



Assessment and Improvement of Reference Evapotranspiration Estimation under Data-Scarce Conditions Using Random Forest Algorithm

M. A. Gholami Sefidkouhi* , **F. Jafari Sayadi**, and **H. Noghankar**

Professor of Water Engineering Department, Crop Engineering Faculty, Sari Agricultural Sciences and Natural Researches University (SANRU). ma.gholami@sanru.ac.ir

Post-Doc of Water Sciences and Engineering, Water Engineering Department, Crop Engineering Faculty, Sari Agricultural Sciences and Natural Researches University (SANRU). f.sayadi@sanru.ac.ir

Ph.D. of Water Sciences and Engineering, Regional Water Company of Mazandaran. noghankar.h@gmail.com

Article Info

Research Article

Received: May 09, 2026

Accepted: September 08, 2026

Keywords: Digital Elevation Model, Empirical Models, FAO Penman-Monteith, Humidity-based, Radiation-based, Temperature-based

Corresponding author's email: ma.gholami@sanru.ac.ir

Abstract

This study aimed to evaluate and improve the estimation of reference evapotranspiration under data-limited conditions using the Random Forest (RF) algorithm. From 17 weather stations, meteorological data of 2000–2025 were used in Mazandaran Province, northern Iran. First, ET_0 was calculated using the standard FAO Penman–Monteith method and eight empirical models classified into temperature-, radiation-, and humidity-based groups. The RF algorithm was then trained using different combinations of meteorological variables and a digital elevation model (DEM) to reproduce and improve ET_0 . Model performance was evaluated using RMSE, NRMSE, MBE, Nash–Sutcliffe efficiency (NS), and Willmott's index of agreement (d). The results showed that radiation-based models tended to overestimate ET_0 , whereas temperature-based models generally underestimated it relative to the reference method. In contrast, the RF models showed improved performance across all scenarios and reduced systematic errors. The highest statistical accuracy was obtained for the radiation-based RF3 scenario incorporating DEM, with an NS value of 0.99 and the lowest RMSE of 0.18. Sensitivity analysis indicated that the importance of input variables varied among scenarios; maximum temperature showed the highest sensitivity in RF1 and RF2, sunshine duration was the most influential variable in RF3, and relative humidity had the dominant effect in RF4. By combining an ensemble of decision trees, the RF algorithm effectively captured non-linear relationships between input variables and ET_0 . Nevertheless, model selection should consider not only statistical accuracy but also data availability and the intended application.

Cite this article: Gholami Sefidkouhi, M. A., Jafari Sayadi, F. and Noghankar, H., 2026. Assessment and Improvement of Reference Evapotranspiration Estimation under Data-Scarce Conditions Using Random Forest Algorithm. *Journal of Water Research in Agriculture*, 40(1), pp. 19-33.



DOI: <https://doi.org/10.22092/jwra.2026.372732.1115>



نشریه پژوهش آب در کشاورزی

<https://wra.areeo.ac.ir>

ارزیابی و بهبود برآورد تبخیر- تعرق مرجع در شرایط محدودیت داده‌های هواشناسی با استفاده از الگوریتم جنگل تصادفی

محمدعلی غلامی سفیدکوهی* , فاطمه جعفری صیادی و هما نوغان‌کار

استاد گروه مهندسی آب، دانشکده مهندسی زراعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری.

ma.gholami@sanru.ac.ir

پژوهشگر پس‌دکتری علوم و مهندسی آب، گروه مهندسی آب، دانشکده مهندسی زراعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری.

f.sayadi@sanru.ac.ir

دکتری علوم و مهندسی آب، شرکت سهامی آب منطقه‌ای مازندران.

noghankar.h@gmail.com

اطلاعات مقاله	چکیده
مقاله پژوهشی	این پژوهش باهدف ارزیابی و بهبود برآورد تبخیر-تعرق مرجع در شرایط محدودیت داده، با استفاده از الگوریتم جنگل تصادفی انجام شد. به این منظور، داده‌های روزانه ۱۷ ایستگاه هواشناسی استان مازندران طی دوره آماری ۲۰۰۰ تا ۲۰۲۵ استفاده شد. ابتدا تبخیر-تعرق مرجع با روش استاندارد فائو پنمن-مانتیت و هشت مدل تجربی در سه گروه دما-محور، تابش - محور و رطوبت -محور محاسبه شد. سپس الگوریتم جنگل تصادفی با استفاده از ترکیب‌های مختلف متغیرهای هواشناسی و مدل رقومی ارتفاع (DEM) برای بازتولید و بهبود برآورد تبخیر-تعرق آموزش داده شد. عملکرد مدل‌ها با استفاده از شاخص‌های RMSE، NRMSE، MBE، NS و d ارزیابی شد. نتایج نشان داد که مدل‌های تابش-محور تمایل به بیش‌برآوردی و مدل‌های دما-محور تمایل به کم‌برآوردی نسبت به روش مرجع دارد. در مقابل، مدل جنگل تصادفی در تمامی سناریوها عملکرد بهتری ارائه داد و توانست خطاهای متداول را کاهش دهد. بیشترین دقت آماری مربوط به سناریوی تابش-محور (RF3) بود که مقدار NS برابر با ۰/۹۹ و کمترین مقدار RMSE برابر با ۰/۱۸ را نشان داد. تحلیل حساسیت نیز نشان داد که اهمیت متغیرهای ورودی به سناریوی مدل وابسته است؛ به‌طوری که دمای بیشینه در RF1 و RF2 و ساعات آفتابی در RF3 بیشترین حساسیت را داشت، درحالی‌که رطوبت نسبی در RF4 اثر غالب را نشان داد. الگوریتم جنگل تصادفی با ایجاد مجموعه‌ای از درخت‌های تصمیم و ترکیب خروجی آن‌ها، قادر به مدل‌سازی روابط غیرخطی بین متغیرهای ورودی و ET_0 است. با این حال، انتخاب مدل مناسب باید علاوه بر دقت آماری، بر اساس میزان دسترسی به داده‌ها و هدف کاربردی مطالعه انجام شود.
واژه‌های کلیدی: تابش-محور، دما-محور، رطوبت-محور، فائو پنمن-مانتیت، مدل رقومی ارتفاع، مدل‌های تجربی	
آدرس ایمیل نویسنده مسئول: ma.gholami@sanru.ac.ir	

استناد: غلامی سفیدکوهی، محمدعلی، جعفری صیادی، فاطمه، و نوغان‌کار، هما، ۱۴۰۵. ارزیابی و بهبود برآورد تبخیر-تعرق مرجع در شرایط محدودیت

داده‌های هواشناسی با استفاده از الگوریتم جنگل تصادفی. نشریه پژوهش آب در کشاورزی، ۴۰ (۱)، صص. ۳۳-۱۹



DOI: <https://doi.org/10.22092/jwra.2026.372732.1115>

مقدمه

تبخیر - تعرق به عنوان یکی از مؤلفه‌های کلیدی چرخه هیدرولوژیکی، نقش مهمی در تبادل آب و انرژی بین سطح زمین و جو ایفا می‌کند و پس از بارش، یکی از مهم‌ترین فرایندهای مؤثر بر توازن منابع آب محسوب می‌شود (تونال و همکاران، ۲۰۲۳). برآورد دقیق این عامل در مطالعاتی نظیر ترازنامه هیدرولوژیکی، طراحی و مدیریت سامانه‌های آبیاری، ارزیابی رشد و عملکرد گیاهان و برنامه‌ریزی منابع آب از اهمیت بالایی برخوردار است (امانی و شفیع‌زاده، ۲۰۲۳؛ تونال و همکاران، ۲۰۲۳). تبخیر - تعرق عاملی پیچیده و به شدت وابسته به شرایط اقلیمی و جغرافیایی است که تحت تأثیر متغیرهایی نظیر دمای هوا، رطوبت نسبی، سرعت باد، تابش خورشیدی و ویژگی‌های سطحی قرار دارد (پریستلی و تیلور، ۱۹۷۲؛ هارگریوز و سامانی، ۱۹۸۵؛ چو و همکاران، ۲۰۱۷؛ وانگ و همکاران، ۲۰۱۹؛ لیو و همکاران، ۲۰۲۱؛ رشید و همکاران، ۲۰۲۱؛ مندل و چاندا، ۲۰۲۳).

در طول دهه‌های اخیر، روش‌های متعددی برای برآورد تبخیر - تعرق شامل مدل‌های تجربی، نظریه و ترکیبی توسعه یافته‌اند که هر یک از نظر دقت و نیاز داده‌ای تفاوت‌های قابل توجهی دارد (طبری و همکاران، ۲۰۱۳؛ کیو و همکاران، ۲۰۱۹؛ ولی‌پور و همکاران، ۲۰۲۰). در این میان، پنمن - مانیتث فائو^۱ به عنوان روش استاندارد و دقیق برآورد تبخیر - تعرق مرجع در طیف وسیعی از شرایط اقلیمی شناخته می‌شود (آلن و همکاران، ۱۹۹۸؛ فنگ و همکاران، ۲۰۱۷؛ بای و همکاران، ۲۰۲۱). با این حال، کاربرد این روش مستلزم دسترسی به مجموعه کاملی از داده‌های هواشناسی است که در بسیاری از مناطق، به ویژه در نواحی با شبکه ایستگاهی محدود، به طور کامل در دسترس نیست. محدودیت در دسترسی به داده‌های هواشناسی یکی از چالش‌های اساسی در برآورد تبخیر - تعرق محسوب می‌شود و موجب شده استفاده از روش‌های جایگزین با نیاز داده‌ای کمتر به یک ضرورت عملی تبدیل

شود. در چنین شرایطی، مدل‌های ساده‌تر مبتنی بر متغیرهای محدود، نظیر روش‌های دما محور، تابش محور و رطوبت محور، به عنوان جایگزین مورد استفاده قرار می‌گیرد، هرچند دقت آن‌ها به شرایط اقلیمی منطقه وابسته بوده و با عدم قطعیت همراه است (چو و همکاران، ۲۰۱۷؛ نوری و همکاران، ۲۰۱۸؛ رشید و همکاران، ۲۰۲۱).

در پاسخ به این محدودیت‌ها، در سال‌های اخیر استفاده از روش‌های داده‌محور و یادگیری ماشین در برآورد تبخیر - تعرق، رشد چشمگیری داشته است. این روش‌ها قادرند روابط غیرخطی میان متغیرهای هواشناسی و تبخیر - تعرق را بدون نیاز به فرضیات سخت‌گیرانه آماری مدل‌سازی کنند و در شرایط کمبود داده عملکرد مناسبی از خود نشان دهند (وانگ و همکاران، ۲۰۱۹؛ لیو و همکاران، ۲۰۲۱؛ لی و همکاران، ۲۰۲۲). در این میان، الگوریتم جنگل تصادفی^۲ به دلیل دقت بالا، امکان کنترل پراکندگی داده‌ای و توانایی کار با داده‌های ناقص، یکی از روش‌های پرکاربرد در مطالعات اخیر برآورد تبخیر - تعرق محسوب می‌شود (ژانگ و همکاران، ۲۰۱۳؛ فنگ و همکاران، ۲۰۱۷؛ کلودوالوز و همکاران، ۲۰۱۹). با وجود پیشرفت‌های صورت گرفته، همچنان یکی از خلأهای مهم پژوهشی، بررسی نظام‌مند عملکرد روش‌های مختلف در شرایط محدودیت داده و ارزیابی تأثیر ساختار و نوع داده‌های ورودی بر دقت مدل‌ها است. به ویژه، مقایسه هم‌زمان روش‌های تجربی و مدل‌های یادگیری ماشین در قالب دسته‌بندی مشخص متغیرهای هواشناسی کمتر مورد توجه قرار گرفته است.

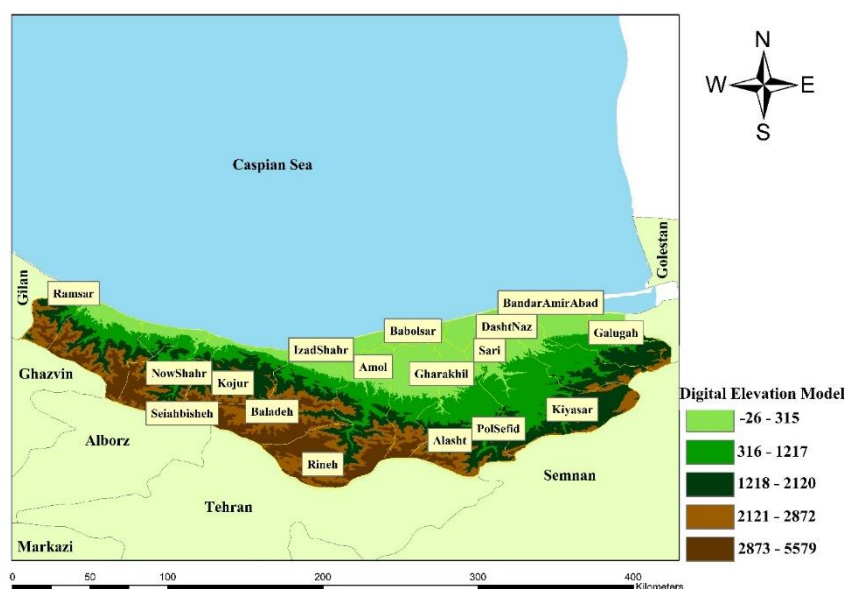
بر این اساس، در پژوهش حاضر، روش‌های برآورد تبخیر - تعرق در چهار دسته شامل روش استاندارد، دما - محور، تابش - محور و رطوبت - محور طبقه‌بندی شد. سپس تبخیر - تعرق مرجع با استفاده از روش استاندارد و روش‌های منتخب در هر دسته محاسبه و عملکرد آن‌ها با استفاده از شاخص‌های آماری RMSE، MBE، NRMSE و NS و d مورد ارزیابی قرار گرفت. در

مواد و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه این پژوهش، استان مازندران با ۱۷ ایستگاه همدید (سینوپتیک) فعال است (شکل ۱). این استان در حاشیه جنوبی دریای خزر و دامنه شمالی رشته کوه البرز واقع شده و به دلیل موقعیت جغرافیایی خاص، دارای تنوع اقلیمی قابل توجهی است. میانگین بارش سالانه در این منطقه حدود ۶۵۰ میلی‌متر و میانگین دمای سالانه حدود ۱۷ درجه سانتی‌گراد گزارش شده است (برارخانپور و همکاران، ۲۰۲۴). از نظر پستی و بلندی، ارتفاع منطقه از حدود ۲۶- متر در نواحی ساحلی تا ۵۵۷۹ متر در قله دماوند متغیر است. این تنوع ارتفاعی در کنار تغییرات اقلیمی، موجب نوسانات قابل توجه در مقدار تبخیر - تعرق در سطح استان شده است؛ بنابراین، انتخاب رویکردهای مناسب برای برآورد دقیق تبخیر - تعرق در این منطقه ضروری هست.

ادامه، به منظور بررسی عملکرد مدل‌های داده‌محور در شرایط محدودیت داده، الگوریتم جنگل تصادفی با استفاده از ترکیب‌های مختلف داده‌ای توسعه یافت. همچنین، به دلیل اهمیت پستی و بلندی در منطقه مورد مطالعه، مدل رقومی ارتفاع (DEM^۱) به عنوان متغیر کمکی در مدل‌ها لحاظ شد. در نهایت، ترکیب‌های داده‌ای مناسب و روش‌های برتر برای برآورد و روندیابی مکانی - زمانی تبخیر - تعرق معرفی شد. نوآوری این پژوهش در ترکیب هم‌زمان طبقه‌بندی فیزیکی روش‌های برآورد تبخیر - تعرق با الگوریتم یادگیری ماشین جنگل تصادفی و استفاده از مدل رقومی ارتفاع در شرایط محدودیت داده‌های هواشناسی است.



شکل ۱- محدوده مورد مطالعه و ایستگاه‌های همدید

Figure 1- Study Area and Senoptic Stations

داده‌های مورد استفاده

در این پژوهش از داده‌های روزانه ایستگاه‌های سینوپتیک شامل دمای بیشینه، دمای کمینه، رطوبت نسبی،

سرعت باد و ساعات آفتابی در دوره آماری ۲۰۰۰-۲۰۲۵ استفاده شد. منظور از محدودیت داده در این پژوهش، وجود داده‌های گمشده در ایستگاه‌های همدید نیست؛ بلکه

^۱ -Digital Elevation Model

$$ET_0 = 0.000939(T_{avg} + 37.5)(T_{max} - T_{min})^{0.424} R_a \quad (3)$$

$$ET_0 = 0.000816(T_{avg} + 33.9)(T_{max} - T_{min})^{0.296} R_a \quad (4)$$

$$R_a = \frac{24 \times 60}{\pi} G_{SC} d_r [\omega_s \sin(\varphi) \sin(\delta) + \cos(\varphi) \cos(\delta) \sin(\omega_s)] \quad (5)$$

که در آن‌ها، ET_0 تبخیر - تفرق مرجع (mm.d^{-1})، T_{avg} میانگین محاسباتی دمای بیشینه و کمینه ($^{\circ}\text{C}$)، T_{max} دمای بیشینه ($^{\circ}\text{C}$)، T_{min} دمای کمینه ($^{\circ}\text{C}$) و R_a تابش خورشیدی در بالای جو روزانه ($\text{MJm}^{-2}\text{d}^{-1}$)، G_{SC} تابش خورشیدی ثابت که معادل 0.082 ($\text{MJm}^{-2}\text{min}^{-1}$) بوده، d_r فاصله نسبی معکوس بین زمین و خورشید، δ انحنای خورشید φ (rad)، عرض جغرافیایی (rad) و ω_s زاویه ساعت غروب خورشید (rad) است.

روش‌های تابش - محور: از سه مدلی که به طور گسترده و با صحت بالا قادر به تخمین تبخیر - تفرق مرجع با استفاده از داده‌های دما و تابش خورشیدی هستند، شامل رابطه Ritchie (Rit)، Priestley-Taylor (P-T) و Makkink (Makk)، به ترتیب روابط ۶، ۷ و ۸ استفاده شد.

$$ET_0 = a_1 \times [0.00387 R_s (0.6 T_{max} + 0.4 T_{min} + 29)] \quad (6)$$

اگر

$$\begin{aligned} 5 < T_{max} \leq 35^{\circ}\text{C} &\rightarrow a_1 = 1.1; \\ T_{max} > 35^{\circ}\text{C} &\rightarrow a_1 = 1.1 + 0.5(T_{max} - 35); \\ T_{max} < 5^{\circ}\text{C} &\rightarrow a_1 = 0.01 \exp(T_{max} + 20). \end{aligned}$$

$$ET_0 = 1.26 \frac{\Delta}{\Delta + \gamma} (R_n - G) \quad (7)$$

$$ET_0 = 0.61 \frac{\Delta}{\Delta + \gamma} (R_s - 0.12) \quad (8)$$

$$R_s = a_s + b_s \frac{n}{N} R_a \quad (9)$$

که در آن‌ها، R_s تابش خورشیدی ($\text{MJm}^{-2}\text{d}^{-1}$)، n مدت‌زمان واقعی غروب خورشید (ساعت)، N بیشینه مدت‌زمان احتمالی غروب خورشید یا طول روز (ساعت)،

استفاده از زیرمجموعه‌های مختلف داده‌های هواشناسی به‌منظور شبیه‌سازی شرایطی است که در بسیاری از ایستگاه‌های غیرهمدید یا مناطق فاقد داده کامل مشاهده می‌شود. از این‌رو، ایستگاه‌های همدید به‌عنوان مرجع کامل داده برای توسعه و ارزیابی سناریوهای مختلف محدودیت داده مورد استفاده قرار گرفت. همچنین به‌منظور در نظر گرفتن اثر پستی و بلندی، مدل رقومی ارتفاع (DEM) نیز به‌عنوان متغیر کمکی وارد مدل‌ها شد.

برآورد تبخیر - تفرق مرجع

برای محاسبه تبخیر - تفرق مرجع، از روش پنمن - مانیت فائو و با جاگذاری مستقیم داده‌ها، استفاده شد. این روش به‌عنوان روش دقیق برآورد تبخیر - تفرق مرجع شناخته می‌شود (فنگ و همکاران، ۲۰۱۷؛ بای و همکاران، ۲۰۲۱) و مبنای مقایسه سایر روش‌ها قرار گرفت:

$$ET_0 = \frac{0.408 \Delta (R_n - G) + \gamma \frac{900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma (1 + 0.34 u_2)} \quad (1)$$

که در آن، ET_0 تبخیر - تفرق مرجع (mm.d^{-1})، R_n تابش خالص در سطح پوشش گیاهی ($\text{MJm}^{-2}\text{d}^{-1}$)، G شار گرما به داخل خاک ($\text{MJm}^{-2}\text{d}^{-1}$)، T متوسط دمای هوا ($^{\circ}\text{C}$)، u_2 سرعت باد در ارتفاع دو متری از سطح زمین (ms^{-1}) و e_a به ترتیب فشار بخار واقعی و اشباع (kPa)، Δ شیب منحنی فشار بخار ($\text{kPa}^{\circ}\text{C}^{-1}$) و γ ضریب رطوبتی ($\text{kPa}^{\circ}\text{C}^{-1}$) است.

طبقه‌بندی روش‌های تجربی تبخیر - تفرق

روش‌های تجربی مورد استفاده در این پژوهش در چهار گروه طبقه‌بندی شد.

روش‌های تجربی دما - محور: مدل Hargreaves (MH-) و Hargreaves-Samani (H-S) 1 & MH-2 اصلاح شده ۱ و ۲ (روابط ۲، ۳ و ۴) که تنها به داده‌های دمای بیشینه و کمینه وابسته هستند.

$$ET_0 = 0.00093(T_{avg} + 17.8)(T_{max} - T_{min})^{0.5} R_a \quad (2)$$

a_s و b_s ثابت رگرسیونی که به ترتیب ۰/۲۵ و ۰/۵ در نظر گرفته می‌شود.

روش‌های رطوبت-محور: از دو رابطه تجربی بر مبنای رطوبت نسبی شامل (R) Romanenko و (S) Schendel (به ترتیب رابطه ۱۰ و ۱۱) که در بسیاری از پژوهش نیز مورد توجه قرار گرفت، استفاده شد.

$$ET_0 = 0.00006(25 + T_{avg})^2(100 - RH) \quad (10)$$

$$ET_0 = 16 \frac{T_{avg}}{RH} \quad (11)$$

که در آن‌ها، RH میانگین درصد رطوبت نسبی هوا است.

مدل یادگیری ماشین جنگل تصادفی

در این پژوهش، برای بررسی اثر محدودیت داده در برآورد تبخیر - تعرق، از الگوریتم جنگل تصادفی (RF) استفاده شد. این الگوریتم به عنوان یک مدل برآورد مستقیم تبخیر - تعرق مرجع به کار گرفته شد؛ بدین صورت که مقادیر محاسبه شده تبخیر - تعرق مرجع به عنوان متغیر هدف در فرایند آموزش مدل تعریف شد و ترکیب‌های مختلفی از متغیرهای هواشناسی به عنوان ورودی مدل مورد استفاده قرار گرفت. این الگوریتم نسبت به داده‌های پرت و پراکنده حساس نبوده و یک روش یادگیری تجمعی مبتنی بر مجموعه‌ای از درخت‌های تصمیم هست که قادر

است تا روابط غیرخطی پیچیده را مدل‌سازی کند. همچنین این روش نیاز به فرض نرمال بودن داده‌ها ندارد و از این نظر برای داده‌های هواشناسی مناسب است. در این مطالعه، چهار سناریوی ورودی مختلف مطابق جدول ۱ تعریف شد که شامل ترکیب‌های متفاوتی از متغیرهای هواشناسی و ارتفاعی است.

به منظور آموزش و ارزیابی مدل جنگل تصادفی، دو روش تقسیم‌بندی تصادفی^۱ و تقسیم‌بندی زمانی^۲ مورد بررسی قرار گرفت. در هر دو روش، داده‌ها به سه مجموعه شامل آموزش، آزمون و اعتبارسنجی تقسیم شد. در روش تقسیم‌بندی تصادفی، ۷۰ درصد داده‌ها به صورت تصادفی برای آموزش، ۲۰ درصد برای آزمون و ۱۰ درصد برای اعتبارسنجی اختصاص یافت. در روش تقسیم‌بندی زمانی نیز همین نسبت‌ها با حفظ توالی زمانی داده‌ها اعمال شد؛ به گونه‌ای که ۷۰ درصد ابتدایی داده‌ها برای آموزش، ۲۰ درصد بعدی برای آزمون و ۱۰ درصد انتهایی برای اعتبارسنجی در نظر گرفته شد. پس از مقایسه عملکرد مدل با دو روش تقسیم‌بندی، روش تصادفی برای تحلیل نهایی انتخاب شد. در هر یک از سناریوها، آموزش مدل جنگل تصادفی به طور مستقل انجام و خروجی آن با مقادیر تبخیر - تعرق مرجع مقایسه شد.

جدول ۱- ترکیب داده‌ای به منظور برآورد تبخیر - تعرق مرجع با استفاده از الگوریتم جنگل تصادفی

نمایه	داده‌های مورد استفاده	مبنای طبقه‌بندی
Index	Used datas	Classification basis
RF1	دمای بیشینه (T_{max}) و کمینه (T_{min})، سرعت باد (u_2)، ساعت آفتابی (N)، رطوبت نسبی (RH) و مدل رقومی زمین (DEM)	استاندارد Standard
RF2	دمای بیشینه (T_{max})، کمینه (T_{min}) و میانگین (T_{avg}) و مدل رقومی زمین (DEM)	دما - محور Temperture Base
RF3	دمای بیشینه (T_{max})، کمینه (T_{min})، ساعت آفتابی (N) و مدل رقومی زمین (DEM)	تابش - محور Radiation Base
RF4	دمای متوسط (T_{avg})، رطوبت نسبی (RH) و مدل رقومی زمین (DEM)	رطوبت - محور Humidity Base

به منظور ارزیابی میزان تأثیر متغیرهای هواشناسی بر برآورد تبخیر - تعرق مرجع، تحلیل حساسیت با استفاده

از ضریب حساسیت نسبی بدون بعد^۳ انجام شد. این روش، امکان مقایسه اثر متغیرهای ورودی با واحدها و دامنه تغییر

متفاوت را بر خروجی مدل فراهم می‌کند (ژو و همکاران، ۲۰۲۰). در این پژوهش، تحلیل حساسیت بر اساس خروجی مدل جنگل تصادفی انجام شد و میزان تغییرپذیری تبخیر - تعرق مرجع در پاسخ به تغییر هر یک از متغیرهای هواشناسی مورد بررسی قرار گرفت. ضریب حساسیت نسبی بدون بعد مطابق رابطه (۱۲) محاسبه شد.

که در آن: S_{vi} ضریب حساسیت نسبی بدون بعد متغیر اقلیمی، ET_0 تبخیر - تعرق مرجع برآوردشده با مدل، V_i متغیر هواشناسی مورد بررسی و $(\partial ET_0 / \partial V_i)$ نرخ تغییر تبخیر - تعرق مرجع نسبت به تغییر متغیر ورودی است. مقادیر مثبت و منفی ضریب حساسیت به ترتیب بیانگر اثر مستقیم و معکوس متغیر بر تبخیر - تعرق مرجع بوده و هرچه مقدار مطلق آن بزرگ‌تر باشد، تأثیر آن متغیر بر خروجی مدل بیشتر است. طبقه‌بندی ضرایب حساسیت مطابق جدول ۲ انجام شد.

برای محاسبه ضریب حساسیت، در هر یک از سناریوهای RF1، RF2، RF3 و RF4، هر متغیر هواشناسی به صورت مستقل در بازه $\pm 10\%$ درصد نسبت به مقدار پایه تغییر داده شد، درحالی‌که سایر متغیرها ثابت نگه داشته شد. سپس تبخیر - تعرق مرجع با مدل جنگل تصادفی مجدداً برآورد و ضریب حساسیت نسبی برای هر متغیر محاسبه شد. این روش اثر تغییر تدریجی هر متغیر بر خروجی مدل را ارزیابی می‌کند و با تحلیل حذف متغیرها^۱ که مبتنی بر حذف کامل یک متغیر است، تفاوت دارد. تحلیل حساسیت تنها برای متغیرهای هواشناسی انجام شد، زیرا مدل رقومی ارتفاع یک متغیر توپوگرافی ثابت است و در طول دوره مطالعه برای هر ایستگاه تغییر زمانی ندارد؛ بنابراین، اعمال تغییرات نسبی و محاسبه ضریب حساسیت برای DEM با روش به‌کاررفته در این پژوهش از نظر مفهومی و فیزیکی توجیه‌پذیر نبود. نقش DEM از طریق مقایسه عملکرد سناریوهای مختلف مدل جنگل تصادفی در بهبود دقت برآورد تبخیر - تعرق مرجع مورد ارزیابی قرار گرفت.

جدول ۲- طبقه‌بندی ضریب حساسیت (Lenhart et al., 2002)

Table 2- Classification of sensitivity coefficient

طبقه	محدوده عددی
Classification	Numerical Range
حساسیت قابل چشم‌پوشی	0 - 0.05
Low Sensitivity	
حساسیت متوسط	0.05 - 0.2
Moderate Sensitivity	
حساسیت بالا	0.2 - 1
High Sensitivity	
حساسیت بسیار بالا	1 <
Very High Sensitivity	

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (ET_{PMF56i} - ET_{Model-i})^2} \quad (13)$$

آماره‌های ارزیابی

روش‌های مختلف تجربی برآورد تبخیر - تعرق مرجع با استفاده از ریشه میانگین مربعات خطا (RMSE)، ریشه میانگین مربعات خطای نرمال‌شده (NRMSE)، میانگین خطای انحراف (MBE)، ضریب کارایی نش - ساتکلیف (NS) و ضریب سازگاری (d) ارزیابی شد.

که بیان‌کننده سازگاری کامل بین مقادیر تخمینی و واقعی است.

نتایج و بحث

وضعیت آب‌وهوایی

نتایج تحلیل داده‌های هواشناسی در دوره ۲۵ ساله (جدول ۳) نشان داد که استان مازندران دارای تنوع قابل توجهی از نظر شرایط اقلیمی است. این تنوع عمدتاً ناشی از تغییرات شدید پستی و بلندی، فاصله از دریا و اختلاف ارتفاع بین مناطق کوهستانی، کوهپایه‌ای، مرکزی و ساحلی است. در میان ایستگاه‌های مورد مطالعه، ایستگاه رینه به دلیل داشتن بیشترین میزان سرعت باد و ساعات آفتابی، بیشترین پتانسیل تبخیر را نشان داد و هم‌زمان کمترین مقدار رطوبت نسبی در این ایستگاه ثبت شد. در مقابل، ایستگاه‌های ساحلی مانند بابلسر، نوشهر و امیرآباد دارای رطوبت نسبی بالاتر و نوسانات دمایی کمتر بودند. از نظر دمایی، بیشترین دمای بیشینه در ایستگاه‌های مرکزی مانند ساری و آمل مشاهده شد، درحالی‌که کمترین مقادیر دما در ایستگاه‌های کوهستانی نظیر بلده و آلاشت ثبت شد. همچنین روند کلی نشان داد که با افزایش ارتفاع، دما و رطوبت نسبی کاهش یافته و در مقابل شدت باد افزایش می‌یابد که با اصول فیزیکی اقلیم کوهستانی همخوانی دارد.

محاسبه تبخیر - تعرق مرجع

نتایج برآورد تبخیر - تعرق مرجع با استفاده از هشت مدل مختلف (جدول ۴) نشان داد که اختلاف قابل توجهی بین روش‌ها در ایستگاه‌های مختلف وجود دارد. به‌طور کلی، روش‌های تابش - محور شامل P-T و Makink بیشترین مقادیر تبخیر - تعرق را در اغلب ایستگاه‌ها برآورد کردند. این رفتار بیانگر تمایل این روش‌ها به بیش‌برآوردی در مقایسه با روش استاندارد پنمن - مانتیث فائو است. در مقابل، روش H-S در اغلب ایستگاه‌ها کمترین مقدار تبخیر - تعرق را ارائه داد که می‌تواند ناشی از عدم استفاده مستقیم از متغیرهای سرعت باد و رطوبت

$$NRMSE = \left(\sqrt{\frac{1}{N} \sum_{i=1}^N (ET_{PMF56i} - ET_{Model-i})^2} / \overline{ET_{PMF56i}} \right) \times 100 \quad (14)$$

$$MBE = \frac{1}{N} \sum_{i=1}^N ET_{Model-i} - ET_{PMF56i} \quad (15)$$

$$NS = 1 - \left[\frac{\sum_{i=1}^N (ET_{PMF56i} - ET_{Model-i})^2}{\sum_{i=1}^N (ET_{PMF56i} - \overline{ET_{PMF56i}})^2} \right] \quad (16)$$

$$d = \left| 1 - \frac{\sum_{i=1}^N (ET_{PMF56i} - ET_{Model-i})^2}{\sum_{i=1}^N (|ET_{Model-i} - \overline{ET_{PMF56i}}| + |ET_{Model-i} - \overline{ET_{Model-i}}|)^2} \right| \quad (17)$$

که در آن‌ها: N تعداد داده‌ها، ET_{PMF56i} i مین برآورد مدل PMF-56 (میلی‌متر در روز)، $\overline{ET_{PMF56i}}$ میانگین برآوردهای مدل PMF-56 (میلی‌متر در روز) و $ET_{Model-i}$ i مین برآورد مدل مورد آزمون (میلی‌متر در روز) و $\overline{ET_{Model-i}}$ میانگین برآوردهای مدل مورد آزمون است. ریشه میانگین مربعات خطا، تفاوت بین مقدار برآورد شده با مدل و مقدار واقعی است. RMSE یک آماره مناسب برای مقایسه خطاهای برآورد در یک مجموعه داده است. هر چه مقدار این آماره کمتر باشد (نزدیک به صفر) مدل تطابق بهتری نسبت به داده‌های مشاهداتی دارد. NRMSE درصد خطای مدل‌ها را نشان می‌دهد. d شاخص سازگاری بوده که نزدیک بودن آن به عدد یک بیانگر نزدیکی مقادیر تخمینی به مقدار واقعی و در نتیجه کارکرد بهتر مدل است. شاخص NS به‌منظور ارزیابی کارایی مدل‌ها استفاده می‌شود. دامنه تغییرات این شاخص بین صفر تا یک است. هر چه شاخص NS به عدد یک نزدیک‌تر باشد، مدل از قدرت برآوردی بیشتری برخوردار است. شاخص MBE با مقایسه اختلاف بین مقادیر تخمینی با مدل و مقادیر واقعی، میزان بیش‌برآوردی یا کم‌برآوردی هر روش را نشان می‌دهد. حد پایین این شاخص صفر است

نسبی باشد. این نتیجه با مطالعات پیشین در اقلیم‌های مشابه
همخوانی دارد و محدودیت این روش در مناطق بادخیز و
متنوع اقلیمی را تأیید می‌کند. به‌طور کلی، نتایج نشان داد که
اختلاف عملکرد مدل‌ها در ایستگاه‌های مختلف ناشی از
ناهمگونی اقلیمی و اثرات خرد اقلیم‌های ناشی از تغییرات
ارتفاعی و فاصله از دریاست.

جدول ۳- مقادیر متوسط روزانه متغیرهای هواشناسی در ایستگاه‌های استان مازندران
Table 3- Daily average of meteorological variables at Mazandaran meteorological stations

ایستگاه Satation	خوشه Cluster	ارتفاع Height	دمای بیشینه T _{max}	دمای کمینه T _{min}	سرعت باد u ₂	رطوبت نسبی RH	ساعات آفتابی N	تبخیر از تشت تبخیر E
		متر m	سانتی‌گراد °C		متر بر ثانیه m.s ⁻¹	درصد %	ساعت hr	میلی‌متر در روز mm.d ⁻¹
بلده	کوهستانی Mountainous	2120	16.68	3.03	3.54	52.49	7.68	5.82
رینه		2130	16.93	6.61	4.38	47.83	8.05	6.09
آلاشت	کوهپایه‌ای Submountainous	1805	15.63	7.06	1.55	66.33	5.95	4.31
سیاه‌بیشه		1855	15.73	6.67	2.51	60.47	5.36	4.32
کجور		1550	18.21	2.63	2.62	59.08	6.71	5.56
پل سفید	میانی Middle	610	20.48	10.68	2.65	68.61	5.06	3.97
کیاسر		12.94	18.63	8.52	2.29	69.92	6.31	4.84
آمل	مرکزی Centric	24	22.34	13.26	1.86	80.17	5.26	3.39
دشت‌ناز		17	22.91	13.05	2.46	78.17	6.13	4.04
ساری		23	23.14	13.71	1.47	77.73	5.54	3.72
قراخیل		15	22.47	12.89	1.72	81.21	5.58	3.71
ایزدشهر		-26	21.79	14.59	3.16	78.75	5.58	3.69
بابلسر	ساحلی Coastal	-21	21.81	14.74	1.84	79.41	5.79	3.54
بندر امیرآباد		-20	21.74	14.20	2.69	81.11	6.64	4.24
رامسر		-20	20.41	14.17	1.88	80.57	4.65	3.46
گلوگاه		-10	22.84	12.86	1.71	74.12	5.51	3.74
نوشهر		-21	20.51	13.69	2.08	81.28	5.31	3.49

جدول ۴- مقادیر متوسط روزانه تبخیر - تعرق مرجع در ایستگاه‌ها بر اساس مدل‌های مختلف در دوره شاخص (میلی‌متر در روز)

Table 4- Daily average of reference evapotranspiration at stations based on different models in the basic period (mm.d⁻¹)

Table 4 - Daily average of reference evapotranspiration at stations based on different models in the basic period (mm/d)													ایستگاه	
رطوبت - محور Humidity Base			تابش - محور Radiation Base				دما - محور Temperture Base				استاندارد Standard		خوشه Cluster	satation
S	R	RF4	Makk	P-T	Rit	RF3	MH-2	MH-1	H-S	RF2	FAO-56	RF1		
3.65	3.84	3.33	4.53	5.46	2.16	2.53	2.37	4.12	3.04	3.39	3.42	3.37	بلده	کوهستانی
5.26	4.64	4.48	6.51	7.02	3.08	2.51	2.32	3.88	2.87	2.94	4.28	4.57	رینه	Mountainous
3.49	3.02	3.25	4.76	5.64	2.27	2.76	2.11	3.42	2.47	3.04	2.47	2.21	آلاشت	کوهپایه‌ای
3.75	3.81	3.67	5.51	6.27	2.63	2.58	2.16	3.57	2.61	2.46	3.21	3.51	سیاه‌بیشه	Submountainous
4.14	3.81	4.01	3.99	5.05	1.96	2.97	2.41	4.11	3.12	3.94	2.89	3.34	کجور	
4.08	3.32	3.69	4.63	5.69	2.39	2.63	2.37	3.92	3.01	3.38	2.91	2.83	پل سفید	میانی
3.58	2.93	3.13	6.04	6.75	3.02	3.04	2.34	3.89	2.94	3.11	3.11	3.56	کیاسر	Middle
3.62	2.34	3.45	5.62	6.81	3.01	3.69	2.41	3.89	3.02	3.19	2.68	2.26	آمل	مرکزی
3.83	2.68	3.23	6.11	7.18	3.35	3.97	2.49	4.09	3.21	3.42	3.16	3.26	دشت‌ناز	Centric
3.93	2.75	2.98	5.85	6.97	3.23	3.85	2.47	4.03	3.16	3.75	2.82	2.58	ساری	
3.58	2.25	3.01	5.81	6.94	3.12	3.81	2.45	3.99	3.11	3.54	2.75	2.16	قراخیل	
3.81	2.57	2.85	3.01	4.41	3.56	3.12	2.26	3.56	2.72	2.84	2.11	2.02	ایزدشهر	ساحلی
3.75	2.47	2.98	6.02	7.29	3.14	4.02	2.01	3.32	2.22	2.76	2.83	2.31	بابلسر	Coastal
3.64	2.32	3.11	6.29	7.53	3.29	4.09	2.31	3.65	2.79	2.89	3.01	3.11	بندر امیرآباد	
3.52	2.26	3.06	5.28	6.49	2.74	2.87	2.13	3.31	2.47	2.77	3.53	2.19	رامسر	
4.07	3.18	3.29	4.69	5.94	2.56	2.95	2.49	4.11	3.22	4.04	2.51	2.11	گلوگاه	
3.43	2.12	3.23	5.58	6.75	2.89	3.04	2.18	3.41	2.57	2.94	2.58	2.36	نوشهر	

RF1: ترکیب استاندارد جنگل تصادفی، FAO-56: پنمن - مانیتث فائو، RF2: ترکیب دما - محور جنگل تصادفی، H-S: هارگریوز - سامانی، MH-1: هارگریوز اصلاح‌شده-۱، MH-2: هارگریوز اصلاح‌شده-۲، RF3: ترکیب تابش - محور جنگل تصادفی، Rit:

ریچی، P-T: پریستلی - تیلور، Makk: مک‌کینک، RF4: ترکیب رطوبت - محور جنگل تصادفی، R: رومنس و S: چندل.

نتایج حاصل از مدل‌های یادگیری ماشین نشان داد که الگوریتم جنگل تصادفی توانایی بالایی در بازتولید تبخیر - تعرق مرجع تحت سناریوهای مختلف داده‌ای دارد. بر اساس شکل ۲، مقدار ضریب نش - ساتکلیف (NS) برای سناریوهای استاندارد (RF1)، دما - محور (RF2)، تابش - محور (RF3) و رطوبت - محور (RF4) به ترتیب برابر با ۰/۹۲، ۰/۹۷، ۰/۹۹ و ۰/۹۱ به دست آمد. این نتایج بیانگر آن است که الگوریتم جنگل تصادفی در تمامی شرایط ورودی، عملکرد مطلوبی در شبیه‌سازی تبخیر - تعرق مرجع دارد. از منظر شاخص‌های آماری، سناریوی RF3 با بیشترین مقدار NS و کمترین مقادیر RMSE و NRMSE، بهترین عملکرد را ارائه داد؛ با این حال، برتری آماری لزوماً به معنای مناسب‌ترین گزینه عملیاتی نیست و انتخاب مدل باید متناسب با سطح دسترسی به داده‌ها و هدف کاربردی صورت گیرد. این موضوع بیانگر نقش مؤثر متغیرهای تابشی در بازنمایی فرایند تبخیر - تعرق است. در مقابل، مدل RF2 نیز با استفاده از داده‌های محدودتر دمایی و مدل رقومی ارتفاع، عملکرد بسیار مناسبی از خود نشان داد و توانست به تعادل مطلوبی میان دقت و نیاز داده‌ای دست یابد. این یافته نشان می‌دهد که کاهش تعداد متغیرهای ورودی لزوماً منجر به افت شدید دقت نمی‌شود، مشروط بر آنکه روابط غیرخطی میان متغیرها به درستی توسط مدل یادگیری ماشین شناسایی شوند.

نتایج شاخص MBE نشان داد که برخی از روابط تجربی دارای سوگیری سامان‌مند هستند؛ به طوری که مدل‌های تابش - محور عمدتاً تمایل به بیش‌برآوردی و مدل‌های دما - محور تمایل به کم‌برآوردی تبخیر - تعرق مرجع دارد. در مقابل، مدل‌های مبتنی بر جنگل تصادفی سوگیری کمتری از خود نشان دادند و مقادیر برآوردی آن‌ها به مقادیر مرجع نزدیک‌تر بود. این موضوع بیانگر توانایی بالاتر مدل‌های داده‌محور در اصلاح خطاهای ساختاری موجود در روابط تجربی است.

نکته مهم در تفسیر نتایج آن است که اگرچه بخش عمده‌ای از مدل‌های مورد بررسی دارای مقادیر NS

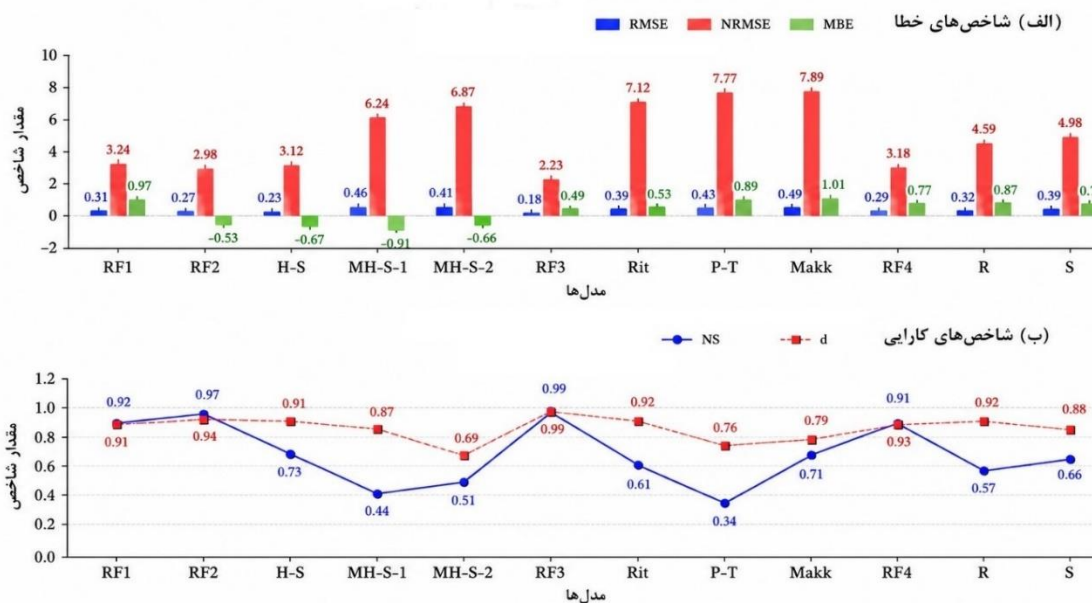
بیش‌تر از ۰/۵۰ بوده و از منظر آماری قابل قبول محسوب می‌شوند، اما معیار انتخاب مدل مناسب تنها به بیشینه‌سازی شاخص‌های دقت محدود نمی‌شود. در شرایط واقعی، عواملی نظیر تعداد متغیرهای موردنیاز، سهولت دسترسی به داده‌ها، هزینه جمع‌آوری اطلاعات و قابلیت تعمیم‌پذیری مدل نیز نقش تعیین‌کننده دارند. برای مثال، مدل (H-S) با وجود استفاده از حداقل داده‌های دمایی، مقدار NS برابر ۰/۷۳ را نشان داد که بیانگر کارایی قابل قبول آن در شرایط محدودیت شدید داده است. به همین ترتیب، RF2 با ثبت NS برابر ۰/۹۷ می‌تواند به عنوان یک گزینه عملیاتی و دقیق در بسیاری از مناطق مورد استفاده قرار گیرد.

نتایج تحلیل حساسیت متغیرهای هواشناسی در سناریوهای مختلف در جدول ۵ ارائه شده است. مطابق نتایج، ضرایب حساسیت متغیرهای دمایی، سرعت باد و ساعات آفتابی در تمامی سناریوها مثبت بوده که بیانگر افزایش تبخیر - تعرق مرجع با افزایش این متغیرها است. در مقابل، ضریب حساسیت منفی رطوبت نسبی نشان می‌دهد که افزایش رطوبت هوا موجب کاهش تبخیر - تعرق مرجع می‌شود. این نتایج با مبانی فیزیکی فرایند تبخیر - تعرق و معادله فائو پنمن - مانیتث سازگار است.

در سناریوی RF1، بیشترین ضریب حساسیت (۰/۵۳) مربوط به دمای بیشینه بود و پس از آن رطوبت نسبی (۰/۳۲-) و سرعت باد (۰/۲۸) قرار گرفتند که نشان‌دهنده نقش غالب دمای بیشینه در واپایش تبخیر - تعرق مرجع در این سناریو است. در سناریوی RF2 نیز دمای بیشینه با ضریب حساسیت ۰/۶۲ بیشترین تأثیر را بر خروجی مدل داشت. در سناریوی RF3، ساعات آفتابی با ضریب حساسیت ۰/۶۱ مؤثرترین متغیر بود و پس از آن دمای بیشینه (۰/۴۴) و دمای کمینه (۰/۱۷) قرار گرفتند. این نتیجه بیانگر آن است که در سناریوی مبتنی بر متغیرهای تابشی، تغییرات انرژی دریافتی نقش تعیین‌کننده‌ای در برآورد تبخیر - تعرق مرجع دارد و علت برتری عملکرد این سناریو نسبت به سایر سناریوها نیز تا حد زیادی به همین موضوع اشاره دارد.

نسبی هر متغیر تابعی از ترکیب متغیرهای ورودی در هر سناریو بوده و هیچ متغیری را نمی‌توان به‌عنوان مؤثرترین عامل در تمامی شرایط معرفی کرد. یافته‌ها نشان داد حساسیت خروجی تبخیر - تعرق مرجع نسبت به متغیرهای مختلف و به‌ویژه عناصر هواشناسی بسته به مکان و زمان متغیر است و با نتایج (یحیوی و همکاران، ۱۴۰۲؛ نصرالهی و همکاران، ۱۴۰۱؛ ژو و همکاران، ۲۰۲۰) همسو هست.

در سناریوی RF4، رطوبت نسبی با ضریب حساسیت ۰/۳۷ - بیشترین اثر را بر خروجی مدل داشت و دمای میانگین نیز با ضریب ۰/۲۳ در رتبه بعدی قرار گرفت. در مجموع، تمامی ضرایب حساسیت در محدوده حساسیت متوسط تا زیاد قرار گرفتند که نشان می‌دهد خروجی مدل جنگل تصادفی نسبت به تغییرات متغیرهای هواشناسی واکنش قابل‌توجهی دارد. همچنین، نتایج نشان داد اهمیت



شکل ۲- مقایسه عملکرد مدل‌های مختلف برآورد تبخیر - تعرق مرجع بر اساس (الف) شاخص‌های خطا و (ب) شاخص‌های کارایی
Figure 2- Comparative performance evaluation of different reference evapotranspiration estimation models based on: (a) error metrics and (b) efficiency criteria

جدول ۵- ضریب حساسیت سناریوهای مختلف به متغیرهای هواشناسی

Table 5- Sensitivity coefficient of meteorological variables in different scenario

متغیرهای هواشناسی Meteorological variables						سناریو Scenario
رطوبت نسبی RH	ساعت آفتابی N	سرعت باد u ₂	دمای میانگین T _{Avg}	دمای کمینه T _{Min}	دمای بیشینه T _{Max}	
-0.32	0.19	0.28	-	0.13	0.53	RF1
-	-	-	0.37	0.25	0.62	RF2
-	0.61	-	-	0.17	0.44	RF3
-0.37	-	-	0.23	-	-	RF4

بهبودی و همکاران (۱۴۰۰) که بر نقش شرایط خرد اقلیمی در تغییرات تبخیر - تعرق تأکید کرده‌اند، همخوانی دارد. مقایسه روش‌های تجربی نشان داد که مدل‌های تابش - محور عمدتاً تمایل به بیش‌برآوردی و مدل‌های دما - محور تمایل به کم‌برآوردی نسبت به روش مرجع دارند. این تفاوت را می‌توان به ساختار متفاوت متغیرهای ورودی

نتایج نشان داد که تبخیر - تعرق مرجع در استان مازندران تحت تأثیر ناهمگنی‌های اقلیمی و توپوگرافی ناشی از تغییرات ارتفاع، فاصله از دریا و توپوگرافی قرار دارد. تفاوت رفتار تبخیر - تعرق در ایستگاه‌های ساحلی، مرکزی و کوهستانی نیز مؤید این موضوع است. این یافته با نتایج طبری و همکاران (۲۰۱۱)، رنجبر و ابراهیمی (۱۴۰۴) و

و میزان توجه هر روش به مؤلفه‌های انرژی و آیرودینامیکی نسبت داد. در مقابل، عملکرد بهتر جنگل تصادفی را می‌توان به توانایی آن در شناسایی روابط غیرخطی و تعامل بین متغیرهای هواشناسی نسبت داد؛ موضوعی که در مطالعات ویسی و همکاران (۲۰۲۵)، گویال و همکاران (۲۰۲۳)، شارما و همکاران (۲۰۲۲)، گراناتا و دی‌نانو (۲۰۲۱) و چن و همکاران (۲۰۲۰) نیز گزارش شده است.

نتایج تحلیل حساسیت نشان داد که اهمیت متغیرهای هواشناسی در سناریوهای مختلف جنگل تصادفی متفاوت است. در سناریوهای RF1 و RF2، دمای بیشینه بیشترین تأثیر را بر ET_0 داشت، درحالی‌که در RF3، ساعات آفتابی مؤثرترین متغیر بود. در RF4 نیز رطوبت نسبی بیشترین حساسیت را نشان داد که بیانگر رابطه معکوس آن با ET_0 است. این نتایج نشان می‌دهد که اهمیت نسبی متغیرهای ورودی به ترکیب داده‌های هر سناریو وابسته است. همچنین، حساسیت بالاتر RF3 نسبت به ساعات آفتابی با عملکرد آماری برتر این سناریو همخوانی دارد و نقش مؤلفه‌های تابشی را در بهبود برآورد ET_0 تأیید می‌کند. با این حال، از آنجاکه DEM در تمامی سناریوها استفاده شده، نمی‌توان سهم مستقل آن را از طریق تحلیل حساسیت تعیین کرد.

از دیدگاه کاربردی، یافته‌های این پژوهش نشان می‌دهد که انتخاب مدل نباید صرفاً بر اساس بالاترین شاخص‌های آماری انجام شود. از نظر آماری RF3 بهترین عملکرد را داشت، اما کاربرد آن به دسترسی به متغیرهای تابشی و سایر داده‌های موردنیاز وابسته است. در شرایطی که داده‌های تابشی یا ساعات آفتابی در دسترس نباشد، RF2 یا سایر سناریوهای ساده‌تر می‌توانند با وجود دقت کمتر، گزینه‌های عملی‌تری باشد؛ بنابراین، انتخاب مدل مناسب باید هم‌زمان بر اساس دقت آماری، نوع و تعداد داده‌های در دسترس و هدف کاربردی مطالعه انجام شود.

مقایسه نتایج این پژوهش با رویکرد بکار گرفته شده در مطالعه ویسی و همکاران (۲۰۲۵) برای شمال

کشور نیز نشان می‌دهد که هر دو رویکرد باهدف کاهش محدودیت داده‌های هواشناسی در برآورد ET_0 توسعه یافته است؛ اما سازوکار جبران کمبود داده در آن‌ها متفاوت است. در رویکرد مبتنی بر داده‌های تحلیل مجدد، پوشش مکانی - زمانی پیوسته امکان استفاده در مناطق فاقد ایستگاه را فراهم می‌کند؛ درحالی‌که رویکرد حاضر با استفاده از داده‌های مشاهداتی و جنگل تصادفی، روابط موجود بین متغیرهای در دسترس و ET_0 را در مقیاس محلی یاد می‌گیرد. از این رو، روش حاضر می‌تواند برای بهبود برآوردهای محلی در مناطق دارای ایستگاه‌های هواشناسی مناسب باشد، درحالی‌که داده‌های تحلیل مجدد می‌توانند مزیت بیشتری برای پوشش منطقه‌ای و نواحی فاقد ایستگاه داشته باشد. در نتیجه، این دو رویکرد را می‌توان مکمل یکدیگر دانست و انتخاب میان آن‌ها باید متناسب با مقیاس مطالعه، هدف کاربردی و سطح دسترسی به داده‌های مشاهداتی صورت گیرد.

نتیجه‌گیری

این پژوهش نشان داد که برآورد تبخیر - تعلق مرجع در یک منطقه با ناهمگنی اقلیمی و توپوگرافی، نیازمند رویکردی متناسب با ساختار داده‌های در دسترس است. یافته اصلی آن است که الگوریتم جنگل تصادفی می‌تواند روابط پیچیده و غیرخطی بین متغیرهای هواشناسی، اطلاعات توپوگرافی و تبخیر - تعلق مرجع را به‌خوبی بازنمایی کند و در مقایسه با روابط تجربی، خطای برآورد را کاهش دهد. در میان سناریوهای بررسی شده، RF3 از نظر آماری عملکرد بهتری را نشان داد و به بیشترین تطابق با مقادیر مرجع دست‌یافت. با این حال، این برتری آماری به‌تنهایی به معنای مناسب بودن این سناریو برای همه مناطق و شرایط داده‌ای نیست؛ زیرا انتخاب مدل کاربردی باید هم‌زمان بر اساس دقت مورد انتظار، نوع و کیفیت داده‌های موجود و امکان دسترسی پایدار به متغیرهای ورودی انجام شود. از این منظر، نتیجه مهم‌تر پژوهش، ارائه یک چارچوب انعطاف‌پذیر برای انتخاب

مدل متناسب با شرایط داده‌ای است، نه معرفی یک مدل واحد برای تمامی مناطق.

باوجود عملکرد مناسب مدل‌های جنگل تصادفی، تعمیم نتایج این پژوهش باید با احتیاط انجام شود. مدل‌ها با استفاده از داده‌های ۱۷ ایستگاه هواشناسی استان مازندران و در محدوده شرایط اقلیمی و توپوگرافی این منطقه آموزش داده شد و ارزیابی به انجام رسید؛ بنابراین، عملکرد آن‌ها در مناطق دارای شرایط اقلیمی متفاوت، دامنه ارتفاعی دیگر یا رژیم‌های تابشی متفاوت ممکن است یکسان نباشد. علاوه بر این، عملکرد مدل‌های داده‌محور به کیفیت، طول دوره و پراکندگی مکانی داده‌های آموزشی وابسته است و استفاده از آن‌ها در مناطق فاقد داده یا دارای کمبود شدید داده، بدون داده‌های آموزشی معتبر و نماینده، نیازمند احتیاط است.

در پژوهش‌های آینده، پیشنهاد می‌شود عملکرد مدل جنگل تصادفی با استفاده از داده‌های مستقل و ایستگاه‌های خارج از محدوده آموزش، در مناطق اقلیمی و توپوگرافی متفاوت ارزیابی شود. همچنین، استفاده از

روش‌های تحلیل اهمیت متغیرها و آزمون حذف متغیرها می‌تواند سهم نسبی هر یک از متغیرهای هواشناسی و توپوگرافی را در بهبود برآورد تبخیر - تعرق مرجع مشخص کند. بررسی سایر الگوریتم‌های یادگیری ماشین، روش‌های انتقال‌پذیری مدل بین مناطق، داده‌های تحلیل مجدد و داده‌های سنجش‌ازدور نیز می‌تواند به ارزیابی جامع‌تر قابلیت تعمیم این رویکرد کمک کند. در نهایت، ترکیب دقت آماری مدل‌ها با معیارهای نیاز داده‌ای و هزینه دسترسی به داده‌ها می‌تواند زمینه را برای توسعه سامانه‌های عملیاتی و قابل‌انتقال برآورد تبخیر - تعرق مرجع فراهم سازد.

تشکر و قدردانی

این تحقیق با حمایت مالی دانشگاه علوم کشاورزی و منابع طبیعی ساری با شماره طرح ۰۱-۱۴۰۴-۰۲ انجام شده است. نویسندگان مراتب قدردانی صمیمانه خود را از این حمایت ابراز می‌دارند.

فهرست منابع

۱. بهزادی کریمی، حسین، مظفری، غلامعلی، امیدوار، کمال و مزیدی، احمد، ۱۴۰۰. چشم‌انداز تغییرات زمانی - مکانی تبخیر و تعرق در حوضه آبریز کارون طی دوره‌های آبی و تحت سناریوهای انتشار گازهای گلخانه‌ای. فصلنامه جغرافیای طبیعی، ۱۳(۵۴): ۸۷-۱۰۸. doi:20.1001.1.20085656.1400.14.54.3.6
۲. رنجبر، آرش و ابراهیمی پاک، نیازعلی، ۱۴۰۴. کمی‌سازی اثر پارامترهای اقلیمی بر تبخیر و تعرق مرجع در استان اردبیل. مجله تحقیقات آب‌ونخاک/ایران، ۵۶ (۸). ۲۲۶۵-۲۲۴۵. doi:10.22059/ijswr.2025.397468.669965
۳. نصرالهی، محمد، ذوالفقاری، علی اصغر و یزدانی، محمدرضا، ۱۴۰۱. تحلیل حساسیت تبخیر و تعرق مرجع در اقلیم مرطوب حاشیه جنوبی دریای خزر. تحقیقات کاربردی خاک، ۱۰(۲): ۷۹-۹۰. doi: 10.30466/asr.2022.121217
۴. یحیوی دیزج، آمنه، اکبری ازیرانی، طیب، خالدی، شهریار و جوان، خدیجه، ۱۴۰۲. بررسی فصلی تبخیر تعرق مرجع و تحلیل حساسیت آن به عناصر هواشناسی در ایران. نشریه آب‌ونخاک، ۳۷(۴): ۶۴۳-۶۵۷. doi: 10.22067/jsw.2023.83328.1307
5. 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. Rome. Italy. 300 p.
6. Amani, S. and Shafizadeh-Moghadam, H. 2023. A review of machine learning models and influential factors for estimating evapotranspiration using remote sensing and ground-based data. *Agricultural Water Management*, 284 (3), 108324.

7. **doi:10.1016/j.agwat.2023.108324**
Bai, Y., Zhang, S., Bhattarai, N., Mallick, K., Liu, Q., Tang, L., Im, J., Guo, L. and Zhang, J. 2021. On the use of machine learning based ensemble approaches to improve evapotranspiration estimates from croplands across a wide environmental gradient. *Agricultural and Forest Meteorology*, 298-299, 108308.
8. **doi:10.1016/j.agrformet.2020.108308**
Bararkhanpour Ahmadi, S., Nadi, M., Mazloom Babanari, S. and Jedariforoughi, A. 2024. Analyzing the effect of climate change on the trend of extreme temperatures along the coast of Mazandaran province based on CMIP6 models. *Journal of Water and Soil Conservation*, 30(4), pp. 1-27. **doi:10.22069/jwsc.2023.21185.3635**
9. Berti, A., Tardivo, G., Chiaudani, A., Rech, F. and Borin, M. 2014. Assessing reference evapotranspiration by the Hargreaves method in north-eastern Italy. *Agricultural Water Management*, 140, pp. 20-25. **doi:10.1016/j.agwat.2014.03.015**
10. Chen, Z., Zhu, Z., Jiang, H. and Sun, S. 2020. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. *Journal of Hydrology*, 591, 125286.
doi:10.1016/j.hydrol.2020.125286
11. Chu, R., Li, M., Shen, S., Islam, A.R.M.T., Cao, W., Tao, S. and Gao, P. 2017. Changes in reference evapotranspiration and its contributing factors in Jiangsu, a major economic and agricultural province of eastern China. *Water*, 9 (7). 486. **doi:10.3390/w9070486**
12. Clodoalves da Silva Júnior, J., Medeiros, V., Garrozia, C., Montenegro, A. and Gonçalves, G.E. 2019. Random forest techniques for spatial interpolation of evapotranspiration data from Brazilian's northeast. *Computers and Electronics in Agriculture*, 166, 105017 ,pp. 105-116. **doi:10.1016/j.compag.2019.105017**
13. Feng, Y., Cui, N., Gong, D., Zhang, Q. and Zhao, L. 2017. Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agricultural Water Management*, 193, pp. 163-173.
doi:10.1016/j.agwat.2017.08.003
14. Goyal, P., Kumar, S. and Sharda, R. 2023. A review of the artificial intelligence (AI) based techniques for estimating reference evapotranspiration: Current trends and future perspectives. *Computers and Electronics in Agriculture*, 209, 107836.
doi:10.1016/co.mpag.2023.107836
15. Granata, F. and Di Nunno, F. 2021. Forecasting evapotranspiration in different climate using ensembles of recurrent neural networks. *Agricultural Water Management*, 255, 107040.
doi:10.1016/j.agwat.2021.107040
16. Hargreaves, G.H. and Samani, Z.A. 1985. Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, 1(2), pp. 96-99.
doi:10.13031/2013.26773
17. Lenhart, T., Eckhardt, K., Fohrer, N. and Frede, H.G. 2002. Comparison of two different approaches of sensitivity analysis. *Physics and Chemistry of the Earth*, 27, pp. 645-654.
18. Li, Y., Wang, W., Wang, G. and Tan, Q. 2022. Actual evapotranspiration estimation over the Tuojiang River Basin based on a hybrid CNN-RF model. *Journal of Hydrology*, 610, 127788. **doi:10.1016/j.jhydrol.2022.127788**
19. Liu, Y., Zhang, S., Zhang, J., Tang, L. and Bai, Y. 2021. Assessment and comparison of six machine learning models in estimating evapotranspiration over croplands using remote sensing and meteorological factors. *Remote Sensing*, 13(19), 3838.
doi:10.3390/rs13193838
20. Mandal, N. and Chanda, K. 2023. Performance of machine learning algorithms for multi-step ahead prediction of reference evapotranspiration across various agro-climatic zones and cropping seasons. *Journal of Hydrology*, 620, 129418.
doi:10.1016/j.jhydrol.2023.129418
21. Nouri, M., Homae, M. and Bannayan, M. 2018. Spatiotemporal reference evapotranspiration changes in humid and semi-arid regions of Iran: Past trends and future projections. *Theory of Applied Climatology*, 133, pp. 361-375. **doi:10.1007/s00704-017-2176-8**

22. Priestley, C.H.B. and Taylor, R.J. 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 100, pp. 81-92.
23. Qiu, R., Liu, C., Cui, N., Wu, Y., Wang, Z. and Li, G. 2019. Evapotranspiration estimation using a modified Priestley-Taylor model in a rice-wheat rotation system. *Agricultural Water Management*, 224, 105755. doi:10.1016/j.agwat.2019.105755
24. Rashid Niaghi, A., Hassanijalilian, O. and Shiri, J. 2021. Estimation of reference evapotranspiration using spatial and temporal machine learning approaches. *Hydrology*, 8(1), 25. <https://doi.org/10.3390/hydrology8010025>
25. Sharma, G., Singh, A. and Jain, S. 2022. A hybrid deep neural network approach to estimate reference evapotranspiration using limited climate data. *Neural Computing and Application*, 34, pp. 4013-4032. doi:10.1007/s00521-021-06661-9
26. Tabari, H., Marofi, S., Aeini, A., Talaei, P.H. and Mohammadi, K. 2011. Trend analysis of reference evapotranspiration in the western half of Iran. *Agricultural and Forest Meteorology*, 151(2), pp. 128–136. doi:10.1016/j.agrformet.2010.09.009
27. Tabari, H., Grismer, M.E. and Trajkovic, S. 2013. Comparative analysis of 31 reference evapotranspiration methods under humid conditions. *Irrigation Sciences*, 31, pp. 107-117. doi:10.1007/s00271-011-0295-z
28. Tunal, U., Tüzel, I.H., Tüzel, Y. and Şenol, Y. 2023. Estimation of actual crop evapotranspiration using artificial neural networks in tomato grown in closed soilless culture system. *Agricultural Water Management*, 284, 108331. doi:10.1016/j.agwat.2023.108331
29. Valipour, M., Bateni, S.M., Gholami Sefidkouhi, M.A., Raeini-Sarjaz, M. and Singh, V.P. 2020. Complexity of forces driving trend of reference evapotranspiration and signals of climate change. *Atmosphere*, 11, pp. 1-26. doi:10.3390/atmos11101081
30. Veysi, S., Zarei, S., Galehban, E. and Tahooni, A. 2025. Leveraging artificial intelligence and machine-learning methods for bias correction of reference evapotranspiration utilizing ERA5 data. *Irrigation and Drainage Engineering*, 151(4), 106647. doi:10.1061/jidedh.irenc-10487
31. 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. *Agricultural Water Management*, 221, pp. 220-230. doi:10.1016/j.agwat.2019.03.027
32. Zhang, D., Liu, X. and Hong, H., 2013. Assessing the effect of climate change on reference evapotranspiration in China. *Stochastic Environmental Research and Risk Assessment*, 27(8), pp. 1871–1881. doi:10.1007/s00477-013-0723-0
33. Zhou, P., Liu, Z. and Cheng, L. 2020. An alternative approach for quantitatively estimating climate variability over China under the effects of ENSO events. *Atmospheric Research*, 238, 104897. doi: 10.1016/j.atmosres.2020.104897