Archive of SID.ir

Water and Soil Management and Modeling

Online ISSN: 2783 - 2546

University of Mohaghegh Ardabili

Monthly prediction of pan evaporation using individual and combined approaches of data mining models in arid regions

Mojtaba Mohammadi[1]*, Hossein Jahantigh[1], Farhad Zolfaghari[1]

[1] Assistant Professor, Department of Desert Management and Control, Faculty of Natural Resource, Higher Educational Complex of Saravan,Saravan, Iran

Abstract

Introduction

Evaporation, the process by which water molecules escape a surface after absorbing sufficient energy to overcome vapor pressure, is a major contributor to water scarcity, especially in arid and semi-arid regions where heat readily facilitates this escape. Accurately estimating evaporation losses is crucial for effective water resource management, crop water demand prediction, and irrigation scheduling. Machine learning (ML) has emerged as a powerful tool for tackling the complex and stochastic nature of environmental problems. ML models excel at identifying relationships between predictor variables and outcomes (predictands), often surpassing traditional methods. However, their performance can vary depending on input factors and climatic conditions. Recently, hybrid techniques that combine multiple models have gained traction in climate and hydrology studies. These techniques leverage the strengths of different approaches within a single algorithm, potentially capturing more complex patterns in data series. This research will explore the potential of various individual ML models and propose a novel hybrid approach for estimating pan evaporation in Sistan and Baluchistan Province.

Materials and Methods

This study investigates pan evaporation simulation and prediction in Sistan and Baluchistan Province, Iran. Synoptic station data (1980-2019) served as model inputs, while pan evaporation measurements from these stations provided the observed values. In this research, in the approach of individual performance of data mining models, eight data mining models were used to simulate and predict evaporation from the pan. In addition to the individual performance approach, the combined VEDL approach was used to provide a hybrid model (a combination of the mentioned eight individual models of deep learning). In this hybrid approach to regression issues, the estimators of all models are averaged to obtain an estimate for a set called vote regressors (VRs). There are two approaches to awarding votes: average voting (AV) and weighted voting (WV). In the case of AV, the weights are equivalent and equal1. A disadvantage of AV is that all of the models in the ensemble are accepted as equally effective; however, this situation is very unlikely, especially if different machine learning algorithms are used. WV specifies a weight coefficient for each ensemble member. The weight can be a floating-point number between Zero and one, in which case the sum is equal to one, or an integer starting at one denoting the number of votes given to the corresponding ensemble member. the weight of each model was selected based on the accuracy of the model's performance using the evaluation criteria obtained from the training implementation section of individual models.

Results and Discussion

The results showed that all the models had very good results in both the training and testing stages. All models exhibited excellent performance during training and testing. The Artificial Neural Network (ANN) achieved the highest accuracy in both phases at the Zahedan station (R^2 = 0.89, RMSE = 45.95 in training; R^2 = 0.96, RMSE = 44.18 in validation). It emerged as the best model for monthly pan evaporation prediction at this station. Other models also performed well, with the Support Vector Machine (SVM) and Random Forest (RF) models achieving R^2 values of 0.89 and 0.88 in training, respectively. Notably, the BART model ranked second in validation (R^2 = 0.96). The Tree Model (TM) had the lowest accuracy (R^2 = 0.84 and 0.93 in training and validation, respectively). Across all stations, ANN, SVM, and RF consistently delivered the best results in both

Archive of SID.ir

Archive of SID.ir

Water and Soil Management and Modeling

Online ISSN: 2783 - 2546

University of Mohaghegh Ardabili

training and testing. In the test phase, the SVM model outperformed others in Khash, Iranshahr, and Chabahar stations (R^2 = 0.94, 0.96, and 0.94, respectively). At the Saravan station, the RF model achieved the highest R^2 (0.94) during testing. To develop a hybrid data mining model, the Voting Ensemble for Deep Learning (VEDL) technique was employed with weighted voting in the training stage. The combined model significantly improved upon the best individual model. RMSE decreased from 45.95 to 33.1, R^2 increased from 0.89 to 0.94, and MAE improved from 32.92 to 23.9. Evaluation using the Taylor diagram further confirmed the superior performance of the VEDL model compared to the individual ANN model.

Conclusion

The excellent performance of the models in the two stages of training and validation is another finding of the research, These results are consistent with the results of researchers who have expressed the appropriate efficiency of machine learning models in estimating evaporation/evaporation and transpiration in different climatic regions of Iran. The results of the combined model showed that the combined model improved the results compared to the best individual model so that the RMSE values increased from 45.95 to 33.1, the R^2 values increased from 0.89 to 0.94, and the MAE value improved from 32.92 to 23.9. The use of the VEDL approach to estimate evaporation from the pan was a new approach that has not been used in past studies. Therefore, according to the results of this research, the proposed deep sensing model is proposed to estimate the evaporation of arid and semi-arid areas for water resources management and agricultural planning.

Keywords: Actual evaporation, Data mining, Simulation, Voting regressors, Sistan and Balouchestan Province

Article Type: Research Article

Acknowledgement

We would like to express our sincere gratitude to the Higher Educational Complex of Saravan for the financial and logistical support that significantly contributed to the research project.

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

Mojtaba Mohammadi: Resources, software, writing, original draft preparation; **Hossein Jahantigh:** Resources, manuscript editing; **Farhad Zolfaghari:** Conceptualization, supervision, final editing.

*Corresponding Author, E-mail: Mohamadi613@gmail.com

Citation: Mohammadi, M., Jahantigh, H., & Zolfahari, F. (2024). Monthly prediction of pan evaporation using individual and combined approaches of data mining models in arid regions. *Water and Soil Management and Modeling*, *4*(2), 227-240. DOI: 10.22098/mmws.2023.12728.1270

Received: 17 April 2023, Received in revised form: 30 April 2023, Accepted: 30 April 2023, Published online: 30 April 2023

Water and Soil Management and Modeling, Year 2024, Vol. 4, No. 2, pp. 227-240

Publisher: University of Mohaghegh Ardabili

© Author(s)

Archive of SID.ir

Archive of SID.ir

مدل‌سازی و مدیریت آب و خاک

شاپا الکترونیکی: ۲۷۸۳-۲۵٤٦

پیش‌بینی ماهانهٔ تبخیر از تشت با استفاده از رویکردهای انفرادی و ترکیبی مدل‌های داده‌کاوی در مناطق خشک

مجتبی محمدی[۱*]، حسین جهانتیغ[۱]، فرهاد ذوالفقاری[۱]

[۱] استادیار، گروه مدیریت و کنترل بیابان، دانشکدهٔ کشاورزی و منابع طبیعی، مجتمع آموزش عالی سراوان، سراوان، ایران

چکیده

برآورد دقیق میزان تبخیر در برنامه‌ریزی صحیح و بهره‌برداری بهینه از پروژه‌های منابع آب و فعالیت‌های کشاورزی بسیار مهم است. تبخیر تحت تأثیر بسیاری از نیروهای محرکه است که با ویژگی غیرخطی، غیرثابت و تصادفی مشخص می‌شوند. چنین عواملی به وضوح مانع از راه‌اندازی مدل‌های پیش‌بینی دقیق می‌شود. از این‌رو، در این مطالعه به پیش‌بینی ماهانهٔ تبخیر از تشت با استفاده از رویکردهای انفرادی و ترکیبی مدل‌های داده‌کاوی در استان سیستان و بلوچستان پرداخته شده است. داده‌های هواشناسی شامل متوسط دمای ماهانه، بیشینه دمای ماهانه، کمینه دمای ماهانه، میانگین سرعت باد، متوسط رطوبت نسبی، بیشینه رطوبت نسبی، کمینه رطوبت نسبی و مجموع بارش ماهانه از ایستگاه‌های سینوپتیک در استان سیستان و بلوچستان به‌عنوان متغیرهای پیش‌بینی‌کننده استفاده شد. در رویکرد عملکرد انفرادی مدل‌های داده‌کاوی، از هشت مدل داده کاوی به‌منظور شبیه‌سازی و پیش‌بینی تبخیر از تشت استفاده شد. از رویکرد ترکیبی VEDL به‌منظور ارائهٔ یک مدل ترکیبی (ترکیب هشت مدل انفرادی داده‌کاوی اشاره شده) با استفاده از روش واپسگرای رای و رای وزنی مورد استفاده قرار گرفت. به‌منظور ارزیابی مدل‌های مورد استفاده در این پژوهش از معیارهای ارزیابی ضریب تعیین (R^2)، ریشهٔ میانگین مربعات خطا (RMSE)، میانگین خطای مطلق و نمودار تیلور استفاده شد. در بین مدل‌های انفرادی مدل شبکهٔ عصبی مصنوعی (ANN) با میزان R^2 برابر ۰/۸۹ و RMSE برابر ۴۵/۹۵ در مرحلهٔ آموزش و میزان R^2 برابر ۰/۹۶ و RMSE برابر ۴۴/۱۸ در مرحلهٔ آزمون، بهترین عملکرد را در هر دو مرحلهٔ آموزش و آزمون داشته است و به‌عنوان بهترین مدل داده‌کاوی در ایستگاه زاهدان به‌منظور پیش‌بینی تبخیر از تشت ماهانه معرفی شد. نتایج اجرای مدل ترکیبی نشان داد که رویکرد ترکیبی مدل‌ها موجب بهبود نتایج نسبت به بهترین مدل انفرادی شده است. به‌طوری‌که مقادیر RMSE از مقدار ۴۵/۹۵ به مقدار ۳۳/۱ و مقادیر R^2 از مقدار ۰/۸۹ به مقدار ۰/۹۴ و مقدار MAE از ۳۲/۹۲ به مقدار ۳۲/۹ بهبود یافته است.

واژه‌های کلیدی: استان سیستان و بلوچستان، تبخیر واقعی، واپسگرای رای، داده‌کاوی، شبیه‌سازی

نوع مقاله: پژوهشی

[*] مسئول مکاتبات، پست الکترونیکی: Mohamadi613@gmail.com

استناد: محمدی، مجتبی، جهانتیغ، حسین، و ذوالفقاری، فرهاد (۱۴۰۳). پیش‌بینی ماهانهٔ تبخیر از تشت با استفاده از رویکردهای انفرادی و ترکیبی مدل‌های داده‌کاوی در مناطق خشک. *مدل‌سازی و مدیریت آب و خاک*، ۴(۲)، ۲۲۷- ۲۴۰.
DOI: 10.22098/mmws.2023.12728.1270

تاریخ دریافت: ۱۴۰۲/۰۱/۲۸، تاریخ بازنگری: ۱۴۰۲/۰۲/۱۰، تاریخ پذیرش: ۱۴۰۲/۰۲/۱۰، تاریخ انتشار: ۱۴۰۲/۰۲/۱۰
مدل‌سازی و مدیریت آب و خاک، سال ۱۴۰۳، دوره ۴، شماره ۲، صفحه ۲۲۷ تا ۲۴۰
ناشر: دانشگاه محقق اردبیلی

© نویسندگان

Archive of SID.ir

Archive of SID.ir

۱- مقدمه

تبخیر به‌عنوان یک فرآیند فیزیکی تعریف می‌شود که در آن مولکول‌های آب در اثر دریافت انرژی کافی بر فشار بخار غلبه و از سطح آب خارج می‌شوند (Malik et al., 2021). از این‌رو تبخیر تلفات را در هر سیستم آبی ایجاد می‌نماید که می‌تواند کمبود آب را در آن سیستم تشدید نماید (Elbeltagi et al., 2023). به‌ویژه در مناطق خشک تا نیمه‌خشک که در آن مولکول‌های آب انرژی گرمایی کافی برای رهایی از سطح را دارند. بنابراین، برآورد دقیق تلفات تبخیر نقش بسیار مهمی در مدیریت بهتر منابع آب، نیاز آبی محصول و برنامه‌ریزی آبیاری دارد (Kushwaha et al., 2021; Mohammadi et al., 2022). در مناطق خشک و نیمه‌خشک، میزان تبخیر بالا معمولاً در فصل تابستان تشدید و موجب کاهش حجم مخازن، رودخانه‌ها، دریاچه‌های طبیعی و بحران کمبود آب می‌شود (Khan et al., 2021). به این ترتیب، تعیین مقدار دقیق تبخیر سطوح آبی برای برنامه‌ریزی و مدیریت صحیح هر پروژه منابع آبی بسیار مهم است (Vishwakarma et al., 2022; Abd-Elaty et al., 2022; Mohammadi et al. 2022). دو روش برای تخمین و برآورد تبخیر وجود دارد. روش‌های مستقیم همانند روش تبخیر تشت[1] (PE) و روش‌های غیرمستقیم مانند انتقال جرم، تعادل آب و انرژی (Lundberg, 1993) و روش‌های پنمن (Zhao et al., 2013) روش تبخیر از تشت کلاس A در سطح جهانی برای تخمین تبخیر استفاده می‌شود، زیرا به خوبی برای تخمین نسبی سطوح تبخیر در مناطق مختلف ویژگی‌های اقلیمی سازگار است (Masoner et al., 2008). با این حال، ماهیت پرهزینه بودن روش کلاس A مانع اساسی کاربرد آن در بسیاری از کشورهای در حال توسعه شده است (Wu et al., 2020).

از سوی دیگر بسیاری از روش‌های غیرمستقیم برآورد تبخیر مانند روش تورنت-وایت، پریستلی-تیلور و پنمن-مانتیث برای برآورد دقیق و قابل اعتماد تبخیر از متغیرهای هواشناسی مانند رطوبت نسبی، ساعات آفتاب، سرعت باد، بارندگی، حداقل، حداکثر و میانگین دما استفاده می‌شود (Elbeltagi et al., 2023). با این حال، به‌دلیل ماهیت پویای متغیرهای هواشناسی که برآمده از ویژگی‌های غیرخطی، غیرثابت و تصادفی آن‌هاست، مستلزم توسعهٔ مدل‌های هوشمند دقیق و قابل اعتماد است که اثرات این ویژگی‌های متغیرهای تبخیر-هواشناسی را حذف یا کم اثر کنند (Kisi et al., 2017; Khan et al., 2019; Naganna et al., 2019; Salih et al., 2020; Kushwaha et al., 2022; Elbeltagi et al., 2023). در سال‌های اخیر روش‌های داده‌کاوی[2] (DM) به‌دلیل قابلیت‌های بالا به‌عنوان یک روش مهم برای پرداختن به جنبه‌های مختلف ارتباط بین پیش‌بینی‌کننده‌ها و پیش‌بینی‌شونده‌ها در بررسی بسیاری از موضوعات مختلف هیدرولوژی و اقلیم‌شناسی مورد استفاده قرار گرفته‌اند. برای نمونه، باران (Adnan et al., 2021; Mohammadi et al., 2022)، جریان آب سطحی (Parisouj et al., 2020; Feng and Tian, 2021)، خشکسالی (Rezaie-Balf et al., 2020; Chen et al., 2020)، کیفیت آب سطحی (Rahman et al., 2020; Mosavi et al., 2021)، آب زیرزمینی (Granata et al., 2020; Granata and Di Nunno, 2021)، تبخیر-تعرق.

در ایران نیز مطالعاتی در زمینهٔ استفاده از روش‌های داده‌کاوی و یادگیری ماشین به‌منظور برآورد و تخمین تبخیر و تبخیر-تعرق صورت گرفته است. در این راستا، Shahabi et al. (2020) به بررسی قابلیت مدل درخت M5 و جنگل تصادفی (RF) در ترکیب با فرآیند پیش پردازشگر موجک به‌منظور برآورد تبخیر-تعرق پتانسیل روزانه و هفتگی در ایستگاه سینوپتیک بابلسر پرداختند. نتایج کارایی رضایت بخش هر دو رویکرد منفرد و ترکیبی در برآورد تبخیر-تعرق دوره‌های زمانی روزانه و هفتگی را نشان داد. در پژوهش دیگری، Sabzevari and Ghanbarpour (2022) مدل‌های تجربی و هوشمند در برآورد تبخیر-تعرق را در ایستگاه الیگودرز با در نظر گرفتن هفت مدل برآورد تبخیر-تعرق مرجع و مدل برنامه‌ریزی بیان ژن (GEP) نسبت به فائو پنمن-مانتیث مورد ارزیابی قرار دادند. نتایج نشان داد که مدل GEP نسبت به سایر مدل‌ها برتری دارد و برای منطقهٔ مورد مطالعه توانایی قابل‌قبولی در برآورد تبخیر-تعرق مرجع را دارد. هم‌چنین، Seyedi et al. (2022) به بررسی عملکرد مدل‌های مختلف شبکهٔ عصبی مصنوعی در برآورد تلفات تبخیر از تشت در محدودهٔ دریاچه سد شهید رجایی شهرستان ساری پرداختند. برای این منظور از ترکیب متنوعی از هشت متغیر هواشناسی به‌عنوان ورودی مدل شبکهٔ عصبی مصنوعی استفاده کردند. نتایج حاصل از شاخص‌های آماری مدل‌ها و دیاگرام پراکنش و میزان تبخیر روزانه برآورد شده و مشاهداتی در مطالعهٔ آن‌ها نشان داد که در مجموع روش شبکهٔ عصبی مصنوعی توانسته تبخیر را با دقت خوبی برآورد نماید.

بررسی مطالعات ذکر شده نشان داد که عملکرد مدل داده‌کاوی در مقایسه با سایر روش‌ها برتر بوده و بسته به عوامل

[1] Pan evaporation

[2] Data mining

Archive of SID.ir

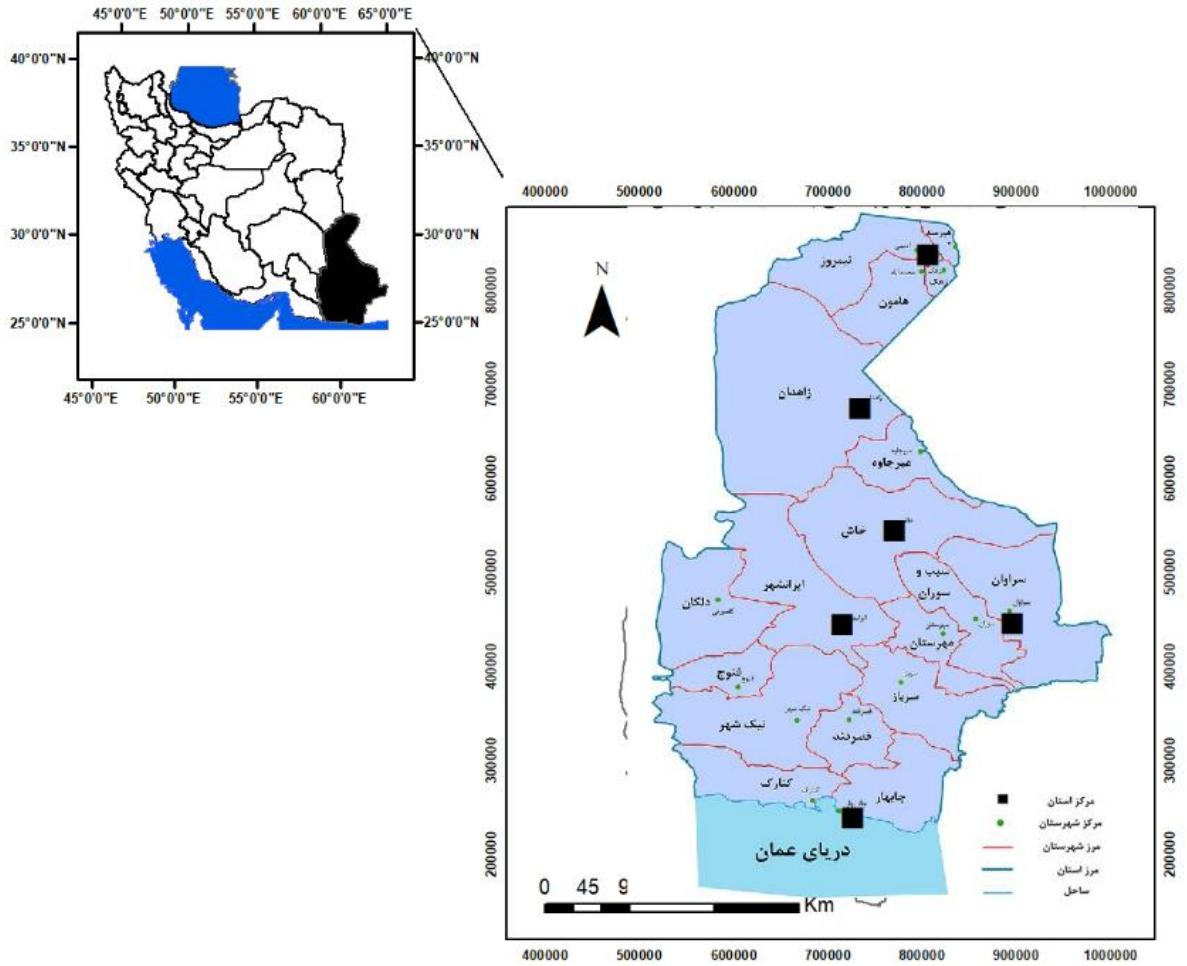
Archive of SID.ir

ورودی در شرایط مختلف اقلیمی، عملکرد آن‌ها متفاوت است. از سوی دیگر بررسی مطالعات گذشته نشان می‌دهد که بیش‌تر پژوهش‌های مدل‌های داده‌کاوی بر مبنای عملکرد انفرادی مدل‌ها استوار بوده‌اند. این در حالی است که تروش‌های ترکیبی که در آن دو یا چند مدل ترکیب و جفت می‌شوند، به‌دلیل ظرفیت آن برای گرفتن الگوهای مختلف در سری داده‌ها با ترکیب ویژگی‌های چند تکنیکی در یک الگوریتم، اخیراً توجه بیش‌تری را در مطالعات اقلیمی و هیدرولوژی به خود جلب کرده است. افزون بر آن با توجه به این واقعیت که هر منطقهٔ اقلیمی با ویژگی‌های خاصی از تصادفی و غیرایستایی همراه است تعمیم قابلیت این مدل‌ها در مناطق مختلف اقلیمی قابل بحث است. بنابراین، بررسی مدل‌های جدید توسعه یافته و بررسی کاربرد آن‌ها برای ویژگی‌های اقلیمی خاص ضروری است. از این‌رو، با توجه به این‌که پهنهٔ سیستان و بلوچستان در شرق ایران دارای اقلیم خشک و فراخشک است و در این پهنه کمیت تبخیر مهم‌ترین و ناشناخته‌ترین مؤلفهٔ بیلان آبی است و بیش‌تر مطالعات صورت گرفته در این پهنه از کشور بر برآورد تبخیر–تعرق پتانسیل/مرجع تمرکز داشته‌اند، و مطالعات کمی به استفاده از رویکردهای سنتی و تجربی تمرکز داشته‌اند، بنابراین، نبود یک پژوهش کامل به‌منظور برآورد تبخیر با استفاده از روش‌های جدید در این پهنه از کشور، یک ضعف اساسی است. از این‌رو، در پژوهش حاضر به بررسی پتانسیل چند گروه از مدل‌های یادگیری ماشین به‌صورت انفرادی و یک رویکرد ترکیبی جدید برای برآورد تبخیر از تشت در پهنهٔ استان پرداخته خواهد شد.

۲- مواد و روش‌ها

۲-۱- منطقهٔ مورد مطالعه

در این پژوهش از داده‌های ایستگاه‌های سینوپتیک استان سیستان و بلوچستان شامل زاهدان، خاش، سراوان، ایرانشهر و چابهار در طول دورهٔ آماری ۲۰۱۹–۱۹۸۰ به‌عنوان ورودی مدل‌ها و از داده‌های اندازه‌گیری تبخیر از تشت این ایستگاه‌ها به‌عنوان مقادیر مشاهداتی تبخیر استفاده شد. شکل ۱ موقعیت ایستگاه سینوپتیک در استان سیستان و بلوچستان و ایران و جدول ۱ مشخصات متغیرهای مورد استفاده در ایستگاه سینوپتیک زاهدان را نشان می‌دهد.



شکل ۱- موقعیت ایستگاه‌های سینوپتیک در استان سیستان و بلوچستان و ایران

Figure 1- Location of Synoptic Stations in Sistan and Baluchestan Province and Iran

Archive of SID.ir

Archive of SID.ir

جدول ۱- مشخصات آماری متغیرهای هواشناسی در ایستگاه سینوپتیک زاهدان

Table 1- Statistical characteristics of meteorological parameters in Zahedan synoptic station

دادهٔ هواشناسی	کمینه	بیشینه	متوسط	انحراف معیار	ضریب تغییرات
متوسط دمای ماهانه (درجه سانتی گراد)	4.3	32.2	20.6	7.6	0.37
بیشینه دمای ماهانه (درجه سانتی گراد)	11	39.3	28.5	7.4	0.26
کمینه دمای ماهانه (درجه سانتی گراد)	-4	22.6	12	6.8	0.57
میانگین سرعت باد (متر بر ثانیه)	1.2	6.17	3.2	0.7	0.23
متوسط رطوبت نسبی (درصد)	9.6	70.3	26.1	12.2	0.47
بیشینه رطوبت نسبی (درصد)	16	90.3	40.7	17	0.42
کمینه رطوبت نسبی (درصد)	4.7	48.7	14.4	7.5	0.52
مجموع بارش ماهانه (میلی متر)	0	85.4	5	10.5	2.08
مجموع ماهانه تبخیر از تشت (میلی متر)	2.20	526.90	275.97	134.08	0.49

۲-۲- رویکردهای عملکرد انفرادی و ترکیبی مدل‌های داده‌کاوی

داده‌کاوی نسل جدیدی از مدل‌های یادگیری است که در سال ۲۰۰۶ ارائه شد (Hinton and Salakhutdinov, 2006). این نوع مدل‌ها با استفاده از سیستم‌های پیچیدهٔ داده محور، نقش کلیدی در راه‌حل‌های کلان داده[1] ایفا می‌کند (Chen and Lin, 2014). این مدل‌ها از یک پیش‌آموزش[2] بدون نظارت و یک راهبرد تنظیم دقیق نظارت‌شده برای یادگیری ویژگی‌های سلسله مراتبی استفاده می‌کنند (Sze et al., 2017). برای اجرای مدل‌های داده‌کاوی دو رویکرد وجود دارد. نخست مبنای عملکرد انفرادی مدل‌ها و دوم رویکرد جدید تکنیک‌های ترکیبی که در آن دو یا چند مدل ترکیب و جفت می‌شوند. در پژوهش حاضر در رویکرد عملکرد انفرادی، با توجه به مطالعات Mohammadi et al. (2021)، Gholami et al. (2022) و Mohammadi et al. (2022) از هشت مدل داده‌کاوی برای شبیه‌سازی و پیش‌بینی تبخیر از تشت استفاده شد. به‌منظور جلوگیری از فزونی مطالب گزیده‌ای از توصیفات مدل‌ها در جدول ۲ آورده شده است. برای مطالعهٔ بیش‌تر می‌توان به پژوهش Mohammadi et al. (2021) مراجعه کرد. شایان ذکر است که تمامی مراحل اجرای مدل‌ها در محیط برنامه‌نویسی R صورت گرفته است.

بر اساس این فرض که مجموعه‌ای از شبکه‌های عصبی به نتایج بهتری نسبت به یک شبکهٔ منفرد منجر می‌شود دو رویکرد ترکیبی مدل داده‌کاوی تحت عنوان SEDL[3] و VEDL[4] را با تکنیک‌های انباشتگی[5] و رای‌گیری[6] پیشنهاد نمودند (Mohammadifar et al., 2023). در پژوهش حاضر افزون بر رویکرد عملکرد انفرادی، از رویکرد ترکیبی VEDL به‌منظور ارائهٔ یک مدل ترکیبی (ترکیب هشت مدل انفرادی داده‌کاوی اشاره شده) استفاده شد. این رویکرد ترکیبی از روش‌های زیادی به‌جای یک مدل واحد برای افزایش عملکرد سیستم استفاده می‌کند. این رویکرد را می‌توان برای موضوعات طبقه‌بندی و رگرسیون با ترکیب نتایج روش‌های متعدد اعمال کرد. برای موضوعات مربوط به رگرسیون، برآوردگرهای همهٔ مدل‌ها برای به‌دست آوردن یک تخمین از مجموعه‌هایی که به‌عنوان واپسگرای رأی[7] (VRs) نام دارند، میانگین‌گیری می‌شوند. دو رویکرد رای متوسط[8] (AV) و رای وزنی[9] (WV) برای اخذ آرا وجود دارد. در رویکرد رای متوسط (AV)، وزن‌ها معادل و برابر یک هستند. یکی از معایب این روش این است که همهٔ مدل‌های موجود در مجموعهٔ ترکیبی، ضریب تأثیر یکسانی دریافت می‌کنند. در رویکرد رای وزنی (WV) هر یک از مدل‌های موجود در مجموعهٔ ترکیبی یک ضریب وزنی را شامل می‌شوند. این ضریب وزن می‌تواند یک عدد اعشاری بین صفر و یک باشد، که در این صورت مجموع ضرایب کل مدل‌های شرکت‌کننده برابر با یک است. در این پژوهش وزن هر مدل بر اساس دقت عملکرد مدل با استفاده از معیارهای ارزیابی به‌دست آمده از بخش اجرای آموزش (train) مدل‌های انفرادی انتخاب شد. برای این منظور وزن هر مدل انفرادی بر اساس رابطهٔ (۱) محاسبه شد (Mohamadifar et al., 2023). در این رابطه، Wm و Ac به‌ترتیب وزن و دقت مدل انفرادی و An جمع دقت تمام مدل‌های انفرادی استفاده شده (مدل‌های ۱ تا ۸) است.

$$W_m = \frac{A_c}{\sum A_n} \qquad (۱)$$

[1] Big data
[2] Pre-training
[3] Stacking-based ensemble deep learning
[4] Voting-based ensemble deep learning
[5] Stacking
[6] Voting
[7] Voting regressors
[8] Average voting
[9] Weighted voting

Archive of SID.ir

Archive of SID.ir

جدول ۲- گزیده‌ای از مدل‌های یادگیری ماشین مورد استفاده

Table 2- A selection of used machine learning models

نام فارسی	نام لاتین	مخفف	پکیج مورد استفاده	منبع
الگوریتم جنگل تصادفی	Random forest	RF	party	Liaw and Wiener (2002)
ماشین بردار پشتیبان	Support Vector Machine	SVM	LiblineaR	Helleputte and Gramme (2017)
شبکهٔ عصبی مصنوعی	Artificial Neural Networks	ANN	brnn	Zema et al. (2020)
ارتقای گرادیان شدید	Extrememe Gradient Bossting	XGBoost	xgboost	Nobre and Neves (2019)
مدل درختی	Tree Model	TM	RWeka	Zahiri and Nezaratian (2020)
مدل اسپلاین چندمتغیرهٔ رگرسیون انطباقی	Multivariate Adaptive Regression Splines	MARS	earth	Zahiri and Nezaratian (2020)
مدل فرایند گوسی	Gaussian Process	GP	Kernlab	Karatzoglou et al. (2019)
مدل بیزی جمعی رگرسیون درختی	Baysian Additive Regression Tree	BART	monomvn	Gramacy (2019)

۲-۳- معیارهای ارزیابی

به‌منظور ارزیابی مدل‌های مورد استفاده در این پژوهش از معیارهای ارزیابی ضریب تعیین (R^2)، ریشهٔ میانگین مربعات خطا[1] (RMSE) و میانگین خطای مطلق[2] (MAE) استفاده شد. شد. در این مطالعه، علاوه‌بر روش‌های ارزیابی نام برده از نمودار تیلور (Taylor, 2001) نیز به‌منظور ارزیابی کارایی مدل‌ها استفاده شد. این نمودار بر مبنای ارتباط هندسی بین ضریب همبستگی، انحراف معیار سری‌های زمانی و RMSE استوار است. روابط و اطلاعات بیش‌تر در مطالعهٔ Taylor (2001) درج شده است.

۳- نتایج و بحث

۳-۱- یافته‌های شبیه‌سازی انفرادی مدل‌ها

همان طورکه بیان شد در این پژوهش از هشت مدل یادگیری ماشین برای شبیه‌سازی و پیش‌بینی مقادیر تبخیر از تشت در ایستگاه سینوپتیک زاهدان استفاده شد. برای این هدف از ۷۰ درصد داده‌ها برای مرحلهٔ آموزش[3] و ۳۰ درصد داده‌ها برای مرحلهٔ آزمون[4] بهره برده شد. در جدول ۳ نتایج ارزیابی کارایی مدل‌های انفرادی در دو مرحلهٔ آموزش و آزمون با استفاده از معیارهای ارزیابی نشان داده شده است. با توجه به جدول ۳ تمامی مدل‌ها در هر دو مرحلهٔ آموزش و آزمون نتایج بسیار خوبی داشته‌اند. در بین مدل‌های انفرادی مدل شبکهٔ عصبی مصنوعی (ANN) با میزان R^2 برابر ۰/۸۹ و RMSE برابر ۴۵/۹۵ در مرحلهٔ آموزش و میزان R^2 برابر ۰/۹۶ و RMSE برابر ۴۴/۱۸ در مرحلهٔ آزمون، بهترین عملکرد را در هر دو مرحلهٔ آموزش و آزمون در ایستگاه سینوپتیک زاهدان داشته است و به‌عنوان بهترین مدل داده‌کاوی در این ایستگاه به‌منظور پیش‌بینی تبخیر از تشت ماهانه معرفی می‌شود. بقیهٔ مدل‌ها نیز عملکرد بسیار مناسبی در هر دو مرحلهٔ آموزش و آزمون داشته‌اند. به طوری‌که در مرحلهٔ آموزش مدل SVM با مقدار R^2 برابر ۰/۸۹، مدل RF با مقدار R^2 برابر ۰/۸۸ در رتبه‌های بعدی قرار دارند و در مرحلهٔ آزمون مدل BART با مقدار R^2 برابر ۰/۹۶ در رتبهٔ دوم قرار دارد. در بین مدل‌ها کم‌ترین دقت مربوط به مدل درختی (TM) با مقدار R^2 برابر ۰/۸۴ و ۰/۹۳ به‌ترتیب برای مرحلهٔ آموزش و آزمون است. هر چند که این مدل نیز با این مقدار معیارهای ارزیابی نیز گروه عملکرد بسیار خوب را شامل می‌شود. خروجی نمودار تیلور نیز این موضوع را تأیید می‌کند. به‌طوری‌که بر اساس این نمودار در هر دو مرحلهٔ آموزش و آزمون، کارایی دقت مدل‌ها نزدیک به هم هستند (شکل‌های ۲ و ۳).

[1] Root mean square error
[2] Mean absolute error
[3] Train
[4] Test

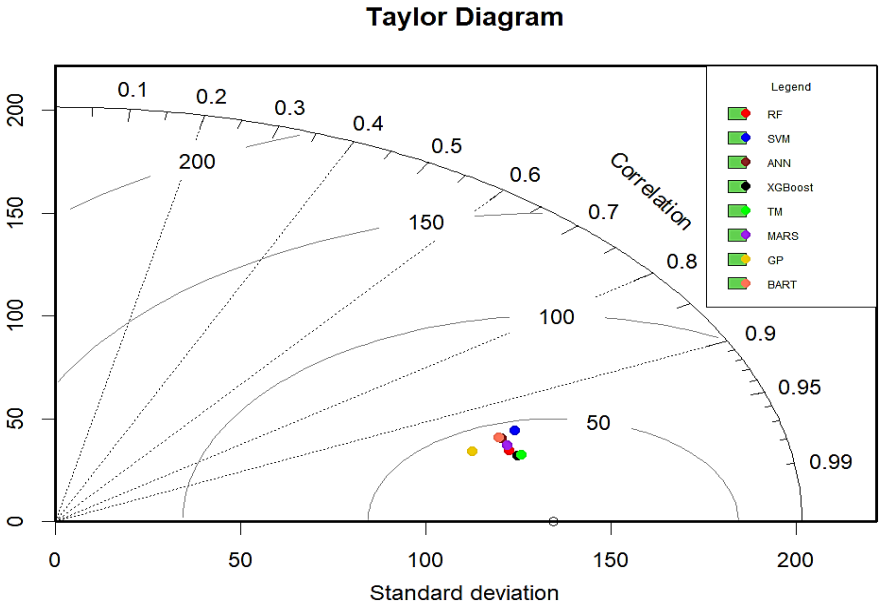
Archive of SID.ir

Archive of SID.ir

جدول ۳- نتایج ارزیابی کارایی مدل‌های انفرادی در دو مرحلهٔ آموزش و آزمون

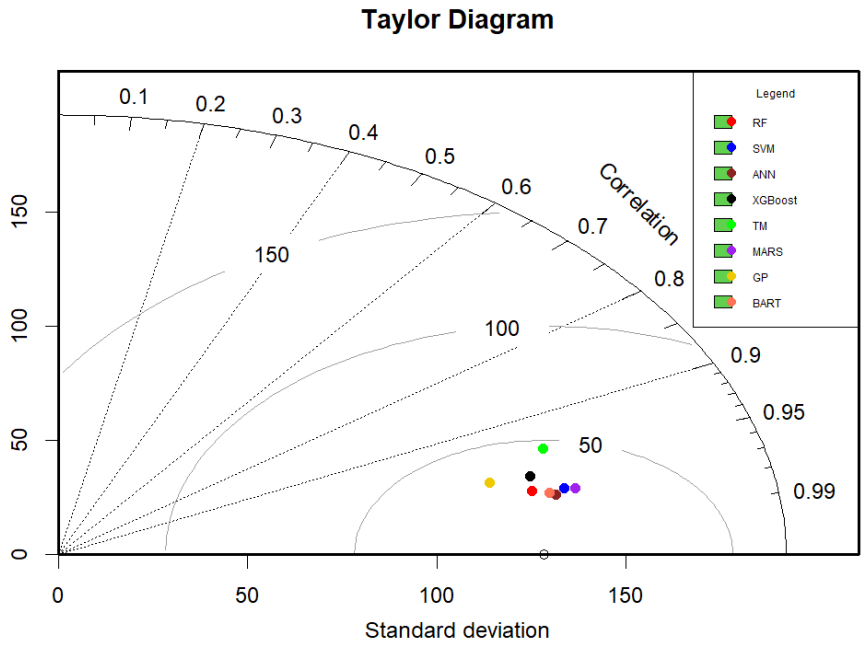
Table 3- The results of evaluating the efficiency of individual models in two stages of training and testing

آزمون			آموزش			مدل انفرادی
MAE	R^2	RMSE	MAE	R^2	RMSE	
42.23	0.95	49.36	34.82	0.88	47.60	RF
41.78	0.95	48.31	33.13	0.89	47.08	SVM
38.26	0.96	44.18	32.92	0.89	45.95	ANN
38.89	0.94	46.85	39.19	0.84	54.44	XGBoost
40.47	0.93	48.57	39.58	0.84	54.89	TM
43.06	0.95	48.12	39.20	0.84	54.62	MARS
42.05	0.95	48.72	35.09	0.87	49.39	GP
40.80	0.96	46.94	32.82	0.88	46.22	BART



شکل ۲- نمودار تیلور در مرحلهٔ آموزش در ایستگاه سینوپتیک زاهدان

Figure 2- Taylor's diagram in the training phase at Zahedan synoptic station



شکل ۳- نمودار تیلور در مرحلهٔ آزمون در ایستگاه سینوپتیک زاهدان

Figure 3- Taylor's diagram in the test phase at Zahedan synoptic station

Archive of SID.ir

Archive of SID.ir

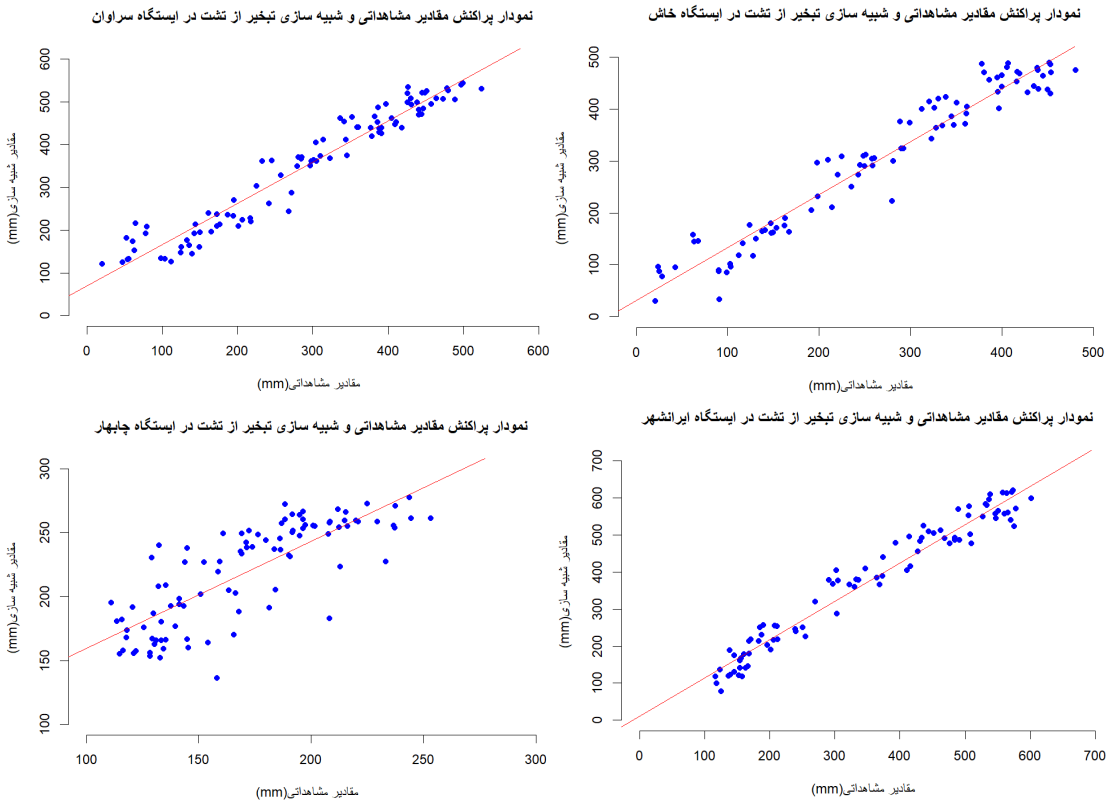
همان طورکه اشاره شد در این پژوهش از ایستگاه‌های سینوپتیک استان استفاده شد که به‌منظور جلوگیری از فزونی مطالب در ادامه فقط نتایج بهترین مدل انفرادی در ایستگاه‌های منتخب در جدول ۴ ارائه شده است. با توجه به جدول ۴، در بین تمامی مدل‌ها، مدل‌های ANN، SVM و RF بهترین عملکرد را در دو مرحلهٔ آموزش و آزمون داشته‌اند. در مرحلهٔ آزمون، مدل SVM با مقادیر R^2 برابر ۰/۹۴، ۰/۹۶ و ۰/۹۴ بهترین عملکرد در ایستگاه‌های خاش، ایرانشهر و چابهار داشته است. در ایستگاه سراوان در مرحلهٔ آزمون مدل RF با مقدار R^2 برابر ۰/۹۴ بهترین عملکرد را در بین مدل‌ها داشته است. عملکرد بسیار عالی مدل‌ها در دو مرحلهٔ آموزش و آزمون نیز از دیگر یافته‌های پژوهش است به‌طوری‌که در ایستگاه‌های خاش و ایرانشهر مدل SVM در مرحلهٔ آموزش نیز عملکرد بسیار مناسبی داشته است.

در حالی‌که در ایستگاه خاش مقادیر R^2 و RMSE مدل SVM به‌ترتیب برابر ۰/۹ و ۴۵/۶۸ و برای ایستگاه چابهار مقادیر R^2 و RMSE مدل SVM به‌ترتیب برابر ۰/۹۳ و ۴۰/۷۷ است. شکل ۴ نمودار پراکنش مقادیر مشاهداتی و شبیه‌سازی شدهٔ مرحلهٔ آزمون ایستگاه‌های مورد مطالعه با استفاده از مدل انفرادی منتخب را نشان می‌دهد. نمودارها بیان‌گر عملکرد بسیار عالی مدل در ایستگاه‌هاست که این نتایج با نتایج مطالعهٔ Shahabi et al. (2021)، Sabzevari and Ghanbarpour (2022)، Mohammadi et al. (2022) و Elbeltagi et al. (2023) هم‌خوانی دارد. این پژوهش‌گران نیز کارایی مناسب مدل‌های یادگیری ماشین در برآورد تبخیر و تبخیر–تعرق در مناطق مختلف اقلیمی ایران را بیان داشته‌اند.

جدول ٤- نتایج آماری به‌دست آمده از عملکرد بهترین مدل انفرادی در ایستگاه‌های مورد مطالعه

Table 4-Statistical results obtained from the performance of the best individual model in the studied stations

ایستگاه	بهترین مدل	آموزش			آزمون		
		RMSE	R^2	MAE	RMSE	R^2	MAE
زاهدان	ANN	45.95	0.89	32.92	44.18	0.96	38.26
خاش	SVM	45.68	0.9	33.77	46.10	0.94	37.18
سراوان	RF	55.31	0.88	44.77	67.72	0.95	59.91
ایرانشهر	SVM	52.11	0.93	41.38	40.93	0.96	32.65
چابهار	SVM	40.77	0.86	30.75	48.03	0.94	42.95



شکل ٤- نمودار پراکنش مقادیر مشاهداتی و شبیه‌سازی شدهٔ مرحلهٔ آزمون با مدل انفرادی منتخب در ایستگاه‌های مورد مطالعه

Figure 4- The distribution diagram of the observed and simulated values of the test stage with the selected individual model in the studied stations

Archive of SID.ir

Archive of SID.ir

۳-۲- یافته‌های شبیه‌سازی با رویکرد ترکیبی مدل‌ها

در ادامه همان طورکه اشاره شد به‌منظور ارائهٔ یک مدل ترکیبی داده‌کاوی از روش VEDL با استفاده از رای‌گیری وزنی بر اساس رابطهٔ (۱) برای مرحلهٔ آموزش استفاده شد. در جدول ۵ مقادیر Wm و Ac که به‌ترتیب وزن و دقت مدل انفرادی در مرحلهٔ آموزش ایستگاه سینوپتیک زاهدان هستند ارائه شده است. در ادامه با رویکرد اشاره شده مدل جدید ترکیبی با استفاده از روش VEDL ارائه و مدل در مرحلهٔ آموزش مجدد اجرا و نتایج بررسی شد. جدول ۶ نتایج آماری بهترین مدل انفرادی و مدل VEDL در ایستگاه زاهدان را نشان می‌دهد. با توجه به جدول، مدل ترکیبی موجب بهبود نتایج نسبت به بهترین مدل انفرادی شده است به‌طوری‌که مقادیر RMSE از مقدار ۴۵/۹۵ به مقدار ۳۳/۱ و مقادیر R^2 از مقدار ۰/۸۹ به مقدار ۰/۹۴ و مقدار MAE از ۳۲/۹۲ به مقدار ۲۳/۹ بهبود یافته است. ارزیابی مدل انفرادی ANN و VEDL با استفاده از نمودار تیلور نیز بیان‌گر عملکرد بهتر مدل VEDL نسبت به مدل انفرادی است (شکل ۵). در شکل ۶ نیز پراکنش مقادیر مشاهداتی و شبیه‌سازی با مدل VEDL که بیان‌گر شبیه‌سازی بسیار مناسب مدل ترکیبی است، ارائه شده است.

جدول ۵- مقادیر دقت مدل انفرادی و وزن هر مدل بر اساس رای‌گیری وزنی در ایستگاه زاهدان

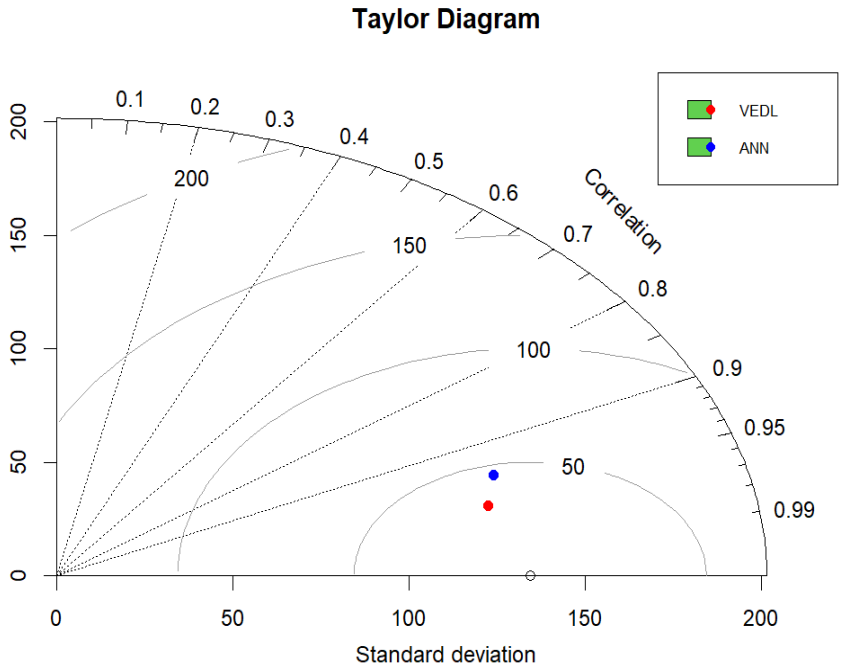
Table 5-Individual model accuracy values and weight of each model based on weighted voting at Zahedan station

BART	GP	MARS	TM	XGBoost	ANN	SVM	RF	
0.88	0.87	0.84	0.84	0.84	0.89	0.89	0.88	Ac
0.13	0.13	0.12	0.12	0.12	0.13	0.13	0.13	Wm

جدول ٦- نتایج آماری بهترین مدل انفرادی و مدل VEDL در ایستگاه زاهدان

Table 6- Statistical results of the best individual model and VEDL model in Zahedan station

روش	مدل	معیار ارزیابی		
		RMSE	R^2	MAE
بهترین مدل انفرادی	ANN	45.95	0.89	32.92
مدل ترکیبی	VEDL	33.1	0.94	23.9

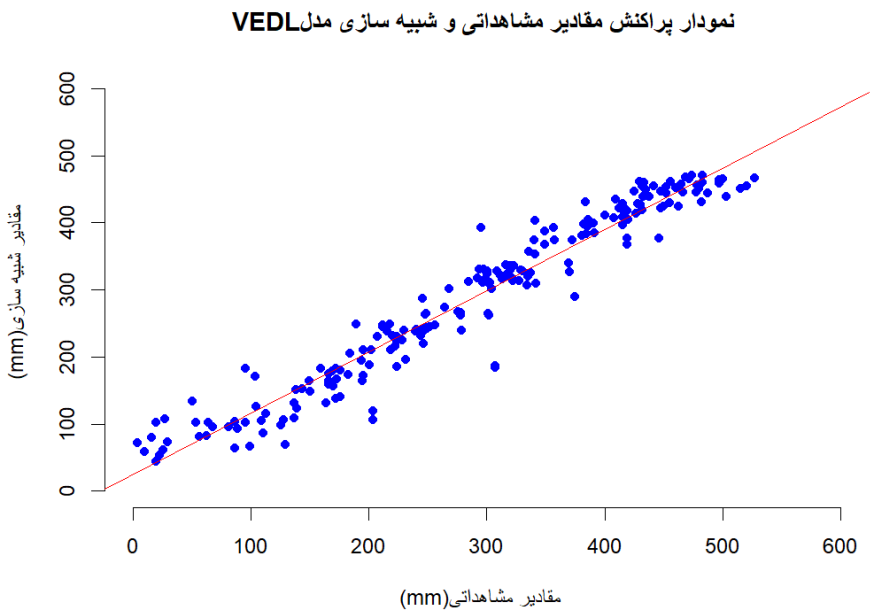


شکل ۵- نمودار تیلور مدل انفرادی و مدل VEDL در مرحلهٔ آموزش

Figure 5- Taylor diagram of the individual model and the VEDL model in the training phase

Archive of SID.ir

Archive of SID.ir



شکل ٦- نمودار پراکنش مقادیر مشاهداتی و شبیه سازی با مدل VEDL در مرحلهٔ آموزش

Figure 6- Distribution chart of observed values and simulation with VEDL model in the training phase

٤- نتیجه‌گیری

تبخیر یکی از مهم‌ترین مؤلفه‌های اقلیمی هر منطقه است. در مناطق خشک و نیمه‌خشک، میزان تبخیر بالا به‌ویژه در فصول گرم سال موجب کاهش حجم مخازن، رودخانه‌ها و دریاچه‌های طبیعی می‌شود و بحران کمبود آب در این مناطق را تشدید می‌نماید. بنابراین، برآورد دقیق تلفات تبخیر نقش بسیار مهمی در مدیریت بهتر منابع آب، کشاورزی و حفاظت آب و خاک دارد. از سوی دیگر در سال‌های اخیر کارایی تکنیک‌های هوش مصنوعی در برآورد و پیش‌بینی پدیده‌های غیرخطی به اثبات رسیده است. از این‌رو، در این پژوهش به بررسی کارایی هشت مدل داده‌کاوی به‌منظور برآورد ماهانهٔ تبخیر از تشت در ایستگاه‌های سینوپتیک استان سیستان و بلوچستان پرداخته شد. سپس از روش VEDL به‌منظور ارائهٔ یک مدل ترکیبی داده‌کاوی پرداخته شد. نتایج نشان داد که در بین تمامی مدل‌ها، مدل‌های ANN، SVM و RF بهترین عملکرد را در دو مرحلهٔ آموزش و آزمون داشته‌اند. در مرحلهٔ آزمون، مدل SVM با مقادیر R^2 برابر ۰/۹۴، ۰/۹۶ و ۰/۹۴ بهترین عملکرد در ایستگاه‌های خاش، ایرانشهر و چابهار داشته است. در ایستگاه سراوان در مرحلهٔ آزمون مدل RF با مقدار R^2 برابر ۰/۹۴ بهترین عملکرد را در بین مدل‌ها داشته است. عملکرد بسیار عالی مدل‌ها در دو مرحلهٔ آموزش و آزمون نیز از دیگر یافته‌های پژوهش است به‌جز ایستگاه چابهار که پراکنش مقداری با ایستگاه‌های دیگر استان تفاوت دارد. از این منظر احتمال وجود تفاوت‌های ساختاری یا جزئی جوی در مناطق ساحلی استان نسبت به مناطق داخلی وجود دارد که می‌توان با طراحی پژوهش‌های گسترده‌تر به آن پرداخت. نتایج نشان داد مدل ترکیبی موجب بهبود نتایج نسبت به بهترین مدل انفرادی شده است به‌طوری‌که مقادیر RMSE از مقدار ۴۵/۹۵ به مقدار ۳۳/۱ و مقادیر R^2 از مقدار ۰/۸۹ به مقدار ۰/۹۴ و مقدار MAE از ۳۲/۹۲ به مقدار ۲۳/۹ بهبود یافته است. استفاده از رویکرد VEDL در برآورد تبخیر از تشت رویکرد جدیدی بود که تاکنون در مطالعات گذشته از آن اسفاده نشده است هرچند که Mohamadifar et al. (2023) در شبیه‌سازی مناطق مستعد فرونشست و Erdebilli and Devrim-İçtenbaş (2022) در پیش‌بینی پسماند پزشکی از این تکنیک اسفاده و کارایی آن را بیان نموده‌اند. از این‌رو، با توجه به نتایج این پژوهش مدل داده‌کاوی پیشنهادی به‌منظور برآورد تبخیر مناطق خشک و نیمه‌خشک برای مدیریت منابع آبی و برنامه‌ریزی‌های کشاورزی پیشنهاد می‌شود. از سویی در این پژوهش به مقایسهٔ عملکرد مدل‌های مختلف در پیش‌بینی تبخیر از تشت پرداخته شد. بنابراین پیشنهاد می‌شود تا تأثیر متغیرها بر دقت پیش‌بینی در پژوهش‌های آینده بررسی شود.

سپاسگزاری

این مقاله از طرح پژوهشی پیش‌بینی تبخیر واقعی با استفاده از مدل‌های محاسبات نرم در ایستگاه‌های سینوپتیک استان سیستان و بلوچستان مصوب مجتمع آموزش عالی سراوان استخراج شده است بدین‌وسیله نویسندگان برخود لازم می‌دانند مراتب سپاسگزاری صمیمانهٔ خود را از مسئولین پژوهشی مؤسسه به عمل آورند.

Archive of SID.ir

Archive of SID.ir

تضاد منافع نویسندگان

نویسندگان این مقاله اعلام می‌نمایند که هیچ‌گونه تضاد منافعی در مورد تألیف یا انتشار این مقاله وجود ندارد.

دسترسی به داده‌ها

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

مشارکت نویسندگان

مجتبی محمدی: مفهوم‌سازی، انجام تحلیل‌های نرم‌افزاری/آماری، نگارش نسخهٔ اولیهٔ مقاله؛ **حسین جهانتیغ:** راهنمایی، ویرایش و بازبینی مقاله، کنترل نتایج؛ **فرهاد ذوالفقاری:** راهنمایی، ویرایش و بازبینی مقاله، کنترل نتایج.

منابع

سبزواری، یاسر، و قنبرپوری، مرادعلی (۱۴۰۱). ارزیابی مدل‌های تجربی و هوشمند در برآورد تبخیروتعرق مرجع: مطالعه موردی الیگودرز. *مهندسی اکوسیستم بیابان*، ۱۱(۳۶)، ۱۷-۳۰. dor:20.1001.1.25386336.1401.11.36.2.4

شهابی، سجاد، آذرپیرا، فریبا، و برزکار، علی (۱۳۹۹). برآورد تبخیر-تعرق روزانه و هفتگی با استفاده از رویکردهای ترکیبی محاسبات نرم. *نشریه آبیاری و زهکشی ایران*، ۱۴(۵)، ۱۵۵۰-۱۵۶۱. dor:20.1001.1.20087942.1399.14.5.5.6

محمدی، مجتبی، فروزان فرد، معصومه، و غلامی، حمید (۱۴۰۱). پیش‌بینی تبخیر از تشت با استفاده از مدل‌های محاسبات نرم در اقلیم فراخشک (مطالعهٔ موردی: دشت سیستان، سیستان و بلوچستان). *مهندسی اکوسیستم بیابان*، ۱۱(۳۶)، ۷۱-۸۲. dor:20.1001.1.25386336.1401.11.36.6.8

محمدی، مجتبی، وقارفرد، حسن، مهدوی نجف آبادی، رسول، دانشکار آراسته، پیمان، و ناظم السادات، سید محمد جعفر (۱۴۰۰). مدل‌سازی بارش-رواناب آبخیزهای مناطق ساحلی در نزدیکی تنگه هرمز با استفاده از روش‌های داده‌کاوی. *تحقیقات آب و خاک ایران*، ۵۲(۲). ۳۱۳-۳۲۷. doi:10.22059/ijswr.2021.309641.668732

References

Abd-Elaty, I., Kushwaha, N.L., Grismer, M.E., Elbeltagi, A., & Kuriqi, A. (2022). Cost-effective management measures for coastal aquifers affected by saltwater intrusion and climate change. *Science of The Total Environment*, *836*, 155656. doi:10.1016/j.scitotenv.2022.155656

Adnan, R.M., Petroselli, A., Heddam, S., Santos, C.A.G., & Kisi, O. (2021). Comparison of different methodologies for rainfall–runoff modeling: machine learning vs conceptual approach. *Natural Hazards*, *105*, 2987-3011. doi: 10.1007/s11069-020-04438-2

Chen, K., Chen, H., Zhou, C., Huang, Y., Qi, X., Shen, R., & Ren, H. (2020). Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data. *Water Research*, *171*, 115454. doi:10.1016/j.watres.2019.115454

Chen, X.W., & Lin, X. (2014). Big data deep learning: challenges and perspectives. *IEEE access*, *2*, 514-525. doi:10.1109/ACCESS.2014.2325029

Elbeltagi, A., Al-Mukhtar, M., Kushwaha, N.L., Al-Ansari, N., & Vishwakarma, D.K. (2023). Forecasting monthly pan evaporation using hybrid additive regression and data-driven models in a semi-arid environment. *Applied Water Science*, *13*(2), 42. doi:10.1007/s13201-022-01846-6

Erdebilli, B., & Devrim-İçtenbaş, B. (2022). Ensemble voting regression based on machine learning for predicting medical waste: a case from Turkey. *Mathematics*, 10(14), 2466. doi:10.3390/math10142466

Feng, K., & Tian, J. (2021). Forecasting reference evapotranspiration using data mining and limited climatic data. *European Journal of Remote Sensing*, 54(sup2), 363-371. doi:10.1080/22797254.2020.1801355

Gholami, H., Mohamadifar, A., Sorooshian, A., & Jansen, J.D. (2020). Machine-learning algorithms for predicting land susceptibility to dust emissions: The case of the Jazmurian Basin, Iran. *Atmospheric Pollution Research*, *11*(8), 1303-1315. doi;10.1016/j.apr.2020.05.009

Granata, F., & Di Nunno, F. (2021). Forecasting evapotranspiration in different climates using ensembles of recurrent neural networks. *Agricultural Water Management*, *255*, 107040. doi:10.1016/j.agwat.2021.107040

Granata, F., Gargano, R., & de Marinis, G. (2020). Artificial intelligence-based approaches to evaluate actual evapotranspiration in wetlands. *Science of The Total Environment*, *703*, 135653. doi:10.1016/j.scitotenv.2019.135653

Archive of SID.ir

Archive of SID.ir

Hinton, G.E., & Salakhutdinov, R.R. (2006). Reducing the dimensionality of data with neural networks. Science, 313(5786), 504-507. doi: 10.1126/science.1127647.

Khan, R.A., El Morabet, R., Mallick, J., Azam, M., Vambol, V., Vambol, S., & Sydorenko, V. (2021). Rainfall Prediction using Artificial Neural Network in Semi-Arid mountainous region, Saudi *Arabia. Ecological Questions*, 32(4), 127-133. doi:10.12775/EQ.2021.038.

Khan, N., Shahid, S., Ismail, T.B., & Wang, X.J. (2019). Spatial distribution of unidirectional trends in temperature and temperature extremes in Pakistan. *Theoretical and Applied Climatology*, *136*, 899-913.

Kisi, O., Mansouri, I., & Hu, J.W. (2017). A new method for evaporation modeling: dynamic evolving neural-fuzzy inference system. *Advances in Meteorology*, *1*, 1-9. doi:10.1155/2017/5356324.

Kushwaha, N.L., Rajput, J., Elbeltagi, A., Elnaggar, A.Y., Sena, D.R., Vishwakarma, D.K., & Hussein, E.E. (2021). Data intelligence model and meta-heuristic algorithms-based pan evaporation modelling in two different agro-climatic zones: a case study from Northern India. *Atmosphere*, *12*(12), 1654. doi:10.3390/atmos12121654

Kushwaha, N.L., Rajput, J., Sena, D.R., Elbeltagi, A., Singh, D.K., & Mani, I. (2022). Evaluation of data-driven hybrid machine learning algorithms for modelling daily reference evapotranspiration. *Atmosphere-Ocean*, *60*(5), 519-540. doi:10.1080/07055900.2022.2087589

Lundberg, A. (1993). Evaporation of intercepted snow-review of existing and new measurement methods. *Journal of Hydrology*, *151*(2-4), 267-290. doi:10.1016/0022-1694(93)90239-6

Malik, A., Tikhamarine, Y., Al-Ansari, N., Shahid, S., Sekhon, H.S., Pal, R.K., Rai, P., Pandey, K., Singh, P., Elbeltagi, A., & Sammen, S.S. (2021). Daily pan-evaporation estimation in diferent agro-climatic zones using novel hybrid support vector regression optimized by Salp swarm algorithm in conjunction with gamma test. *Engineering Applications of Computational Fluid Mechanics*, *15*, 1075-1094. doi:10.1080/19942060.2021.1942990

Masoner, J.R., Stannard, D.I., & Christenson, S.C. (2008). Differences in evaporation between a floating pan and class a pan on land 1. *Journal of the American Water Resources Association*, *44*(3), 552-561. doi:10.1111/j.1752-1688.2008.00181.x

Mohammadi, M., Forozanfard, M., & Gholami, H. (2022). Predicting pan evaporation in a hyper-arid climate using soft computing models: A Case Study of Sistan Plain, Sistan-Baluchistan, Iran. *Desert Ecosystem Engineering*, *11*(36), 71-82. doi:10.22052/deej.2021.11.36.43.[In Persian]

Mohammadi, M., Vagharfard, H., Mahdavi Najafabadi, R., Daneshkar Arasteh, P., & Nazemosadat, M.J. (2021). Rainfall-runoff modelling of coastal watersheds near Hormuz strait using data mining. *Iranian Journal of Soil and Water Research*, *52*(2), 313-327. doi: 10.22059/ijswr.2021.309641.668732. [In Persian]

Mohammadifar, A., Gholami, H., & Golzari, S. (2023). Stacking-and voting-based ensemble deep learning models (SEDL and VEDL) and active learning (AL) for mapping land subsidence. *Environmental Science and Pollution Research*, *30*(10), 26580-26595.

Mosavi, A., Sajedi Hosseini, F., Choubin, B., Taromideh, F., Ghodsi, M., Nazari, B., & Dineva, A.A. (2021). Susceptibility mapping of groundwater salinity using machine learning models. *Environmental Science and Pollution Research*, *28*, 10804-10817. doi:10.1007/s11356-020-11319-5.

Naganna, S.R., Deka, P.C., Ghorbani, M.A., Biazar, S. M., Al-Ansari, N., & Yaseen, Z.M. (2019). Dew point temperature estimation: application of artificial intelligence model integrated with nature-inspired optimization algorithms. *Water*, *11*(4), 742. doi:10.3390/w11040742

Parisouj, P., Mohebzadeh, H., & Lee, T. (2020). Employing machine learning algorithms for streamflow prediction: a case study of four river basins with different climatic zones in the United States. *Water Resources Management*, *34*, 4113-4131. doi: 10.1007/s11269-020-02659-5

Rahman, A.S., Hosono, T., Quilty, J.M., Das, J., & Basak, A. (2020). Multiscale groundwater level forecasting: Coupling new machine learning approaches with wavelet transforms. *Advances in Water Resources*, *141*, 103595. doi:10.1016/j.advwatres.2020.103595

Rezaie-Balf, M., Attar, N.F., Mohammadzadeh, A., Murti, M.A., Ahmed, A.N., Fai, C.M., & El-Shafie, A. (2020). Physicochemical parameters data assimilation for efficient improvement of water quality index prediction: Comparative assessment of a noise suppression hybridization approach. *Journal of Cleaner Production*, *271*, 122576. doi:10.1016/j.jclepro.2020.122576

Sabzevari, Y., & Ghanbarpouri, M. (2022). Evaluation of experimental and intelligent models in estimation of reference evapotranspiration: Case Study Aligodarz. *Desert Ecosystem Engineering*, *11*(36), 17-30. doi: 10.22052/deej.2023.248181.0 [In Persian]

Archive of SID.ir

Archive of SID.ir

Salih, S.Q., Sharafati, A., Ebtehaj, I., Sanikhani, H., Siddique, R., Deo, R.C., & Yaseen, Z.M. (2020). Integrative stochastic model standardization with genetic algorithm for rainfall pattern forecasting in tropical and semi-arid environments. *Hydrological Sciences Journal*, *65*(7), 1145-1157. doi:10.1080/02626667.2020.1734813

Seyedi, S.N., Fazloula, R., Masoudian, M., & Kia, I. (2022). Evaluation the performance of different models of artificial neural network in estimating evaporation losses from pan around the Shahid Rajaei Dam Lake. *Irrigation and Water Engineering*, *13*(2), 179-196. doi:10.22125/iwe.2022.162631

Shahabi, S., Azarpira, F., & Barzkar, A. (2020). Estimation of daily and weekly evapotranspiration using hybrid approaches of soft computing. *Iranian Journal of Irrigation & Drainage*, *14*(5), 1550-1561. dor:20.1001.1.20087942.1399.14.5.5.6. [In Persian]

Sze, V., Chen, Y.H., Yang, T.J., & Emer, J.S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, *105*(12), 2295-2329. doi:10.48550/arXiv.1703.09039

Taylor, K.E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of geophysical research: atmospheres*, *106*(D7), 7183-7192. doi:10.1029/2000JD900719

Vishwakarma, D.K., Pandey, K., Kaur, A., Kushwaha, N.L., Kumar, R., Ali, R., & Kuriqi, A. (2022). Methods to estimate evapotranspiration in humid and subtropical climate conditions. *Agricultural Water Management*, *261*, 107378. doi:10.1016/j.agwat.2021.107378

Wu, L., Huang, G., Fan, J., Ma, X., Zhou, H., & Zeng, W. (2020). Hybrid extreme learning machine with meta-heuristic algorithms for monthly pan evaporation prediction. *Computers and electronics in agriculture*, *168*, 105115. doi:10.1016/j.compag.2019.105115

Zhao, L., Xia, J., Xu, C.Y., Wang, Z., Sobkowiak, L., & Long, C. (2013). Evapotranspiration estimation methods in hydrological models. *Journal of Geographical Sciences*, *23*, 359-369. doi:10.1007/s11442-013-1015-9.

Archive of SID.ir