکنون مطالعات زیادی در ایران و

¹Artificial Neural Networks

²Gene expression programming

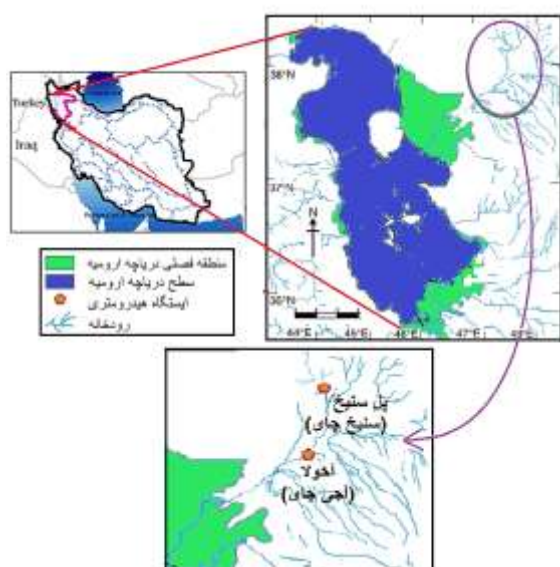
³Support Vector Machine

⁴Boosted Tree

⁵Wavelet Support Vector machine

⁶Wavelet Artificial Neural Network

انتهای دشت به دلیل شیب بسیار کم سطح زمین به چند شاخه تقسیم می‌شوند تا در نهایت به دریاچه ارومیه تخلیه شود. رودخانه آجی‌چای ۲۶۵ km طول دارد و در مناطق مرکزی استان در جریان می‌باشد (Bigdeli et al. 2023). مساحت حوزه آبخیز این رودخانه ۹۲۰۰ km² و حداکثر دبی آن ۴۰/۶ m³/s می‌باشد. بر روی آجی‌چای و سرشاخه‌های آن ۳۴ ایستگاه آب‌سنجی و هیدرومتری نصب شده است (Aji Chay - Wikipedia). طبق شکل (۱)، مطالعه حاضر بر روی آخرین ایستگاه‌های منتهی به دریاچه ارومیه (ایستگاه آخولا و پل سنیخ) بر روی رودخانه آجی‌چای و سنیخ‌چای انجام شد.



شکل ۱- موقعیت جغرافیایی ایستگاه‌های آخولا و پل سنیخ
Fig. 1 Geographical location of Akhula and Pole Senikh stations

۲-۲- داده‌های مورد استفاده

داده‌های مورد استفاده برای مدل‌سازی در این پژوهش، شامل داده‌های ماهانه بارش، میانگین دما و دبی خروجی از دو ایستگاه مورد مطالعه می‌باشد که از ۲۴ ایستگاه باران‌سنجی و سینوپتیک از سال ۱۳۵۵ تا ۱۳۹۷ و از شرکت آب منطقه‌ای و سازمان هواشناسی استان آذربایجان شرقی تهیه شد. صرفاً بر اساس مشاهده تغییر روند موجود در سری زمانی بارش و رواناب در سال ۱۳۷۴، داده‌های ۴۳ ساله به دو دوره ۱۳۷۴-۱۳۵۵ و ۱۳۹۷-۱۳۷۵ تقسیم شدند. این تغییر روند در مطالعات مختلفی از جمله Nazeri Tahroudi et al. (2017) و Fazli Fard et al. (2019) مشاهده می‌شود. نکته قابل ذکر رودخانه‌های زیر حوضه آجی‌چای می‌باشد که تمام این رودخانه‌های فصلی یا دائمی به شاخه اصلی رودخانه

به ترتیب مدل‌های RF، MLP و SVR¹ بالاترین عملکرد را در پیش‌بینی جریان رودخانه هنزه داشتند. همچنین نتایج نشان داد که مدل‌های مبتنی بر الگوریتم‌های یادگیری ماشینی به‌ویژه مدل RF می‌توانند با دقت بسیار بالایی در پیش‌بینی جریان رودخانه به کار گرفته شوند. همچنین در زمینه استفاده از مدل‌های هوشمند برای پیش‌بینی جریان می‌توان به مطالعات Adnan et al. (2018) و Shafeizadeh (2019) اشاره کرد.

درواقع سؤال اساسی پژوهش این است که آیا برای شبیه‌سازی فرایند بارش-رواناب، در شرایط کمبود داده و اطلاعات، این مدل‌ها می‌تواند جوابگوی پیش‌بینی بارش و رواناب در این دو ایستگاه باشد. همچنین کدامیک از مدل‌ها در دو ایستگاه آخر دشت تبریز (منتهی به دریاچه ارومیه) عملکرد بهتری در شبیه‌سازی بارش - رواناب دارند. از اینرو، شبیه‌سازی فرایند بارش-رواناب در ایستگاه‌های هیدرومتری پایانی دشت تبریز (آخولا و پل سنیخ) در استان آذربایجان شرقی با استفاده از نرم‌افزار STATISTICA و ایجاد مدل بارش-رواناب با روش‌های ماشین بردار پشتیبان و درخت تقویت‌شده هدف این پژوهش می‌باشد.

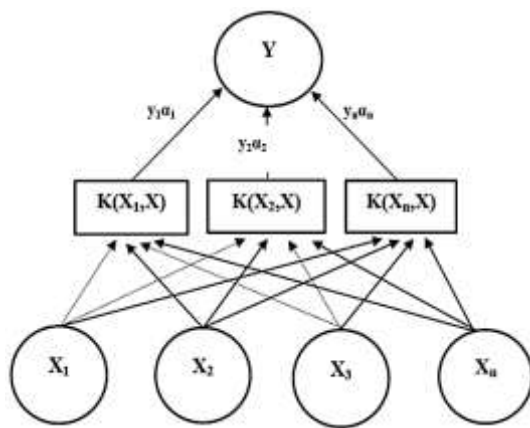
۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

منطقه مورد مطالعه در دشت تبریز در شمال غربی ایران و در استان آذربایجان شرقی واقع شده است. مساحت زیر حوضه تبریز ۵۴۲۸ km² است که یکی از زیر حوضه‌های اصلی حوضه دریاچه ارومیه می‌باشد. اقلیم غالب منطقه تبریز نیمه‌خشک بوده و در فصل مرطوب، میانگین بارش سالانه ۲۳۰/۷ mm و میانگین دمای روزانه ۱۲/۱ °C است (Tabriz - Wikipedia). این زیر حوضه از شمال به حوزه آبخیز اهر چای، از جنوب به زیر حوضه آذرشهر و از شرق به زیر حوضه بیلوردی دوزدوزان و از غرب به دریاچه ارومیه محدود می‌شود. رودخانه آجی‌چای رودخانه دائمی در منطقه مورد مطالعه است و سایر رودخانه‌ها به این رودخانه اصلی می‌پیوندند. از ابتدای زیر حوضه آجی‌چای، رودخانه‌های بیوک‌چای، تاجیارچای، رازلیق‌چای، ونق‌چای و اوجان‌چای و در دشت تبریز نهندچای، مهران‌رود، گمناب‌چای و سنیخ‌چای از سمت راست و رودخانه‌های سردرود و عنصررود، ليقوان‌چای از سمت چپ به آجی‌چای پیوسته و در

¹Support Vector Regression

مهم‌ترین وظایف مدل ماشین‌های بردار پشتیبان، طبقه‌بندی و خطی‌سازی یا رگرسیون داده‌ها است. ماشین بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است. روشی است که هم برای گروه‌بندی و هم برای تخمین و برآورد تابع برازش داده‌ها در مسائل رگرسیون به کار می‌رود به‌طوری که کمترین خطا در گروه‌بندی داده‌ها یا تابع برازش رخ دهد. SVM الگوریتمی است که نوع خاصی از مدل‌های خطی را می‌یابد که حداکثر حاشیه ابر صفحه را حاصل می‌کنند. به نزدیک‌ترین نقاط آموزشی به حداکثر حاشیه ابر صفحه، بردارهای پشتیبان گفته می‌شود و تنها از این بردارها (نقاط) برای مشخص کردن مرز بین طبقات استفاده می‌شود. شکل (۳) فرآیند روش SVM را نشان می‌دهد (Seyedian et al. 2014).



شکل ۳- فرآیند مدل SVM (Seyedian et al. 2014)

Fig. 3 SVM model process (Seyedian et al. 2014)

در مدل رگرسیونی SVM باید وابستگی تابع متغیر وابسته Y به مجموعه‌ای از متغیرهای مستقل X تخمین زده شود. در این مطالعه متغیرهای مستقل، بارش و رواناب قبلاً ثبت شده و تابع متغیر وابسته، داده رواناب فعلی تعریف شده‌اند. برای اتصال سه لایه ورودی، پنهان و خروجی و ساخت مدل با بالاترین عملکرد، از کرنل تابع شعاعی استفاده گردید که در ادامه تعریف خواهد شد. فرض بر این است که مانند دیگر مسائل رگرسیونی، رابطه بین متغیرهای وابسته و مستقل با استفاده از یک تابع معین f به‌علاوه مقداری اغتشاش (خطای مجاز یا ε) مشخص می‌شود، در واقع هدف پیدا کردن فرم تابعی برای $f(x)$ است که در رابطه‌های (۲) و (۳) عنوان شده‌اند (Yoon et al. 2011):

$$f(x) = W^T \phi(x) + b \quad (2)$$

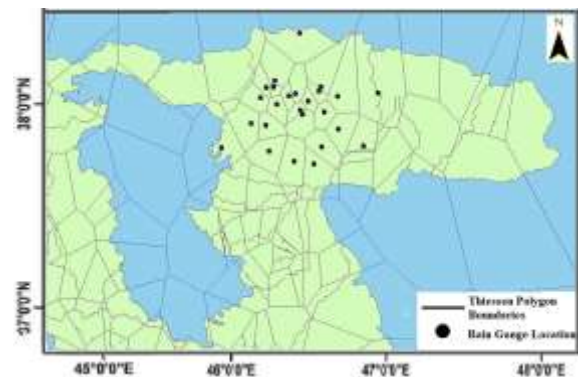
$$y = f(x) + \text{noise} \quad (3)$$

آبی‌چای می‌ریزند. در نهایت دبی رودخانه آبی‌چای و سنج‌چای در دو ایستگاه آخولا و پل سنج ثبت می‌شوند.

۲-۳- میانگین بارش از طریق روش چندضلعی تیسن
چندضلعی‌های تیسن، که سلول‌های ورونوی^۱ هستند، جهت تعریف چندضلعی بارش معادل مساحت سطح A_i برای وزن کردن بارش هر باران‌سنج I_i استفاده می‌شود. نهایتاً، بارش معادل با وزن مساحت با استفاده از رابطه (۱) محاسبه شد (Aurenhammer et al. 2013).

$$r_{th}(t) = \frac{\sum_{i=1}^n r_i(t) \cdot A_i}{\sum_{i=1}^n A_i} \quad (1)$$

از نرم‌افزار GIS که برای توصیف پارامترهای جغرافیایی حوضه استفاده می‌شود. معمولاً می‌توان برای محاسبه بارش معادل از رابطه (۱) استفاده کرد. به این ترتیب با استفاده از ایستگاه‌های باران‌سنجی موجود بارش متوسط در مقیاس ماهانه و سالانه برای هر زیر حوضه بر اساس چندضلعی‌های تیسن تعیین شد.



شکل ۲- تقسیم‌بندی چندضلعی تیسن حوضه دریاچه ارومیه شرقی

Fig.2 Thiessen polygon partitioning of the Eastern Lake Urmia Basin

همان‌طور که در شکل (۲) نشان داده شده است، چندضلعی‌های تیسن حوضه دریاچه ارومیه شرقی با استفاده از سیستم اطلاعات جغرافیایی کوانتومی (www.qgis.org) تولید شدند (Shekhar and Xiong 2007).

۲-۴- روش پژوهش

۲-۴-۱- الگوریتم ماشین بردار پشتیبان

ماشین‌های بردار پشتیبان همانند دیگر روش‌های هوش مصنوعی بر مبنای الگوریتم داده‌کاوی عمل می‌کنند.

^۱Voronoi

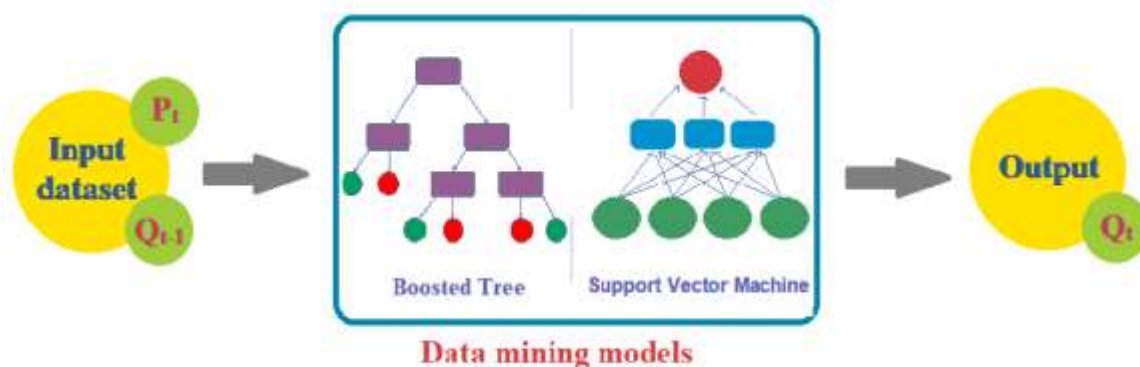
پیوسته و حساسیت به اندازه داده‌های نمونه دارد. اما وقتی درخت پیچیده‌تر می‌شود، مشکل اضافه برآزش به وجود می‌آید که نشان می‌دهد الگوریتم یک مدل بیش از حد تنظیم شده برای داده‌های ورودی ایجاد می‌کند (Friedman et al. 1999).

همان‌طور که اشاره شد برای شبیه‌سازی بارش - رواناب، ابتدا بارش متوسط زیر حوضه تبریز با روش پلیگون‌بندی محاسبه شد. سپس به همراه دبی خروجی از ایستگاه آخولا و پل سنخ به همراه تأخیر زمانی یک‌روزه وارد دو مدل SVM (تابع کرنل) و BT شدند. برای این منظور از ۷۰٪ داده‌ها برای مرحله آموزش و ۳۰٪ داده‌ها برای مرحله آزمون استفاده شد. لازم به ذکر است که انتخاب داده‌ها در این نرم‌افزار برای مجموعه آموزش کاملاً به صورت اتفاقی انتخاب شده‌اند و لزوماً داده‌های منتخب برای مجموعه آموزش دو مدل SVM و BT با یکدیگر برابر نمی‌باشند. سپس مجموعه آموزشی بارش به عنوان متغیر پیش‌بینی کننده و مجموعه آموزشی رواناب به عنوان متغیر هدف انتخاب شدند. چندین ترکیب ورودی رواناب و بارندگی برای مدل‌سازی مورد ارزیابی قرار گرفت. ورودی‌ها مقادیر بارش (P_t) و دبی (Q_{t-1}) ماهانه قبلاً ثبت شده (P_t ، Q_{t-1}) را ارائه می‌دهند و خروجی مربوط به داده‌های رواناب فعلی (Q_t) است، که زیرنویس t نشان‌دهنده مرحله زمانی است. همان‌طور که در شکل (۴) مشاهده می‌شود دو ترکیب ورودی از داده‌های P و Q ایجاد و مدل SVM و BT برای مدل‌سازی بارش-رواناب برای تعیین ترکیب ورودی بهینه استفاده شد.

بردار ضریب W و ثابت b و تابع کرنل ϕ در این رابطه با حل مسئله بهینه‌سازی محدب و با استفاده از روش لاگرانژین و اصل کمینه‌سازی خطای ساختاری به دست می‌آیند. در مطالعات بارش-رواناب عمدتاً از تابع پایه شعاعی استفاده می‌شود. زیرا ماشین بردار پشتیبان با این نوع تابع در مقایسه با سایر توابع از دقت بیشتری در شبیه‌سازی این پدیده برخوردار است (Adamowski 2013).

۲-۴-۲- درخت تقویت‌شده

روش درخت تقویت‌شده یک روش ناپارامتریک است که برای مدل‌سازی داده‌های پیوسته و گسسته مناسب است. این روش قابل استفاده برای رگرسیون و طبقه‌بندی با استفاده از متغیرهای پیوسته و گسسته است و یک ابزار قوی برای داده‌کاوی و پیش‌بینی محسوب می‌شود. ارزیابی مدل با استفاده از داده‌های آزمون صورت می‌گیرد و دقت پیش‌بینی معیاری ارزیابی مهم است. الگوریتم درخت تقویت‌شده معمولاً در مسائل رگرسیون و پیش‌بینی متغیرهای پیوسته استفاده می‌شود. در صورتی که تعداد کلاس‌ها بیش از ۱۰۰ باشد، طبقه‌بندی دقیق‌تری نیاز دارد و زمان محاسباتی بیشتری ممکن است طول بکشد. روش درخت تقویت‌شده به عنوان یک رویکرد یادگیری مجموعه‌ای عمل می‌کند و شباهت زیادی به الگوریتم جنگل تصادفی دارد. در طبقه‌بندی گسترده، درختان دودویی ایجاد می‌شود و مجموعه داده را به دو نمونه تقسیم می‌کند. نقطه تقسیم‌بندی داده‌ها در هر گره با استفاده از روش درخت تقویت‌شده تعیین می‌شود. انتخاب معیارهای توقف ضروری است زیرا یک درخت ساده محدودیت‌هایی در متغیرهای



شکل ۴- شماتیک ورودی و خروجی مدل‌های SVM و BT

Fig. 4 Schematic of SVM and BT model inputs and outputs

۲-۵- معیارهای ارزیابی

برای ارزیابی مدل‌ها و مقایسه کارایی آنها، در این پژوهش از معیارهای جذر میانگین مربعات خطا^۱، ضریب همبستگی^۲، ضریب تعیین^۳ و ضریب کارایی نش- ساتکلیف^۴ استفاده گردید. در ادامه، معادلات مربوطه با شماره‌های (۴)، (۵)، (۶) و (۷) ارائه شده است (Kumar et al. 2019, Shiri et al. 2016).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \quad (4)$$

$$r = \frac{n(\sum_{i=1}^n O_i P_i) - (\sum_{i=1}^n O_i)(\sum_{i=1}^n P_i)}{\sqrt{(n \sum_{i=1}^n O_i^2 - (\sum_{i=1}^n O_i)^2)(n \sum_{i=1}^n P_i^2 - (\sum_{i=1}^n P_i)^2)}} \quad (5)$$

$$R^2 = \frac{[\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})]^2}{\sum_{i=1}^n (O_i - \bar{O})^2 \sum_{i=1}^n (P_i - \bar{P})^2} \quad (6)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (P_i - \bar{P})^2} \quad (7)$$

که، n تعداد داده‌ها، P_i داده‌های مشاهداتی، $\bar{Q}$ میانگین داده‌های مشاهداتی و Q_i داده‌های محاسباتی می‌باشد. هر چقدر مقدار R^2 به عدد یک نزدیکتر باشد، برآورد و برازش را بهتر نشان می‌دهد. ضریب کارایی نش- ساتکلیف می‌تواند از منفی بی‌نهایت تا ۱ تغییر کند، هرچه مقادیر به عدد یک نزدیکتر باشد، مدل برازش بهتری دارد و عدد یک برازش عالی را نشان می‌دهد. $RMSE$ جذر میانگین مربع خطاهای داده‌های محاسباتی و مشاهداتی را بیان می‌کند. هر چه مقدار این عدد کمتر باشد، عملکرد مدل در شبیه‌سازی داده‌ها بهتر است.

امروزه یکی از ابزارهای تعیین تغییرهای اقلیمی تحلیل روند می‌باشد. روش‌های مختلفی برای تعیین روند وجود دارند که به دو دسته روش پارامتریک و ناپارامتریک تقسیم می‌شوند. در اصل آزمون‌های پارامتریک در صورتی که داده‌ها به صورت نرمال باشند، قوی‌تر از آزمون‌های ناپارامتریک هستند. حال - آنکه شرایط داده‌های غیرنرمال آزمون ناپارامتریک قوی‌تر از پارامتریک است (Sabouhi and Soltani 2008). از آنجایی که توزیع تعدادی از سری‌های اقلیمی از جمله بارش و دبی نرمال نیست، در چنین حالت‌هایی روش من- کندال مناسب‌تر است. تحلیل روند در این مطالعه با استفاده از

آزمون ناپارامتریک من-کندال انجام شد. آزمون من- کندال ابتدا توسط Mann (1945) مطرح و سپس توسط Kendall (1975) بسط و توسعه یافت.

این روش بطور متداول و گسترده‌ای در تحلیل روند سری‌های زمانی هیدرولوژیکی و هواشناسی بکار گرفته می‌شود (Lettenmaier 1994). از نقاط قوت این روش می‌توان به مناسب بودن کاربرد آن برای سری‌های زمانی که از توزیع آماری خاصی پیروی نمی‌کنند، اشاره کرد. اثرپذیری ناچیز این روش از مقادیر حدی که در برخی از سری‌های زمانی مشاهده می‌شوند نیز از دیگر مزیت‌های استفاده از این روش است (Turgay and Ercan 2005). فرض صفر این آزمون بر تصادفی بودن و وجود نداشتن روند در سری داده‌ها دلالت دارد و پذیرش فرض یک (رد فرض صفر) دال بر وجود روند در سری داده‌ها می‌باشد. آزمون من- کندال با تعریف آماره S به صورت رابطه (۸) تعریف می‌شود.

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sign}(X_i + X_j) \quad (8)$$

که در آن X_i و X_j مقادیر مرتب شده‌ی نمونه و n تعداد نمونه است. مقدار $\text{Sign}(x_j - x_i)$ برای $0 < (x_j - x_i)$ برابر با منفی یک، $0 = (x_j - x_i)$ برابر با صفر و برای $0 > (x_j - x_i)$ برابر با یک است. مقدار نمره استاندارد شده Z و واریانس آماره S نیز از رابطه (۹) محاسبه می‌شود.

$$Z_c = \begin{cases} \frac{S-1}{\sqrt{\text{Var}(S)}} & , S > 0 \\ 0 & , S = 0 \\ \frac{S+1}{\sqrt{\text{Var}(S)}} & , S < 0 \end{cases} \quad (9)$$

مقادیر مثبت Z نشان‌دهنده‌ی روندهای افزایشی و مقادیر منفی Z نشان‌دهنده‌ی روندهای کاهش‌ی است. در شرایطی که $S=0$ باشد داده‌ها فاقد روند بوده و هیچ روند معنی‌دار یا غیر معنی‌داری در داده‌ها وجود ندارد که این شرایط ایده‌آل برای داده‌های هیدرولوژیکی بندرت ممکن است اتفاق بیفتد. در این پژوهش سطح‌های معنی‌داری $P=0.05$ به کار گرفته شده است (Barani et al. 2019).

۳- یافته‌ها و بحث

۳-۱- سری زمانی بارش و رواناب

سری زمانی مربوط به بارش و رواناب ایستگاه آخولا در شکل (۵-الف و ۵-ب) و پل سنخ (۵-ج و ۵-د) در دو دوره مطالعاتی قبل و بعد از سال ۱۳۷۴ نشان داده شده است.

¹ Root-mean-square deviation

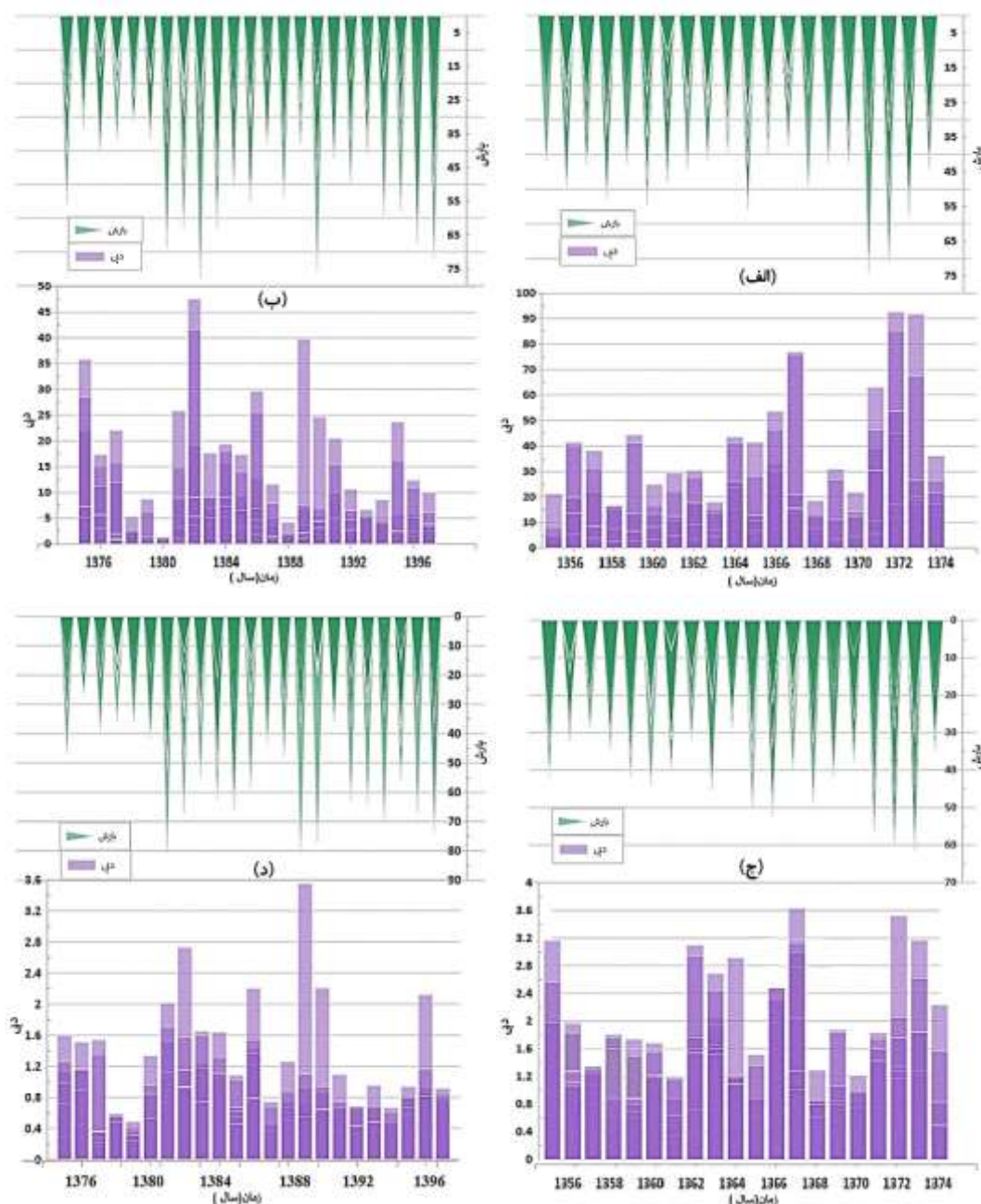
² Correlation coefficient

³ Coefficient of determination

⁴ Nash-Sutcliffe efficiency

بنابراین بارش سالانه در سال های ۱۳۵۵ تا ۱۳۹۷ دارای روند نبوده و سری زمانی بارش نوسانی است و روند در این دوره زمانی طبیعی می باشد. بررسی شکل (۵) همچنین نشان داد که در هر دو ایستگاه آخولا و پل سنیخ، بارش روند مشخصی نداشته و در برخی سال ها روند افزایشی و در برخی دیگر روند کاهشی داشته است.

مقدار آماره من کندال متغیر بارش در دو دوره مطالعاتی برای ایستگاه آخولا (الف) و (ب) به ترتیب برابر با ۰/۰۴۱ و ۰/۰۳۴ و برای ایستگاه پل سنیخ (ج) و (د) برابر با ۰/۰۳۹ و ۰/۰۲۸ تعیین شد که در سطح ۹۵٪ معنادار نمی باشد. این موضوع مؤید این است که بارش سالانه در این سال ها برای دو ایستگاه مورد مطالعاتی از نظر آماری معنادار نیست.



شکل ۵- سری زمانی متوسط بارش و دبی خروجی: الف- آخولا (۱۳۵۵-۱۳۷۴)، ب- آخولا (۱۳۷۵-۱۳۹۷)، ج- پل سنیخ (۱۳۷۴-۱۳۵۵) و د- پل سنیخ (۱۳۷۵-۱۳۹۷)

Fig. 5 Time series plots of the observed rainfall and runoff values: a) Akhula (1976-1994), b) Akhula (1995-2019), c) Pole Senikh (1976-1994) and d) Pole Senikh (1995-2019)

بارندگی های ماهانه کمتر از ۵ mm در سطح احتمال ۱٪ افزایش معنی داری داشته و در برخی محدوده ها روند

Salehi Babil et al. (2017) گزارش کردند که نوسانات بارش در زیر حوضه های دریاچه ارومیه وجود دارد و در

رودخانه‌ها نام برد. (Nazeri Tahroudi et al. (2018) روند کاهشی داده‌های سطح آب دریاچه ارومیه را یک سال پس از روند کاهشی داده‌های جریان گزارش کرد. مطالعات زیادی برای بررسی تغییرات روند جریان رودخانه و میزان بارش حوضه دریاچه ارومیه انجام شده است و همه آنها تأیید می‌کنند که کاهش رواناب رودخانه‌های حوضه دریاچه ارومیه چشمگیر بوده است (Hosseini-Moghari et al. 2020 ، Bashirian et al. 2020).

۳-۲- پارامترهای آماری

جهت بررسی ایستگاه‌های آخولا و پل سنیخ در دو دوره، آمار بارندگی و رواناب (میانگین، حداقل، حداکثر) برای دوره ۱۳۷۴-۱۳۵۵ و دوره ۱۳۹۷-۱۳۷۵ در جدول (۱) ارائه شده است. مدل‌های BT و SVM (تابع کرنل) با داده‌های واقعی آموزش و آزمون شدند. همان‌طور که مشاهده می‌شود نتایج مربوط به دوره اول مطالعاتی (۱۳۷۴-۱۳۵۵) در ردیف اول و نتایج مربوط به دوره دوم مطالعاتی (۱۳۹۷-۱۳۷۵) داخل پرانتز و زیر نتایج دوره اول نوشته شده است. طبق جدول (۱) در هر دو دوره، در ستون داده کل، ایستگاه آخولا بیشترین میانگین بارش و رواناب را داشته و کمترین مقادیر میانگین بارندگی و رواناب به ایستگاه پل سنیخ مربوط می‌شود. همان‌طور که اشاره شد به دلیل اینکه داده‌های منتخب دو مجموعه آموزش برای هر دو مدل متفاوت می‌باشد، طبیعتاً اعداد به دست آمده برای پارامترهای آماری این دو مدل، با یکدیگر تفاوت دارد.

جدول ۱- پارامترهای آماری داده‌های بارش و رواناب در مدت ۱۳۷۴-۱۳۵۵ و ۱۳۹۷-۱۳۷۵

Table 1 Statistical parameters of the rainfall and runoff data during 1976-1994 and (1995-2019)

Station	Parameter	Entire data 1976-1994 (1995-2019)		Training set		Test set	
		P (mm/month)	Q (m ³ /s)	BT	SVM	BT	SVM
Akhula	Mean	21.97 (22.40)	11.74 (5.07)	11.32 (4.53)	10.46 (5.73)	11.13 (5.97)	11.90 (3.38)
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00
	Maximum	(0.81)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
		74.31	92.50	76.85	76.85	53.85	75.48
		(77.66)	(47.47)	(47.47)	(47.47)	(41.65)	(36.15)
Pole senikh	Mean	18.97 (20.39)	0.86 (0.52)	0.85 (4.53)	0.86 (0.54)	0.93 (0.51)	0.91 (0.41)
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00
	Maximum	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
		61.38	3.63	3.63	3.86	3.86	3.70
		(81.25)	(3.55)	(3.55)	(2.73)	(2.21)	(2.23)

۳-۳- مدل های ماشین بردار پشتیبان و درخت تقویت شده

به منظور نشان دادن قابلیت مدل ها برای مدل سازی بارش-رواناب، در این بخش فرآیند اعتبارسنجی مدل ها مورد بحث قرار می گیرد. مقادیر معیار ارزیابی کمی عملکرد مدل های SVM با تابع کرنل و BT در ایستگاه هیدرومتری آخولا و پل سنیخ در مدل های کاربردی در این مطالعه تحلیل و مقایسه شد. ضریب همبستگی، ضریب تعیین، کارایی نش سائکلیف و جذر میانگین مربعات خطا بر اساس مجموعه داده های

آموزش و آزمون محاسبه شد. نتایج نشان می دهد که هر دو مدل در هر دو دوره مطالعاتی عملکرد خوبی دارند. با توجه به نتایج جدول (۲)، در هر دو دوره مطالعاتی در ایستگاه آخولا مدل SVM نسبت به مدل BT عملکرد بهتری داشته است و با دقت خوبی توانسته دبی جریان و بارش ماهانه را شبیه سازی کند. همچنین در این دوره ها در ایستگاه پل سنیخ مدل BT کارایی بهتری را نسبت به مدل SVM نشان می دهد.

جدول ۲- عملکرد مدل های SVM و BT در مجموعه های آموزشی و آزمون در مدت ۱۳۹۷-۱۳۵۵

Table 2 Performance of SVM and BT models in the training and testing sets during 1976-2019

Table 2: Performance of SVM and BT models in the training and testing sets during 1976-2019									
Station	Entire data (1976-2019)	Cross Correlation		R ²		NSE		RMSE	
		BT	SVM	BT	SVM	BT	SVM	BT	SVM
1976-1994									
Akhula	Train	0.90	0.90	0.82	0.81	0.81	0.81	6.86	6.59
	Test	0.81	0.89	0.65	0.80	0.64	0.72	8.13	8.09
Pole senikh	Train	0.84	0.83	0.70	0.69	0.70	0.69	0.45	0.46
	Test	0.83	0.82	0.69	0.67	0.69	0.67	0.51	0.54
1995-2019									
Akhula	Train	0.79	0.83	0.62	0.68	0.62	0.67	4.23	4.53
	Test	0.69	0.75	0.48	0.57	0.47	0.55	6.00	3.65
Pole senikh	Train	0.80	0.80	0.65	0.64	0.64	0.64	0.31	0.29
	Test	0.77	0.76	0.59	0.58	0.58	0.57	0.35	0.32

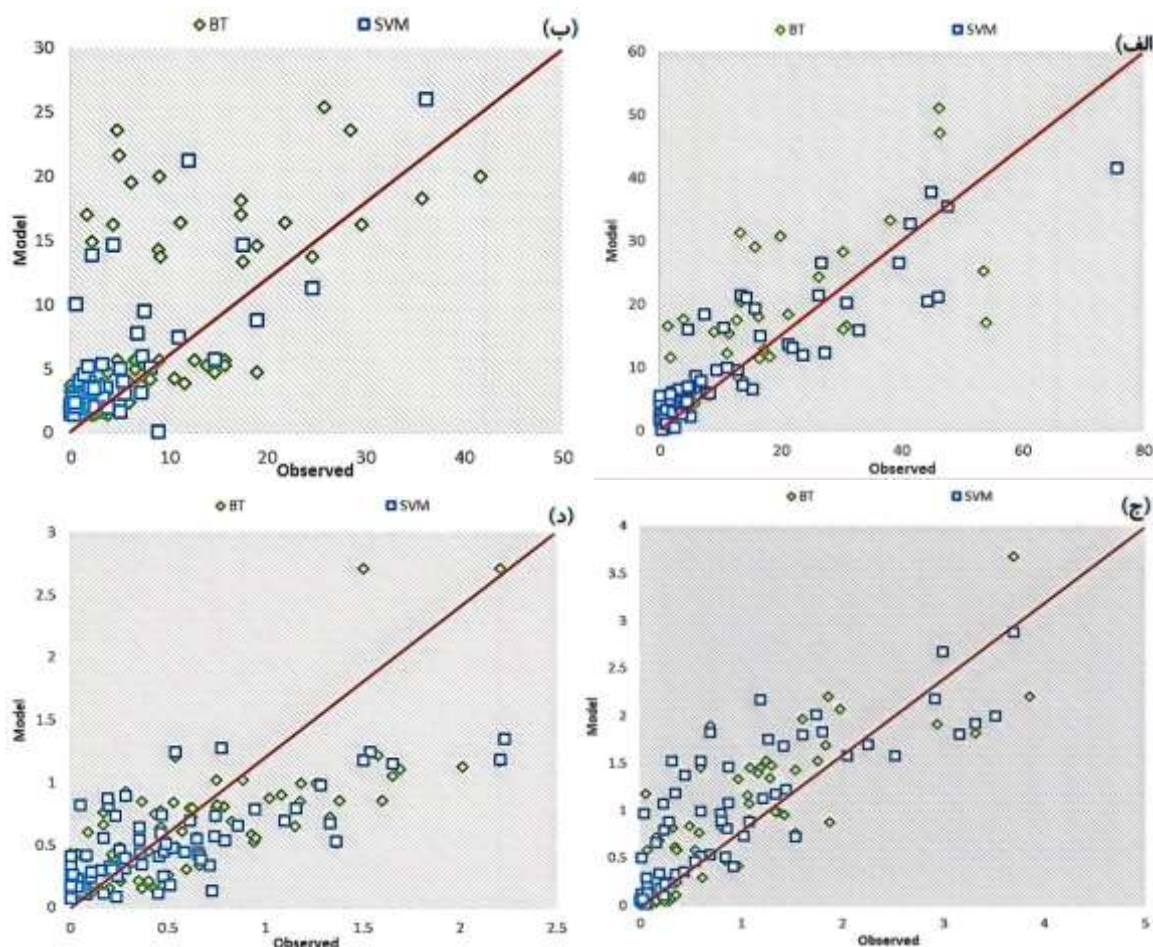
همان طور که در جدول (۲) (دوره اول) به وضوح دیده می شود همبستگی متقابل در مدل SVM در ایستگاه آخولا برای داده های آموزشی ۰/۹۰ و در داده های آزمون ۰/۸۹ به دست آمده است. عدد R² نیز در این مدل برای داده های آموزشی و آزمون به ترتیب برابر با ۰/۸۱ و ۰/۸۰ تعیین شد. شاخص کارایی نش سائکلیف داده های آموزشی و آزمون به ترتیب برابر است با ۰/۸۱ و ۰/۷۲ و آخرین شاخص که در این مطالعه مورد بررسی قرار گرفته است خطای جذر میانگین مربعات می باشد و مقدار آن برای داده های آموزشی و آزمون به ترتیب برابر با ۶/۵۹ و ۸/۰۹ به دست آمده است. اما در این دوره برای ایستگاه پل سنیخ، مقدار ضریب همبستگی برای دوره های آموزشی و آزمون مدل BT برابر با ۰/۸۴ و ۰/۸۳ ثبت گردید. مقدار R² برای این مدل، برای داده های آموزشی ۰/۷۰ و برای داده های آزمون ۰/۶۹ تعیین شد. شاخص NSE برای این مدل و دوره های آموزشی و آزمون، ۰/۷۰ و ۰/۶۹ تعیین شد. اما آخرین شاخص خطای جذر میانگین مربعات به ترتیب برابر با ۰/۴۵ و ۰/۵۱ برای مجموعه آموزشی و آزمون ثبت گردید. اما در دوره دوم (۱۳۹۷-۱۳۷۵) و برای ایستگاه آخولا، مقدار CC برابر با

۰/۸۳ و ۰/۷۵ (آموزشی و آزمون) به دست آمد. همچنین مقدار R² برای دوره های آموزشی و آزمون، ۰/۶۸ و ۰/۵۷ و مقدار شاخص NSE برای مجموعه آموزشی و آزمون ۰/۶۷ و ۰/۵۵ تعیین گردید. آخرین پارامتر خطای جذر میانگین مربعات برای دوره آموزش و آزمون به ترتیب برابر با ۴/۵۳ و ۳/۶۵ به دست آمد. اما مقادیر مدل BT برای ایستگاه پل سنیخ و برای پارامتر CC و مجموعه آموزشی و آزمون به ترتیب ۰/۸۰ و ۰/۷۷، پارامتر R²، ۰/۶۵ و ۰/۵۹، آماره NSE برابر با ۰/۶۴ و ۰/۵۸ و آخرین پارامتر RMSE به ترتیب با ۰/۳۱ و ۰/۳۵ تعیین گردید. نتایج نشان داد که در دوره دوم، به دلیل کاهش چشمگیر جریان آب خروجی از دو ایستگاه زیر حوضه تبریز نسبت به دوره اول، ارتباط بین بارش و رواناب کاهش یافته است. بنابراین، در دوره دوم، عملکرد مدل ها در تمامی شاخص های ارزیابی کاهش یافته است. با توجه به تأثیر ارتباط بارش و رواناب بر مدل ها، نتایج در دوره دوم منطقی به نظر می آید.

همچنین، نتایج در داده های آزمون نیز نشان می دهد که شاخص هایی مانند CC، R²، NSE و RMSE برای بارش و رواناب در داده های آزمون قابل قبول هستند و می توان

دانست. به عبارت دیگر، می‌توان از این مدل‌ها برای پیش‌بینی رواناب در آینده استفاده کرد.

مدل‌های ماشین بردار پشتیبان و درخت تقویت‌شده را برای ایستگاه‌های آخولا و پل سنیخ زیر حوضه تبریز مناسب



شکل ۶- نمودارهای پراکنده رواناب شبیه‌سازی شده با استفاده از SVM و BT: الف- آخولا (۱۳۷۴-۱۳۵۵)، ب- آخولا (۱۳۹۷-۱۳۷۵)، ج- پل سنیخ (۱۳۷۴-۱۳۵۵) و د- پل سنیخ (۱۳۹۷-۱۳۷۵)

Fig. 6 Verification scatter plots of simulated runoff using SVM and BT: a) Akhula (1976-1994), b) Akhula (1995-2019), c) Pole Senikh (1976-1994) and d) Pole Senikh (1995-2019)

از SVM می‌باشد. از طرف دیگر فاصله داده‌ها از محور یک‌به‌یک در مدل منتخب ایستگاه آخولا (مدل SVM) و همین‌طور در مدل منتخب ایستگاه پل سنیخ (مدل BT) در دوره دوم مطالعاتی نسبت به دوره اول بیشتر می‌باشد که در مقایسه با مطالعات انجام شده در این منطقه می‌توان به مطالعات (Bigdeli et al. 2023) اشاره کرد. در این مطالعه مدل‌سازی بارش- رواناب برای حوضه آجی چای با دو مدل RF و ANN انجام شد. نتایج نشان داد که برای این منطقه که شامل ۶ ایستگاه بود مدل شبکه عصبی مصنوعی کارایی بالاتری نسبت به جنگل تصادفی نشان داد. در مطالعه Nourani et al. (2009) برای پیش‌بینی رواناب ماهانه و روزانه در حوضه ليقوان‌چای از سه مدل هوشمند (شبکه‌های عصبی مصنوعی، سیستم استنتاج فازی و سیستم عصبی-

همچنین در شکل (۶) نمودار پراکنش مقادیر مشاهداتی در برابر مقادیر شبیه‌سازی شده دبی‌های ماهانه، با استفاده از مدل‌های ماشین بردار پشتیبان و درخت تقویت‌شده برای دو ایستگاه آخولا و پل سنیخ در دو دوره مطالعاتی نشان داده شده است. در شکل (۶)، در آخولا (۶-الف) و (۶-ب) به‌طور نسبی حالت بیش تخمینی دیده می‌شود. حال آنکه در قسمت پل سنیخ (۶-ج) بیش تخمینی و در پل سنیخ (۶-د) کم تخمینی اتفاق افتاده است. همچنین شکل (۶) قرار گرفتن بیشتر نقاط روی یا نزدیکی خط نیمساز ($Y=X$) و دقت بیشتر این مدل‌ها در شبیه‌سازی بارش- رواناب را نمایش می‌دهد. پراکندگی داده‌ها در ایستگاه آخولا در هر دو دوره مطالعاتی و برای مدل SVM کمتر از BT بوده و در ایستگاه پل سنیخ نیز، پراکندگی داده‌ها برای مدل BT کمتر

آجی چای و پل سنیخ در ایستگاه آخولا و پل سنیخ، نشان داد که روند کاملاً نزولی برای جریان آب برای رودخانه‌ها برای دو دوره مطالعاتی ثبت شده است.

از دلایل مهم کاهش ورود دبی رودخانه‌ها به دریاچه ارومیه، در ایستگاه‌های مورد مطالعه، استفاده از جریان آب در بخش کشاورزی و صنعت می‌باشد. از دلایل دیگر می‌توان به احداث سد های متعدد بر روی رودخانه حوضه دریاچه ارومیه اشاره کرد. با توجه به عملکرد خوب این مدل‌ها در شبیه سازی بارش-رواناب، می‌توان برای سایر ایستگاه‌های حوضه دریاچه ارومیه از روش‌های یادگیری ماشین ذکر شده استفاده کرد.

سیاسگزاری

این مقاله با استفاده از داده‌های دریافتی از شرکت آب منطقه‌ای و سازمان هواشناسی استان آذربایجان شرقی نگارش شده است. از این رو نویسندگان مقاله از سازمان‌های نامبرده کمال تشکر را دارند.

دسترسی به داده‌ها

داده‌های این پژوهش از (شرکت آب منطقه‌ای و سازمان هواشناسی استان آذربایجان شرقی) اخذ شده است که بعد از اخذ مجوز از (شرکت آب منطقه‌ای و سازمان هواشناسی استان آذربایجان شرقی) قابل ارائه می‌باشد.

تضاد منافع نویسندگان

نویسندگان این مقاله اعلام می‌دارند که، هیچ‌گونه تضاد منافی در رابطه با نویسندگی و یا انتشار این مقاله ندارند.

References

- 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. *Land Reclam.*, 45 (1). DOI: [10.29252/jwmr.8.16.22](https://doi.org/10.29252/jwmr.8.16.22)
- Adnan, R. M., Yuan, X., Kisi, O., Adnan, F., & Mehmood, A. (2018). Stream flow forecasting of poorly gauged mountainous watershed by Least Square Support Vector Machine, Fuzzy Genetic Algorithm and M5 Model Tree using climatic data from Nearby Station. *Water Resour Manag*, 32, 4469-4486. DOI: [10.1007/s11269-018-2033-2](https://doi.org/10.1007/s11269-018-2033-2)
- Aurenhammer, F., Klein, R., & Lee, D. T. (2013). Voronoi Diagrams and Delaunay Triangulations. World Scientific Publ. Co., Singapore, p. 337. DOI: [10.1142/8685](https://doi.org/10.1142/8685)
- Barani, N., & Karami, A. (2019). Annual trend analysis of climate parameters of temperature and precipitation in decuple agroecology regions of Iran. *Environm. Sci.*, 17(4), 75-90. DOI: [10.29252/envs.17.4.75](https://doi.org/10.29252/envs.17.4.75)

فازی (تطبیقی) استفاده کردند. نتایج سه مدل مذکور را با نتایج به دست آمده از روش‌های رگرسیون خطی و مدل سری زمانی ARIMA مقایسه کردند که مدل سازی فازی (تطبیقی و استنتاجی) بیشترین مقدار نش- ساتکلیف و کمترین خطا را در پی داشت.

۴- نتیجه گیری

در این پژوهش کارایی مدل‌های هوشمند ماشین بردار پشتیبان و درخت تقویت شده به منظور مدل سازی بارش-رواناب ایستگاه‌های پل سنیخ و آخولا برای بازه زمانی ۴۳ ساله و در دو دوره مطالعاتی (۱۳۷۴-۱۳۵۵) و (۱۳۹۷-۱۳۷۵) مورد بررسی قرار گرفت. نتایج به دست آمده به شرح زیر می‌باشد:

۱- در هر دو دوره مطالعاتی برای ایستگاه آخولا مدل SVM با تابع کرنل کارایی بالاتری نسبت به مدل BT نشان داد و در ایستگاه پل سنیخ نیز برای دو دوره مطالعاتی، مدل BT عملکرد بهتری را نسبت به مدل SVM ارائه کرد.

۲- مقدار همبستگی متقابل برای مجموعه آزمون و در دو دوره مطالعاتی اول و دوم ایستگاه آخولا به ترتیب برابر با ۰/۸۹ و ۰/۷۵ تعیین و همین مقدار آزمون برای دوره‌های مطالعاتی اول و دوم ایستگاه پل سنیخ برابر با ۰/۸۳ و ۰/۷۷ به دست آمد.

۳- با توجه به نمودارهای سری زمانی بارش و رواناب طی ۴۳ سال و همچنین نتایج حاصل از آزمون من- کندال، برای هر دو ایستگاه و برای دو دوره مطالعاتی، روند مشخصی برای بارش دیده نشد، درواقع نمودارها نشان دادند که بارش در این مناطق به صورت نوسانی بوده است. اما نتایج آزمون من کندال و نمودارهای سری زمانی برای دبی رودخانه‌های

- Bashirian, F., Rahimi D., Movahedi S., & Zakerinejad, R. (2020). Water level instability analysis of Urmia Lake Basin in the northwest of Iran. *Arab. J. Geosci.*, 13(4), 1-14. DOI: [10.1007/s12517-020-5207-1](https://doi.org/10.1007/s12517-020-5207-1)
- Bigdeli, Z., Majnooni Heris, A., Delirhasannia, R., & Karimi, S. (2023). Rainfall-Runoff Modeling of Aji Chai Basin using random forest and artificial neural network models. *New Res. Sustain. Water Eng.*, 1(2), 27-42. DOI: [10.22103/nrswe.2023.20278.1013](https://doi.org/10.22103/nrswe.2023.20278.1013)
- Dibike, Y. B., Velickov, S., Solomatine, D. P., & Abbott, M. B. (2001). Model induction with support vector machines: introduction and applications. *J. Comput. Civ. Eng.*, 15(3), 208-216. DOI: [10.1061/\(ASCE\)0887-3801\(2001\)15:3\(208\)](https://doi.org/10.1061/(ASCE)0887-3801(2001)15:3(208))
- Fazli Fard, P., Sheikh, V., Sadoddin, A., & Hessari, B. (2019). Study of trends in precipitation and stream flow of the lake Urmia Basin during the past four decades. *Water Soil Sci.*, 29(4), 27-41. [In Persian].
- Friedman, N., Goldszmidt, M., & Wyner, A. (1999). Data analysis with Bayesian networks: A bootstrap approach. Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence. Burlington (MA): Morgan Kaufmann Publishers. DOI: [10.1061/\(ASCE\)0887-3801\(2001\)15:3\(208\)](https://doi.org/10.1061/(ASCE)0887-3801(2001)15:3(208))
- Ghorbani, M., Azani, A., & Mahmoudi Vanolya, S. (2015). Rainfall-Runoff Modeling Using Hybrid Intelligent Models Iran. *Water Resour. Manag.*, 11(2), 146-150 [In Persian]. DOI: [20.1001.1.2008479.1400.52.5.18.2](https://doi.org/10.1001.1.2008479.1400.52.5.18.2)
- Hosseini -Moghari, S. M., Araghinejad, S., Tourian, M. J., Ebrahimi, K., & Döll, P. (2020). Quantifying the impacts of human water use and climate variations on recent drying of Lake Urmia basin: the value of different sets of spaceborne and *in situ* data for calibrating a global hydrological model. *Hydrol Earth Syst. Sci.*, 24(4), 1939-1956. DOI: [10.5194/hess-24-1939-2020](https://doi.org/10.5194/hess-24-1939-2020)
- <https://en.wikipedia.org/wiki/Tabriz>
- https://en.wikipedia.org/wiki/Aji_Chay
- Hussain, D., & Khan, A. A. (2020). Machine learning techniques for monthly river flow forecasting of Hunza River, Pakistan. *ESIN*, 13, 939-949. DOI: [10.1007/s12145-020-00450-z](https://doi.org/10.1007/s12145-020-00450-z)
- Kendall, M. G. (1975). Rank correlation methods. fourth ed. Charles Griffin, London.
- Kisi, O., & Parmar, K. S. (2016). Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution. *J. Hydrol.*, 534, 104-112. DOI: [10.1016/j.jhydrol.2015.12.014](https://doi.org/10.1016/j.jhydrol.2015.12.014)
- Kumar, S., Roshni, T. and Himayoun, D. (2019). A comparison of emotional neural network (ENN) and artificial neural network (ANN) approach for rainfall-runoff modelling. *C. E. J.*, 5(10), 2120-2130. DOI: [10.28991/cej-2019-03091398](https://doi.org/10.28991/cej-2019-03091398)
- Lee, S., Kim, J. C., Jung, H. S., Lee, M. J., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomat. Nat. Hazard. Risk*, 8(2), 1185-1203. DOI: [10.1080/19475705.2017.1308971](https://doi.org/10.1080/19475705.2017.1308971)
- Lorrai, M., & Sechi, M. G. (1995). Neural nets for modeling rainfall-runoff transformation. *Water Resour. Manag.*, 9, 299 - 313. DOI: [10.1007/BF00872489](https://doi.org/10.1007/BF00872489)
- Mann, H. B. (1945). Non-parametric test against trend. *Econom.*, 13, 245-259.
- Nazeri Tahroudi, M., Ahmadi, F., & Khalili, K. (2017). Evaluation the trend and trend chang point of Urmia Lake Basin precipitation. *Journal of water and soil (Agricultural sciences and technology)*, 31(2), 644-659 [In Persian]. DOI: [10.22067/JSW.V31I2.55338](https://doi.org/10.22067/JSW.V31I2.55338)
- Nazeri Tahroudi, M., Ahmadi, F., & Khalili, K. (2018). Impact of 30 years changing of river flow on Urmia Lake Basin. *AJCE*, 2(1), 115-122. DOI: [10.22060/AJCE.2018.14520.5481](https://doi.org/10.22060/AJCE.2018.14520.5481)
- Nourani, V., Keynezhad, M., & Makani L. (2009). Using adaptive neuro-fuzzy inference system rainfall-runoff modeling. *J. Civil Environ. Eng.*, 39, 75-81 [In Persian].
- Phomcha, P., Wirojanagud, P., Vangpaisal, T., & Thaveevouthti, T. (2011). Suitability of SWAT model for simulating of monthly streamflow in Lam Sonthi Watershed. *J. Indust. Technol.*, 7(2), 49- 56.
- Poursalehi, F., Khasheisiuki, A., & Hashemi, S. R. (2022). Investigating the performance of random forest algorithm in predicting water

- table fluctuations compared with two models of decision tree and artificial neural network (Case study: unconfined aquifer of Birjand plain). *Iran. J. Ecohydrol.*, 8(4), 961-974 [In Persian]. DOI: [10.22059/IJE.2022.327263.1526](https://doi.org/10.22059/IJE.2022.327263.1526)
- Sabouhi, R., & Soltani, S. (2008). Analysis of the climate trend in the major cities of Iran. *J. Agric. Sci. Technol.*, 12(46). [In Persian]. DOI: [10.22059/JPHGR.2021.301111.1007528](https://doi.org/10.22059/JPHGR.2021.301111.1007528)
- Salehi Babil, S., Zeinalzadeh, K., & Hessari, B. (2017). The changes in the frequency of daily precipitation in Urmia Lake basin, Iran. *Theor. Appl. Climatol.*, 1- 10. DOI: [10.1007/s00704-017-2177-7](https://doi.org/10.1007/s00704-017-2177-7)
- Seyedian, S. M., Soleimani, M., & Kashani, M. (2014). Predicting streamflow using data-driven model and time series. *Iran. J. Ecohydrol.*, 1(3), 167-179. [In Persian]. DOI: [10.22059/IJE.2014.54219](https://doi.org/10.22059/IJE.2014.54219)
- Shafeizadeh, M., Fathian, H., & Nikbakht Shahbazi, A. (2019). Continuous rainfall-runoff simulation by artificial neural networks based on efficient input variables selection using partial mutual information (PMI) algorithm. *Iran-Water Resour. Res.*, 15(2), 144-161 [In Persian]. DOI: [20.1001.1.17352347.1398.15.2.12.1](https://doi.org/20.1001.1.17352347.1398.15.2.12.1)
- Shekhar, S., & Xiong, H. (2007). Encyclopedia of GIS. Springer Science & Business Media, New York, USA, p. 1370.
- Shiri, J., Shamshirband, S., Kisi, O., Karimi, S., Bateni, S. M., Hosseini Nezhad, S. H., & Hashemi, A. (2016). Prediction of water-level in the Urmia Lake using the extreme learning machine approach. *Water Resour. Manag.*, 30, 5217-5229. DOI: [10.1007/s11269-016-1480-x](https://doi.org/10.1007/s11269-016-1480-x)
- Worland, S. C., Farmer, W. H., & Kiang, J. E. (2018). Improving predictions of hydrological low-flow indices in ungauged basins using machine learning. *Environ. Model. Software*, 101, 169-182. DOI: [10.1016/j.envsoft.2017.12.021](https://doi.org/10.1016/j.envsoft.2017.12.021)
- Yoon, H., Jun, S. C., Hyun, Y., Bae, G. O., & Lee, K. K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *J. Hydrol.*, 396, 128-138. DOI: [10.1016/j.jhydrol.2010.11.002](https://doi.org/10.1016/j.jhydrol.2010.11.002)
- Zarei, M., Zandi, R., & Naemitabar, M. (2022). Assessment of flood occurrence potential using data mining models of support vector machine, chaid and random forest (Case study: Frizi watershed). *JWMR*, 13 (25) ,133-144. [In Persian]. DOI: [10.52547/jwmr.13.25.133](https://doi.org/10.52547/jwmr.13.25.133)

How to cite this paper:

Bigdeli, Z., Majnooni-Heris, A., Delirhasannia, R. and Karimi, S. (2023). Application of support vector machine and boosted tree algorithm for rainfall-runoff modeling (Case study: Tabriz Plain). *Environ. Water Eng.*, 9(4), 532-547. DOI: [10.22034/ewe.2023.366913.1816](https://doi.org/10.22034/ewe.2023.366913.1816).