



reference Evaluation of three data mining methods to estimate evapotranspiration in Zanjan province

Azadeh Sedaghat^{1✉}, Niaz Ali Ebrahimi Pak², Arash Tafteh³,
Seyedeh Narges Hosseini⁴

1. Corresponding Author, Soil and Water Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. E-mail: Azadehsedaghat65@gmail.com
2. Soil and Water Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. E-mail: nebrahimipak@yahoo.com
3. Soil and Water Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. E-mail: arash_tafteh@yahoo.com
4. Soil and Water Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. E-mail: narges_eau@yahoo.com

Article Info

Article type: Research Article

Article history:

Received: Dec. 24, 2022

Revised: Jan. 14, 2023

Accepted: Jan. 25, 2023

Published online: Feb. 20, 2023

Keywords:

Evapotranspiration,
Random Forest,
Artificial Neural Network,
Penmanmonteith model,
Support Vector Machine.

ABSTRACT

Introduction

Reference evapotranspiration (ET₀), a complex hydrological variable affecting crop water requirements and irrigation scheduling, is defined by a number of climatic factors that have an impact on water and energy balances. On the basis of accurate climatic data, conventional methods for calculating ET₀ include a variety of empirical approaches. But there are a lots of locations where different climatic information might not be available for ET₀ estimation.

Objective:

The objective of this study is to evaluate different data mining methods to estimate ET₀ with limited meteorological data. This study aims to answer the question: can reference evapotranspiration be estimated without reducing accuracy, regardless of the availability of all variables? In this research, the accuracy of data mining methods in estimating ET₀ with respect to the plant water demand system (FAO Penman-Monteith standard method) was evaluated

Materials and methods:

Data such as sunshine hour, air temperature, wind speed, and relative humidity from thirteen climatology stations in the Zanjan province over a ten-year period (2010-2021) were collected. The ET₀ was calculated using the FAO56 Penman-Mantith method on a daily time scale (as reference method) and the estimated values obtained by data mining methods (Artificial Neural Network (ANNs), Random Forest (RF) and Support Vector Machine (SVM)) were evaluated. The data from each station were divided into two sets: training (two-thirds of the data) and testing (one-third of the data) in order to calibrate and validate the proposed methods. Finally, based on NRMSE, RMSE, MBE, and EF criteria, the generalizability of the aforementioned methods for estimating ET₀ was examined.

Results and discussion:

According to the results, ANNs performed better than SVM and RF methods. The mean values of, RMSE, EF and NRMSE criteria for the ANNs method in the training and testing steps were 0.49, 0.94 and 0.14, respectively. The mean values of these criteria for RF method in the training step were 0.49, 0.94 and 0.14 and in the testing step was 0.52, 0.94 and 0.15, respectively. The mean values of these criteria for the SVM method for both (training and testing) steps were 0.52, 0.94 and 0.15, respectively.

The average air temperature is the most significant and effective parameter to estimate ET₀, according to more than 92 percent (12 stations) of the results obtained from two ANNs and RF methods. The sunshine hours is the second-most crucial and useful input in estimating ET₀, according to more than 84 percent (11 stations) of the results. As a result, using four meteorological variables such as average air temperature, average relative humidity, wind speed, and sunshine hours as input, excellent performance can be achieved. The NRMSE values obtained from ET₀ estimation did not exhibit regular variations with the average values of parameters (temperature, humidity, wind speed, sunshine hours, slope percentage).

Conclusion: It was found that the average air temperature was the most crucial and useful parameter as a result of the sensitivity analysis of the ANNs method and the Predictor Importance of the RF method. According to the current study, Pari and Zanjan stations outperformed than the other stations in Zanjan province, probably due to their plainer conditions. The results of the current study will help to estimate ET₀ for semi-arid climates where ET₀ is critical for agricultural water resource management.

Cite this article: Sedaghat, A., Ebrahimi Pak, N. A., Tafteh, A., & Hoseini, S. N. (2023). Evaluation of three data mining methods to estimate reference evapotranspiration in Zanjan province. *Iranian Journal of Soil and Water Research*, 53 (12), 2739-2757. <https://doi.org/10.22059/ijswr.2023.352890.669419>

© The Author(s).

Publisher: University of Tehran Press.

DOI: <https://doi.org/10.22059/ijswr.2023.352890.669419>



ارزیابی سه روش داده کاوی برای تخمین تبخیر تعرق مرجع در استان زنجان

آزاده صداقت^۱؛ نیازعلی ابراهیمی پاک^۲؛ آرش تافته^۳؛ سیده نرگس حسینی^۴۱. نویسنده مسئول، موسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران. رایانامه: Azadehsedaghat65@gmail.com۲. موسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران. رایانامه: nebrahimipak@yahoo.com۳. موسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران. رایانامه: arash_tafteh@yahoo.com۴. موسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران. رایانامه: narges_eau@yahoo.com

چکیده

اطلاعات مقاله

تبخیر تعرق مرجع (ETO) یک متغیر هیدرولوژیکی پیچیده است که براساس متغیرهای هواشناسی مختلف که بر تعادل آب و انرژی تأثیر می گذارند، تعریف می شود. به طور معمول ETO با روش های تجربی مختلف بر مبنای داده های اقلیمی دقیق محاسبه می شود. هدف از این مطالعه ارزیابی روش های مختلف داده کاوی که شامل شبکه های عصبی مصنوعی (ANNs)، جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) می باشد جهت تخمین ETO با حداقل ورودی های هواشناسی است. در این مطالعه، مدل FAO-56 Penman-Monteith (FPM) به عنوان مدل استاندارد در نظر گرفته شد. داده های هواشناسی استفاده شده در این پژوهش از ۱۳ ایستگاه استان زنجان در دوره زمانه ده ساله (۱۳۹۹-۱۳۸۹) می باشد. نتایج نشان داد که روش ANNها بهتر از دو روش SVM و RF عمل کرد. میانگین مقادیر آماره های EF، RMSE و NRMSE برای روش ANNها در هر دو مرحله آموزش و آزمون برابر و به ترتیب برابر است با ۰/۴۹، ۰/۹۴ و ۰/۱۴. مقادیر میانگین آماره های مذکور برای روش RF در مرحله آموزش برابر با ۰/۴۹، ۰/۹۴ و ۰/۱۴ و در مرحله آزمون برابر با ۰/۵۲، ۰/۹۴ و ۰/۱۵ بود. همچنین مقادیر میانگین این آماره ها برای روش SVM در هر دو مرحله آموزش و آزمون برابر و به ترتیب برابر است با ۰/۵۲، ۰/۹۴ و ۰/۱۵ شد. بیش از ۹۲ درصد (۱۲ ایستگاه) نتایج بدست آمده از دو روش ANNها و RF نشان دادند که میانگین دما مهم ترین و موثرترین پارامتر در تخمین ETO می باشد. همچنین بیش از ۸۴ درصد (۱۱ ایستگاه) نتایج نشان دادند که ساعت آفتابی دومین ورودی مهم و موثر در تخمین ETO می باشد. بنابراین می توان عملکرد عالی را با استفاده از چهار متغیر هواشناسی (میانگین دما، میانگین رطوبت نسبی، سرعت باد و ساعت آفتابی) به عنوان ورودی با روش های ANNها، RF و SVM به دست آورد. نتایج پژوهش حاضر می تواند به تخمین ETO برای مناطقی که مدیریت آب کشاورزی در آنجا ضروری می باشد، کمک کند.

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۴۰۱/۱۰/۳

تاریخ بازنگری: ۱۴۰۱/۱۰/۲۴

تاریخ پذیرش: ۱۴۰۱/۱۱/۵

تاریخ انتشار: ۱۴۰۱/۱۲/۱

واژه های کلیدی:

تبخیر تعرق، جنگل تصادفی، شبکه عصبی مصنوعی، ماشین بردار پشتیبان، مدل پنمن مانتیث.

استناد: صداقت؛ آزاده، ابراهیمی پاک؛ نیازعلی، تافته؛ آرش، حسینی؛ سیده نرگس (۱۴۰۱). ارزیابی سه روش داده کاوی برای تخمین تبخیر تعرق مرجع در استان زنجان،

مجله تحقیقات آب و خاک ایران، ۵۳ (۱۲)، ۲۷۵۷-۲۷۳۹. <https://doi.org/10.22059/ijswr.2023.352890.669419>

© نویسندگان.

ناشر: مؤسسه انتشارات دانشگاه تهران.

DOI: <https://doi.org/10.22059/ijswr.2023.352890.669419>

مقدمه

تبخیر و تعرق (ET) فرآیند از دست رفتن آب از خاک و سطوح محصول به اتمسفر توسط فرآیندهای تبخیر و تعرق است. تخمین تبخیر تعرق مرجع (ET₀) گام مهمی در مطالعاتی مانند هیدرولوژیکی، برنامه ریزی آبیاری و مدیریت منابع آب است. نیاز آبی محصولات را می‌توان از ET₀ محاسبه کرد (Ferreira *et al.*, 2019). تغییرات اقلیمی که در سطح جهان در حال رخ دادن است و تاثیر زیادی روی مقدار ET₀ دارد. نتایج تاثیر تغییر اقلیم بر روی مقدار ET₀ نشان دهنده آن است کاهش مقدار ET₀ در استرالیا (Roderick *et al.*, 2002)، و ایران (Tabari and Talaei, 2014) عمدتاً ناشی از سرعت باد ناشی از مناطق است، درحالی‌که در هند (Bandyopadhyay *et al.*, 2009) رطوبت نسبی و در چین حداکثر دما (Zhou *et al.*, 2014) و فشار بخار (Su *et al.*, 2021) مهم‌ترین عامل موثر بر مقدار ET₀ است. ET₀ را می‌توان با استفاده از روش‌های مستقیم (لایسیمترها) اندازه‌گیری کرد، که اندازه‌گیری‌های دقیقی را ارائه می‌دهد و معمولاً در توسعه و اعتبار سنجی روش‌های دیگر استفاده می‌شود (Allen *et al.*, 2011; Fen and Tian, 2021). با این حال، با توجه به هزینه و پیچیدگی یک لایسیمتر، استفاده از آن معمولاً محدود به تحقیقات است (Allen *et al.*, 2011). بنابراین، استفاده از مدل‌های ریاضی مبتنی بر داده‌های هواشناسی، رویکرد مناسب‌تری برای کاربردهای عملی است (Fen and Tian, 2021).

سازمان غذا و کشاورزی (FAO) استفاده از معادله FAO-56 Penman-Monteith (FPM) را به‌عنوان یک روش مرجع برای تخمین ET₀ توصیه می‌کند (Allen *et al.*, 1998). این معادله، با وجود داشتن دقت شناخته شده در مقیاس جهانی به تعداد زیادی داده‌های هواشناسی (حداکثر و حداقل دمای هوا، رطوبت نسبی، تابش خورشیدی و سرعت باد) نیاز دارد. اما متأسفانه تعداد محدودی از ایستگاه‌های هواشناسی وجود دارد که اطلاعات کاملی از این متغیرهای آب و هوایی را داشته باشد (Panaitescu *et al.*, 2014). تبخیر تعرق یک پدیده غیرخطی، پویا و پیچیده است (Wu *et al.*, 2019)، و استخراج فرمولی که بتواند تمام فرآیندهای فیزیکی درگیر را نشان دهد دشوار است (Wen *et al.*, 2015). در نتیجه ضروری است که برای مدل‌سازی این پدیده از روش‌های قوی و غیرخطی استفاده شود. در این راستا، روش‌های داده‌کاوی ایده خوبی برای مدل‌سازی ET₀ هستند. روش‌های داده‌کاوی در بسیاری از مطالعات برای حل مسائل پیچیده و غیرخطی استفاده شده است. بنابراین، روش‌های داده‌کاوی به دلیل ظرفیت بالایشان برای رسیدگی به مسائل پیچیده و غیرخطی، جایگزین مناسبی برای روش‌ها سنتی هستند (Ferreira *et al.*, 2019; Rahimikhoob, 2014). در سال‌های اخیر، چندین مطالعه به برآورد ET₀ با استفاده از روش‌های داده‌کاوی (مانند شبکه عصبی مصنوعی (ANNs)^۱ و ماشین بردار پشتیبانی (SVM)^۲) انجام شده است (Feng *et al.*, 2018; Kisi and Alizamir, 2018; Mattar, 2018; Mehdizadeh *et al.*, 2018). در این مطالعات، اغلب عملکرد SVM عملکرد خوبی برای تخمین ET₀ نشان داده است (Fan *et al.*, 2018; Wen *et al.*, 2015).

در زمینه تخمین تبخیر تعرق به کمک روش‌های داده‌کاوی پژوهش‌های مختلفی در داخل و خارج کشور انجام شده است. سپهری و همکاران (۱۳۹۹) به بررسی و تحلیل حساسیت روش‌های مبتنی بر شبکه عصبی در برآورد تبخیر و تعرق مرجع در استان تهران پرداختند. نتایج نشان داد مدل ANNs تکنیک بسیار مناسبی برای تحلیل تبخیر و تعرق مرجع می‌باشد. بیدآبادی و همکاران (۱۴۰۰) به برآورد تبخیر و تعرق مرجع با استفاده از مدل ANNs و مدل نرو فازی در اقلیم خشک و نیمه خشک (استان کرمان) پرداختند. نتایج نشان داد دو مدل فوق عملکرد بهتری نسبت به مدل‌های تجربی دارند و دقت مدل نرو فازی در مقایسه با سایر مدل‌ها بیشتر است. ذرتی پور و همکاران (۱۳۹۸) به محاسبه تبخیر و تعرق پتانسیل به کمک مدل ANNs، عصبی فازی، درختی M5 پرداختند. نتایج نشان داد هر سه مدل دقت قابل قبولی در محاسبه تبخیر تعرق دارند.

Yassin *et al.*, (2016) با استفاده از ANNs و برنامه‌ریزی بیان ژن (GEP) در آب‌وهوای خشک تخمین زدند. نتایج نشان داد که هشت مدل ET₀ ایجاد شده با استفاده از روش ANNs کمی دقیق‌تر از مدل‌های GEP بودند. (Allawi *et al.*, 2019) در بررسی توانایی شبکه عصبی مصنوعی بر پایه شعاعی (RBF-NN) و روش‌های رگرسیون بردار پشتیبانی (SRV) در تهیه یک مدل پیش‌بینی مقدار ET₀ در منطقه گرمسیری در مالزی نشان دادند که در مقایسه با مقادیر تبخیر تعرق اندازه‌گیری شده، دقت روش RBF-NN بالاتر از SRV بود. (Ferreira *et al.*, 2019) با استفاده از داده‌های اقلیمی محدود با استفاده از الگوریتم MARS و معادله FPM، ET₀ مدل‌سازی کردند.

^۱. Artificial Neural Networks

^۲. Support Vector Machine



مدل MARS در تمام سناریوها عملکرد برتر را نشان داد. در مدل‌هایی که از ساعت آفتابی استفاده شد، بهترین عملکرد داشتند و پس از آن بهترین عملکرد را مدل‌هایی نشان دادند که از رطوبت نسبی و سرعت باد استفاده کردند. (Wang *et al.*, 2019) مدل‌های تبخیرتغرق را با داده‌های محدود بر اساس GEP و RF^۱ در چین بررسی کردند. نتایج نشان داد که مدل‌های ET₀ مبتنی بر RF کمی بهتر از مدل‌های مبتنی بر GEP عمل کردند. (Fan *et al.*, 2019) با استفاده از داده‌های هواشناسی محلی و خارجی و با بکارگیری از روش‌های M5، RF، و معادلات تجربی مکینک، طبری، هارگریوز-سامانی و ترابرت در مناطق مرطوب، ET₀ تخمین زدند. نتایج نشان داد که هر سه مدل، برآوردهای ET₀ روزانه بهتری را نسبت به مدل‌های تجربی با استفاده از متغیرهای ورودی یکسان تولید کردند. (Algetawee and Alshama, 2021) تبخیرتغرق مرجع را به کمک روش ANNs برآورد کردند. برای محاسبه ET₀ از روش فائو پنمن مونیتث استفاده گردید. نتایج نشان داد که روش ANNs دقت قابل قبولی در تخمین ET₀ دارد. همچنین مشخص گردید تغییرات آب و هوایی باعث افزایش ET₀ می‌گردد.

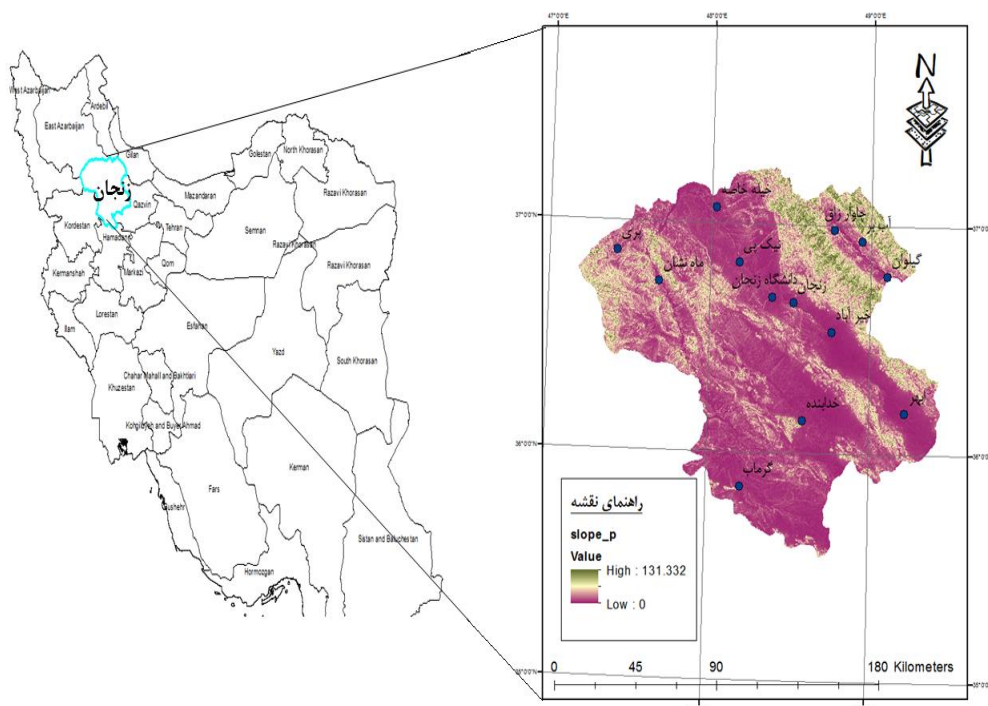
با توجه به این که ایران جزء کشورهای است که دارای اقلیم خشک و نیمه‌خشک می‌باشد، که هر ساله با محدودیت‌های گسترده منابع آب به‌خصوص در بخش کشاورزی مواجه است (Tabari *et al.*, 2014). در بخش کشاورزی تأمین نیاز آبی محصولات برای بالاتر بردن کارایی استفاده از آب و جلوگیری از مصرف مازاد آن و دسترسی به حداکثر عملکرد، ضروری است. ET₀ با دقت بالایی برآورد شود. بنابراین هدف از پژوهش حاضر تجزیه و تحلیل ET₀ با استفاده از داده کاوی می‌باشد. در این پژوهش، هدف پاسخ به این سؤال است: آیا می‌توان بدون توجه به در دسترس بودن همه متغیرها، ET₀ را بدون کاهش دقت تخمین زد؟ برای حل این مشکل، از مجموعه داده با دوره زمانی ده‌ساله که از ۱۳ ایستگاه هواشناسی مستقر در استان زنجان اخذ گردید، استفاده شد.

مواد و روش‌ها

موقعیت جغرافیایی استان زنجان و مشخصات ایستگاه‌های هواشناسی

استان زنجان در شمال باختر فلات ایران قرار گرفته است. موقعیت جغرافیایی استان منطبق بر عرض جغرافیایی ۳۵ درجه و ۳۵ دقیقه تا ۳۷ درجه و ۱۵ دقیقه شمال و طول جغرافیایی ۴۷ درجه و ۱۵ دقیقه تا ۴۵ درجه و ۲۵ دقیقه شرقی از نصف‌النهار گرینویچ است. از نظر شکل ظاهری زمین شامل دو منطقه کوهستانی و جلگه‌ای (دشت) می‌باشد. بیشتر مساحت این استان را ترکیبی از مناطق مرتفع و جلگه‌ای یا دشت تشکیل می‌دهد (<https://www.znu.ac.ir/museum/zanjan-locality>). استان زنجان به لحاظ داشتن تنوع نقاط ارتفاعی از یک سو و دیگر تأثیرپذیری از چند توده هوایی خزری، مدیترانه‌ای و صحرای مرکزی از طرف دیگر، صاحب اقلیم‌ها و اکوسیستم‌های متنوعی شده است. این استان در بیش‌تر از ۷۰ درصد از مناطق خود آب و هوای نیمه خشک فراسرد و در ۳۰ درصد باقی مناطق از تنوع اقلیمی و آب و هوایی برخوردار است. میزان بارندگی سالانه استان زنجان حدود ۳۲۳ میلی‌متر برآورد شده است (<https://magirans.com>). موقعیت جغرافیایی استان زنجان در شکل ۱ نشان داده شده است. مشخصات ۱۳ ایستگاه هواشناسی مورد مطالعه در پژوهش حاضر در جدول ۱ آمده است.

^۱. Random Forest



شکل ۱. موقعیت جغرافیایی و نقشه درصد شیب استان زنجان

جدول ۱- مشخصات ایستگاه‌های مورد مطالعه

استان	شهرستان	نام ایستگاه	نوع ایستگاه	طول جغرافیایی	عرض جغرافیایی	ارتفاع
۱ زنجان	طارم	گیلوان	کلیماتولوژی	۴۹/۱	۳۶/۷۸	۳۰۱
۲ زنجان	طارم	چاوار زاق	کلیماتولوژی	۴۸/۷۷	۳۶/۹۸	۴۸۱
۳ زنجان	طارم	آب بر	سینوپتیک	۴۸/۹۴	۳۶/۹۳	۶۲۵
۴ زنجان	ماه نشان	ماه نشان	سینوپتیک	۴۷/۶۸	۳۶/۷۴	۱۲۸۴
۵ زنجان	زنجان	خیله خاصه	کلیماتولوژی	۴۸/۰۳	۳۷/۰۷	۱۳۲۵
۶ زنجان	زنجان	نیک پی	کلیماتولوژی	۴۸/۱۸	۳۶/۸۳	۱۴۱۵
۷ زنجان	خرمدره	ابهر	کلیماتولوژی	۴۹/۲۲	۳۶/۱۸	۱۵۵۲
۸ زنجان	زنجان	دانشگاه زنجان	سینوپتیک	۴۸/۳۹	۳۶/۶۸	۱۵۷۹
۹ زنجان	خدابنده	گرماب	سینوپتیک	۴۸/۲۱	۳۵/۸۴	۱۶۳۲
۱۰ زنجان	زنجان	زنجان	سینوپتیک	۴۸/۵۲	۳۶/۶۶	۱۶۶۰
۱۱ زنجان	ماه نشان	پری	کلیماتولوژی	۴۷/۴۲	۳۶/۸۷	۱۷۴۷
۱۲ زنجان	سلطانیه	خیر آباد	سینوپتیک	۴۸/۷۵	۳۶/۵۳	۱۷۶۷
۱۳ زنجان	خدابنده	خدابنده	سینوپتیک	۴۸/۵۸	۳۶/۱۴	۱۸۸۷

روش محاسبه تبخیر و تعرق مرجع با مدل فائو - پنمن - مانیتث (FPM)

رابطه ۱، را (Allen et al., (1998 با نام مدل فائو - پنمن - مانیتث (FPM)، به عنوان بهترین روش برآورد ET_0 در مقایسه با دیگر روش‌ها ارائه کرده است:

$$ET_0 = \frac{0.408D(R_n - G) + \gamma \left(\frac{900}{T + 273} \right) U_2 (e_s - e_a)}{D + \gamma(1 + 0.34U_2)} \quad \text{رابطه ۱}$$

ET_0 تبخیر تعرق مرجع ($mm\ d^{-1}$)، D : شیب منحنی فشار بخار ($kpa^\circ C^{-1}$)، R_n : تشعشع خالص ($MJ\ m^{-2}\ d^{-1}$)، G : جریان حرارتی خاک ($MJ\ m^{-2}\ d^{-1}$)، U_2 : سرعت باد در ارتفاع ۲ متر ($m\ s^{-1}$)، $(e_s - e_a)$: کمبود فشار بخار در ارتفاع ۲ متر (kPa)، T : دمای متوسط در



ارتفاع ۲ متری (درجه سانتی‌گراد)، ۹۰۰: ضریبی برای گیاه مرجع ($J^{-1} \text{ Kg K d}^{-1}$)، γ : ثابت سایکرومتریک ($\text{kpa}^\circ\text{C}^{-1}$)، ۰/۳۴: ضریب باد برای گیاه مرجع (s m^{-1}).

مدل FPM که مقادیر ET_0 را محاسبه کرده است به‌عنوان معیار اندازه‌گیری برای آموزش و آزمون روش‌های مختلف یادگیری ماشین در نظر گرفته شد. ET_0 محاسبه شده با مدل FPM از سامانه نیاز آب متعلق به موسسه تحقیقات خاک و آب کشور استخراج شده است (<http://www.swri.ir/en-S/DouranPortal/1/page/Home>). پارامترهای مورد استفاده در این پژوهش شامل میانگین دمای هوای روزانه (T^{C})، میانگین رطوبت نسبی، متوسط سرعت وزش باد^۱ در ارتفاع دو متری ($U_2, \text{m s}^{-1}$) و ساعت آفتابی^۲ ($R_s, \text{MJ m}^{-2} \text{d}^{-1}$) می‌باشد.

توصیفی از روش‌های داده کاوی

مدل‌سازی مقدار ET_0 با سه روش شبکه عصبی مصنوعی (ANNs)، ماشین بردار پشتیبان (SVM) و جنگل تصادفی (RF) با نرم افزار STATISTICA 12 انجام شد. روش ANN یک روش طبقه‌بندی غیر پارامتری است که بر فرض توزیع نرمال داده‌ها بستگی ندارد (Dixon and Candade, 2008). یکی از انواع مهم شبکه‌های عصبی، شبکه پرسپترون چندلایه (MLP)^۳ است که شامل یک لایه ورودی، یک یا چند لایه مخفی و یک لایه خروجی می‌باشد. روش ANN استفاده شده در این پژوهش شامل دو نوع شبکه عصبی پرسپترون چند لایه (MLP) و تابع پایه‌ی شعاعی (RBF)^۴ می‌باشد. اساس شبکه عصبی مرکب (هر دو نوع شبکه) توسط Picton (2000) و Hagan *et al.* (1996) ارائه شده است. برای کاهش خطا در مدل‌سازی، شبکه عصبی در ۱۰۰ تکرار تنظیم شد. میانگین ۵۰ خروجی از ۱۰۰ تکرار که خطای کمتری دارند، به‌عنوان خروجی نهایی برای تخمین ET_0 انتخاب شد.

نسل بعدی از روش‌های هوش مصنوعی، ماشین بردار پشتیبان (SVM)، نام دارد که نسبت به روش ANN سریع‌تر آموزش دیده و نتایج دقیق‌تری را ارائه می‌دهد. از طرفی این مدل کارایی مناسبی در کار با داده‌های غیرخطی نسبت به سایر مدل‌ها از خود نشان داده است (Lin *et al.*, 2009). روش SVM، روش بسیار کارا و مفید برای حل مسائل داده مینا است که میان داده‌های ورودی و خروجی رابطه برقرار می‌کند (Gill *et al.*, 2007). این روش همچنین یک روش طبقه‌بندی دودویی^۵ است که با تعیین یک صفحه تفکیک کننده بهینه در فضای داده‌های آموزشی، کلاس‌های مختلف را با حداکثر جدایی بین آنها، تفکیک می‌کند. فرا صفحه‌ای که حداکثر حاشیه را بین دو کلاس فراهم می‌نماید، فرا صفحه بهینه و داده‌های نزدیک‌تر به فرا صفحه، بردارهای پشتیبان نامیده می‌شوند (Petropoulos *et al.*, 2015). از جمله قابلیت‌های ماشین بردار پشتیبان، توانایی غلبه بر مشکل توزیع غیرخطی داده‌های آموزشی است. در این حالت با استفاده از توابع کرنل، داده‌ها به فضایی با بعد بزرگتر انتقال می‌یابند که در آن تفکیک‌پذیری بهتر انجام می‌شود (Petropoulos *et al.*, 2015). جذابیت ماشین‌های بردار پشتیبان در قابلیت آنها در حداقل کردن ریسک ساختاری یا خطای طبقه‌بندی هنگام حل مسئله طبقه‌بندی است (Aslami *et al.*, 2015). این روش دارای چهار نوع کرنل شامل خطی^۶، تابع پایه شعاعی^۷، سیگموئید^۸ و چند جمله‌ای^۹ می‌باشد (Achirul Nanda *et al.*, 2018). در این پژوهش، روش SVM از نوع کرنل، هسته تابع پایه شعاعی (RBF) که شامل عامل گاما (γ) می‌باشد. روش SVM دارای عامل‌های تنظیم‌کننده شامل C و nu هسته است.

جنگل تصادفی یا به‌اختصار RF پتانسیل بسیار زیادی برای تبدیل به یک روش محبوب برای طبقه‌بندی آینده را دارد، زیرا کارایی آن قابل مقایسه با دیگر روش‌های گروهی از جمله Bagging و Boosting می‌باشد (Kulkarni and Sinha, 2014). به‌عنوان یک الگوریتم گروهی، RF چندین درخت تصمیم مختلف به‌عنوان طبقه‌بندهای پایه تولید می‌کند و رأی اکثریت را برای ترکیب با نتایج درختان اصلی اعمال می‌کند. پارامترهای کلیدی برای مدل جنگل تصادفی، تعداد درختان و تعداد متغیرهای پیشگو می‌باشند. براساس این دو پارامتر، درخت تصمیم به بزرگترین اندازه ممکن رویانده و بدون هرس رها می‌شود (Rodriguez-Galiano *et al.*, 2012). اصول کلی تکنیک‌های

1. Wind Speed
2. Solar Radiation
3. Multilayer perceptron's
4. Radial basis function
5. Binary
6. Linear
7. Radial Basis Function
8. Sigmoid
9. Polynomial

آموزش درخت‌ها بر پایه‌ی این فرض است که دقت آن‌ها از دیگر الگوریتم‌های آموزشی بالاتر است. چون ترکیبی از چند مدل پیش‌بینی، دقیق‌تر از یک مدل می‌باشد و گروه‌ها در حالی که نقاط ضعف طبقه‌ها را کاهش می‌دهد، همزمان قدرت مجموعه‌های منفرد و منحصر به‌فرد طبقه‌ها را نیز بیشتر می‌کند (Kotsiantis and Pintelas, 2004). اساس روش RF توسط (Breiman, 2001) ارائه شده است. الگوریتم‌های ارزیابی شده در آموزش خود دارای مشکلات متفاوتی هستند. همه روش‌های داده‌کاوی دارای مزایا و معایب خاصی هستند. الگوریتم‌های مبتنی بر RF دشواری کمتری در آموزش دارند. روش SVM بر اساس انواع هسته‌های مختلفی که دارند و براساس آنها ترکیب پارامترهایی که باید بهینه شوند، متفاوت است. ANNs پتانسیل تبدیل شدن به یک الگوریتم طبقه‌بندی پرکاربرد را دارد. با این حال، باید به شدت تأکید کرد که نمی‌توان در مورد برتری یک روش برای غلبه بر همه نوع مشکلات ادعا کرد، زیرا عملکرد روش‌ها برای مجموعه داده‌های مختلف، متفاوت می‌باشد (Boateng et al., 2020).

فرآیندهای پیش‌پردازش جهت ایجاد مدل

فرضیه‌های پیش‌پردازش و رگرسیون داده‌ها، از جمله تشخیص داده‌های پرت‌ها^۱، آزمون نرمال کردن داده‌ها، بررسی چندخطی بودن برای همه عامل‌ها اعمال شد (Berry, 1993). فرض صحت رگرسیون خطی، بررسی هم‌راستایی بین متغیرهای مستقل است. همچنین در تمام مدل‌ها این مسئله از طریق آماره عامل تورم واریانس (VIF)^۲ بررسی شد. متغیرهایی که VIF آن‌ها کمتر از ۱۰ باشد مشکل هم‌راستایی نخواهند داشت (Hocking, 2013). که در این تحقیق VIF تمام ورودی‌ها و مقایسه میانگین بین نتایج‌های سه روش بکار رفته در این پژوهش با SPSS.26 بررسی شد. در این پژوهش دوسوم داده‌ها (۶۶ درصد داده‌ها) برای مرحله آموزش و یک‌سوم از داده‌ها (تقریباً ۳۴ درصد داده‌ها) برای مرحله آزمون انتخاب شده است.

معیارهای ارزیابی

ارزیابی عملکرد روش‌های مختلف ANNs، RF و SVM در تخمین مقادیر ET_0 از آماره‌های مجذور میانگین مربعات خطا ($RMSE^3$)، میانگین انحراف خطای (MBE^4)، مجذور میانگین خطای نرمال شده (NRMSE) و ضریب کارایی مدل (EF) بهره‌گیری شد (رابطه‌های ۲-۵). از لحاظ آماری هرچقدر مقدار RMSE کمتر باشد، طبیعتاً مدل برآورد کننده دارای دقت بیشتری خواهد بود (Adab et al., 2020). MBE عمدتاً برای نشان دادن اریب در مدل مورد استفاده قرار می‌گیرد و مقدار مثبت آن نشان‌دهنده بیش برآوردی و مقدار منفی آن کم‌برآوردی را در تخمین مقدار ET_0 را نشان می‌دهد (Gopinathan, 1988). از NRMSE به‌عنوان شاخص خطای پیش‌بینی مدل استفاده می‌شود. بنابراین اگر مقدار NRMSE کمتر از ۱۰٪ باشد، نتیجه تخمین عالی، بین ۱۰ تا ۲۰٪، خوب؛ اگر بین ۲۰ تا ۳۰٪، قابل قبول؛ و برای مقادیر بیشتر از ۳۰، ضعیف می‌باشد (Sandhu and Irmak., 2020). همچنین مقدار EF از منفی بی‌نهایت برای بدترین برآورد تا یک برای تطابق کامل تغییر کرده و کفایت و کارایی مدل را نشان می‌دهد و بالای ۰/۵ عملکرد قابل قبول مدل را نشان می‌دهد (کریمی‌پور و بنی‌طالبی، ۱۳۹۹).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}} \quad \text{رابطه ۲}$$

$$NRMSE = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2}}{\bar{X}} \quad \text{رابطه ۳}$$

$$MBE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i) \quad \text{رابطه ۴}$$

$$EF = 1 - \frac{\sum_{i=1}^N x_i - y_i^2}{\sum_{i=1}^N x_i - \bar{X}^2} \quad \text{رابطه ۵}$$

در رابطه‌های ۵-۲، N تعداد داده‌ها، x_i : داده‌های اندازه‌گیری شده، \bar{X} میانگین داده‌های اندازه‌گیری شده، y_i : داده‌های تخمینی، و \bar{Y} میانگین داده‌های تخمینی می‌باشد.

نتایج و بحث

توصیف آماری

ویژگی‌های آماری متغیرهای بکار رفته در پژوهش حاضر و تعداد داده‌های استفاده شده در هر ایستگاه در جدول ۲ آمده است. جدول ۲ تغییرات پارامترهای بکار رفته در این تحقیق را از سال ۱۳۸۹ تا ۱۳۹۹ را نشان می‌دهد. کلیه داده‌های ورودی و خروجی در ۱۳ ایستگاه منتخب دارای دامنه تغییرات بالایی (فاصله بین کمترین و بیشترین مقدار) می‌باشند. هنگامی که تغییرات داده‌های ورودی و خروجی زیاد باشد، نشان دهنده این است که قابلیت تعمیم‌پذیری مدل افزایش و هنگامی که تغییرپذیری کم است، قابلیت تعمیم‌پذیری آن نیز کاهش می‌یابد. به منظور نشان دادن پراکندگی نسبی داده‌ها به میانگین، از ضریب تغییرات (CV) استفاده شد. انتخاب داده‌ها برای آموزش و آزمون به صورت تصادفی انجام شد، بنابراین محدوده تغییرات و میانگین متغیرها برای داده‌های آموزش و آزمون مشابه است. بیشترین CV برای میانگین دما (حد اکثر دما: ۲۹/۱۵ و حداقل دما: ۱۸/۹-)، میانگین رطوبت نسبی (حد اکثر: ۱۰۰ و حداقل: ۰) و ساعت آفتابی (حد اکثر: ۱۴/۵ و حداقل: ۰) را ایستگاه خیرآباد، بیشترین CV برای سرعت باد (حد اکثر: ۳/۸۳ و حداقل: ۰) را ایستگاه نیک پی و بیشترین CV برای ET_0 (حداقل: ۹/۵۱ و حداقل: ۰/۳۱) را ایستگاه خدابنده دارد (جدول ۲). بنابراین بیشترین تغییرات (FPM) ET_0 در ایستگاه خدابنده که در جنوب استان زنجان قرار دارد، مشاهده شده است.

ارزیابی عملکرد روش‌های مختلف در برآورد مقدار ET_0

در این پژوهش مقدار ET_0 روزانه در ۱۳ ایستگاه هواشناسی استان زنجان با سه روش ANNs، RF و SVM برآورد شد، و سپس عملکرد روش‌ها باهم مورد مقایسه قرار گرفت.

ویژگی‌های روش ANNs

با توجه به جدول ۳، روش ANNs بکار برده شده در این پژوهش دارای سه لایه داده ورودی، میانی و خروجی بودند. از توابع فعال‌سازی Sine، Tanh، Logistic، Gaussian، Exponential و Identity در لایه‌های میانی و خروجی استفاده شد (جدول ۳). از ۵۰ شبکه انتخاب شده در هر مدل، شامل هر دو نوع شبکه RBF و MLP هستند، اما در هر ۱۳ ایستگاه، تعداد MLP استفاده شده در ایجاد مدل بیشتر از RBF است. نتایج نشان داد که شبکه عصبی MLP نسبت به RBF در برآورد مقدار ET_0 از عملکرد بهتری برخوردار است. در این پژوهش تعداد سلول‌ها (نورون‌های) پنهان شبکه عصبی از ۳ تا ۳۰ متغیر است (جدول ۳). Identity و Logistic و به ترتیب تابع فعال‌سازی غالب در لایه پنهان و خروجی بوده و بهترین تابع برای توصیف روابط بین متغیرهای ورودی و ET_0 می‌باشند.

ویژگی‌های روش SVM

مشخصات روش SVM که در این مطالعه استفاده شد، در جدول ۴ آمده است. در این پژوهش نوع کرنل، از هسته تابع پایه شعاعی (RBF) استفاده شده که شامل پارامتر گاما (γ) است. در روش SVM پارامترها و هسته مورد استفاده باید بهینه شود در غیر این صورت به دقت مطلوب نمی‌رسد. عملکرد روش SVM با انتخاب فرآیندهای مناسب مانند مقدار ϵ ، C و γ تعیین می‌شود (Yapnik 2000). اگر پارامتر C مقدار بزرگی باشد، دقت، در مرحله آموزش بالاست در حالی که معمولاً در مرحله آزمون دقت پایین می‌آید. همچنین، اگر پارامتر C مقدار کوچک باشد، دقت نامطلوب است و مدل را بی‌اعتبار می‌کند. پارامتر هسته گاما (γ) معمولاً تأثیر بیشتری نسبت به پارامتر C ، بر نتیجه مدل‌سازی دارد. اگر مقدار پارامتر C بزرگ انتخاب شود منجر به بیش‌برازشی می‌شود، در حالی که مقدار کوچک آن منجر به کم‌برازشی می‌شود.

جدول ۲. ویژگی‌های آماری متغیرهای مورد مطالعه

چولگی	CV (%)	SD	میانگین	بیشینه	کمینه	چولگی	CV	SD	میانگین	بیشینه	کمینه
آب بر (تعداد: ۴۰۱۸)						ابه‌ر (تعداد: ۳۶۴۵)					
-۳/۹	۳۵/۴	۱۸/۱۲	۸/۶۲	۴۷/۵۸	-۰/۱	-۱۰/۹	۳۲/۲	۱۳/۲۷	۹/۱۹	۶۸/۷۴	-۰/۱۳
۱۰/۵	۹۷	۵۶/۶۹	۱۵/۱۶	۲۶/۷۴	۰/۱۹	۱۸	۹۹	۵۲/۲۷	۱۶/۲۴	۳۱/۰۷	۰/۳۲
.	۵/۵۳	۱/۴۱	۰/۶۲	۴۴/۲۲	۱/۰۷	.	۳/۹۳	۱/۵۲	۰/۶۱	۳۹/۸۹	۰/۶۲
.	۱۴/۵۴	۸/۹۴	۳/۵۶	۳۹/۸۴	-۰/۹۳	.	۱۳/۲	۸/۳۱	۳/۴۹	۴۱/۹۲	-۰/۹۳
-۰/۳۷	۱۰/۸۳	۳/۶۷	۲/۱۲	۵۷/۸۲	۰/۲۳	-۰/۳۵	۸/۹۲	۳/۴۴	۲	۵۸/۰۶	۰/۲۴
چوار رازق (تعداد: ۴۰۱۱)						پری (تعداد: ۴۰۱۸)					
-۳/۹	۳۵/۴	۱۸/۳	۸/۶۸	۴۷/۴۴	-۰/۱۲	-۱۷/۳	۳۴/۱	۱۴/۹۹	۹/۸۶	۶۵/۷۸	-۰/۱۴
۱۰/۵	۹۷	۵۵/۶۹	۱۵/۳۴	۲۷/۵۴	۰/۲۴	۱۱	۹۶/۵	۴۹/۷۸	۱۶/۸۳	۳۳/۸۱	۰/۲۴
.	۳/۰۶	۱/۳۵	۰/۵۱	۳۸/۰۹	۰/۳	.	۸/۴۵	۱/۵۴	۰/۵۱	۳۳/۵	۳/۱۷
.	۱۴/۵۵	۸/۶۲	۳/۹۱	۴۵/۳۷	-۰/۴	.	۱۳/۹	۷/۹۲	۳/۲۳	۴۰/۷۳	-۰/۳۷
-۰/۳۵	۹/۵۸	۳/۶۴	۲/۰۷	۵۶/۹۳	۰/۲۱	-۰/۴۱	۹/۱۶	۳/۷۱	۲/۲۴	۶۰/۳۳	۰/۲۹
خدابنده (تعداد: ۴۰۱۸)						خیرآباد (تعداد: ۳۶۵۲)					
-۱۴/۵	۳۱/۸	۱۱/۹۴	۹/۸۶	۸۲/۵۴	-۰/۱۴	-۱۸/۹	۲۹/۱۵	۱۰/۲۲	۹/۱۳	۸۹/۳۳	-۰/۱۶
۱۱/۵	۹۶/۵	۵۱/۲۵	۱۹/۴۹	۳۸/۰۲	۰/۱۹	.	۱۰۰	۵۰/۹۶	۲۰/۸۵	۴۰/۹۲	-۰/۴۳
.	۹/۲۷	۱/۴۸	۰/۷۳	۴۹/۰۹	۳/۲۴	.	۳/۵۴	۱/۵۵	۰/۶۳	۴۰/۵۱	۰/۰۵
.	۱۴/۰۷	۸/۳۵	۳/۶۶	۴۳/۸	-۰/۸۲	.	۱۴/۵	۸/۳	۴/۱۱	۴۹/۵	-۰/۵۹
-۰/۳۱	۹/۵۱	۳/۳۱	۲/۰۴	۶۱/۶۸	۰/۲۴	-۰/۳۳	۸/۸۶	۳/۲۹	۲/۰۳	۶۱/۵۹	۰/۴۲
دانشگاه زنجان (تعداد: ۴۰۱۸)						خیله خاصه (تعداد: ۴۰۱۸)					
-۱۵/۷	۳۱/۱	۱۲/۸۵	۹/۱۶	۷۱/۳	-۰/۹۱	-۸/۳	۳۱/۱	۱۳/۲۹	۹/۰۸	۶۸/۳۳	-۰/۰۴
۱۸/۵	۹۸	۵۴/۲۵	۱۴/۷۶	۳۷/۲	-۰/۶۱	.	۹۹	۵۲/۱۵	۱۷/۰۴	۳۲/۶۸	-۰/۶
-۰/۳۷	۳/۰۸	۱/۵۵	۰/۵۲	۳۴/۰۶	-۱/۳۴	.	۲/۸	۱/۱۸	۰/۴۷	۳۹/۵۶	۰/۵۲
.	۱۳/۷	۷/۹۳	۲/۶۷	۴۶/۳	-۰/۵۶	.	۱۳/۵	۸/۲۴	۳	۳۶/۴۵	-۰/۳۹
-۰/۴۲	۸/۲	۳/۴۴	۲/۰۱	۵۸/۶	-۱/۱۹	-۰/۰۵	۸/۰۳	۳/۲۸	۲/۰۳	۶۱/۷۵	۰/۲۹
گرماب (تعداد: ۴۰۱۸)						زنجان (تعداد: ۴۰۱۸)					
-۲۲/۳	۳۱/۹	۱۲/۶۶	۹/۶۲	۷۶/۰۳	-۰/۲	-۱۵/۷	۳۱/۱	۱۳/۸	۹/۱۲	۷۱/۲۳	-۰/۱۳
۹/۵	۹۹	۴۹/۹۱	۱۸/۸۴	۳۷/۷۴	۰/۲۲	۱۸/۵	۹۹/۵	۵۴/۵۱	۱۵/۰۲	۲۷/۵۶	۰/۱۸
.	۱۱/۲۱	۱/۵	۰/۷۴	۴۸/۹۷	۲/۴۷	-۰/۲۹	۳/۲	۱/۵۲	۰/۵۳	۳۴/۸	۰/۲۹
.	۱۴/۴۴	۸/۲۶	۳/۵۳	۴۲/۶۶	-۰/۳۸	.	۱۳/۷	۷/۹۱	۳/۷۱	۴۶/۸۶	-۰/۶۳
-۰/۲۴	۱۰/۷۸	۳/۵۲	۲/۰۹	۵۹/۵۴	۰/۳۱	-۰/۴۲	۸/۱۹	۳/۴۳	۲/۰۲	۵۸/۹	۰/۲۷
ماه نشان (تعداد: ۴۰۱۸)						گیلوان (تعداد: ۴۰۱۸)					
-۱۷/۳	۳۴/۱	۱۵/۲۴	۹/۸۴	۶۴/۵۵	-۰/۱۵	-۳/۹	۳۵/۴	۱۸/۱	۸/۶۳	۴۷/۶۹	-۰/۰۹
۱۱	۹۶/۵	۴۹/۸	۱۶/۸۵	۳۳/۸۳	۰/۲۴	۱۰/۵	۹۷	۵۶/۲۷	۱۵/۱۳	۲۶/۸۸	۰/۱۸
.	۳/۹۶	۱/۳۷	۰/۵۵	۴۰/۰۸	۰/۶۶	.	۴/۱۶	۱/۳۸	۰/۵۶	۴۰/۷۸	۰/۵
.	۱۳/۴	۷/۴۲	۳/۴۶	۴۶/۵۵	-۰/۴۴	.	۱۴/۵۳	۸/۲۸	۳/۷۱	۴۴/۸۳	-۰/۴۳
-۰/۳۸	۸/۹۶	۳/۵	۲/۱	۵۹/۹۳	۰/۲۸	-۰/۳۵	۱۰/۴۱	۳/۶۳	۲/۱۱	۵۸/۲	۰/۲۵
نیک پی (تعداد: ۳۹۹۲)											
						-۱۸/۹	۳۴	۱۲/۵۶	۹/۶۹	۷۷/۱۳	-۰/۰۸
						۵/۵	۱۰۰	۵۲/۲۹	۱۶/۵	۳۱/۵۵	۰/۲۴
						.	۳/۸۳	۱/۲۳	۰/۶۴	۵۲/۰۷	۰/۶۵
						.	۱۴/۲	۸/۱۵	۲/۹۸	۳۶/۵۹	-۰/۳۳
						-۰/۳۱	۱۰/۹۹	۳/۳۴	۲/۰۶	۶۱/۵۹	۰/۵۴

SD= Standard Deviation
CV= Coefficient of variation



جدول ۳. ویژگی‌های روش ANNs

ایستگاه‌ها	نوع شبکه		تعداد سلول پنهان	تابع لایه خروجی							تابع لایه پنهان			
	MLP	RBF		Logistic	Exponential	Sine	Identity	Tanh	Exponential	Gaussian	Identity	Logistic	Sine	Tanh
آب پر	۴۰	۱۰	۳-۳۰	۶	۸	۹	۲۰	۷	۶	۱۰	۳	۱۰	۱۰	۱۱
ابهر	۴۵	۵	۳-۲۹	۹	۸	۹	۱۷	۷	۱۳	۵	۹	۱۰	۷	۶
پری	۴۷	۳	۳-۳۰	۷	۱۰	۱۲	۱۳	۸	۱۰	۳	۹	۱۴	۶	۸
چاوار زاق	۳۵	۱۵	۳-۲۹	۷	۵	۵	۲۶	۷	۱۲	۱۵	۴	۸	۷	۴
خدابنده	۴۶	۴	۳-۲۸	۶	۱۰	۸	۱۴	۱۲	۱۱	۴	۶	۱۲	۷	۱۰
خیرآباد	۴۸	۲	۳-۲۹	۱۰	۵	۱۵	۱۴	۶	۸	۲	۱	۱۲	۶	۲۱
دانشگاه زنجان	۴۹	۱	۳-۲۱	۱۱	۱۲	۹	۱۴	۴	۱۱	۱	۱	۱۴	۱۰	۱۳
خیله خاصه	۴۳	۷	۳-۲۸	۱۱	۸	۹	۱۳	۹	۱۰	۷	۵	۵	۷	۱۶
زنجان	۴۳	۷	۳-۲۸	۵	۱۲	۴	۱۴	۱۵	۴	۷	۲	۱۳	۸	۱۶
گرماب	۴۵	۵	۳-۳۰	۱۱	۸	۱۰	۱۳	۸	۱۳	۵	۹	۹	۳	۱۱
گیلوان	۳۸	۱۲	۳-۳۰	۶	۷	۷	۲۲	۸	۴	۱۲	۵	۹	۶	۱۴
ماه نشان	۴۳	۷	۳-۲۸	۱۰	۷	۱۰	۱۹	۴	۹	۷	۴	۱۲	۷	۱۱
نیک پی	۴۶	۴	۳-۲۸	۹	۱۴	۹	۱۲	۶	۱۴	۴	۴	۱۵	۵	۸

جدول ۴. مشخصات روش SVM به کار برده شده در این پژوهش

ایستگاه‌ها	γ	nu	C	تعداد بردارهای پشتیبان	خطای اعتبارسنجی ^۱
آب پر	۰/۲۳	۰/۱	۲	۷۲۳ (۶۷۲ bounded)	۰/۰۱۲
ابهر	۰/۲۵	۰/۱	۲	۶۶۱ (۶۱۱ bounded)	۰/۰۱۳
پری	۰/۲۳	۰/۱	۲	۷۲۱ (۶۷۸ bounded)	۰/۰۱۲
چاوار زاق	۰/۲۰	۰/۱	۲	۸۹۱ (۸۴۰ bounded)	۰/۰۲۲
خدابنده	۰/۰۱	۰/۱	۲	۶۴۴ (۶۰۷ bounded)	۰/۲۴۷
خیرآباد	۰/۲۵	۰/۲	۶	۲۸۶ (۲۵۷ bounded)	۰/۰۲۲
دانشگاه زنجان	۰/۲۵	۰/۱	۷	۳۷۰ (۳۲۴ bounded)	۰/۰۲۰
خیله خاصه	۰/۲۵	۰/۱	۲	۷۳۷ (۶۹۰ bounded)	۰/۰۱۴
زنجان	۰/۲۵	۰/۳	۹	۴۷۰ (۳۴۴ bounded)	۰/۰۱۴
گرماب	۰/۲۵	۰/۱	۲	۶۷۰ (۶۴۴ bounded)	۰/۰۲۲
گیلوان	۰/۲۳	۰/۱	۴	۵۴۳ (۴۶۵ bounded)	۰/۰۱۲
ماه نشان	۰/۲۵	۰/۱	۲	۶۹۸ (۶۵۰ bounded)	۰/۰۱۴
نیک پی	۰/۲۵	۰/۱	۸	۳۷۱ (۲۹۷ bounded)	۰/۰۱۳

هرچه مقدار پارامتر nu بیشتر باشد تعداد بردارهای پشتیبان کمتر و دقت تخمین کمتر می‌شود (Vapnik, 2000). در این پژوهش تعداد بردارهای پشتیبانی بین ۸۹۱-۲۸۶، مقدار nu بین ۰/۱-۰/۳ و مقدار C بین ۲-۹ می‌باشد (جدول ۴). Cherkassky and Ma (2004) در تحقیقات خود مقادیر ۰/۲ و ۱۴/۹ به ترتیب برای ϵ و C گزارش کردند. Bayat et al., (2021) در تحقیقات خود مقادیر ۳۶، ۰/۲ و ۱۰ را به ترتیب برای تعداد بردار پشتیبان، ϵ و C گزارش کردند. Taftah et al., (2022) مقادیر ۷۱۷، ۰/۱ و ۲ را به ترتیب برای تعداد بردار پشتیبان، ϵ و C در تحقیقات خود گزارش کردند. صداقت و همکاران (۱۴۰۰) در تحقیقاتی مقادیر پارامتر nu بین ۰/۵-۰/۰۱، پارامتر C بین ۵۰-۱ و پارامتر γ بین ۰/۱۳-۰/۳۵ را گزارش کردند. بنابراین با توجه به پژوهش‌هایی که انجام شده است مقادیر پارامترهای روش SVM در محدوده مناسب قرار دارد، و تخمین ET_0 با دقت مناسبی انجام گرفته است.

ویژگی‌های روش RF

یکی از مهم‌ترین ویژگی‌های روش RF، ساده‌سازی مدل از طریق تعیین اهمیت پارامترها^۱ (تخمین‌گرها) می‌باشد. نتایج اهمیت پارامترها برای ایجاد مدل با روش RF، در جدول ۵ ارائه شده است. این روش پارامترها را براساس سهم پارامتر در دقت تخمین و کاهش خطا، مقدار ۰ تا ۱ می‌دهد (Sedaghat *et al.*, 2022). آنالیز حساسیت، مقدار تغییرات متغیر خروجی را با تغییر هر متغیر ورودی نشان می‌دهد (Minasny and McBratney, 2002). جدول ۵ اهمیت پارامترهای ورودی مدل که از روش RF و آنالیز حساسیت ورودی‌های مدل که از روش ANNs بدست آمده است را نشان می‌دهد. نتایج جدول ۵ نشان داد بیش از ۹۲ درصد (۱۲ ایستگاه) نتایج بدست آمده از هر دو روش (RF و ANNs)، میانگین دما مهم‌ترین و موثرترین پارامتر در تخمین ET₀ می‌باشد. همچنین بیش از ۸۴ درصد (۱۱ ایستگاه) نتایج هر دو روش نشان داد که ساعت آفتابی دومین ورودی مهم و موثر در در تخمین ET₀ می‌باشد. سرعت باد و میانگین رطوبت در سطح اهمیت سوم و چهارم قرار دارند. نتایج این پژوهش مشابه نتایج تحقیقات Shi (2019) و Ning (2017) می‌باشد. همچنین، نتایج این پژوهش مخالف نتایج تحقیقات Nie *et al.*, (2022) می‌باشد، علت این تفاوت احتمالاً به دلیل شرایط جغرافیایی مناطق مورد مطالعه می‌باشد، که میانگین ارتفاع در تحقیقات Nie *et al.*, (2022) حدود ۴۸۱ متر و دارای آب و هوای معتدل قاره‌ای موسمی می‌باشد اما میانگین ارتفاع ایستگاه‌های مورد استفاده در این پژوهش حدود ۱۳۲۷ متر و دارای آب و هوای نیمه خشک فراسرد می‌باشد.

جدول ۵. فاکتور اهمیت پارامترها و آنالیز حساسیت ورودی‌های مدل

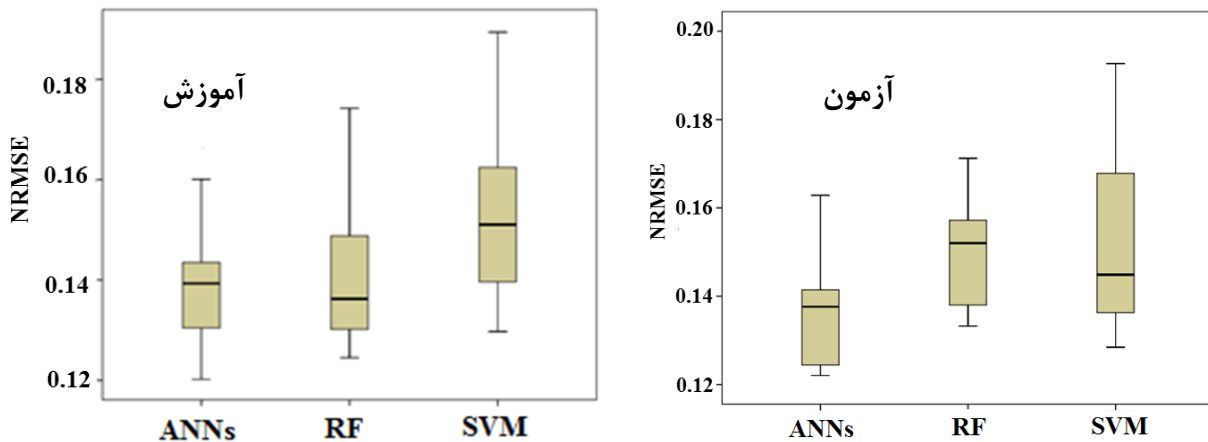
ایستگاه‌ها	ANNs (آنالیز حساسیت)			RF (اهمیت پارامتر)		
	میانگین دما	ساعت آفتابی	سرعت باد	میانگین دما	ساعت آفتابی	سرعت باد
چوار زاق	۶/۵۹	۱/۷۳	۱/۴۹	۱	۰/۴۵	۰/۳۱
آب پر	۷/۳۳	۲/۰۲	۱/۸۲	۱	۰/۵۵	۰/۳۴
ابه‌ر	۱۰/۹۸	۲/۳۸	۱/۷۷	۱	۰/۶۱	۰/۵۴
خداینده	۱۰/۴۳	۲/۸۹	۱/۷۱	۱	۰/۶۳	۰/۶۱
خیرآباد	۶/۵۳	۱/۲۳	۱/۷۵	۱	۰/۴۹	۰/۴۹
خیله خاصه	۶/۷۰	۳/۳۳	۱/۷۳	۱	۰/۹۸	۰/۴۶
دانشگاه زنجان	۷/۰۲	۴/۲۶	۲۱/۹۹	۱	۰/۸۰	۰/۶۵
زنجان	۹/۱۸	۳/۱۸	۱/۴۹	۱	۰/۶۷	۰/۶۱
گرماب	۹/۲۷	۱/۸۸	۲/۱۶	۱	۰/۶۶	۰/۶۵
گیلوان	۸/۰۹	۲/۱۹	۱/۷۷	۱	۰/۵۱	۰/۳۷
ماه نشان	۹/۶۰	۲/۱۱	۲/۰۶	۱	۰/۵۳	۰/۵۲
نیک پی	۶/۳۹	۲/۰۶	۲/۲۵	۱	۰/۷۴	۰/۶۳
پری	۱۰/۳۱	۲/۶۶	۱/۳۴	۱	۰/۷۵	۰/۷۴

علت اهمیت بیشتر میانگین دما نسبت به سایر پارامترهای ورودی احتمالاً به دلیل است که ارتفاع استان زنجان از سطح دریای زیاد می‌باشد (جدول ۱)، با افزایش ارتفاع سبب کاهش دما و به تبع آن کاهش مقدار ET₀ و افزایش رطوبت نسبی می‌شود (رضیعی و همکاران، ۱۳۸۴). Liang *et al.*, (2010) گزارش کردند که پارامترهای اقلیمی نظیر میانگین و حداکثر دمای هوا، رطوبت نسبی و ساعت آفتابی از مهم‌ترین متغیرهای اقلیمی مؤثر بر مقدار ET₀ بوده و با داشتن این متغیرها، مقدار ET₀ برای دوره‌های آینده قابل محاسبه و پیش‌بینی خواهد بود.

مقایسه میانگین نتایج بین ایستگاه‌های مختلف بر اساس آماره NRMSE با سه روش در شکل ۲ نشان داده شده است. با توجه به جدول ۶ و شکل ۲، نتایج مقایسه سه روش در برآورد مقدار ET₀ مشاهده می‌شود که در مرحله آموزش به‌طور نسبی دقت دو روش RF و ANNs نزدیک به هم و در مرحله آزمون دقت روش ANNs نسبت به دو روش دیگر بهتر بوده است، اما به‌طور کلی روش ANNs نسبت به دو روش دیگر عملکرد بالاتری داشته است، که با نتایج تحقیقات Ayaz *et al.*, (2021) مطابقت دارد. اما با انجام مقایسه میانگین از آماره NRMSE مشاهده شده دقت در سه روش تفاوت معنی‌دار باهم نداشتند (نتایج مقایسه میانگین نشان داده نشده است). مقایسه میانگین نتایج حاصل از سه روش براساس آماره NRMSE در ۱۳ ایستگاه هواشناسی استان زنجان در هر دو مرحله آموزش و آزمون در

1 . Important parameter

شکل ۲ ارائه شده است. با توجه به شکل ۲ مشاهده می‌شود در مرحله آزمون میانگین خطای روش ANNs از دو روش دیگر کمتر و میانگین خطای روش SVM نسبت به دو روش دیگر بیشتر می‌باشد.

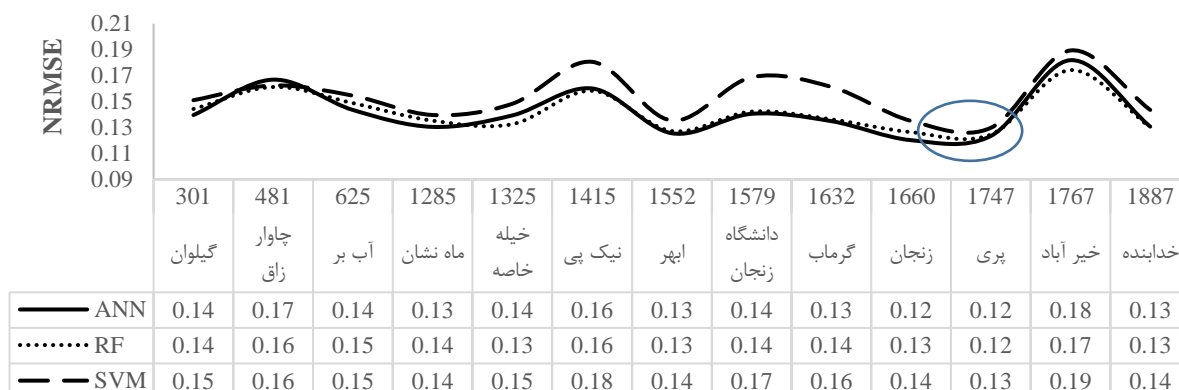


شکل ۲. مقایسه میانگین نتایج سه روش بکار برده شده براساس آماره NRMSE

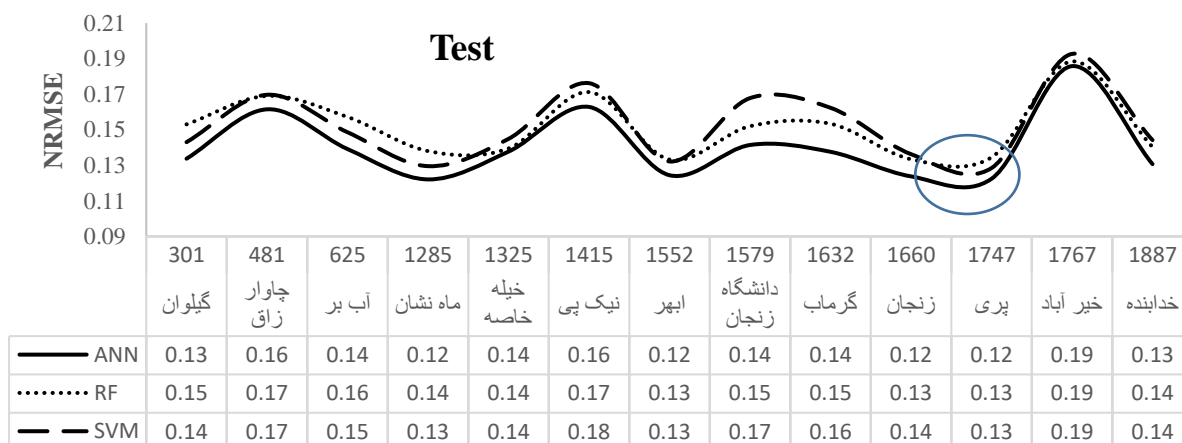
نتایج نشان می‌دهد که دقت سه روش به کار رفته در این پژوهش در پری و زنگان نسبت به سایر مناطق استان بیشتر و میانگین خطای روش‌های بکار رفته در این دو ایستگاه نسبت به بقیه کمتر می‌باشد (شکل ۳). اما میانگین خطای روش‌های بکار رفته در خیرآباد نسبت به سایر ایستگاه‌ها بیشتر و در نتیجه دقت روش‌های بکار رفته در این ایستگاه پایین‌تر است. با توجه به شکل ۳ مشاهده می‌شود که با افزایش ارتفاع از سطح دریا، به‌طور نسبی میانگین خطا در تخمین مقدار ET_0 افزایش می‌یابد. شرایط منطقه‌ای پری و زنگان به‌صورت جلگه‌ای (دشت) می‌باشد، که در نتیجه در دشت‌ها با افزایش تابش خورشید میزان دما افزایش یافته و بر میزان ET_0 اثر خواهد گذاشت، که این موضوع با پژوهش‌های گودرزی و همکاران (۱۳۹۷) و مهرآذر و همکاران (۱۳۹۷) مطابقت دارد. همانطور که در شکل ۱ و ۳ مشاهده می‌شود پری و زنگان در شمال ولی ایستگاه خیرآباد در جنوب شرقی استان زنگان قرار دارند، در این مناطق با وجود اینکه ارتفاع از سطح دریا آن‌ها بالا است، اما به نظر می‌رسد ارتفاع تاثیر چندانی در دقت تخمین ET_0 نداشته است. با توجه به جدول ۲ مشاهده می‌شود بیشترین ضریب تغییرات عامل‌های هواشناسی مانند میانگین دما، میانگین رطوبت نسبی و ساعت آفتابی مربوط به ایستگاه خیرآباد است که نشان‌دهنده وجود تغییرات شدید پارامترهای هواشناسی در منطقه‌ای که این ایستگاه قرار دارد، می‌باشد که احتمالاً به این دلیل دقت تخمین ET_0 در ایستگاه خیرآباد پایین و بیشترین مقدار NRMSE را دارد (شکل ۳). این نتایج با تحقیقات Adnan et al., (2020) مطابقت دارد، که گزارش کردند زمانی که دما، تابش خورشیدی و سرعت باد افزایش می‌یابد و فشار بخار کاهش می‌یابد، حداکثر افزایش در ET_0 مشاهده شد.

در زنگان، در روش ANNs مقادیر آماره‌های RMSE و NRMSE به‌ترتیب برای آموزش برابر با ۰/۴۱ و ۰/۱۲، و برای آزمون برابر با ۰/۴۲ و ۰/۱۲ می‌باشد. در روش RF مقادیر آماره‌های RMSE و NRMSE به‌ترتیب برای آموزش برابر با ۰/۴۳ و ۰/۱۲، و برای آزمون برابر با ۰/۴۵ و ۰/۱۲ می‌باشد. در روش SVM مقادیر این آماره‌ها به‌ترتیب برای آموزش برابر با ۰/۴۶، ۰/۹۵ و ۰/۱۳، و برای آزمون برابر با ۰/۴۷ و ۰/۱۳ می‌باشد (جدول ۶). در ایستگاه پری، در روش ANNs مقادیر آماره‌های RMSE و NRMSE به‌ترتیب برای آموزش برابر با ۰/۴۵ و ۰/۱۲، و برای آزمون برابر با ۰/۴۵ و ۰/۱۲ می‌باشد. در روش RF مقادیر این آماره‌ها به‌ترتیب برای آموزش برابر با ۰/۴۶ و ۰/۱۳، و برای آزمون برابر با ۰/۴۸ و ۰/۱۳ می‌باشد. در روش SVM هم مقادیر این آماره‌ها به‌ترتیب برای آموزش برابر با ۰/۴۸ و ۰/۱۳، و برای آزمون برابر با ۰/۴۸ و ۰/۱۳ می‌باشد (جدول ۶). در ایستگاه خیرآباد، در روش ANNs مقادیر آماره‌های RMSE و NRMSE به‌ترتیب برای آموزش برابر با ۰/۴۸ و ۰/۱۳، و برای آزمون برابر با ۰/۴۸ و ۰/۱۳ می‌باشد. در روش RF مقادیر این آماره‌ها به‌ترتیب برای آموزش برابر با ۰/۴۸ و ۰/۱۳، و برای آزمون برابر با ۰/۴۸ و ۰/۱۳ می‌باشد. در روش SVM هم مقادیر این آماره‌ها به‌ترتیب برای آموزش برابر با ۰/۴۸ و ۰/۱۳، و برای آزمون برابر با ۰/۴۸ و ۰/۱۳ می‌باشد (جدول ۶).

Train



Test



ارتفاع (متر)

شکل ۳. مقایسه NRMSE سه روش در برآورد ET_0

همانطور که گفته شد با افزایش ارتفاع، میانگین خطا به طور قابل توجهی افزایش می‌یابد (شکل ۳ و جدول ۶). تغییرات تشعشع خورشیدی به‌عنوان تابعی از عرض جغرافیایی، در فصول مختلف سال متفاوت است. به طوری که در تمامی فصول با افزایش عرض جغرافیایی از میزان تشعشع ورودی به زمین کاسته می‌شود. بنابراین انتظار می‌رود که مقدار ET_0 در عرض‌های جغرافیایی بالاتر، کمتر از عرض‌های جغرافیایی پایین‌تر به دست آید، از جمله دلایل این امر می‌تواند مرتبط با در نظر گرفتن عوامل فیزیوگرافی از قبیل عرض جغرافیایی و ارتفاع از سطح دریا در محاسبات مقدار ET_0 باشد، که این عوامل از کنترل‌کننده‌های مؤثر مقدار ET_0 در هر منطقه می‌باشد (پورمحمدی و همکاران، ۱۳۸۹). بنابراین، این متغیرها نیز به شدت تحت تأثیر شرایط منطقه‌ای مانند توپوگرافی و ارتفاع قرار می‌گیرند (Alexandris and Proutsos, 2020). نظر به اینکه ارتفاع نقش مهمی در تعدیل تغییرات اقلیمی در مناطق کوهستانی دارد (Liu et al., 2021; Yang et al., 2019)، اما نتایج پژوهش حاضر نشان می‌دهند که اثر ارتفاع بر تغییرات مقدار ET_0 نامشخص است. در این پژوهش میانگین پارمترهای سرعت باد، ساعت آفتابی، میانگین دما و رطوبت هر ایستگاه و همچنین درصد شیب در هر ایستگاه با مقادیر NRMSE بدست آمده از سه روش بکار رفته در این پژوهش بررسی گردید (نتایج نشان داده نشده است). در این بررسی مشاهده شد که مقادیر خطا با هیچ کدام از پارامترها (سرعت باد، ساعت آفتابی، میانگین دما، رطوبت و درصد شیب) روند افزایش یا کاهش منظمی نداشته است و تغییرات در ایستگاه‌های مختلف متغیر بوده است.

جدول ۶. نتایج روش‌های مختلف در برآورد مقدار ET_o در ایستگاه‌های هواشناسی

		Train				Test			
		RMSE	MBE	NRMSE	EF	RMSE	MBE	NRMSE	EF
آب بر	ANN	۰/۵۲	۰/۰۰۹-	۰/۱۴	۰/۹۴	۰/۵۱	۰/۰۲۵-	۰/۱۴	۰/۹۴
	RF	۰/۵۴	۰/۰۰۱	۰/۱۵	۰/۹۳	۰/۵۸	۰/۰۱۴-	۰/۱۶	۰/۹۳
	SVM	۰/۵۶	۰/۰۴۳	۰/۱۵	۰/۹۳	۰/۵۵	۰/۰۲۷	۰/۱۵	۰/۹۳
چاوار زاق	ANN	۰/۶۰	۰/۰۰۳-	۰/۱۷	۰/۹۱	۰/۵۸	۰/۰۲۳-	۰/۱۶	۰/۹۱
	RF	۰/۵۸	۰/۰۰۳-	۰/۱۶	۰/۹۲	۰/۶۱	۰/۰۲۳-	۰/۱۷	۰/۹۱
	SVM	۰/۵۹	۰/۰۰۲-	۰/۱۶	۰/۹۱	۰/۶۱	۰/۰۲۳-	۰/۱۷	۰/۹۰
گیلوان	ANN	۰/۵۰	۰/۰۱۵	۰/۱۴	۰/۹۴	۰/۴۹	۰/۰۲۳	۰/۱۳	۰/۹۴
	RF	۰/۵۲	۰/۰۰۰	۰/۱۴	۰/۹۴	۰/۵۶	۰/۰۱۹	۰/۱۵	۰/۹۳
	SVM	۰/۵۴	۰/۰۴۳	۰/۱۵	۰/۹۳	۰/۵۲	۰/۰۴۶	۰/۱۴	۰/۹۴
ماه نشان	ANN	۰/۴۵	۰/۰۰۴	۰/۱۳	۰/۹۹	۰/۴۳	۰/۰۱۱-	۰/۱۲	۰/۹۵
	RF	۰/۴۷	۰/۰۰۵-	۰/۱۳	۰/۹۹	۰/۴۸	۰/۰۱۳-	۰/۱۴	۰/۹۴
	SVM	۰/۴۹	۰/۰۱۱	۰/۱۴	۰/۹۴	۰/۴۵	۰/۰۰۳	۰/۱۳	۰/۹۵
خیله خاصه	ANN	۰/۴۵	۰/۰۲۵	۰/۱۴	۰/۹۵	۰/۴۶	۰/۰۱۲-	۰/۱۴	۰/۹۵
	RF	۰/۴۳	۰/۰۰۰	۰/۱۳	۰/۹۵	۰/۴۷	۰/۰۲۸-	۰/۱۴	۰/۹۴
	SVM	۰/۴۸	۰/۰۶۶	۰/۱۵	۰/۹۴	۰/۴۹	۰/۰۲۵	۰/۱۴	۰/۹۴
نیک پی	ANN	۰/۵۳	۰/۰۰۱-	۰/۱۶	۰/۹۳	۰/۵۵	۰/۰۲۷-	۰/۱۶	۰/۹۳
	RF	۰/۵۳	۰/۰۰۳-	۰/۱۶	۰/۹۳	۰/۵۸	۰/۰۳۴-	۰/۱۷	۰/۹۲
	SVM	۰/۶۰	۰/۰۱۷-	۰/۱۸	۰/۹۱	۰/۵۹	۰/۰۳۶-	۰/۱۸	۰/۹۲
ابه‌ر	ANN	۰/۴۳	۰/۰۰۵	۰/۱۳	۰/۹۵	۰/۴۳	۰/۰۱۰-	۰/۱۲	۰/۹۵
	RF	۰/۴۳	۰/۰۰۴-	۰/۱۳	۰/۹۵	۰/۴۶	۰/۰۲۰-	۰/۱۳	۰/۹۵
	SVM	۰/۴۶	۰/۰۰۳	۰/۱۳	۰/۹۵	۰/۴۶	۰/۰۱۶-	۰/۱۳	۰/۹۵
دانشگاه زنجان	ANN	۰/۴۸	۰/۰۰۳-	۰/۱۴	۰/۹۴	۰/۴۹	۰/۰۱۹-	۰/۱۴	۰/۹۴
	RF	۰/۴۸	۰/۰۰۲	۰/۱۴	۰/۹۴	۰/۵۲	۰/۰۱۵-	۰/۱۵	۰/۹۳
	SVM	۰/۵۷	۰/۰۸۵	۰/۱۷	۰/۹۱	۰/۵۸	۰/۰۵۹	۰/۱۷	۰/۹۱
گرماب	ANN	۰/۴۷	۰/۰۰۳	۰/۱۳	۰/۹۴	۰/۴۸	۰/۰۲۳	۰/۱۴	۰/۹۵
	RF	۰/۴۸	۰/۰۰۲-	۰/۱۴	۰/۹۵	۰/۵۴	۰/۰۰۸	۰/۱۵	۰/۹۳
	SVM	۰/۵۷	۰/۰۳۶	۰/۱۶	۰/۹۳	۰/۵۷	۰/۰۶۷	۰/۱۶	۰/۹۲
زنجان	ANN	۰/۴۱	۰/۰۰۶-	۰/۱۲	۰/۹۶	۰/۴۲	۰/۰۰۶	۰/۱۲	۰/۹۶
	RF	۰/۴۳	۰/۰۰۱-	۰/۱۳	۰/۹۵	۰/۴۶	۰/۰۰۰	۰/۱۳	۰/۹۵
	SVM	۰/۴۶	۰/۰۱۵-	۰/۱۳	۰/۹۵	۰/۴۷	۰/۰۰۴-	۰/۱۴	۰/۹۵
پری	ANN	۰/۴۵	۰/۰۰۰	۰/۱۲	۰/۹۶	۰/۴۶	۰/۰۰۴-	۰/۱۲	۰/۹۶
	RF	۰/۴۶	۰/۰۰۰	۰/۱۲	۰/۹۶	۰/۵۰	۰/۰۱۱-	۰/۱۳	۰/۹۵
	SVM	۰/۴۸	۰/۰۰۳-	۰/۱۳	۰/۹۵	۰/۴۸	۰/۰۰۶-	۰/۱۳	۰/۹۶
خیر آباد	ANN	۰/۶۰	۰/۰۰۱-	۰/۱۸	۰/۹۱	۰/۶۰	۰/۰۱۳	۰/۱۹	۰/۹۱
	RF	۰/۵۸	۰/۰۰۳-	۰/۱۷	۰/۹۲	۰/۶۱	۰/۰۰۳-	۰/۱۹	۰/۹۱
	SVM	۰/۶۲	۰/۰۲۶-	۰/۱۹	۰/۹۰	۰/۶۳	۰/۰۱۷-	۰/۱۹	۰/۹۰
خدابنده	ANN	۰/۴۳	۰/۰۰۶	۰/۱۳	۰/۹۶	۰/۴۴	۰/۰۰۴	۰/۱۳	۰/۹۵
	RF	۰/۴۳	۰/۰۰۵-	۰/۱۳	۰/۹۶	۰/۴۷	۰/۰۰۷-	۰/۱۴	۰/۹۵
	SVM	۰/۴۷	۰/۰۱۷	۰/۱۴	۰/۹۵	۰/۴۹	۰/۰۱۶	۰/۱۴	۰/۹۵

با توجه به جدول ۶، مقادیر آماره MBE در مرحله آموزش و آزمون اکثراً مقادیر کوچک و نزدیک به صفر می‌باشد. بنابراین با توجه به کم بودن مقادیر آماره MBE، نمی‌توان گفت مدل کم‌برآوردی یا بیش برآوردی قابل‌توجهی داشته است و در نتیجه می‌توان از آن چشم‌پوشی کرد (Singh et al., 2020). در جدول ۶ مشاهده می‌شود آماره EF در همه ایستگاه‌ها در هر دو مرحله آموزش و آزمون بالای ۰/۹۰ می‌باشد که نشان از صحت و کارایی بالای روش‌های استفاده در این پژوهش در تخمین مقدار تبخیر تفرق مرجع می‌باشد. مقادیر

این آمارها نشان می‌دهد انتخاب عوامل موثر در تخمین مقدار ET_0 در این تحقیق مناسب بوده و مقدار ET_0 برآورد شد قابل اعتماد است. با توجه به نتایج پژوهش حاضر می‌توان گفت با توسعه روش‌های یادگیری ماشین می‌توان با تعداد کمی از داده‌های هواشناسی، ET_0 را با دقت بالا تخمین زد که با نتایج محققانی مانند (Feng et al., 2017) و (Shiri 2017) مطابقت دارد.

نتیجه‌گیری

در این مطالعه سعی شد بهترین روش برای تخمین ET_0 در صورت عدم وجود همه داده‌های آب و هوایی که برای مدل FPM نیاز هست، در شرایط آب و هوایی سرد و خشک استان زنجان، ارائه شود. در پژوهش حاضر، با مقایسه مقادیر پیش‌بینی شده توسط روش‌های مورد مطالعه (ANNs، RF و SVM) و مقادیر حاصله از مدل فائو-پنمن - مانیتیت توسط پارامترهای آماری، مشخص شد که روش ANNs با داشتن کمترین مقدار خطا در مقایسه با روش‌های RF و SVM، دقت به مراتب بالاتری را از خود نشان داده است. به عبارت دیگر، می‌توان با استفاده از پارامترهای هواشناسی میانگین دمای هوای روزانه، میانگین رطوبت نسبی، سرعت باد و ساعت آفتابی و با استفاده از روش ANNs، مقدار ET_0 را با دقت بالایی پیش‌بینی نمود. نتایج نشان داد که شبکه عصبی MLP نسبت به RBF در برآورد مقدار ET_0 از عملکرد بهتری برخوردار است و Identity و Logistic و به ترتیب تابع فعال‌سازی غالب در لایه پنهان و خروجی شبکه عصبی مصنوعی بوده است. نتایج آنالیز حساسیت روش ANNs و فاکتور اهمیت پارامتر روش RF، نشان داد که میانگین دما مهمترین و موثرترین پارامتر در تخمین ET_0 می‌باشد. نتایج پژوهش حاضر نشان داد که دقت سه روش به کار رفته در این پژوهش در ایستگاه‌های پری و زنجان نسبت به سایر مناطق استان بیشتر است که احتمالاً به دلیل شرایط جلگه‌ای بودن این مناطق و جهت شیب این مناطق به سمت شمال باشد. با توجه به نتایج این پژوهش مقادیر NRMSE بدست آمده از تخمین ET_0 با میانگین هیچ کدام از پارامترها (میانگین دما، میانگین رطوبت، سرعت باد، ساعت آفتابی و درصد شیب) تغییرات منظمی نداشتند. بنابراین، با توجه به نتایج این پژوهش، حتی اگر تمام اطلاعات پارامترها در یک ایستگاه خاص در دسترس نباشد، می‌توان از چهار پارامتر هواشناسی که ترکیبی از میانگین دمای هوای روزانه، میانگین رطوبت نسبی، سرعت باد و ساعت آفتابی می‌باشد، استفاده کرد و ET_0 را با دقت قابل قبول تخمین زد. در نتیجه می‌توان گفت ET_0 نه تنها تحت تأثیر عوامل مختلف هواشناسی قرار می‌گیرد، بلکه ممکن است تحت تأثیر برهمکنش بین عوامل مختلف هواشناسی هم تغییر کند. علاوه بر این، ET_0 ممکن است تحت تأثیر فعالیت‌های انسانی از جمله آلودگی هوا هم قرار گیرد. بنابراین در آینده تحقیق بیشتر برای تجزیه و تحلیل سهم تغییرات ET_0 در استان زنجان و همچنین در سایر مناطق ایران مورد نیاز خواهد بود.

"هیچ‌گونه تعارض منافع بین نویسندگان وجود ندارد"

منابع

- بیدآبادی، م، بابازاده، ح، شیرینی، ج، و صارمی، ع. (۱۴۰۰). برآورد تبخیر و تعرق گیاه مرجع با استفاده از ANN و ANFIS در اقلیم نیمه خشک و خشک، آبیاری و زهکشی/ایران، ۱۵ (۶): ۱۴۲۰-۱۴۱۲.
- پورمحمدی، سمانه؛ ملکی نژاد، حسین؛ و رحیمیان، محمدحسن (۱۳۸۹). بررسی نقش عوامل فیزیوگرافی بر روی پارامترهای دمایی مؤثر بر تبخیر و تعرق (مطالعه موردی: استان یزد). خشکبوم، ۱ (۲): ۹-۱۹.
- ذرتی پور، ا؛ نیسی، ل؛ گلابی، م؛ بزاز، ا؛ و ذرتی پور، ا. (۱۳۹۸). شبیه‌سازی و مقایسه ی تبخیر و تعرق پتانسیل به روش‌های شبکه عصبی مصنوعی، نروفازی و درخت تصمیم‌گیری M5 (مطالعه موردی: ایستگاه سینوپتیک شیراز). تحقیقات منابع آب ایران، ۱۵ (۱): ۳۷۱-۳۶۵.
- رضیعی، طیب؛ دانش کار آراسته، پیمان؛ و ثقفیان، بهرام (۱۳۸۴). بررسی روند بارندگی سالانه در مناطق خشک و نیمه خشک مرکزی و شرقی ایران. مجله آب و فاضلاب، ۱۶ (۲): ۷۳-۸۱.
- سپهری، س، عباسی، ف، زارعی، ق. و نخجوانی مقدم، م. م. (۱۳۹۹). بررسی و تحلیل حساسیت روش‌های مبتنی بر شبکه‌های عصبی مصنوعی در برآورد تبخیر و تعرق گیاه مرجع آبیاری و زهکشی/ایران، ۱۴ (۶): ۲۰۸۹-۲۰۹۹.
- صدافت، ا؛ شعبانپور شهرستانی، م؛ نوروزی، ع.؛ فلاح نصرت آباد، ع.ر؛ و بیات، ح. (۱۴۰۰). استفاده از شاخص‌های طیفی در برآورد رطوبت سطحی خاک براساس الگوریتم یادگیری ماشین. تحقیقات آب و خاک ایران، ۵۲ (۱۲): ۳۰۱۸-۳۰۰۱.
- کریمی پور، احمدرضا؛ و بنی طالبی، گلنوش (۱۳۹۹). تحلیل حساسیت داده‌های هواشناسی در برآورد تبخیر و تعرق مرجع با حداقل داده‌های هواشناسی



با استفاده از مدل‌های موجک-عصبی-فازی، ANN و ANFIS. نشریه حفاظت منابع آب و خاک (علمی-پژوهشی)، ۹(۳)، ۴۷-۷۲.
گودرزی، مسعود؛ صلاحی، برومند؛ و حسینی، سید اسعد (۱۳۹۷). برآورد میزان تبخیر و تعرق در ارتباط با تغییرات اقلیمی در حوزه ی آبخیز دریاچه ارومیه. مجله علوم ومهندسی آبخیزداری ایران، ۱۲(۴۱)، ۱-۱۲.
مهرآذر، آیدا؛ مساح بوانی، علیرضا؛ مشعل، محمود؛ و رحیمی خوب، حدیثه (۱۳۹۷). بررسی اثرات تغییر اقلیم بر بخش کشاورزی دشت هشتگرد با تاکید بر عدم قطعیت مدل های AOGCM گزارش ارزیابی پنجم IPCC. نشریه علوم و مهندسی آبیاری، ۴۱(۳)، ۴۵-۵۹.

REFERENCES

- Achirul Nanda, M., Boro Seminar, K., Nandika, D., and Maddu, A. (2018). A comparison study of kernel functions in the support vector machine and its application for termite detection. *Information*, 9(1), 5.
- Adab, H., Morbidelli, R., Saltalippi, C., Moradian M., and Ghalhari, G. A. F. (2020). Machine learning to estimate surface soil moisture from remote sensing data. *Water*, 12 (11), 3223.
- Adnan, S., Ullah, K., and Ahmed, R. (2020). Variability in meteorological parameters and their impact on evapotranspiration in a humid zone of Pakistan. *Meteorological Applications*, 27(1), e1859.
- Alexandris, S., and Proutsos., N. (2020). How significant is the effect of the surface characteristics on the Reference Evapotranspiration estimates. *Agric. Water Manag.*, 237, 106181.
- Algretawee, H., and Alshama, G. (2021). Modeling of Evapotranspiration (ET_o) in a Medium Urban Park within a Megacity by Using Artificial Neural Network (ANN) Model. *Periodica Polytechnica Civil Engineering*. 65(4): 1260–1268.
- Allen, R. G., Pereira, L. S., Howell, T. A., and Jensen, M. E. (2011). Evapotranspiration information reporting: I. Factors governing measurement accuracy. *Agricultural Water Management*, 98(6), 899-920.
- Allen, R. G., Pereira, L.S., Raes, D., and Smith, M. (1998). Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56. *Fao, Rome*, 300(9), D05109.
- Aslami, F., Ghorbani, A., Sobhani, B., and Panahandeh, M. (2015). Comparing artificial neural network, support vector machine and object-based methods in preparation land use/cover maps using landSat-8 images. *Journal of RS and GIS for Natural Resources*, 6(3), 1-14.
- Ayaz, A., Rajesh, M., Singh, S. K., and Rehana, S. (2021). Estimation of reference evapotranspiration using machine learning models with limited data. *AIMS Geosciences*, 7(3), 268-290.
- Bandyopadhyay, A., Bhadra, A., Raghuwanshi, N. S., and Singh, R. (2009). Temporal trends in estimates of reference evapotranspiration over India. *Journal of Hydrologic Engineering*, 14(5), 508-515.
- Bayat, H., Ebrahimzadeh, G., and Mohanty, B.P. (2021) Investigating the capability of estimating soil thermal conductivity using topographical attributes for the Southern Great Plains, USA. *Soil and Tillage Research*, 206, 104811.
- Berry, W.D.(1993). *Understanding Regression Assumptions*. Sage Publications, London.
- Bidabadi, M., Babazadeh, H., Shiri, J., and Saremi, A. (2022). Estimation of Reference Crop Evapotranspiration Using ANN and ANFIS in Semi-Arid and Dry Climates. *Iranian Journal of Irrigation & Drainage*, 15(6), 1412-1420.(*in Persian*).
- Boateng, E. Y., Otoo, J., and Abaye, D. A. (2020). Basic tenets of classification algorithms K-nearest-neighbor, support vector machine, random forest and neural network: a review. *Journal of Data Analysis and Information Processing*, 8(4), 341-357.
- Boser, B.E., Guyon, I.M., and Vapnik, V.N. (1992). A training algorithm for optimal margin classifiers. In D.Haussler, editor, 5th Annual ACM Workshop on COLT, pages 144-152, Pittsburgh, PA.
- Breiman, L. (2001). Random forests. *Machine Learn.* 45: 5–32.
- Cherkassky V, Ma Y (2004) Practical selection of SVM parameters and noise estimation for SVM regression. *Neural networks* 17, 113-126.
- Dixon, B., and Candade, N. (2008). Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?. *International Journal of Remote Sensing*, 29(4), 1185-1206.
- Fan, J., Ma, X., Wu, L., Zhang, F., Yu, X., and Zeng, W. (2019). Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agric Water Manag.*, 225,105758.
- Feng, K., and Tian, J. (2021). Forecasting reference evapotranspiration using data mining and limited climatic data. *European Journal of Remote Sensing*, 54(sup2), 363-371
- Feng, Y., Peng, Y., Cui, N., Gong, D., and Zhang, K. (2017). Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Comput Electron Agric*, 136, 71–78.

- Feng, Y., Peng, Y., Cui, N., Gong, D., and Zhang, K. (2017). Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Comput. Electron. Agric.* 136, 71–78.
- Ferreira, L.B., França, F., Oliveira, R.A., De, I.E., and Filho, F. (2019). Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM ; a new approach. *J Hydrol.*
- Gill, M., Kemblowski, M.W., and McKee, M. (2007). Soil moisture data assimilation using support vector machines and ensemble Kalman filter 1. *JAWRA. Journal of the American Water Resources Association*, 43(4), 1004-1015.
- Gopinathan, K.K. (1988). A general formula for computing the coefficients of the correlation connecting global solar radiation to sunshine duration. *Solar energy*, 41, 499-502
- Goudarzi, M., Salahi, B., and Hosseini, S. A. (2018). Estimation of evapotranspiration rate due to climate change in the Urmia Lake Basin. *Iranian Journal of Watershed Management Science and Engineering*, 12(41), 1-12. (in Persian).
- Hagan, M.T., Demuth, H.B., and Beale, M.H. (1996). *Neural Network*. Design PWS Publishing Co.
- Hocking, R.R. (2013). *Methods and Applications of Linear Models: Regression and The Analysis of Variance*. John Wiley & Sons.
- Karimipour, A., and Banitalebi, G. (2020). Sensitivity analysis of meteorological data in estimating reference evapotranspiration with the minimum data using wavelet-neuro-fuzzy, ANN and ANFIS models. *Journal of Soil and Water Resources Conservation*, 9(3), 47-72. (in Persian).
- Kisi, O., and Alizamir, M., (2018). Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: wavelet extreme learning machine vs wavelet neural networks. *Agric. For. Meteorol.* 263, 41–48.
- Kotsiantis, S., and Pintelas, P. (2004). Combining bagging and boosting. *Journal of Computational Intelligence*. 1(4): 324–33.
- Kulkarni, V.Y., and Sinha P.K. (2014). Effective learning and classification using random forest algorithm. *International Journal of Engineering and Innovative Technology (IJEIT)*, 3, p. 267-273.
- Liang, L., Lijuan, L., and Qiang, L. (2010). Temporal variation of reference evapotranspiration during 1961-2005 in the Taoer river basin of Northeast China. *Agricultural and Forest Meteorology*, 150, 298-306.
- Lin G., Chen G., Huang P., and Chou Y. (2009). Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods. *Journal of hydrology*. Vol. 372, 17–29.
- Liu, W., Yang, L., Zhu, M., Adamowski, J. F., Barzegar, R., Wen, X., and Yin, Z. (2021). Effect of elevation on variation in reference evapotranspiration under climate change in northwest china. *Sustainability*, 13(18), 10151.
- Mattar, M.A., (2018). Using gene expression programming in monthly reference evapotranspiration modeling: a case study in Egypt. *Agric. Water Manage.* 198, 28–38.
- Mehdizadeh, S. (2018). Estimation of daily reference evapotranspiration (ET_o) using artificial intelligence methods: offering a new approach for lagged ET_o data-based modeling. *J. Hydrol.* 559, 794–812.
- Mehrazar, A., Massah Bavani, A., Mashal, M., and Rahimikhoob, H. (2018). Assessment of climate change impacts on agriculture of the Hashtgerd Plain with emphasis of AR5 models uncertainty. *Irrigation Sciences and Engineering*, 41(3), 45-59. (in Persian).
- Minasny, B., and McBratney, A. B. (2002). The neuro-m method for fitting neural network parametric pedotransfer functions. *Soil Sci. Soc. Am. J.* 66 (2), 352– 361.
- Nie, T., Yuan, R., Liao, S., Zhang, Z., Gong, Z., Zhao, X., and Jiang, H. (2022). Characteristics of Potential Evapotranspiration Changes and Its Climatic Causes in Heilongjiang Province from 1960 to 2019. *Agriculture*, 12(12), 2017.
- Ning, T.T. (2017). *Spatial-Temporal Variation of Evapotranspiration in the Loess Plateau under Budyko Framework and Its Attribution Analysis*; Research Center for Soil and Water Conservation and Eco-Environment of the Ministry of Education; Chinese Academy of Sciences: Beijing, China.
- Pal, M. (2006). M5 model tree for land cover classification. *International Journal of Remote Sensing*, 27(4), 825-831
- Panaitescu, L., Ilie, C., Lungu, M. L., Popescu, M., Lungu, D., and Nita, S. (2014). Modern approach to the phenomenon of drought and aridity in Central and South Dobrudja. *Journal of Environmental Protection and Ecology*, 15(1), 110-122.
- Petropoulos, G. P., Ireland, G., and Barrett, B. (2015). Surface Soil Moisture Retrievals from Remote Sensing: Current Status, Products & Future Trends, *Physics and Chemistry of the Earth*, 83 (84), 36-56.



- Picton, P. (2000). *Neural Networks*, 2nd edn. Palgrave, New York.
- Poormohammadi, S., Malekinezhad, H., and Rahimian, M. H. (2010). Investigating the role of physiographical factors on temperature-related parameters affecting evapotranspiration (Case study: Yazd province). *Journal of Arid Biome*, 1(2), 9-19. (in Persian).
- Rahimikhoob, A. (2014). Comparison between m5 model tree and neural networks for estimating reference evapotranspiration in an arid environment. *Water resources management*, 28(3), 657-669.
- Raziei, T., Daneshkar Arasteh, P., and Saghafian, B. (2005). Annual rainfall trend analysis in arid and semi-arid regions of central and eastern Iran. *Water and Wastewater*, 54, 73-81. (in Persian).
- Roderick, M.L., and Farquhar, G.D. (2002). The cause of decreased pan evaporation over the past 50 years. *Science*, 298(5597), 1410-1411.
- Rodriguez-Galiano, V., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson P., and Jeganathan, C. (2012). Random Forest classification of Mediterranean land covers using multi-seasonal imagery and multi-seasonal texture. *Journal of Remote Sensing of Environment*, 121: 93-107.
- Sandhu, R., and Irmak, S. (2020). Performance assessment of Hybrid-Maize model for rainfed, limited and full irrigation conditions. *Agricultural Water Management*, 242, 106402.
- Sedaghat, A., Shahrestani, M.S., Noroozi, A.A., Nosratabad, A.F., and Bayat, H. (2022). Developing pedotransfer functions using Sentinel-2 satellite spectral indices and Machine learning for estimating the surface soil moisture. *Journal of Hydrology*, 127423.
- Sedaghat, A., Shabanpour, M., Noroozi, A., Fallah Nosratabad, A., and Bayat, H. (2022). The use of spectral indices to estimate soil surface moisture using machine learning algorithms. *Iranian Journal of Soil and Water Research*, 52(12), 3001-3018. (in Persian).
- Sepehri, S., Abbasi, F., Zarei, G., and Nakhjavanimoghaddam, M. M. (2021). Investigation of Artificial Neural Network Based Models and Sensitivity Analysis for Reference Evapotranspiration Estimating. *Iranian Journal of Irrigation & Drainage*, 14(6), 2089-2099. (in Persian).
- Shi, Y. (2019). Climate Change on the Tibetan Plateau and Its Impact on Potential Evapotran-Spiration. *Beijing Forestry University: Beijing, China*.
- Shiri J. (2017). Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. *Agric Water Manag*, 188, 101-114
- Singh A., Haghverdi, A., Öztürk, H.S., and Durner, W. (2020). Developing Pseudo Continuous Pedotransfer Functions for International Soils Measured with the Evaporation Method and the HYPROP System: I. The Soil Water Retention Curve. *Water*, 12, 3425.
- Su, J.W., Zhang, X.L., and Shen, B. (2021). Spatio-temporal variation characteristics and influencing factors of potential evapotranspiration in Heilongjiang. *Heilongjiang Water Conserv. Sci. Technol*, 49, 1-8.
- Tabari, H., and Talaei, P. H. (2014). Sensitivity of evapotranspiration to climatic change in different climates. *Global and Planetary Change*, 115, 16-23.
- Tabari, H., Hosseinzadeh Talaei, P., and Willems, P. (2014). Links between Arctic Oscillation (AO) and inter-annual variability of Iranian evapotranspiration, *Quaternary International*, 345, 148-157.
- Tafteh, A., Davatgar, N., and Sedaghat, A. (2022). Estimation of important points on soil water retention curve (SWRC): comparison experimental-physical models and data mining technique. *Arabian Journal of Geosciences*, 15(10), 1-13.
- Vapnik, V.N. (1998). *Statistical Learning Theory*. Wiley, New York. 736 pp.
- Wang, S., Lian, J., Peng, Y., Hu, B., and Chen H. (2019). Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. *Agric Water Manag*, 221, 220-30.
- Wen, X., Si, J., He, Z., Wu, J., Shao, H., & Yu, H. (2015). Support-vector-machine-based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions. *Water resources management*, 29(9), 3195-3209.
- Wu, L., Zhou, H., Ma, X., Fan, J., and Zhang, F. (2019). Daily reference evapotranspiration prediction based on hybridized extreme learning machine model with bio-inspired optimization algorithms: Application in contrasting climates of China. *J Hydrol*, 577, 123960.
- Xavier, F., Tanaka, A. K., and Amorim, F. A. (2016). Application of data science techniques in evapotranspiration estimation, (Master's thesis).
- Yang, Y., Chen, R., Song, Y., Han, C., Liu, J., and Liu, Z. (2019). Sensitivity of potential evapotranspiration to meteorological factors and their elevational gradients in the Qilian Mountains, northwestern China. *J. Hydrol*, 568, 147-159

- Yassin, MA., Alazba, A.A., and Mattar, MA. (2016). Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agric Water Manag*, 163, 110–24.
- Zhou, B. R., Li, F. X., Xiao, H. B., Hu, A. J., and Yan, L. D. (2014). Temporal and spatial differentiation characteristics of potential evapotranspiration and climate attribution in the source region of the Three Rivers. *J. Nat. Resour*, 29, 2068-2077.
- Zorati Pur, E., Neisi, L., Golabi, M., Bazaz, A., and Zoratipur, A. (2019). Simulation and Comparison of Potential Evapotranspiration by Artificial Neural Networks, ANFIS (Fuzzy Neural Network) and Decision Making M5 (Case Study; Synaptic Station of Shiraz). *Iran-Water Resources Research*, 15(1), 365-371.(*in Persian*).