



## Investigating the effectiveness of transfer functions based on machine learning methods for predicting reference evaporation and transpiration (Case study: Bushehr)

Halimeh Piri<sup>1\*</sup> | Mojtaba Mobaraki<sup>2</sup>

1. Corresponding Author, Associate Professor, Department of Water Engineering, Faculty of Water and Soil, University of Zabol, Zabol, Iran. Email: [h\\_piri2880@uoz.ac.ir](mailto:h_piri2880@uoz.ac.ir)

2. Graduate student, Department of Water Engineering, Faculty of Water and Soil, University of Zabol, Zabol, Iran. Email: [m\\_mobaraki@yahoo.com](mailto:m_mobaraki@yahoo.com)

---

### ARTICLE INFO

**Article type:**  
Research Article

**Article History:**  
Received October 23, 2023  
Revised November 22, 2023  
Accepted December 06, 2023  
Published online 12 December 2023

**Keywords:**  
*cubist model,*  
*random forest,*  
*step by step regression,*  
*support vector machine.*

---

### ABSTRACT

Accurate calculation of reference evapotranspiration is one of the basic tasks to achieve optimal management of water resources. In this research, the evapotranspiration of Bushehr reference was calculated by the Fau Penman Mantis method. Then, temperature methods (Hargreaves Samani and Blaney Criddle) and radiation methods (modified Mc King, Tork and Prestley Taylor) were also used to calculate evaporation transpiration. The results obtained from these methods were compared with the combined Fau Penman Mantis method. The results showed that among the two temperature methods, the Hargreaves Samani method and among the radiation methods, the Prestley Taylor method had closer results to the combined Fau Penman Mantis method. Artificial intelligence, support vector machine, random forest and cubist models were also used to estimate reference evaporation transpiration. The data used included minimum, maximum and average temperature, relative humidity, sunshine hour and wind speed during a thirty year statistical period from 1370 to 1400. In order to check the results of the mentioned models, the standard evaluation criteria of error mean square, absolute mean error and  $R^2$  explanation coefficient were used. The results showed that all three models were highly accurate in simulating evaporation transpiration. Cubist model with higher  $R^2$  (0.95), the lowest mean squared error (0.87) and the lowest absolute mean error (0.38) was chosen as the best method for evaporation transpiration.

---

**Cite this article:** Piri, H. & Mobaraki, M. (2023). Investigating the effectiveness of transfer functions based on machine learning methods for predicting reference evaporation and transpiration (Case study: Bushehr). *ECO HYDROLOGY*.10 (3), 421-434. Doi: [doi.org/10.22059/ije.2023.368738.1775](https://doi.org/10.22059/ije.2023.368738.1775)



© Halimeh Piri, Mojtaba Mobaraki      **Publisher:** University of Tehran Press.  
DOI: <http://doi.org/10.22059/ije.2023.368738.1775>

---



## بررسی کارایی توابع انتقالی مبتنی بر روش‌های یادگیری ماشین برای پیش‌بینی تبخیر و تعرق مرجع (مطالعه موردی: بوشهر)

حلیمه پیری<sup>۱\*</sup> | مجتبی مبارکی<sup>۲</sup>

۱. دانشیار گروه مهندسی آب، دانشکده آب و خاک، دانشگاه زابل. رایانامه: [h\\_piri2880@uoz.ac.ir](mailto:h_piri2880@uoz.ac.ir)  
 ۲. دانش‌آموخته کارشناسی ارشد، گروه مهندسی آب، دانشکده آب و خاک، دانشگاه زابل. رایانامه: [m\\_mobaraki@yahoo.com](mailto:m_mobaraki@yahoo.com)

### اطلاعات مقاله

### چکیده

#### نوع مقاله:

پژوهشی

#### تاریخ‌های مقاله:

تاریخ دریافت: ۱۴۰۲/۰۸/۰۱

تاریخ بازنگری: ۱۴۰۲/۰۹/۰۱

تاریخ پذیرش: ۱۴۰۲/۰۹/۱۵

تاریخ انتشار: ۱۴۰۲/۰۹/۲۱

#### کلیدواژه:

جنگل تصادفی،

رگرسیون گام‌به‌گام،

ماشین بردار پشتیبان،

مدل کیوبیست.

محاسبه دقیق تبخیر-تعرق مرجع یکی از گام‌های اساسی برای رسیدن به مدیریت بهینه منابع آب است. در این تحقیق تبخیر-تعرق مرجع بوشهر به روش فائو-پنمن-مانتیت محاسبه شد. سپس، از روش‌های روش‌های دمایی (هارگریوز سامانی و بلانی کرایدل) و روش‌های تشعشی (مک‌کینگ اصلاح‌شده، تورک و پرستلی تیلور) نیز برای محاسبه تبخیر-تعرق استفاده شد. نتایج به‌دست‌آمده از این روش‌ها با روش ترکیبی فائو-پنمن-مانتیت مقایسه شد. نتایج نشان داد از بین دو روش دمایی روش هرگریوز سامانی و از بین روش‌های تشعشی روش پرستلی تیلور نتایج نزدیک‌تری به روش ترکیبی فائو-پنمن-مانتیت داشتند. از مدل‌های هوش مصنوعی، مدل ماشین بردار پشتیبان، جنگل تصادفی و کیوبیست نیز برای تخمین تبخیر-تعرق مرجع استفاده شد. داده‌های مورد استفاده شامل دمای حداقل، حداکثر و متوسط، رطوبت نسبی، ساعت آفتابی و سرعت باد طی یک دوره آماری سی‌ساله از سال ۱۳۷۰ تا ۱۴۰۰ بود. برای بررسی نتایج مدل‌های یادشده از معیارهای ارزیابی مجذور میانگین مربعات خطا، میانگین مطلق خطا و ضریب تبیین  $R^2$  استفاده شد. نتایج نشان داد هر سه مدل دقت زیادی در شبیه‌سازی تبخیر-تعرق داشتند. مدل کیوبیست با داشتن  $R^2$  بالاتری (۰/۹۵)، کمترین مجذور میانگین خطا (۰/۸۷) و کمترین میانگین مطلق خطا (۰/۳۸) به عنوان روش برتر برای تبخیر-تعرق انتخاب شد.

استناد: پیری، حلیمه و مبارکی، مجتبی (۱۴۰۲). بررسی کارایی توابع انتقالی مبتنی بر روش‌های یادگیری ماشین برای پیش‌بینی تبخیر و تعرق مرجع (مطالعه موردی: بوشهر). *آکوهیدرولوژی*، ۱۰ (۳) ۴۲۱-۴۳۴.

DOI: <http://doi.org/10.22059/ije.2023.368738.1775>

© حلیمه پیری، مجتبی مبارکی. ناشر: مؤسسه انتشارات دانشگاه تهران.

DOI: <http://doi.org/10.22059/ije.2023.368738.1775>



## مقدمه

تبخیر- تعرق یکی از مؤلفه‌های مهم چرخه هیدرولوژی است که تخمین دقیق آن برای مدیریت بهینه منابع آب ضروری است. تخمین میزان دقیق تبخیر- تعرق نه تنها در برنامه‌ریزی آبیاری و تعیین دوره آبیاری، بلکه در مطالعات مربوط به مدل‌سازی بیلان آب هر ناحیه از اهمیت فراوانی برخوردار است. درک اهمیت اجزای بیلان آب از قبیل تبخیر- تعرق و قابلیت نفوذ همراه با برنامه آبیاری مشخص، در توسعه استراتژی‌های مدیریت آبیاری بسیار کاربردی است [۱]. روش‌های مختلفی برای برآورد تبخیر- تعرق وجود دارد. تبخیر- تعرق واقعی گیاه از حاصل ضرب تبخیر- تعرق مرجع در یک ضریب گیاهی به دست می‌آید. بنابراین، به دست آوردن تبخیر- تعرق مرجع ضروری است. تبخیر- تعرق مرجع به روش‌های فیزیکی، یا تجربی به دست می‌آید. اندازه‌گیری‌های مستقیم تبخیر- تعرق گیاهی به‌ندرت در دسترس هستند [۲]. از این‌رو، مدل‌هایی بر اساس، داده‌های هواشناسی در سطح جهانی توسعه یافته‌اند که با اصول فیزیکی، معادلات تجربی و یا ترکیبی از این دو روش تبخیر- تعرق را برآورد می‌کنند [۳]. روش‌های تجربی بر اساس آنالیزهای آماری و مشاهدات ایستگاه‌های هواشناسی و معمولاً برای یک منطقه یا شرایط آب‌وهوایی یا اقلیمی خاص ارائه می‌شوند [۴]. مدل‌های فیزیکی بیشتر هدفشان شبیه‌سازی اصول بنیادی مثل تعادل انرژی و جابه‌جایی جرم است. برای مثال تابش خورشیدی، رطوبت نسبی، سرعت باد تعدادی از داده‌های مهم برای محاسبه تبخیر تعرق مرجع بر اساس مدل‌های فیزیکی هستند [۵]. از بین روش‌های مختلف برآورد تبخیر- تعرق مرجع، معادله فائو- پنمن- مانیتث به عنوان روش مناسب برای تخمین تبخیر- تعرق مرجع در مناطق و اقلیم‌های متفاوت معرفی شده است. با این‌وجود، این معادله نیاز به داده‌های اولیه فراوان (دمای هوا، تابش خورشید، سرعت باد و رطوبت نسبی) دارد. به همین دلیل، استفاده از روش‌های تجربی ساده‌تر مانند تورنت وایت، هارگریوز، پرستلی تیلور و... توصیه شده است. تحقیقات نشان می‌دهد دقت روش‌های تخمین تبخیر- تعرق در مناطق مختلف متفاوت است و استفاده از این معادلات در شرایط اقلیمی مختلف نیاز به واسنجی و ارزیابی دارد. در سال‌های اخیر مطالعات زیادی در زمینه برآورد تبخیر- تعرق مرجع انجام شده است. نتایج تحقیقات نشان داد در اقلیم سرد و خشک تبریز، از بین روش‌های تجربی بلانی کرایدل، جنسن هیز، هارگریوز و تورنت وایت، روش هارگریوز نزدیک‌ترین برآورد را به روش ترکیبی فائو- پنمن- مانیتث داشت [۵]. تحقیقات روش‌های تبخیر- تعرق در شمال شرق هند و مقایسه آن‌ها با روش فائو- پنمن- مانیتث نشان داد روش تورک، ایرماک، بلانی کرایدل و فائو ۲۴ بیشترین نزدیکی را با روش فائو- پنمن- مانیتث داشتند [۶]. در مطالعات مشابهی به ارزیابی ۲۴ مدل تبخیر- تعرق گیاه مرجع در اقلیم‌های مختلف ایران پرداخته شد. نتایج تحقیق نشان داد در اقلیم سرد و نیمه‌خشک، سرد معتدل، گرم و خشک و سرد و خشک و گرم و مرطوب به‌ترتیب شامل شهرهای ارومیه، شهرکرد، زاهدان، مشهد و اهواز مدل مک‌کینگ اصلاح‌شده، در اقلیم گرم و مرطوب بندرعباس مدل والیانتراس [۵] و در اقلیم‌های مدیترانه‌ای، نیمه‌مرطوب و مرطوب شامل شهرهای گرگان، ساری و رشت مدل دروگرز و آلن برآورد بهتر و نزدیک‌تری به روش فائو- پنمن- مانیتث داشتند [۷]. از آنجا که تخمین تبخیر و تعرق از اهمیت زیادی در مهندسی منابع آب دارد و داده‌های لایسیمیتری نیز در بسیاری از مناطق و اقلیم‌های مختلف ایران در دسترس نیست، بنابراین ارزیابی دقت معادلات ارائه‌شده بر اساس منطقه و اقلیم می‌تواند نقش مهمی در مدیریت منابع آب داشته باشد. لذا یکی از اهداف این تحقیق، مقایسه و ارزیابی دقت ۴ روش برآورد تبخیر- تعرق مرجع هارگریوز، پرستلی تیلور، مک‌کینگ و والیانتراس برای اقلیم بوشهر و ارائه مدلی که نزدیک‌ترین برآورد را به مدل فائو- پنمن- مانیتث دارد، است.

نتایج تحقیقات نشان می‌دهد مدل‌های خطی، به دلیل تغییرپذیری زیاد مؤلفه تبخیر- تعرق عملکرد مناسبی در برآورد آن نداشته و باید از مدل‌های غیرخطی استفاده کرد که خود نیازمند برآورد پارامترهای زیادی خواهد بود [۸]. اساساً تبخیر- تعرق یک فرایند پیچیده و غیرخطی است و از این‌رو روش‌های هوش مصنوعی به عنوان یک روش جدید که قابلیت زیادی در مدل‌سازی فرایندهای پیچیده و غیرخطی برخوردار است، می‌تواند برای مدل‌سازی تبخیر- تعرق به کار گرفته شود. از جمله روش‌های هوش مصنوعی می‌توان به روش‌های یادگیری ماشین اشاره کرد. روش‌های یادگیری ماشین به عنوان بخشی از دانش هوش مصنوعی در عصر فناوری و اطلاعات به طور فزاینده‌ای در زمینه‌های مختلف علمی توسعه یافته است [۹]. از جمله مدل‌های یادگیری

ماشین می‌توان به مدل‌های ماشین بردار پشتیبان<sup>۱</sup>، جنگل تصادفی<sup>۲</sup> و کیوبیست<sup>۳</sup> اشاره کرد. ماشین بردار پشتیبان یک طبقه‌بندی‌کننده دوتایی غیرآماری است که در سال‌های اخیر بسیار مورد توجه قرار گرفته است. رویکرد این مدل به این صورت است که در فاز آموزش، تلاش می‌شود که مرز تصمیم‌گیری به گونه‌ای انتخاب شود که حداقل فاصله آن با هر یک از دسته‌های مورد نظر ماکزیمم شود [۱۰]. از مزایای این روش آموزش نسبتاً ساده آن است، به خلاف شبکه‌های عصبی مصنوعی در ماکزیمم‌های محلی گیر نمی‌افتد، برای داده‌های با ابعاد بالا تقریباً خوب جواب می‌دهد [۱۱]. مدل جنگل تصادفی به عنوان روشی توسعه‌یافته برای بهبود و به‌سازی روش‌های درخت طبقه‌بندی و درخت رگرسیون به کار گرفته می‌شود. این مدل دقت و قابلیت انطباق و سازگاری تقویت‌شده دارد و در انجام محاسبات از سرعت قابل قبولی برخوردار است. همچنین، در انجام فرایند مدل‌سازی با جنگل تصادفی نیازی به نرمال‌سازی داده‌ها وجود ندارد [۱۲]. مدل کیوبیست نوعی الگوریتم پیشرفته از روش درخت رگرسیون است که مدل‌های مختلفی را از داده‌های آموزشی ایجاد می‌کند. هر مدل متشکل از چندین قاعده است و هر قاعده در یک یا چندین شرط خلاصه می‌شود. وقتی همه شرایط یک قاعده تأمین می‌شود، رابطه خطی مربوطه برای پیش‌بینی متغیر هدف به کار گرفته می‌شود. قواعد مدل بر اساس روند کاهشی اهمیت به وسیله الگوریتم کیوبیست رتبه‌بندی می‌شود [۱۳]. تحقیقات زیادی در زمینه استفاده از مدل‌های هوشمند در پیش‌بینی تخریب-تعرق در مناطق مختلف انجام شده است. در بررسی کاربرد سیستم‌های استنتاج فازی عصبی و ماشین بردار پشتیبان برای برآورد تخریب-تعرق ماهانه این نتیجه بیان شد که ماشین بردار پشتیبان کارایی بهتری نسبت به مدل استنتاج فازی-عصبی داشت [۱۴]. از روش‌های درخت تصمیم، یادگیری عمیق و جنگل تصادفی برای برآورد تخمین تعرق مرجع روزانه در منطقه سیستان استفاده شد. نتایج تحقیق نشان داد مدل یادگیری عمیق به عنوان بهترین مدل در بین مدل‌های انتخابی است [۲].

با توجه به اینکه تخریب-تعرق بعد مکانی و زمانی دارد، هدف دیگر از این تحقیق تعیین و ارزیابی بهترین مدل هوش مصنوعی تحت سناریوهای مختلف حداقل داده اقلیمی و مقایسه آن با استفاده از شاخص‌های آماری با روش‌های تجربی معتبر و ارائه کاربردی‌ترین سناریوها در منطقه بوشهر است.

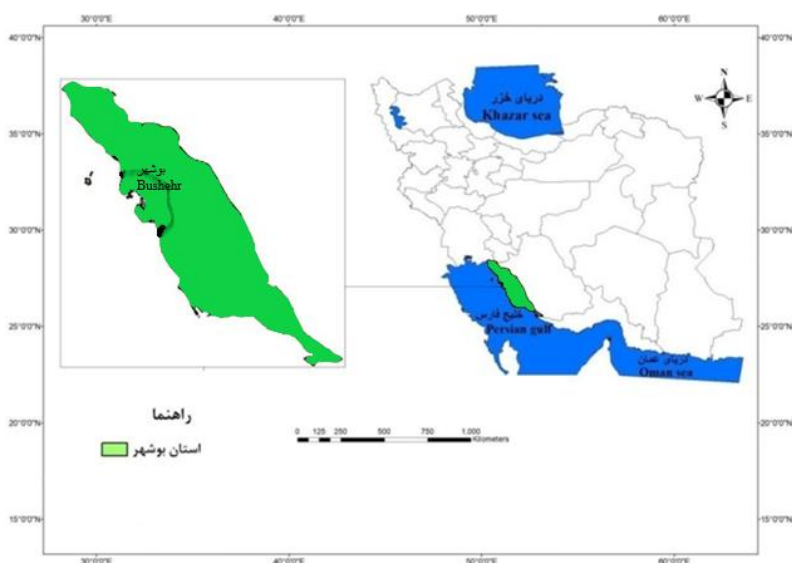
## مواد و روش‌ها

### منطقه مطالعه شده

استان بوشهر با مساحتی حدود ۲۷۶۵۳ کیلومتر مربع در حاشیه خلیج فارس به صورت نواری نسبتاً باریک در فاصله بین ۲۷ درجه و ۱۴ تا ۳۰ درجه و ۱۶ دقیقه عرض شمالی و ۵۰ درجه و ۶ دقیقه تا ۵۲ درجه و ۵۸ دقیقه طول شرقی از نصف النهار گرینویچ قرار گرفته است. این استان از شمال، به استان خوزستان و قسمتی از استان کهگیلویه و بویر احمد؛ از جنوب، به خلیج فارس و قسمتی از هرمزگان؛ از شرق، به استان فارس و از غرب به خلیج فارس محدود است. شهر بوشهر با ۵۰ درجه و ۴۹ دقیقه عرض شمالی و ۲۸ درجه و ۵۸ دقیقه طول شرقی و در ارتفاع ۹ متر از سطح دریا در بخش مرکزی استان بوشهر، از سمت شمال، غرب و جنوب به خلیج فارس محدود شده و در جنوب غرب کشور ایران واقع شده است و از نظر اقلیمی آب‌وهوای گرم و مرطوب دارد [۱۵]. شکل ۱ موقعیت جغرافیایی استان بوشهر و شهر بوشهر را نشان می‌دهد.

در این تحقیق ابتدا داده‌های هواشناسی شهر بوشهر شامل داده‌های رطوبت نسبی، دمای حداکثر، حداقل، میانگین، ساعت آفتابی و سرعت باد در ارتفاع ۲ متری از سطح زمین در بازه زمانی ۳۰ ساله ۱۳۷۰-۱۴۰۰ از سازمان هواشناسی تهیه شد.

1. support vector regression model  
2. Random Forest  
3. Cubist



شکل ۱. منطقه مورد مطالعه

### بر آورد تبخیر- تعرق مرجع

در این تحقیق برای محاسبه تبخیر- تعرق مرجع از معادلات تشعشی (مک‌کینگ اصلاح‌شده، تورک و پرستلی تیلور)، معادلات دمایی هارگریوز سامانی و بلانی کریدل و معادله ترکیبی فائو- پنمن- مانتیث استفاده شد. به علت نبود داده‌های لایسیمیتری از روش فائو- پنمن- مانتیث برای ارزیابی و اعتبارسنجی مدل‌ها استفاده شد. ابتدا داده‌های روزانه تبخیر و تعرق مرجع به روش فائو- پنمن- مانتیث محاسبه شد و سپس، مقادیر تبخیر و تعرق با معادلات و مدل‌های ارائه‌شده در جدول ۱ به دست آمد.

### معادله فائو- پنمن- مانتیث

در این معادله  $ET_0$ : تبخیر و تعرق مرجع ( $l^1 \text{ mm day}$ )،  $R_n$ : تابش خالص ورودی به سطح گیاه ( $l^2 \text{ day MJ m}$ )،  $G$ : شار گرمای خاک ( $l^2 \text{ day MJ m}$ )،  $T_{\text{mean}}$ : میانگین دمای هوای روزانه ( $^0C$ )،  $U_2$ : میانگین روزانه سرعت باد در ارتفاع دومتری از سطح زمین ( $l^1 \text{ m s}$ )،  $e_s$ : فشار بخار اشباع ( $KPa_a$ )،  $e_a$ : فشار بخار واقعی ( $KPa_a$ )،  $\Delta$ : شیب منحنی فشار بخار اشباع ( $l^1 KPa^0C$ )،  $\gamma$ : ضریب سایکرومتری ( $l^1 KPa^0C$ )،  $C_n$ : ضریبی برای گیاه مرجع که مقدار آن ۹۰۰ است ( $l^1 KJ^0K$ )،  $C_d$  و  $l^1 \text{ day}$  و ضریب باد برای گیاه مرجع که مقدار آن ۰/۳۴ است ( $l^1 \text{ s m}$ ).

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{C_n}{T + 273} U_2 (e_s - e_a)}{\Delta + \gamma(1 + C_d U_2)} \quad (1)$$

جدول ۱. روابط مربوط به مدل‌های تجربی برآورد تبخیر- تعرق مرجع

معادلات	روش	شماره	نوع روش
$ET_0 = 0.408 \times 0.0023 Ra (T_{\text{mean}} + 17.8) (T_{\text{max}} - T_{\text{min}})^{0.5}$	هارگریوز سامانی (۱۹۸۵)	۲	دمایی
$ET_0 = P (0.46 T_{\text{mean}} + 8.17)$	بلانی کریدل (۱۹۷۷)	۳	
$ET_0 = 0.7 \frac{\Delta}{\Delta + \gamma} \frac{R_s}{\lambda}$	مک‌کینگ اصلاح‌شده (۱۹۵۷)	۴	تشعشی
$ET_0 = 0.013 \left( \frac{T_{\text{mean}}}{T_{\text{mean}} + 15} \right) \left( \frac{23.88 R_s + 50}{\lambda} \right)$	تورک (۱۹۶۱)	۵	
$ET_0 = 1.26 \frac{\Delta}{\Delta + \gamma} \frac{R_n - G}{\lambda}$	پرستلی تیلور (۱۹۷۲)	۶	

در روابط ۲ تا ۶: Ra تابش فراجوی ( $Mj m^{-2} day^{-1}$ )، Tmean: متوسط دمای روزانه هوا ( $^{\circ}C$ )، Tmax: دمای بیشینه ( $^{\circ}C$ )، Tmin: دمای کمینه ( $^{\circ}C$ )،  $Rs$  ( $day Mj m^{-2}$ )، P: ساعت آفتابی (%، ضریب  $\lambda$  در دمای ۲۰ درجه سانتیگراد برابر  $2/45 MJ m^{-2} day^{-1}$ )،  $Rn$ : تابش خالص ورودی به سطح گیاه ( $MJ m^{-2} day^{-1}$ ) و  $\Delta$ : شیب منحنی فشار بخار اشباع ( $KPa^{\circ}C$ )<sup>۱</sup> است. در تمامی روابط مقدار ETO برحسب ( $mm day^{-1}$ ) است.

بعد از اینکه تبخیر- تعرق مرجع به روش‌های یادشده محاسبه شد، برای ارزیابی دقت تخمین هر یک از روش‌ها، مقادیر محاسبه‌شده از هریک از روش‌ها با مقادیر محاسبه‌شده روش ترکیبی فائو- پنمن- مانتیت توسط آماره‌های خطا مقایسه شدند.

### مدل‌های هوش مصنوعی

بعد از محاسبه تبخیر- تعرق به روش مینا (فائو- پنمن- مانتیت)، از بین کل داده‌های موجود ۸۰ درصد برای آموزش و ۲۰ درصد برای آزمون مدل‌های هوش مصنوعی مورد استفاده قرار گرفتند. برای بررسی تأثیر متغیرهای مختلف هواشناسی و تعیین الگوهای ورودی به هر مدل از روش رگرسیون گام‌به‌گام استفاده شد. سپس، با استفاده از نتایج رگرسیون گام‌به‌گام ترکیب‌های مختلفی از داده‌های هواشناسی به عنوان ورودی مدل‌ها و مقادیر محاسبه‌شده متناظر به عنوان خروجی مدل‌ها در نظر گرفته شد و شبکه آموزش داده شد. نتایج حاصل از این ترکیب‌ها مورد ارزیابی قرار گرفت و بهترین ترکیب برای برآورد تبخیر- تعرق در نظر گرفته شد.

### رگرسیون گام‌به‌گام

رگرسیون گام‌به‌گام سه نوع است: (۱) پس‌رو، (۲) پیش‌رو و (۳) مرحله‌ای (ترکیبی از هر دو روش پیش‌رو و پس‌رو). روش پس‌رو، مدلی را که در آن همه متغیرها حضور دارند در نظر می‌گیرد و با ملاک قرار دادن یکی از معیارهای سنجش متغیرهای بی‌تأثیر در مدل را حذف می‌کند. روش پیش‌رو مدلی را که تنها در آن عرض از مبدأ وجود دارد را در نظر گرفته و با ملاک قرار دادن یکی از معیارهای سنجش تأثیرگذارترین متغیر را به مدل اضافه می‌کند. روش مرحله‌ای ترکیبی از روش‌های پس‌رو و پیش‌رو است که در آن، در هر مرحله هم‌زمان با حذف یک متغیر از مدل بر اساس معیار سنجش، وجود همان متغیر در مدل بر اساس معیار سنجش بررسی می‌شود [۱۶]. با این روش می‌توان به طور هم‌زمان به تحلیل و بررسی چند متغیر مختلف پرداخت. رگرسیون گام‌به‌گام درواقع، رابطه بین چند متغیر پیش‌بین با متغیر پاسخ مورد نظر را بیان می‌کند.

### ماشین بردار پشتیبان

مفهوم اولیه ماشین‌های بردار پشتیبان اولین بار توسط محقق روسی به نام واپنیک در سال ۱۹۹۵ مطرح شد. این مفهوم روش جدیدی را بر مبنای طبقه‌بندی ارائه می‌دهد و یکی از ویژگی‌های آن طراحی دسته‌بندی‌کننده با حداکثر تعمیم است [۱۷]. هدف ماشین‌های بردار پشتیبان، یافتن یک تابع پیوسته با مقدار حقیقی است که بتواند به‌ازای مقادیر ورودی، پیش‌بینی دقیقی از مقدار خروجی داشته باشد [۱۸]. به طور کلی، ماشین بردار پشتیبان نوعی سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که توسط واپنیک ارائه شده است [۱۹]. ماشین بردار پشتیبان روشی است که هم برای گروه‌بندی و هم برای تخمین و برآورد تابع برازش داده‌ها در مسائل رگرسیون به کار می‌رود، به طوری که کمترین خطا در گروه‌بندی داده‌ها با تابع برازش رخ می‌دهد. SVM الگوریتمی است که نوع خاصی از مدل‌های خطی را می‌یابد که حداکثر حاشیه ابر صفحه را حاصل می‌کند، به نزدیک‌ترین نقاط آموزشی به حداکثر حاشیه ابر صفحه، بردارهای پشتیبان گفته می‌شود و تنها از این بردارها برای مشخص کردن مرز بین طبقات استفاده می‌شود. در یک مدل رگرسیونی SVM لازم است وابستگی تابع متغیر وابسته (y) به مجموعه‌ای از متغیرهای مستقل (x) تخمین زده شود. فرض بر این است که مانند دیگر مسائل رگرسیونی، رابطه بین متغیرهای وابسته و مستقل توسط یک تابع معین f به علاوه یک مقدار اضافی نویز مشخص می‌شود (رابطه ۲):

$$y = f(x) + \text{noise} \quad (1)$$

بنابراین، موضوع اصلی پیدا کردن فرم اصلی تابع f است که بتواند به صورت صحیح، موارد جدیدی را که SVM تا کنون تجربه نکرده است، پیش‌بینی کند. در روش SVM E از توابع هسته‌ای (کرنل) به منظور جداسازی صفحات و بهینه‌سازی فواصل

موجود بین داده‌های مشاهداتی و تخمینی استفاده می‌شود، اما در مدل SVM  $V$  از روش برنامه‌ریزی غیرخطی برای کاهش خطای پیش‌بینی استفاده می‌شود. در این مطالعه مدل SVM  $\varepsilon$  به دلیل کاربرد گسترده آن در مسائل رگرسیونی استفاده شد. برای این مدل، تابع خطا به صورت رابطه ۳ تعریف می‌شود:

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

تابع خطای یادشده لازم است که با توجه به محدودیت‌های زیر حداقل شود (روابط ۴-۶):

$$W^T \phi(x_i) + b - y_i \leq \varepsilon + \varepsilon_i \quad (4)$$

$$y_i - W^T \phi(x_i) + b \leq \varepsilon + \varepsilon_i \quad (5)$$

$$\varepsilon_i \times \varepsilon^* \geq 0 \quad (6)$$

در این روابط  $C$  ثابت گنجایش،  $W$  بردار ضریب،  $W^T$  ترانزپوز بردار ضریب،  $\varepsilon^*$  و  $\varepsilon_i$  ضرایب کمبود،  $b$  ضریب ثابت،  $n$  الگوی آموزش مدل و  $\phi$  تابع کرنل است.

ماشین‌های بردار پشتیبان برای حل مسائل غیرخطی، ابعاد مسئله را از طریق توابع کرنل تغییر می‌دهند. انتخاب کرنل برای SVM به حجم داده‌های آموزشی و ابعاد بردار ویژگی بستگی دارد. به بیان دیگر، باید با توجه به این پارامترها تابع کرنلی را انتخاب کرد که توانایی آموزش برای ورودی‌های مسئله را داشته باشد. در عمل کرنل خطی به کار برده می‌شود [۲۰].

### جنگل تصادفی

جنگل تصادفی یک روش گروهی است که چند الگوریتم درختی را برای تولید یک پیش‌بینی مکرر از هر پدیده ترکیب می‌کند. جنگل تصادفی می‌تواند الگوهای پیچیده را یاد بگیرد و ارتباط غیرخطی بین متغیرهای توضیحی و متغیرهای وابسته را در نظر بگیرد. همچنین، می‌تواند انواع مختلف داده‌ها را به علت نبود توزیع نرمال داده‌های استفاده شده، در تجزیه و تحلیل بگنجانند و ترکیب کند. مدل جنگل تصادفی برای بسیاری از مجموعه داده‌ها، دسته‌بندی را با سرعت بالایی انجام می‌دهد و به خلاف مدل‌های کلاسیک چون رگرسیون که تنها بر پایه یک مدل تکیه دارند با استفاده از صدها و هزاران درخت از اطلاعات بیشتری در داده‌ها استفاده می‌کند تا بتوان استنباط بهتری از متغیرها داشت. این روش یک تکنیک مدل ناپارامتری و متعلق به خانواده روش‌های دسته‌جمعی است [۲۱]. جنگل تصادفی با استفاده از مجموعه‌ای از درخت‌ها با در نظر گرفتن  $n$  داده مشاهده‌ای مستقل ساخته می‌شود (رابطه ۷):

$$(Y_i, X_i), i = 1, \dots, n \quad (7)$$

این الگوریتم از جمله دسته‌بندی‌هایی است که روش Bagging را استفاده می‌کند. این روش ترکیبی از چندین درخت تصمیم است که در ساخت آن چندین نمونه بوت استرپ از داده‌ها شرکت دارند و در ساخت هر درخت به طور تصادفی تعدادی از متغیرهای ورودی شرکت می‌کنند. با استفاده از روش بوت استرپ به تعداد زیاد (به عنوان مثال ۲۰۰۰ مرتبه) نمونه‌های  $n$  تایی از مجموعه داده‌های مشاهداتی اولیه، نمونه‌برداری همراه با جای‌گذاری می‌شوند. طی فرایند نمونه‌گیری حدود یک سوم از داده‌ها نمونه‌گیری نمی‌شوند و به عنوان نمونه خارج از کیسه (از این داده‌ها برای تعیین متغیرهای مهم و همچنین برآورد ناریب خطا استفاده می‌شود) در نظر گرفته می‌شوند. سپس، روی هر نمونه بوت استرپ یک درخت گسترش داده می‌شود. طی فرایند ساخت درخت در هر شاخه، از بین تمام  $M$  متغیر مستقل به صورت تصادفی  $m$  متغیر برای تقسیم شدن انتخاب می‌شود. برای حالت رگرسیونی نسبت  $m/M$  برابر یک سوم است و برای کلاسه‌بندی برابر با  $m = \sqrt{M}$  پیشنهاد شده است. پس از ساخت تمام درخت داده‌های تست به درخت معرفی می‌شود و به تعداد درخت‌ها برای بردار ورودی یک خروجی به دست می‌آید. با میانگین‌گیری این خروجی‌ها، خروجی نهایی مدل و با در نظر گرفتن توزیع تجربی خروجی‌ها مقادیر صدک‌ها و دامنه عدم قطعیت محاسبه می‌شود. روش جنگل تصادفی زمانی که تعداد مشاهدات در مقایسه با تعداد پیش‌بینی‌کننده‌ها نسبتاً کم باشد، یک روش پیش‌بینی کارآمد است [۲۲].

## کیویست

مدل کیویست توسعه یافته مدل درختی M5 کویلین است. یک درخت تا جایی رشد می کند که گره های انتهایی تشکیل یک سری مدل های رگرسیونی بدهند. این مدل های رگرسیونی مبتنی بر پیش بینی کننده هایی است که در تفکیک های درخت (توسعه شاخه و گره های درخت) استفاده شده اند. در هر مرحله از تشکیل درخت مدل های خطی حد واسطی هم تشکیل می شود. پیش بینی نهایی با استفاده از مدل رگرسیون خطی در گره انتهایی درخت انجام می شود، اما برای پیرایش نتیجه نهایی از مدل های خطی گره های پیشین درخت تشکیل شده، استفاده می شود [۲۳]. مدل کیویست از نزدیک ترین همسایه ها برای اصلاح پیش بینی مبتنی بر قاعده استفاده می کند. در مدل کیویست، مدل های تشکیل شده توسط دو مدل خطی به صورت رابطه ۸ بیان می شوند:

$$\hat{y}_{par} = a \times \hat{y}_{kid} + (1-a) \hat{y}_{par} \quad (8)$$

که در این رابطه  $\hat{y}_{par}$  پیش بینی از گره والد و  $\hat{y}_{kid}$  پیش بینی از گره فرزند است. ضریب  $a$  هم از رابطه ای شامل واریانس پیش بینی های والد و فرزند و اختلاف آن ها به دست می آید.

مدل کیویست از طرح کمیتی استفاده می کند. یادگیرنده های ضعیف به صورت تکرار شونده آموزش داده می شوند و به مجموعه قبلی اضافه می شوند تا در نهایت یک درخت طبقه بندی یا رگرسیون قوی حاصل شود. یادگیرنده های ضعیف در حین اضافه شدن به مجموعه، وزن دهی می شوند که این وزن دهی معمولاً بر اساس میزان دقت در نتایج درختان طبقه بندی یا رگرسیون است [۲۴]، اما در طرح کمیتی در مدل کیویست درخت های تکراری به ترتیب ایجاد می شوند. درخت اول که بر اساس قواعد یاد شده تولید شده و درخت های بعدی با استفاده از نسخه های تنظیم شده بر اساس نتایج مجموعه آموزشی ایجاد می شوند. اگر مدل یک مقدار را بیش از حد پیش بینی کرد، پاسخ در مدل بعدی به سمت مقدار کمتر تنظیم می شود. وزن هر مرحله برای هر کمیته برای میانگین پیش بینی های هر درخت مدل استفاده نمی شود. پیش بینی نهایی میانگین ساده پیش بینی های هر درخت در مدل است [۲۴].

## آماره های ارزیابی

برای ارزیابی صحت و اعتبار مدل های مورد استفاده از شاخص های آماری ریشه میانگین مربعات خطا (RMSE)، ضریب تبیین ( $R^2$ )، و میانگین مطلق خطا (MAE) استفاده شد (روابط ۹-۱۱).

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - x_i^*)^2}{\sum_{i=1}^n (x_i - \bar{x}_i)^2} \quad (9)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - x_i^*)^2} \quad (10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x_i^*| \quad (11)$$

در این روابط  $x_i$  مقدار اندازه گیری شده،  $x_i^*$  مقادیر برآورد شده و  $n$  تعداد نمونه است.

## نتایج و بحث

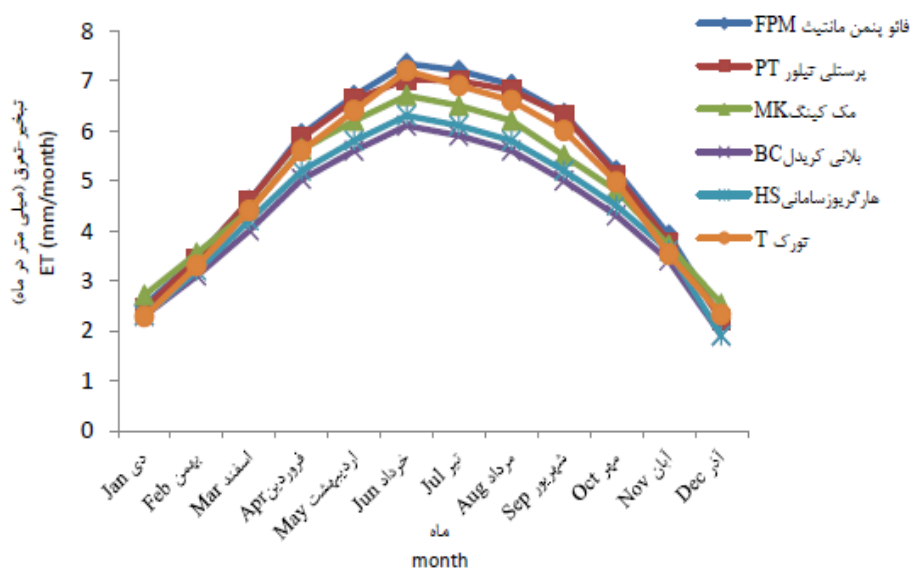
نتایج حاصل از مقایسه آماری تبخیر- تعرق محاسبه شده به روش های تجربی با نتایج روش ترکیبی فائو- پنمن- مانیت در جدول ۲ ارائه شده است. همان گونه که در جدول ۲ مشاهده می شود، در گروه دمایی روش هارگریوز سامانی با داشتن بالاترین ضریب تبیین (۰/۷۲) و کمترین مقدار جذر خطای میانگین مربعات (۰/۳۸) نسبت به روش بلانی کریدل برآورد بهتری از تبخیر- تعرق مرجع داشته است. محققان دیگر نیز در تحقیق خود بیان داشتند که از بین روش های دمایی تبخیر- تعرق، روش هارگریوز سامانی بهترین روش برای برآورد تبخیر- تعرق در دشت قزوین است [۲۵]. از بین روش های تشعشی روش پرستلی تیلور با داشتن ضریب تبیین ۰/۹۱ و جذر خطای میانگین مربعات ۰/۱۶ برآورد بهتری از تبخیر- تعرق مرجع نسبت به سایر روش های این گروه داشته است. به طور کلی، با توجه به نتایج به دست آمده از جدول ۲ می توان بیان کرد روش های تشعشی در منطقه مورد مطالعه (بوشهر) برآورد بهتری از تبخیر- تعرق مرجع داشتند.



جدول ۲. مقایسه روش‌های مختلف تجربی برآورد تبخیر- تعرق مرجع با روش ترکیبی فائو- پنمن- مانیت

نوع روش	روش	R <sup>2</sup>	RMSE	MAE
دمايي	هارگريوز سامانی (۱۹۸۵)	۰/۷۲	۰/۳۸	۰/۹۴
	بلانی کریدل (۱۹۷۷)	۰/۶۹	۰/۵۶	۰/۹۸
تشعشعی	مک کینگ اصلاح‌شده (۱۹۵۷)	۰/۷۷	۰/۳۱	۰/۸۲
	تورک (۱۹۶۱)	۰/۸۳	۰/۲۹	۰/۷۱
	پرستلی تیلور (۱۹۷۲)	۰/۹۱	۰/۱۶	۰/۲۹

مقادیر میانگین ماهانه تبخیر- تعرق محاسبه‌شده به روش ترکیبی فائو- پنمن- مانیت با هر یک از روش‌های تجربی در شکل ۲ نشان داده شده است. همان‌طور که در شکل ۲ نیز مشاهده می‌شود، مقادیر محاسبه‌شده با روش‌های تجربی تشعشعی (پرستلی تیلور و تورک) انطباق بیشتری با مقادیر محاسبه‌شده فائو- پنمن- مانیت دارند. بیشترین مقدار تبخیر- تعرق ماهانه محاسبه‌شده با روش فائو- پنمن- مانیت ۷/۳۴ میلی‌متر در ماه خرداد بود که با مقادیر محاسبه‌شده روش پرستلی تیلور (۷/۰۳ میلی‌متر در ماه) و تورک (۷/۲ میلی‌متر در ماه) تفاوت چندانی نداشت. ارزیابی مدل‌های تجربی و هوشمند در تخمین تبخیر- تعرق مرجع در شهرکرد نشان داد از بین مدل‌های تجربی، مدل‌های پرستلی تیلور و هارگريوز عملکرد بهتر و مدل جنسن هیز ضعیف‌ترین عملکرد را داشتند [۲۶]. تحقیقات نشان می‌دهد بررسی روش‌های مختلف برآورد تبخیر- تعرق مرجع تا حد زیادی محلی است و بستگی شدیدی به اقلیم منطقه مورد پژوهش دارد؛ گرچه نتایج در یک اقلیم یکسان هم ممکن هست متفاوت باشد [۲۷].



شکل ۲. مقایسه مقادیر میانگین تبخیر- تعرق ماهانه محاسبه‌شده با روش‌های تجربی و روش ترکیبی فائو- پنمن- مانیت

### پارامترهای ورودی به مدل‌های هوش مصنوعی بر اساس رگرسیون گام‌به‌گام

جدول ۳ پارامترهای آماری داده‌های آموزش و آزمون مورد استفاده طی دوره آماری را نشان می‌دهد. پارامترهای متوسط دما (Tmean)، حداقل دما (Tmin)، حداکثر دما (Tmax)، رطوبت نسبی متوسط (RH)، سرعت باد (U2) و ساعت‌های آفتابی (N) به عنوان پارامترهای ورودی و تبخیر- تعرق مرجع به عنوان پارامتر خروجی مدل هستند. برای بررسی تأثیر پارامترهای ورودی به مدل‌ها همچون حداقل، حداکثر و متوسط دما، رطوبت نسبی، ساعت آفتابی و سرعت باد در برآورد تبخیر- تعرق مرجع و تعیین الگوی ورودی به مدل‌ها، تحلیل رگرسیونی گام‌به‌گام بین این پارامترها انجام شد. نتایج رگرسیون گام‌به‌گام در جدول ۴ نشان داده شده است.

جدول ۳. میانگین پارامترهای آماری داده‌های استفاده‌شده در دوره آماری مورد مطالعه

ساعت آفتابی (ساعت)	حداقل دما	حداکثر دما درجه سانتی‌گراد	متوسط دما	رطوبت نسبی (درصد)	سرعت باد (متر بر ثانیه)
۹/۶	۷	۴۳	۲۶	۶۸	۲/۶

جدول ۴. نتایج آماری رگرسیون گام‌به‌گام

پارامتر	سطح معناداری	t	$\beta$
ساعت آفتابی	.	۳۷/۵۴	۰/۱۵
حداقل دما	.	۱۳/۲۵	۰/۱۱
حداکثر دما	.	۳۹/۴۱	۰/۶۳
متوسط دما	.	۲۴/۷۴	۰/۲۶
رطوبت نسبی	.	۲۶/۸۵	۰/۱۴
سرعت باد	.	۷۶/۲۷	۰/۴۵

همان‌طور که در جدول ۴ مشاهده می‌شود، تأثیر همه پارامترها در سطح یک درصد بر تبخیر- تعرق مرجع معنادار بوده است. مقادیر  $\beta$  نشان می‌دهد حداکثر دما با ضریب ۰/۶۳ بیشترین تأثیر را بر تبخیر- تعرق مرجع دارد. سرعت باد، متوسط دما، ساعت آفتابی، رطوبت نسبی و حداقل دما در رتبه‌های بعدی تأثیرگذاری قرار دارند. پژوهش‌های مشابه نیز همچنین نشان دادند حداکثر دما و سرعت باد بیشترین تأثیر را بر تبخیر- تعرق دارد [۲۸ و ۲۹]. با توجه به نتایج به‌دست‌آمده از رگرسیون گام‌به‌گام و معناداری همه پارامترها بر تبخیر- تعرق مرجع، همه پارامترها به عنوان ورودی مدل‌های هوشمند در نظر گرفته شدند. جدول ۵ الگوهای مختلف از ورود داده‌ها به مدل‌های هوش مصنوعی را نشان می‌دهد.

پس از طراحی سناریوها برای مدل‌های مختلف، خروجی این مدل‌ها دریافت و با شاخص‌های آماری مورد ارزیابی قرار گرفت. جدول ۶ نتایج ارزیابی را برای سه مدل مورد بررسی (کیوبیست، ماشین بردار پشتیبان و جنگل تصادفی) نشان می‌دهد.

جدول ۵. الگوهای مختلف ورودی، برای تخمین مقدار تبخیر- تعرق مرجع در مدل‌های مورد مطالعه

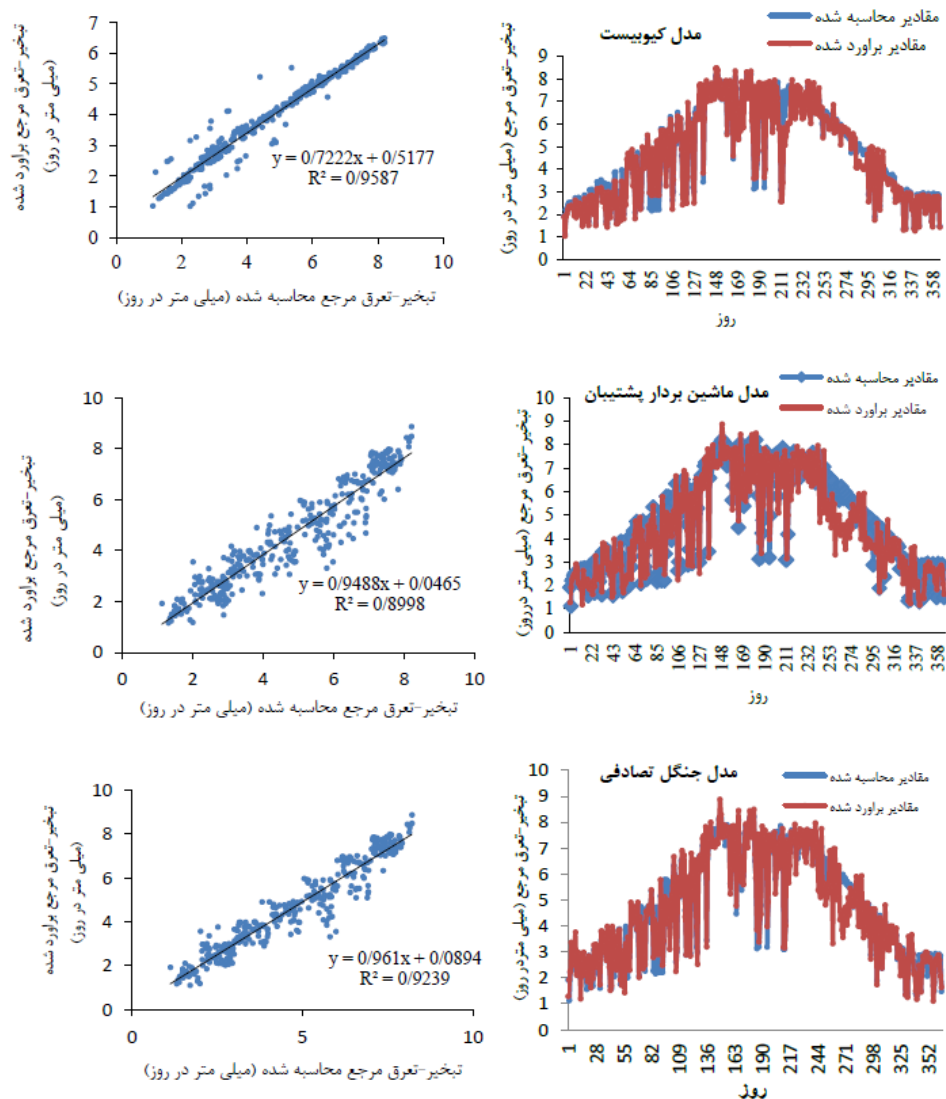
سناریو	پارامترهای ورودی	خروجی
۱	میانگین دما	تبخیر- تعرق مرجع
۲	میانگین دما	تبخیر- تعرق مرجع
۳	میانگین دما	تبخیر- تعرق مرجع
۴	میانگین دما	تبخیر- تعرق مرجع
۵	میانگین دما	تبخیر- تعرق مرجع
۶	میانگین دما	تبخیر- تعرق مرجع
۷	میانگین دما	تبخیر- تعرق مرجع
۸	میانگین دما	تبخیر- تعرق مرجع
۹	میانگین دما	تبخیر- تعرق مرجع
۱۰	میانگین دما	تبخیر- تعرق مرجع

جدول ۶. شاخص‌های آماری ارزیابی مدل‌های هوش مصنوعی مورد مطالعه

مدل	ماشین بردار پشتیبان			جنگل تصادفی			کیوبیست		
	R2	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE
۱	۰/۴۱	۱/۷۱	۱/۳۶	۰/۵۶	۱/۴۳	۰/۶۸	۰/۵۳	۱/۳۴	۰/۵۳
۲	۰/۴۳	۱/۶۷	۱/۳۱	۰/۵۲	۱/۳۸	۰/۶۳	۰/۵۷	۱/۲۸	۰/۵۷
۳	۰/۴۶	۱/۶۲	۱/۲۷	۰/۵۸	۱/۳۱	۰/۵۸	۰/۶۱	۱/۲۲	۰/۶۱
۴	۰/۴۹	۱/۵۶	۱/۲۳	۰/۶۳	۱/۲	۰/۵۴	۰/۶۸	۱/۱۷	۰/۶۸
۵	۰/۵۲	۱/۵۱	۱/۱۶	۰/۶۷	۱/۱۷	۰/۵۱	۰/۷۱	۱/۱۱	۰/۷۱
۶	۰/۶۲	۱/۲۴	۰/۸۶	۰/۹۲	۰/۹۴	۰/۳۸	۰/۹۵	۰/۸۷	۰/۹۵
۷	۰/۶۸	۱/۱۸	۰/۷۹	۰/۷۷	۰/۹۸	۰/۴۱	۰/۸۶	۰/۹۱	۰/۸۶
۸	۰/۸۹	۱/۰۳	۰/۷۹	۰/۷۵	۰/۹۹	۰/۴۴	۰/۸۳	۰/۹۵	۰/۸۳
۹	۰/۵۷	۱/۴۷	۱/۱۲	۰/۷۱	۱/۱۳	۰/۵۳	۰/۷۴	۱/۰۳	۰/۷۴
۱۰	۰/۵۲	۱/۵۳	۱/۱۶	۰/۶۷	۱/۱۷	۰/۵۸	۰/۶۹	۱/۱۴	۰/۶۹

نتایج ارزیابی مدل‌های هوش مصنوعی (جدول ۶) نشان داد در روش ماشین بردار پشتیبان الگوی ۸ با داشتن کمترین میانگین مطلق خطا (۰/۷۹)، میانگین مربعات خطا (۱/۰۳) و بیشترین ضریب تبیین (۰/۸۹) توانست از سایر مدل‌ها برآورد بهتری از تبخیر-تعرق مرجع داشته باشد. در روش‌های کیوبیست و جنگل تصادفی الگوی ۶ با داشتن به ترتیب مقدار ۰/۳۸ و ۰/۵۲ میانگین مطلق خطا، ۰/۸۷ و ۰/۹۴ میانگین مربعات خطا و ۰/۹۲ و ۰/۹۵ ضریب تبیین بهترین برآورد را از تبخیر-تعرق مرجع داشتند. همچنین، مقایسه بین مدل‌ها (جدول ۶) نشان داد اضافه کردن پارامترهای هواشناسی تا الگوی ۶ در مدل‌های کیوبیست و جنگل تصادفی باعث افزایش ضریب تبیین و کاهش آماره‌های خطا شد و عملکرد روش‌ها بهبود یافت. اضافه کردن ساعت آفتابی و رطوبت نسبی در روش‌های جنگل تصادفی و کیوبیست آماره‌های خطا را افزایش و ضریب تبیین را کاهش داد. تخمین تبخیر-تعرق مرجع روزانه در دشت سیستان نشان داد بیشترین حساسیت مدل جنگل تصادفی به پارامترهای دمای هوا است و سایر پارامترهای هواشناسی تأثیر زیادی بر مدل ندارند [۲]. بنابراین، با توجه به نتایج به دست آمده از تحقیق حاضر می‌توان یک الگوی مشخصی از داده‌های هواشناسی برای برآورد تبخیر-تعرق مرجع در بوشهر معرفی کرد.

منحنی پراکندگی مقادیر برآوردشده تبخیر-تعرق مرجع در برابر مقادیر محاسبه‌شده مدل ترکیبی فائو-پنمن-مانتیت در شکل ۳ نشان داده شده است.



شکل ۳. مقادیر برآوردشده و محاسبه‌شده تبخیر-تعرق مرجع با استفاده از روش‌های هوش مصنوعی

با توجه به اینکه در محاسبه تبخیر- تعرق پارامترهای هواشناسی زیادی در نظر گرفته می‌شود، انتخاب پارامترهای مؤثر و حذف برخی از آن‌ها برای کاهش متغیرهای ورودی، می‌تواند کمک زیادی در برآورد سریع و کم‌هزینه تبخیر- تعرق در مدل‌سازی‌ها داشته باشد. همان‌طور که در شکل ۳ مشاهده می‌شود، انطباق زیادی بین داده‌های محاسبه‌شده با روش ترکیبی فائو- پنمن- مانتیث و روش‌های هوش مصنوعی وجود دارد. بنابراین، می‌توان از این مدل‌ها برای برآورد تبخیر- تعرق مرجع در زمانی که داده‌های محدودی در دسترس است، استفاده کرد.

### نتیجه‌گیری

برآورد تبخیر- تعرق مرجع به دلیل برنامه‌ریزی‌های منابع آب از اهمیت خاصی برخوردار است، از این‌رو برآورد دقیق آن برای محققان مهم بوده و روش‌های مختلفی برای بهبود عملکرد مدل‌های پیش‌بینی تبخیر تعرق ارائه کرده‌اند. در این تحقیق ابتدا تبخیر- تعرق با استفاده از معادله ترکیبی فائو- پنمن- مانتیث محاسبه شد. سپس، تبخیر- تعرق از روش‌های اقلیمی (هرگزیوز سامانی و بلانی کریدل) و روش‌های تشعشعی (مک‌کینگ اصلاح‌شده، تورک و پرستلی تیلور) که نیاز به داده‌های اقلیمی کمتری دارند نیز محاسبه شد. نتایج به‌دست‌آمده از این روش‌ها با روش ترکیبی فائو- پنمن- مانتیث مقایسه شد. نتایج نشان داد از بین دو روش دمایی روش هرگزیوز سامانی و از بین روش‌های تشعشعی روش پرستلی تیلور نتایج نزدیک‌تری به روش ترکیبی فائو- پنمن- مانتیث داشتند. بنابراین، در شهر بوشهر می‌توان در شرایط کمبود داده اقلیمی از این دو روش برای محاسبه تبخیر- تعرق مرجع استفاده کرد. در ادامه از روش‌های هوش مصنوعی ماشین بردار پشتیبان، جنگل تصادفی و کیوبیست با ده الگوی مختلف برای برآورد تبخیر- تعرق مرجع استفاده شد. نتایج نشان داد روش‌های هوش مصنوعی توانایی بالایی برای برآورد تبخیر- تعرق دارند. از بین روش‌های مختلف روش کیوبیست با الگوی ۶ با داشتن کمترین آماره‌های خطا و بیشترین ضریب تبیین برآورد بهتری از تبخیر- تعرق مرجع داشت. از بین پارامترهای هواشناسی پارامتر حداکثر دما و سرعت باد بیشترین تأثیر را بر تبخیر- تعرق داشت. با توجه به نقش مهم تبخیر- تعرق مرجع در مطالعات کشاورزی و طرح‌های مدیریت منابع آب نتایج حاصل از این تحقیق می‌تواند راه حل مناسبی برای محاسبه و برآورد این پارامتر مهم هیدرولوژیکی به مدیران و طراحان ارائه دهد.

### سپاسگزارى

این پژوهش با حمایت مالی دانشگاه زابل کد پژوهانه 1837 GR UOZ IR انجام شده است.

## منابع

- [1]. Bethune, M., Selle, B., Wang, Q. Understanding and predicting deep percolation under surface irrigation. *Journal of Water Resour.* 2008: 15(4). 120-134.
- [2]. Siasser, H., and Dindarlou, A. Estimation of daily reference evaporation and transpiration using deep learning model, random forest and decision tree (case study: Sistan plain). *Iranian Water Research*, 2019: 14(1):108-99. [Persian].
- [3]. Maeda, E.E., Wiberg, D.A., and Pellikka, P.K.E. Estimating reference evapotranspiration using sensing empirical models in a region with limited data availability in Kenya. *Applied Geography*. 2010: 31: 251-258
- [4]. Bos, M.G., Kselik, R.G., Allen, K., Molden, D.J. *Water Requirements for Irrigation and the Environment*. Springer, 2009: 186p.
- [5]. Panahi, S. F. Rezvanizadeh and Samadianfard, S. Evaluation and comparison of experimental methods for estimation of reference evapotranspiration in Tabriz station. The first international conference on Iran's natural hazards and environmental crises, solutions and challenges, 2016: 9 p. [Persian]
- [6]. Pandey, P.K., P.P. Dabral., and Pandey, V. Evaluation of reference evapotranspiration methods for the northeastern region of India. *International Soil and Water Conservation Research*, 2016: 4(1), 52-63.
- [7]. Piri, H. and Poozan., M. Evaluation of 24 evapotranspiration models in different climates of Iran. *Ecohydrology*, 2019: 6(3):611-622.
- [8]. Wang, Z., P. Xie., C. Lai., X. Chen., X. Wu., Z. Zeng and J. Li. Spatiotemporal variability of reference evapotranspiration and contributing climatic factors in China during 1961–2013. *Journal of Hydrology*, 2017: 544, 97-108.
- [9]. Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B., and Gräler, B. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, 2018: 6, e5518.
- [10]. Pai, P., F. and W. C. Hong. A recurrent support vector regression model in rainfall forecasting. *Hydrological Process*. 2007: 21:819-827.
- [11]. Kihani, A., Akhundali, A., and Fathian, H. Uncertainty analysis of parameters of SVM model for estimation of suspended sediment load and bed in Sierra Karaj station with Monte Carlo simulation method. *Iran Water and Soil Research (Agricultural Sciences of Iran)*, 2021: 52(1), 195-212. [Persian].
- [12]. Breiman L. Application and analysis of random forests and machine learning. *Journal of Water Management*. 2001; 15(1): 5-32
- [13]. Kuhn, M., Weston, S., Keefer, C., Coulter, N., and Quinlan, R. Cubist: Rule and Instance based Regression Modeling, R package version 0.4.2.1. (<https://cran.r-project.org/web/packages/Cubist/Cubist.pdf>. Last access date: 3 May 2023)
- [14]. Ahmadi, F., Aysham, S., Khalili, K., and Beahmanesh, J. performance evaluation of artificial neural network (ANN) and support vector machine (SVM) (in estimating daily evaporation values) (Case study: Tabriz and Maragheh meteorological stations). *Soil and Water*, 2016: 1(49):151-168.
- [15]. Hajari, Z., Naserzadeh, M.H., Tagvi Guderzi, S. Preparation of Persian Gulf ecotourism calendar based on bioclimatic indicators of MEMI model (case study: Bushehr). *Scientific Quarterly Journal of Tourism Management Studies*, 2018:14(46): 282-245. [Persian].
- [16]. Hastie, T. and Pregibon, J. *Generalized linear models*. Eberly College of Science, London. 1992
- [17]. Burges, C.J. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 1998: 2: 121-167
- [18]. Norouzi Ghoshbalagh, H., Nadiri, A., Asghari Moghaddam, A. and Qarahkhani, M. Comparison of the efficiency of artificial neural networks, fuzzy logic and random forest in estimating the aquifer transfer capability of Malekan plain. *Echo Hydrology*, 2018, 5(3): 739-751. [Persian]
- [19]. Nosrati Karizak, F., Movahedi Naeni, S.A., and Hezarjaribi, A. Using Artificial Neural Networks to estimate saturated hydraulic conductivity from easily available soil properties. *J. Soil Manage. Sust. Prod.* 2012, 2(1): 95-110. [Persian].
- [20]. Wosten, J.H.M., Pachepsky, Ya.A. and Rawls, W.J. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *J. Hydrol.* 2001, 251:123–150.
- [21]. Breiman L. Application and analysis of random forests and machine learning. *Journal of Water Management*. 2001; 15(1): 5-32
- [22]. Siassar, H. and Honar, T. The application of support vector machine, chaid and random forest models in estimating daily reference transpiration evaporation in the north of Sistan and Baluchistan province. *Iran Irrigation and Drainage*. 2018 ;2(13):378-388. [Persian].
- [23]. Quinlan, J.R. Learning with continuous classes. P 343-348, In: *Proceedings of 5th Australian conference on artificial intelligence*. World Scientific. Singapore, 1992.
- [24]. Zhou, Z. *Ensemble Methods: Foundations and Algorithms*. Chapman and Hall/CRC. 23p. 2012.

- [25]. Nazari, R., Kaviani, A. Evaluation of potential evaporation and transpiration methods and evaporation pan with lysimeter values in a semi arid climate (case study, Qazvin Plain). *Ecohydrology*, 2015, 3(1):19 30. [Persian]
- [26]. Hosseini Vardanjani, S.M., Ganji Khoram Del, N., and Khalt Abadi Farahani, A.H. Evaluation and sensitivity analysis of different methods of daily reference evaporation and transpiration estimation in a cold climate. *Applied Water Science Research*, 2014, 1(2):29 40. [Persian].
- [27]. Tavakoli, A., Hero b., Davari, K., and Ansari, H. Estimation of reference evapotranspiration in data deficient conditions (case study: North Khorasan province), *Journal of Agricultural Sciences and Techniques and Natural Resources*, 2012: 65: 222 211. [Persian].
- [28]. Sabzevari, Y., Parsai, A., and Haqi Abi, A.H. Modeling and estimation of daily evaporation and transpiration of a reference plant with soft computing models (case study: Aliguderz station), 2023:13(52):292 306. [Persian].
- [29]. Hosseini Vardanjani, S.M., Ganji Khoram Del, N., and Khalt Abadi Farahani, A.H. Evaluation of experimental and intelligent models in estimation of reference evaporation and transpiration in the conditions of minimum climatic data; A case study of Kurd city. *Water and Irrigation Engineering*, 2015: 25(7):141 128. [Persian].