نشریه آبیاری و زهکشی ایران
شماره۴، جلد ۲۰، مهر - آبان ۱۴۰۵، ص. ۵۶۵-۵۴۷

Iranian Journal of Irrigation and Drainage
No.4, Vol. 20, Oct.-Nov. 2026, p. 547-565

مقاله علمی-پژوهشی

بهبود پیش‌بینی تبخیرتعرق مرجع با تلفیق داده‌های سنجش از دور و الگوریتم‌های یادگیری ماشین در شمال ایران

زهرا روستا[۱]، سعید صمدیان فرد[۲*]، رضا دلیرحسن‌نیا[۳]، صدرا کریم زاده[۴]

تاریخ دریافت: ۱۴۰۴/۰۸/۲۲ تاریخ پذیرش: ۱۴۰۵/۰۳/۱۸

چکیده

در مواجهه با چالش‌های ناشی از تغییرات اقلیمی و افزایش تقاضای آب، پیش‌بینی دقیق تبخیرتعرق مرجع (ET_0) به عنوان یکی از ارکان مدیریت پایدار منابع آبی اهمیت یافته است. مدل‌های کلاسیک مانند فائو-پنمن-مانتیث، با وجود اعتبار علمی، در مناطق فاقد داده‌های کامل هواشناسی با کاهش دقت روبرو هستند. این پژوهش با هدف ارتقای دقت پیش‌بینی ET_0، از تلفیق داده‌های سنجش از دور (ET_{MODIS}) و الگوریتم‌های یادگیری ماشین بهره گرفته است. دو مدل تقویت گرادیانی شدید (XGB) و نسخه بهبودیافته آن (HXGB) توسعه یافته‌اند که مدل دوم از قابلیت تعمیم بالاتری برخوردار است. داده‌های هواشناسی و پارامترهای ماهواره‌ای مربوط به ایستگاه‌های رامسر و بندرانزلی در بازه ۲۰۰۱ تا ۲۰۲۳ برای آموزش و ارزیابی مدل‌ها استفاده شد. نتایج نشان داد مدل HXGB در هر دو ایستگاه عملکرد بهتری نسبت به مدل پایه دارد. در بندرانزلی، سناریوی هشتم که شامل تمام ورودی‌ها بود با ترکیب داده‌های هواشناسی و ET_{MODIS}، RMSE را به ۰/۱۱ میلی‌متر در روز کاهش داد. در رامسر نیز همین سناریو به RMSE برابر با ۰/۱۹ میلی متر در روز دست یافت. این ترکیب داده‌ای موجب افزایش حساسیت مدل نسبت به تغییرات مکانی-زمانی و بهبود چشمگیر دقت پیش‌بینی شد.

واژه‌های کلیدی: بندرانزلی، تصاویر ماهواره‌ای، رامسر، سناریوسازی، سنجنده MODIS.

مقدمه

در دهه‌های اخیر، تغییرات اقلیمی، رشد جمعیت و افزایش تقاضا برای آب، فشار بی‌سابقه‌ای بر منابع آبی وارد کرده‌اند. این فشارها نه‌تنها در مناطق خشک و نیمه‌خشک، بلکه در مناطق مرطوب و بسیار مرطوب نیز به‌وضوح قابل مشاهده‌اند (صداقت و همکاران، ۱۴۰۱). مناطقی که پیش‌تر به‌عنوان مناطق امن آبی شناخته می‌شدند، اکنون با چالش‌هایی نظیر نوسانات بارشی، افزایش دما و تغییرات الگوهای تبخیرتعرق مواجه‌اند. این تحولات، ضرورت بازنگری در شیوه‌های مدیریت منابع آب را دوچندان کرده‌اند و نیاز به ابزارهای دقیق، قابل اعتماد، و مکانی‌محور برای پایش و پیش‌بینی مؤلفه‌های کلیدی چرخه هیدرولوژیکی را برجسته‌تر ساخته‌اند. یکی از مهم‌ترین مؤلفه‌های چرخه هیدرولوژیکی، تبخیرتعرق مرجع (ET_0) است که به‌عنوان شاخصی از تقاضای جوی برای آب، نقش محوری در برنامه‌ریزی‌های آبی، کشاورزی و زیست‌محیطی ایفا می‌کند. ET_0 بیانگر میزان آبی است که تحت شرایط ایده‌آل از سطح زمین و پوشش گیاهی مرجع به اتمسفر منتقل می‌شود. این شاخص نه‌تنها در تعیین نیاز آبی گیاهان و طراحی سیستم‌های آبیاری مؤثر کاربرد دارد، بلکه به‌عنوان ابزاری برای پایش خشکسالی، ارزیابی پایداری کشاورزی و مدل‌سازی‌های هیدرولوژیکی و اقلیمی نیز مورد استفاده قرار می‌گیرد (Hussain et al., 2023). در واقع ET_0 یکی از بزرگ‌ترین اجزای خروجی بیلان آبی در حوضه‌های آبریز است و پیش‌بینی دقیق آن می‌تواند به تخصیص بهینه آب بین مصارف مختلف، به‌ویژه در شرایط کمبود منابع، کمک شایانی کند. اهمیت این شاخص زمانی دوچندان می‌شود که بدانیم افزایش ET_0 در شرایط کاهش بارش، نشانه‌ای از خشکسالی و نیاز به اقدامات مدیریتی فوری

۱- دانشجوی دکتری، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، ایران
۲- دانشیار، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، ایران
۳- دانشیار، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، ایران
۴- دانشیار گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده برنامه‌ریزی و علوم محیطی، دانشگاه تبریز، ایران
(*- نویسنده مسئول: Email: s.samadian@tabrizu.ac.ir)
DOI: 10.22034/idj.2026.559252.2644

است (Guo et al., 2023). همچنین به‌طور مستقیم تحت تأثیر پارامترهای هواشناسی نظیر دما، رطوبت نسبی، ساعات آفتابی، سرعت باد و به مقدار کمتر فشار هوا و بارش قرار دارد (Nie et al., 2022، Luo et al., 2021). این پارامترها با تعامل پیچیده و پویا، نوسانات مکانی و زمانی ET_0 را شکل می‌دهند. افزایش دما و تابش موجب افزایش تبخیر، در حالی‌که رطوبت بالا و فشار زیاد آن را کاهش می‌دهد (Liu et al., 2022). در بسیاری از مناطق، به‌ویژه در کشورهایی با اقلیم متغیر و منابع آبی محدود، پایش ET_0 به‌عنوان ابزاری برای تصمیم‌گیری در زمینه الگوهای کشت، زمان‌بندی آبیاری و تخصیص منابع آبی به بخش‌های مختلف اقتصادی مورد استفاده قرار می‌گیرد (Zhou et al., 2017 ،Afzaal et al., 2020). به‌علاوه ET_0 در مدل‌سازی‌های اقلیمی و هیدرولوژیکی نیز نقش کلیدی دارد. چرا که به‌عنوان ورودی اصلی در مدل‌های بارش-رواناب، مدل‌های آب‌های زیرزمینی و مدل‌های شبیه‌سازی تغییرات کاربری اراضی مورد استفاده قرار می‌گیرد.

روش‌های پیش‌بینی ET_0 به‌طور کلی به دو دسته اصلی تقسیم می‌شوند: روش‌های مستقیم و روش‌های غیرمستقیم. روش‌های مستقیم شامل اندازه‌گیری واقعی میزان آب خروجی از خاک یا گیاه هستند که از جمله آن‌ها می‌توان به سنجش رطوبت خاک و استفاده از لایسیمتر اشاره کرد. لایسیمترها به‌عنوان دقیق‌ترین ابزار اندازه‌گیری مستقیم، امکان پایش تبخیرتعرق در شرایط کنترل‌شده را فراهم می‌کنند. در مقابل روش‌های غیرمستقیم از داده‌های هواشناسی، روابط تجربی، تصاویر ماهواره‌ای و الگوریتم‌های یادگیری ماشینی برای تخمین ET_0 بهره می‌برند. روش‌هایی مانند فائو-پنمن-مانتیث، هارگریوز-سامانی و بلانی-کریدل از جمله مدل‌های تجربی رایج هستند که با استفاده از پارامترهای اقلیمی و ضرایب گیاهی، تبخیرتعرق را پیش‌بینی می‌کنند (جوزی و همکاران، ۱۴۰۲). همچنین استفاده از تشت تبخیر و تصاویر ماهواره‌ای مانند داده‌های MODIS[1] امکان پایش تبخیرتعرق در مقیاس‌های وسیع و با وضوح زمانی بالا را فراهم کرده‌اند.

با وجود اهمیت بالای ET_0، روش‌های سنتی پیش‌بینی آن با چالش‌های متعددی مواجه‌اند. مدل‌هایی مانند فائو-پنمن-مانتیث، که به‌عنوان استاندارد جهانی در محاسبه ET_0 شناخته می‌شوند، برای عملکرد دقیق نیازمند مجموعه‌ای کامل از داده‌های هواشناسی زمینی شامل دما، رطوبت نسبی، سرعت باد، و ساعات آفتابی هستند (یحیوی دیزج و همکاران، ۱۴۰۲). در مناطق دورافتاده، کوهستانی یا با زیرساخت‌های ضعیف، دسترسی به چنین داده‌هایی دشوار بوده و منجر به عدم دقت در پیش‌بینی‌ها می‌شود. همچنین داده‌های نقطه‌ای ایستگاه‌های زمینی نماینده مناسبی برای شرایط متغیر مکانی در مناطق ناهمگن نیستند. این محدودیت‌ها باعث شده‌اند که در بسیاری از مناطق به‌ویژه در کشورهای در حال توسعه، پیش‌بینی ET_0 با عدم قطعیت بالا همراه باشد و تصمیم‌گیری‌های مبتنی بر آن با ریسک مواجه شوند. در پاسخ به این چالش‌ها، رویکردهای نوین مبتنی بر سنجش از دور و یادگیری ماشینی، به‌ویژه یادگیری عمیق، در سال‌های اخیر توسعه یافته‌اند. سنجش از دور با فراهم آوردن داده‌هایی با پوشش مکانی وسیع و وضوح زمانی بالا، امکان پایش پارامترهای سطح زمین و اتمسفر را فراهم می‌کند. داده‌های ماهواره‌ای مانند محصولات سنجنده MODIS، با ارائه شاخص‌هایی نظیر دمای سطح زمین (LST)[2]، شاخص تفاوت نرمال پوشش گیاهی (NDVI)[3]، شاخص EVI، و ضریب آلبیدو اطلاعات ارزشمندی را در اختیار پژوهشگران قرار می‌دهند که می‌توانند به‌عنوان ورودی مدل‌های پیش‌بینی ET_0 مورد استفاده قرار گیرند (صادقی و احمدی، ۱۴۰۱). از سوی دیگر الگوریتم‌های یادگیری ماشینی و یادگیری عمیق، با توانایی بالا در مدل‌سازی روابط پیچیده و غیرخطی میان متغیرها، ابزارهایی قدرتمند برای تحلیل داده‌های بزرگ و متنوع محسوب می‌شوند. این الگوریتم‌ها قادرند الگوهای پنهان در داده‌ها را شناسایی کرده و پیش‌بینی‌هایی با دقت بالا ارائه دهند، حتی در شرایطی که داده‌های ورودی ناقص یا پراکنده باشند. یادگیری عمیق به‌ویژه با استفاده از شبکه‌های عصبی بازگشتی مانند LSTM[4]، امکان تحلیل داده‌های زمانی و استخراج وابستگی‌های بلندمدت میان متغیرها را فراهم می‌کند. این ویژگی‌ها باعث شده‌اند که مدل‌های یادگیری عمیق به‌عنوان جایگزینی مؤثر برای روش‌های سنتی در پیش‌بینی ET_0 مطرح شوند. علاوه بر این، ترکیب داده‌های سنجش از دور با الگوریتم‌های یادگیری عمیق، زمینه را برای ارائه پیش‌بینی‌های مکانی پیوسته و دقیق از ET_0 فراهم می‌سازد. برخلاف داده‌های نقطه‌ای ایستگاه‌های زمینی، داده‌های ماهواره‌ای امکان پوشش مناطق وسیع با وضوح مکانی بالا را دارند و می‌توانند تغییرات مکانی ET_0 را به‌صورت دقیق‌تری نمایش دهند. این ویژگی به‌ویژه در مناطق ناهمگن، کوهستانی یا با کاربری‌های متنوع زمین، اهمیت بالایی دارد و می‌تواند به تصمیم‌گیری‌های مکانی‌محور در مدیریت منابع آب کمک کند.

در کنار مزایای علمی، استفاده از داده‌های ماهواره‌ای و مدل‌های یادگیری عمیق از نظر اقتصادی نیز مقرون‌به‌صرفه است. جمع‌آوری داده‌های زمینی نیازمند زیرساخت‌های گسترده، نیروی انسانی متخصص و هزینه‌های نگهداری بالاست. در حالی‌که داده‌های ماهواره‌ای به‌صورت رایگان یا با هزینه اندک در دسترس هستند و می‌توانند به‌صورت خودکار و پیوسته به‌روزرسانی شوند. همچنین،

1- Moderate Resolution Imaging Spectroradiometer

2- Land Surface Temperature

3- Normalized Difference Vegetation Index

4- Long Short-Term Memory

مدل‌های یادگیری عمیق پس از آموزش اولیه، قادرند به‌صورت خودکار و با سرعت بالا پیش‌بینی‌های جدید ارائه دهند، که این امر در شرایط اضطراری مانند خشکسالی یا بحران‌های آبی بسیار ارزشمند است. با توجه به این مزایا، پژوهش‌های متعددی در سال‌های اخیر به بررسی کارایی مدل‌های یادگیری ماشینی و داده‌های سنجش از دور در پیش‌بینی ET_0 پرداخته‌اند. این مطالعات نشان داده‌اند که مدل‌هایی مانند XGBoost[1]، LSTM، ELM و CNN قادرند با دقت بالا ET_0 را پیش‌بینی کنند و در بسیاری از موارد عملکرد بهتری نسبت به مدل‌های سنتی داشته‌اند. با این حال بیشتر این مطالعات در مناطق خشک یا نیمه‌خشک انجام شده‌اند و مناطق مرطوب و بسیار مرطوب کمتر مورد توجه قرار گرفته‌اند. این در حالی است که ویژگی‌های اقلیمی خاص این مناطق، مانند رطوبت نسبی بالا، پوشش گیاهی متراکم و نوسانات تابشی می‌توانند بر دقت مدل‌سازی تأثیرگذار باشند.

بررسی مطالعات پیشین نشان می‌دهد که بخش عمده‌ای از پژوهش‌ها در زمینه پیش‌بینی تبخیر تعرق مرجع بر مناطق خشک و نیمه‌خشک متمرکز بوده‌اند و کمتر به شرایط اقلیمی مرطوب و بسیار مرطوب توجه شده است. در این مطالعات، استفاده از مدل‌های کلاسیک یا الگوریتم‌های یادگیری ماشین غالباً با محدودیت‌هایی نظیر وابستگی به داده‌های کامل هواشناسی یا عدم توجه به اثرات مکانی و زمانی همراه بوده است. نوآوری پژوهش حاضر در دو محور اصلی نهفته است: نخست، تمرکز بر مناطق شمالی کشور با ویژگی‌های اقلیمی خاص (رطوبت نسبی بالا، پوشش گیاهی متراکم و نوسانات تابشی) که کمتر مورد توجه قرار گرفته‌اند. دوم، تلفیق داده‌های سنجش از دور (ET_{MODIS}) با داده‌های زمینی و توسعه الگوریتم ترکیبی HXGB که توانایی بالاتری در شناسایی روابط غیرخطی و کاهش خطا نسبت به مدل‌های پایه دارد. این رویکرد، علاوه بر ارتقای دقت پیش‌بینی ET_0، امکان تحلیل حساسیت مدل‌ها در شرایط واقعی و محدودیت‌های داده‌ای را فراهم ساخته و جایگاه علمی پژوهش را نسبت به مطالعات مشابه متمایز می‌کند.

در این پژوهش با هدف ارتقای دقت پیش‌بینی ET_0 در مناطق مرطوب و بسیار مرطوب کشور، از دو مدل یادگیری ماشین پیشرفته با نام‌های HXGB[2] و XGB استفاده شده است. این مدل‌ها با بهره‌گیری از داده‌های سنجنده MODIS و داده‌های هواشناسی زمینی، در ایستگاه‌های منتخب رامسر و بندرانزلی مورد ارزیابی قرار گرفته‌اند. سناریوهای مختلفی برای ترکیب داده‌ها طراحی شده‌اند که در آن‌ها ترکیبات متفاوتی از داده‌های ماهواره‌ای و زمینی به‌کار گرفته شده است. نوآوری این پژوهش در تلفیق داده‌های سنجش از دور با مدل‌های یادگیری عمیق و تمرکز بر مناطق مرطوب کشور نهفته است. مناطقی که به‌دلیل ویژگی‌های خاص اقلیمی، کمتر مورد توجه مطالعات پیشین قرار گرفته‌اند و مدل‌سازی دقیق ET_0 در آن‌ها با چالش‌های بیشتری همراه است. همچنین، این تحقیق با طراحی سناریوهای کم‌داده، عملکرد مدل‌ها را در شرایط واقعی و محدودیت‌های داده‌ای مورد آزمون قرار داده و نشان داده است که مدل‌های پیشنهادی می‌توانند جایگزینی مؤثر یا مکملی قدرتمند برای روش‌های سنتی باشند.

مواد و روش ها

منطقه مورد مطالعه

استان گیلان با مساحتی حدود ۱۴۰۴۴ کیلومتر مربع در شمال ایران و در حاشیه‌ی غربی دریای خزر واقع شده است. این استان از شمال به دریای خزر و کشور جمهوری آذربایجان، از شرق به استان مازندران، از جنوب به استان قزوین و از غرب به استان اردبیل محدود می‌شود. مرکز استان شهر رشت است. گیلان به‌عنوان یکی از مرطوب‌ترین استان‌های کشور شناخته می‌شود و در بسیاری از مناطق آن، میانگین بارندگی سالانه به بیش از ۱۵۰۰ میلی‌متر می‌رسد. پوشش گیاهی متراکم، رودخانه‌های متعدد، تالاب‌های طبیعی و رطوبت نسبی بالا از ویژگی‌های بارز اقلیمی این منطقه هستند که آن را به بستری مناسب برای مطالعات هیدرولوژیکی و اقلیمی تبدیل کرده‌اند. شهر بندرانزلی در مجاورت تالاب بین‌المللی انزلی واقع شده و به‌دلیل موقعیت جغرافیایی خاص خود، نقش مهمی در مطالعات مرتبط با تبخیرتعرق، زهکشی و مدیریت منابع آب ایفا می‌کند. بندرانزلی در مختصات جغرافیایی ۳۷ درجه و ۲۸ دقیقه عرض شمالی و ۴۹ درجه و ۲۷ دقیقه طول شرقی قرار دارد. این شهر دارای اقلیم بسیار مرطوب خزری با تابستان‌هایی گرم و مرطوب و زمستان‌هایی ملایم است. میانگین بارندگی سالانه بندرانزلی حدود ۱۸۹۲ میلی‌متر گزارش شده و رطوبت نسبی هوا در این منطقه بین ۷۱ تا ۹۷ درصد متغیر است. این شرایط اقلیمی خاص، همراه با پوشش گیاهی متراکم و مجاورت با تالاب، موجب شده است که تبخیرتعرق در این منطقه تحت تأثیر شدید پارامترهای جوی و زیستی قرار گیرد (Najafi and Alizadeh., 2023).

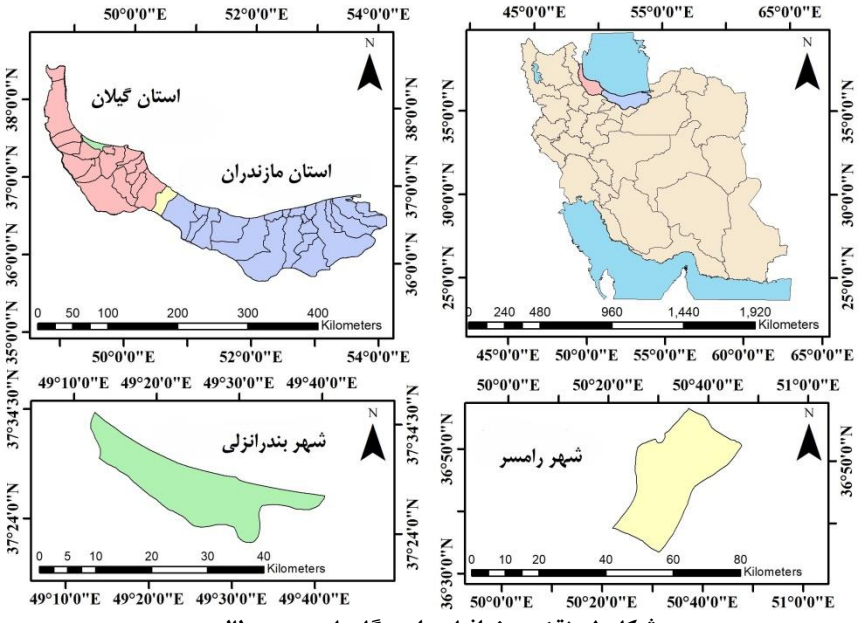
از طرف دیگر استان مازندران با وسعتی حدود ۲۳۷۵۶ کیلومتر مربع، معادل ۱/۴۶ درصد از مساحت کل کشور، در کرانه‌ی جنوبی دریای خزر واقع شده است. این استان از شمال به دریای خزر، از شرق به استان گلستان، از جنوب به استان‌های سمنان، تهران و البرز و از غرب به استان گیلان محدود می‌شود. مرکز استان شهر ساری است. مازندران دارای اقلیم مرطوب خزری با تابستان‌های معتدل و زمستان‌های نسبتاً سرد است. این استان به‌دلیل برخورداری از منابع آبی فراوان، خاک حاصلخیز، پوشش گیاهی متنوع و دسترسی به دریا

1- Extreme Gradient Boosting
2- Hybrid Extreme Gradient Boosting

یکی از قطب‌های کشاورزی و گردشگری کشور به‌شمار می‌رود. شهر رامسر در بخش غربی استان مازندران و در مجاورت استان گیلان واقع شده است. این شهر به‌دلیل موقعیت جغرافیایی خاص خود، در منطقه‌ای با رطوبت بالا، بارندگی قابل‌توجه و پوشش گیاهی متراکم قرار دارد. رامسر در مختصات جغرافیایی ۳۶ درجه و ۵۶ دقیقه عرض شمالی و ۵۰ درجه و ۴۰ دقیقه طول شرقی واقع شده است. اقلیم رامسر نیز مرطوب خزری، با تابستان‌هایی معتدل و مرطوب و زمستان‌هایی نسبتاً سرد است. میانگین دمای سالانه این شهر حدود ۱۷/۵ درجه سانتی‌گراد و میانگین بارندگی سالانه آن در حدود ۸۸۹ میلی‌متر گزارش شده است. وجود ایستگاه هواشناسی فعال در این منطقه، همراه با تنوع اقلیمی و زیستی، شرایط مناسبی برای مطالعات تبخیرتعرق و مدل‌سازی‌های اقلیمی فراهم کرده است (Dadashi Roudbari et al., 2020).

داده‌های مورد استفاده

در تحقیق حاضر، از داده‌های هواشناسی روزانه شامل ساعات آفتابی، سرعت باد، دمای بیشینه، کمینه و میانگین، رطوبت نسبی بیشینه، کمینه و میانگین، در بازه زمانی ۲۳ ساله (۲۰۰۱-۲۰۲۳) برای دو ایستگاه هواشناسی بندرانزلی و رامسر، که از سامانه رسمی هواشناسی ایران (IRIMO) استخراج شد، استفاده شده است. این داده‌ها به‌عنوان ورودی مدل‌های یادگیری ماشین برای پیش‌بینی ET_0 به‌کار گرفته شده‌اند. انتخاب دو شهر بندرانزلی و رامسر در این پژوهش با هدف بررسی عملکرد مدل‌های یادگیری ماشین در شرایط اقلیمی نسبتاً مشابه اما با تفاوت‌های جزئی در الگوهای بارندگی، دما، رطوبت نسبی و پوشش گیاهی صورت گرفته است. این تفاوت‌ها امکان تحلیل حساسیت مدل‌ها نسبت به متغیرهای محیطی و ارزیابی قابلیت تعمیم آن‌ها را فراهم می‌سازد. به‌ویژه در مدل‌سازی تبخیرتعرق مرجع، شناخت دقیق از رفتار پارامترهای هواشناسی در مناطق مرطوب و بسیار مرطوب اهمیت بالایی دارد. داده‌های تبخیر تعرق ماهواره‌ای از محصول رسمی ET_{MODIS} استخراج شدند. این محصول مبتنی بر الگوریتم استاندارد MOD16 ناسا توسعه یافته و با استفاده از داده‌های سطح زمین، پوشش گیاهی و شرایط اتمسفری، مقادیر تبخیر تعرق واقعی را در مقیاس پیکسلی ارائه می‌دهد. داده‌ها از طریق کدنویسی در سامانه Google Earth Engine فراخوانی شدند. موقعیت جغرافیایی دقیق ایستگاه‌های مورد مطالعه با استفاده از نرم‌افزارهای ArcGIS و ArcMap ترسیم شده‌اند. شکل ۱ موقعیت جغرافیایی ایستگاه‌های مورد مطالعه را نشان می‌دهد و در جدول ۱ مشخصات جغرافیایی ایستگاه‌های مورد مطالعه ارائه شده است. همچنین، برای تحلیل همبستگی بین پارامترهای هواشناسی و تبخیرتعرق سنجنده MODIS، ضرایب همبستگی در جدول ۲ برای هر ایستگاه مورد بررسی قرار گرفته‌اند.



شکل ۱- نقشه جغرافیایی ایستگاه‌های مورد مطالعه

جدول ۱- مشخصات جغرافیایی ایستگاه‌های مورد مطالعه

ایستگاه	طول جغرافیایی		عرض جغرافیایی		ارتفاع از سطح دریا
رامسر	′۴۰	°۵۰	′۵۶	°۳۶	-۲
بندرانزلی	′۲۷	°۴۹	′۲۸	°۳۷	-۲۶

جدول ۲- ضرایب همبستگی بین پارامتر تبخیرتعرق فائو-پنمن-مانتیث هر ایستگاه با سایر پارامترهای هواشناسی

ایستگاه	حداقل دما (^0C)	حداکثر دما (^0C)	دمای میانگین (^0C)	رطوبت میانگین (%)	حداقل رطوبت (%)	حداکثر رطوبت (%)	سرعت باد (m/s)	ساعات آفتابی (h)	تبخیرتعرق MODIS (mm)
رامسر	۰/۶۶۹	۰/۸۱۵	۰/۷۵۸	-۰/۷۱۱	-۰/۶۰۸	-۰/۶۸۰	-۰/۲۱۲	۰/۶۶۴	۰/۶۶۴
بندرانزلی	۰/۷۲۰	۰/۸۳۰	۰/۷۹۱	-۰/۷۱۹	-۰/۶۶۹	-۰/۵۹۳	-۰/۳۸۵	۰/۷۵۲	۰/۷۰۴

این جدول به‌عنوان بخشی از ارزیابی اولیه داده‌ها ارائه شده است تا روابط میان متغیرهای هواشناسی و تبخیرتعرق مشخص شود و کفایت داده‌ها پیش از ورود به تحلیل اصلی بررسی گردد و هدف آن نشان دادن روند پیش‌پردازش، بررسی کیفیت داده‌ها و به نوعی آنالیز حساسیت داده‌ها بوده است. در بررسی همبستگی بین پارامترهای هواشناسی و تبخیرتعرق سنجنده MODIS، مشاهده می‌شود که دما در هر دو ایستگاه نقش بسیار مهمی در افزایش ET_0 ایفا می‌کند. در بندرانزلی، همبستگی بین دمای حداکثر و ET_0 برابر با ۰/۸۳۰ و در رامسر برابر با ۰/۸۱۵ است که نشان‌دهنده تأثیر قوی دمای روزانه بر شدت تبخیرتعرق در هر دو منطقه است. همچنین دمای میانگین در بندرانزلی (۰/۷۹۱) نسبت به رامسر (۰/۷۵۸) همبستگی بالاتری دارد، که بیانگر پایداری بیشتر گرمای روزانه در بندرانزلی و تأثیر آن بر افزایش تبخیرتعرق است. دمای حداقل نیز در بندرانزلی با ضریب ۰/۷۲۰ نسبت به رامسر با ضریب ۰/۶۶۹ تأثیر بیشتری دارد، که می‌تواند ناشی از نوسانات شبانه و تفاوت در الگوهای حرارتی باشد. در مقابل، رطوبت نسبی به‌عنوان عامل بازدارنده تبخیرتعرق، همبستگی منفی قابل توجهی در هر دو ایستگاه دارد. به‌گونه‌ای که رطوبت میانگین در بندرانزلی با ضریب ۰/۷۱۹- و در رامسر با ۰/۷۱۱- نشان می‌دهد افزایش رطوبت محیط موجب کاهش ET_0 می‌شود. حداقل رطوبت نیز در بندرانزلی (۰/۶۶۹-) تأثیر منفی بیشتری نسبت به رامسر (۰/۶۰۸-) دارد، که احتمالاً به دلیل کاهش رطوبت نسبی در ساعات گرم روز و افزایش تبخیر در آن بازه زمانی است. حداکثر رطوبت در رامسر با ضریب ۰/۶۸۰- نسبت به بندرانزلی با ضریب ۰/۵۹۳- تأثیر منفی بیشتری دارد، که می‌تواند به پوشش گیاهی متراکم‌تر یا شرایط اشباع‌تر هوا در رامسر مرتبط باشد. ساعات آفتابی نیز به‌عنوان یکی از محرک‌های اصلی تبخیرتعرق، در بندرانزلی با ضریب ۰/۷۵۲- و در رامسر با ضریب ۰/۶۶۴ همبستگی مثبت بالایی دارد. این تفاوت نشان می‌دهد که در بندرانزلی ساعات آفتابی نقش مؤثرتری در تحریک فرآیند تبخیرتعرق ایفا می‌کند. از سوی دیگر سرعت باد در بندرانزلی با ضریب ۰/۳۸۵- نسبت به رامسر با ضریب ۰/۲۱۲- همبستگی منفی بیشتری دارد. موضوعی که برخلاف انتظار است زیرا باد معمولاً موجب افزایش تبخیر می‌شود اما در مناطق مرطوب هم‌زمانی باد با رطوبت بالا ممکن است موجب کاهش گرادیان فشار بخار و در نتیجه کاهش تبخیر گردد. در نهایت همبستگی داده‌های MODIS در بندرانزلی با ضریب ۰/۷۰۴ نسبت به رامسر با ضریب ۰/۶۶۴ بهتر ارزیابی شده است. بنابراین می‌توان گفت که در هر دو ایستگاه دما و ساعات آفتابی محرک‌های اصلی ET_0 هستند، در حالی‌که رطوبت نسبی نقش بازدارنده دارد. تفاوت‌های جزئی در ضرایب همبستگی بین دو ایستگاه نشان‌دهنده تأثیر شرایط اقلیمی محلی بر شدت و الگوی تبخیرتعرق است و این تحلیل می‌تواند در انتخاب ورودی‌های بهینه برای مدل‌های یادگیری ماشین و تنظیم وزن‌دهی به پارامترها در الگوریتم‌هایی مانند XGB و LSTM مؤثر باشد.

روش‌شناسی و توسعه مدل‌های یادگیری ماشین

در این پژوهش، با هدف تخمین دقیق ET_0 در مناطق مرطوب و بسیار مرطوب شمال کشور، از ترکیب داده‌های سنجش از دور حاصل از سنجنده MODIS و داده‌های هواشناسی روزانه ایستگاه‌های رامسر و بندرانزلی استفاده شده است. برای مدل‌سازی تبخیرتعرق، دو الگوریتم قدرتمند یادگیری ماشین شامل XGB و نسخه ترکیبی و بهینه‌شده آن HXGB توسعه داده شدند. مدل XGB به‌عنوان یکی از دقیق‌ترین الگوریتم‌های تقویتی مبتنی بر درخت تصمیم، توانایی بالایی در مدل‌سازی روابط غیرخطی، مدیریت داده‌های ناقص و جلوگیری از بیش‌برازش دارد. این مدل با استفاده از ساختار گرادیان تقویتی، درخت‌های تصمیم را به‌صورت مرحله‌ای آموزش داده و خطای پیش‌بینی را به‌طور پیوسته کاهش می‌دهد. تنظیمات دقیق پارامترهایی مانند نرخ یادگیری، عمق درخت‌ها، تعداد تکرارها و وزن‌دهی به متغیرها، XGB را به گزینه‌ای مناسب برای مدل‌سازی تبخیرتعرق در مناطق با تنوع اقلیمی بالا تبدیل کرده است. مدل

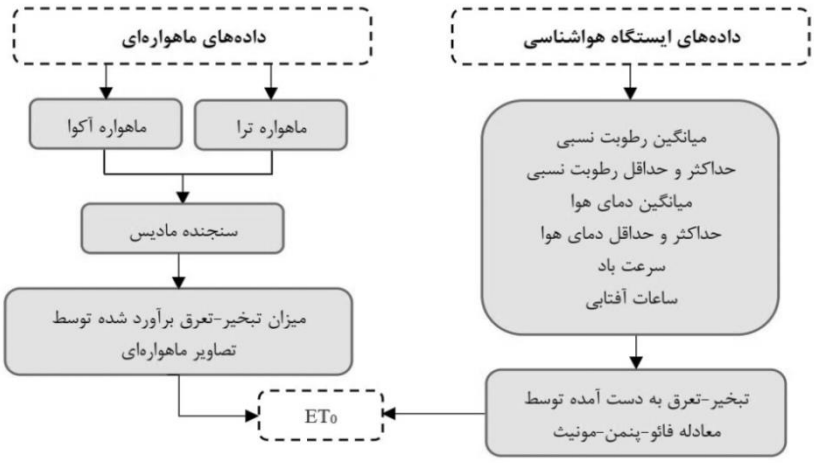
HXGB نیز با هدف ارتقای دقت پیش‌بینی و افزایش پایداری مدل در شرایط داده‌ای متنوع طراحی شده است. در این مدل، ساختار پایه XGB با تنظیمات خاص و ترکیب داده‌های چندمنبعی (ماهواره‌ای و زمینی) تقویت شده است. استفاده از HXGB امکان تحلیل دقیق‌تر وابستگی‌های مکانی و زمانی بین پارامترهای هواشناسی و تبخیرتعرق را فراهم می‌سازد و به‌ویژه در مناطق مرطوب با نوسانات رطوبتی بالا، عملکرد بهتری نسبت به مدل پایه نشان داده است.

ورودی مدل‌ها و مراحل پیش‌پردازش

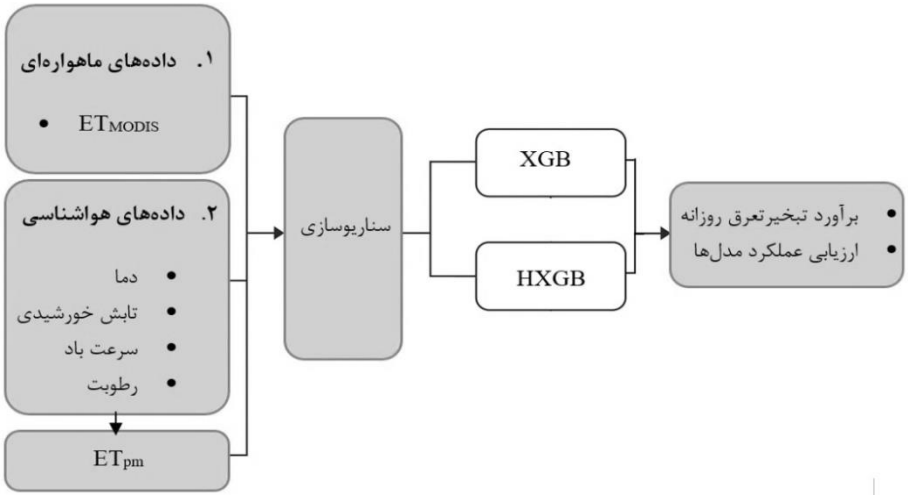
داده‌های ورودی مدل‌ها شامل مجموعه‌ای جامع از پارامترهای هواشناسی روزانه نظیر دمای بیشینه، کمینه و میانگین، ساعات آفتابی، سرعت باد، رطوبت نسبی بیشینه، کمینه و میانگین، و همچنین تبخیرتعرق استخراج‌شده از داده‌های ماهواره‌ای بودند. داده‌های هواشناسی از سامانه رسمی هواشناسی ایران و داده‌های ماهواره‌ای از طریق کدنویسی در سامانه Google Earth Engine استخراج شدند. این داده‌ها در بازه زمانی ۲۰۰۱ تا ۲۰۲۳ برای دو ایستگاه رامسر و بندرانزلی گردآوری شدند و خروجی مدل‌ها نیز مقادیر ET_0 روزانه برحسب میلی‌متر در روز بود. تمامی مدل‌ها از یک روند پیش‌پردازشی یکسان پیروی کردند تا کیفیت داده‌ها حفظ شده و با چارچوب‌های یادگیری ماشین سازگار شوند. در مرحله نخست، داده‌های مفقود با استفاده از میانگین هر ویژگی و به‌کمک فیلتر کالمن در محیط Python تکمیل شدند. سپس برای تبدیل داده‌ها به فرمت روزانه، از روش اسپیلاین مکعبی استفاده شد. به‌منظور کاهش اثر تفاوت مقیاس متغیرها و ارتقاء توان تعمیم‌دهی مدل‌ها، تمامی ویژگی‌ها با استفاده از استانداردساز کتابخانه scikit-learn نرمال‌سازی شدند، به‌گونه‌ای که داده‌ها حول مقدار صفر با واریانس واحد متمرکز گردیدند. مجموعه داده‌ها با حفظ ترتیب زمانی به نسبت ۸۰ درصد برای آموزش و ۲۰ درصد برای آزمون تقسیم شد تا ارزیابی عملکرد مدل‌ها به‌صورت معتبر انجام گیرد. پس از پیش‌پردازش، تحلیل همبستگی بین پارامترها با استفاده از ضریب پیرسون انجام شد تا روابط بین متغیرهای هواشناسی و ET_0 به‌طور دقیق مشخص شود. بر اساس نتایج این تحلیل، سناریوهای مختلفی تعریف گردید و مدل‌ها تحت این سناریوها اجرا شدند تا تأثیر ترکیب داده‌های زمینی و ماهواره‌ای بر دقت پیش‌بینی ET_0 بررسی شوند. شکل ۲ فلوچارت روش پیش‌بینی ET_0 در مناطق مورد مطالعه را نشان می‌دهد و شکل ۳ روند کلی انجام پژوهش شامل مراحل جمع‌آوری داده‌ها، پیش‌پردازش، آموزش مدل‌ها، اعتبارسنجی و ارزیابی عملکرد را به‌صورت شماتیک نمایش می‌دهد.

معادله فائو-پنمن-مانتیث

معادله فائو-پنمن-مانتیث به عنوان استانداردی برای پیش‌بینی ET_0 شناخته می‌شود و تاریخچه آن به تلاش‌های اولیه در محاسبه نیاز آبی گیاهان برمی‌گردد. معادله پنمن در سال ۱۹۴۸ توسط هارولد پنمن معرفی شد. این معادله به عنوان یک روش ترکیبی برای پیش‌بینی تبخیر از سطح آب و خاک طراحی شده بود و به تدریج به عنوان یک ابزار مهم در مدیریت منابع آب شناخته شد (Hargreaves, 2003). در سال ۱۹۶۵ معادله مانتیث به این معادله اضافه شد که به بهبود دقت آن کمک کرد. این ادغام به ایجاد یک مدل جامع‌تر برای پیش‌بینی تبخیرتعرق منجر شد (Allen et el., 1998).



شکل ۲- فلوچارت روش پیش‌بینی ET_0 در مناطق مورد مطالعه



شکل ۳- الگوریتم روند کلی انجام پژوهش

در سال ۱۹۹۰ سازمان خواربار و کشاورزی ملل متحد این معادله را به عنوان استاندارد جهانی برای پیش‌بینی تبخیرتعرق مرجع معرفی کرد و در سال ۱۹۹۸ نسخه به‌روزرسانی شده‌ای از آن منتشر شد که به فائو-پنمن-مانتیث معروف است (Gelcer et al., 2010). معادله فائو-پنمن-مانتیث به داده‌های متعددی از جمله دما، رطوبت نسبی، سرعت باد و ساعات آفتابی نیاز دارد. این ویژگی باعث می‌شود که این معادله در مناطقی که داده‌های هواشناسی کامل در دسترس نیست کمتر مورد استفاده قرار گیرد (Yoder et al., 2005، Trajkovic, 2007). به دلیل نیاز به داده‌های زیاد برخی از محققان به استفاده از روش‌های ساده‌تر و تجربی مانند معادله هارگریوز روی آورده‌اند که تنها به داده‌های دما نیاز دارد. هرچند که این روش‌ها معمولاً دقت کمتری دارند (Hargreaves and Samani., 1985). همچنین این معادله به عنوان یکی از دقیق‌ترین روش‌ها برای پیش‌بینی تبخیرتعرق شناخته می‌شود و در مقایسه با سایر روش‌ها معمولاً نتایج بهتری ارائه می‌دهد (Nikam et al., 2014). رابطه ۱ معادله فائو-پنمن-مانتیث را نشان می‌دهد.

$$ET_0 = \frac{0.408\Delta(R_n-G)+\gamma\frac{900}{T_m+273}U_2(e_s-e_a)}{\Delta+\gamma(1+0.34U_2)} \quad (۱)$$

ET_0 تبخیرتعرق گیاه مرجع (میلی‌متر بر روز)، Δ شیب منحنی فشار بخار اشباع (کیلوپاسکال بر درجه سلسیوس)، R_n تابش خالص خورشیدی (مگاژول برمتر مربع بر روز)، G چگالی شار حرارتی خاک (مگاژول برمتر مربع بر روز)، γ ثابت سایکرومتری (کیلوپاسکال بر درجه سلسیوس)، T_m متوسط دمای هوا در ارتفاع ۲ متری (درجه سلسیوس)، U_2 متوسط سرعت باد در یک شبانه روز در ارتفاع دو متری از سطح زمین (متر بر ثانیه)، e_s فشار بخار اشباع (کیلوپاسکال) و e_a فشار بخار واقعی (کیلوپاسکال) می‌باشند.

مبانی سنجش از دور و سنجنده MODIS

سنجش از دور به عنوان روشی برای استخراج اطلاعات از سطح زمین بدون تماس مستقیم، بر پایه تابش الکترومغناطیسی عمل می‌کند و امکان شناسایی و ارزیابی ویژگی‌های سطح زمین را فراهم می‌سازد. اصول فیزیکی سنجش از دور بر این واقعیت استوار است که انرژی خورشیدی پس از رسیدن به سطح زمین بخشی جذب، بخشی بازتاب و بخشی عبور می‌کند. هر پدیده سطحی مانند جنگل، دریاچه یا مناطق شهری الگوی بازتاب طیفی خاص خود را دارد و همین تفاوت‌ها توسط سنجنده‌ها ثبت می‌شود. تصاویر ماهواره‌ای مهم‌ترین خروجی سنجش از دور هستند که اطلاعات چندبعدی از سطح زمین را در باندهای مختلف طیفی ارائه می‌دهند. این تصاویر می‌توانند در باندهای مرئی، مادون قرمز نزدیک، مادون قرمز حرارتی یا حتی امواج مایکروویو ثبت شوند. ویژگی‌هایی مانند دقت مکانی، دقت زمانی و دقت طیفی تعیین‌کننده کیفیت و کارایی این تصاویر هستند. به عنوان نمونه، تصاویر با دقت مکانی بالا جزئیات بیشتری از سطح زمین را آشکار می‌کنند و تصاویر با دقت زمانی بالا امکان پایش مداوم تغییرات را فراهم می‌سازند (Huang et al., 2013). سنجنده‌ها ابزار اصلی جمع‌آوری داده‌های الکترومغناطیسی هستند که بر روی ماهواره‌ها یا سکوهای پروازی نصب می‌شوند. سنجنده‌ها به دو دسته فعال و غیرفعال تقسیم می‌شوند. سنجنده‌های غیرفعال مانند MODIS تنها انرژی بازتاب‌شده خورشید را ثبت می‌کنند، در حالی که سنجنده‌های فعال مانند رادارها خود امواجی به سطح زمین ارسال کرده و بازتاب آن‌ها را اندازه‌گیری می‌کنند. در میان سنجنده‌های مختلف، MODIS یکی از پرکاربردترین‌هاست که بر روی ماهواره‌های Terra و Aqua نصب شده است. این سنجنده توانایی تصویربرداری در ۳۶ باند طیفی مختلف را دارد و دقت مکانی آن بسته به نوع باند بین ۲۵۰ متر تا یک کیلومتر متغیر است. MODIS تقریباً

کل سطح زمین را روزانه دو بار پوشش می‌دهد و همین ویژگی آن را به ابزاری ارزشمند برای پایش‌های محیطی و اقلیمی در مقیاس جهانی تبدیل کرده است. داده‌های خام MODIS پیش از تحلیل نیازمند پیش‌پردازش هستند. این مراحل شامل تصحیح هندسی برای تطبیق تصویر با مختصات واقعی زمین، تصحیح رادیومتریک برای حذف نویزها و تبدیل داده‌ها به واحدهای استاندارد و تصحیح اتمسفری برای کاهش اثر بخار آب و ذرات معلق جوی است. داده‌های MODIS مورد استفاده در این پژوهش از طریق کدنویسی در محیط Google Earth Engine استخراج شدند. این سامانه داده‌های MODIS را به‌صورت پیش‌پردازش‌شده (شامل تصحیحات هندسی، رادیومتریک و اتمسفری استاندارد) در اختیار پژوهشگران قرار می‌دهد. بنابراین نیازی به انجام مجدد این مراحل در نرم‌افزارهای جانبی نبود و داده‌ها پس از فراخوانی مستقیم در قالب واحدهای استاندارد برای تحلیل آماده شدند. پس از این مراحل، داده‌ها آماده تولید شاخص‌های تحلیلی مانند NDVI یا LST می‌شوند و می‌توانند در مدل‌سازی‌های اقلیمی و محیطی مورد استفاده قرار گیرند (Gonenc et al., 2019).

ساختار و اجزای تابع هدف درمدل XGBoost

الگوریتم XGBoost یکی از پیشرفته‌ترین و کارآمدترین روش‌های یادگیری ماشین در دسته مدل‌های تقویتی است که با استفاده از درخت‌های تصمیم‌گیری به‌صورت تجمیعی، دقت و تعمیم‌پذیری بالایی را ارائه می‌دهد. این الگوریتم بر پایه ایده تقویت گرادیان عمل می‌کند. به این معنا که هر مدل جدید، خطاهای باقی‌مانده مدل‌های قبلی را یاد می‌گیرد و اصلاح می‌کند. یکی از ویژگی‌های شاخص XGBoost، ساختار تابع هدف آن است که علاوه بر تابع زیان، شامل یک بخش منظم‌کننده نیز هست. بخش منظم‌کننده، با جریمه دادن به مدل‌های پیچیده‌تر، از رشد بی‌رویه درخت‌ها جلوگیری کرده و خطر بیش‌برازش را کاهش می‌دهد. این ویژگی به‌ویژه در کار با داده‌های دارای نویز یا ساختارهای پیچیده، مزیت قابل‌توجهی ایجاد می‌کند. مزیت‌های دیگر XGBoost شامل کارایی محاسباتی بالا به دلیل پشتیبانی از پردازش موازی، قابلیت مدیریت داده‌های بزرگ و پراکنده و انعطاف در انتخاب انواع توابع زیان است. همچنین این الگوریتم به‌خوبی با داده‌های غیرخطی سازگار می‌شود و توانایی شناسایی و ترکیب خودکار ویژگی‌ها را دارد. این قابلیت‌ها باعث شده است که XGBoost در حوزه‌های متنوعی از جمله پیش‌بینی سری‌های زمانی، طبقه‌بندی، رگرسیون و به‌ویژه مدل‌سازی پدیده‌های محیطی مانند منابع آب و هواشناسی، عملکرد قابل اتکایی داشته باشد (Wiens et al., 2025). الگوریتم XGBoost یک پیاده‌سازی بهینه از تقویت گرادیان براساس درخت‌های تصمیم گیری است که تابع هدف کلی آن به صورت زیر است (رابطه ۲):

$$\sum_{k=1}^{k} \Omega(f_k) + \sum_{i=1}^{n} l(y_i, \hat{y}_i) = \mathcal{L}(\theta) \qquad (۲)$$

که در آن l تابع زیان (مثلا مربع خطا) و $\Omega(f)$ دوره منظم کننده برای کنترل پیچیدگی هر درخت است و معمولا به فرم زیر تعریف می‌شود (رابطه ۳):

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda\|\omega\|^2 \qquad (۳)$$

T تعداد برگ‌ها و ω بردار وزن‌های برگ‌ها است. مزیت‌های XGBoost شامل کاربرد محاسباتی، کنترل بیش برازش از طریق منظم سازی و قابلیت کار با متغیرهای غیر خطی است.

نسخه HXGB برخی پارامترهای کلیدی مانند نرخ یادگیری، تعداد درخت‌ها و عمق درخت‌ها به‌صورت خودکار و بهینه‌تر تنظیم می‌شوند. این بهینه‌سازی موجب کاهش بیش‌برازش، افزایش سرعت همگرایی و بهبود دقت پیش‌بینی در داده‌های پیچیده می‌گردد. در واقع HXGB را می‌توان نسخه ارتقایافته XGBoost دانست که با استفاده از تنظیمات پیشرفته و روش‌های جستجوی فراپارامتر، عملکرد مدل در شرایط داده‌های بزرگ و متنوع بهبود یافته است.

معیارهای ارزیابی دقت

در این پژوهش، به‌منظور ارزیابی دقت مدل‌های پیش‌بینی ET_0 و مقایسه نسبی نتایج حاصل از آن‌ها با مقادیر اندازه‌گیری‌شده، از مجموعه‌ای از معیارهای آماری معتبر استفاده شد. این معیارها شامل میانگین مربعات خطا (MSE)، جذر میانگین مربعات خطا (RMSE)، RMSE نرمال‌شده (NRMSE)، ضریب همبستگی (R^2)، ضریب نش–ساتکلیف (NSE) و شاخص توافق ویلموت (WI) هستند که هر یک جنبه‌ای خاص از عملکرد مدل را مورد سنجش قرار می‌دهند. ضریب همبستگی میزان ارتباط خطی بین مقادیر مشاهده‌شده و پیش‌بینی‌شده را نشان می‌دهد. مقدار نزدیک به ۱ بیانگر همبستگی قوی و عملکرد مطلوب مدل است. RMSE به‌عنوان معیار خطای مطلق، میانگین اندازه خطاها را محاسبه می‌کند و به خطاهای بزرگ وزن بیشتری می‌دهد. بنابراین مقدار پایین آن نشان‌دهنده دقت بالای مدل است. NRMSE نیز نسخه نرمال‌شده RMSE است که امکان مقایسه بین مدل‌ها در مقیاس‌های مختلف را فراهم می‌سازد. شاخص ویلموت میزان تطابق کلی بین داده‌های مشاهده‌شده و شبیه‌سازی‌شده را ارزیابی می‌کند و مقدار نزدیک به ۱ نشان‌دهنده توافق بسیار خوب بین دو مجموعه داده است. ضریب نش–ساتکلیف عملکرد مدل را با مقایسه واریانس خطاها نسبت به واریانس داده‌های واقعی می‌سنجد. مقدار ۱ بیانگر تطابق کامل و مقدارهای منفی نشان‌دهنده عملکرد ضعیف‌تر از میانگین مشاهدات است. در ادامه، روابط ریاضی این معیارها ارائه شده‌اند (روابط ۴ تا ۹).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2 \quad (۴)$$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(O_i-\bar{O}_i)(P_i-\bar{P}_i)}{\sum_{i=1}^{n}(O_i-\bar{O})^2\sum_{i=1}^{n}(P_i-\bar{P}_i)^2}\right]^2 \quad (۵)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i-O_i)^2}{n}} \quad (۶)$$

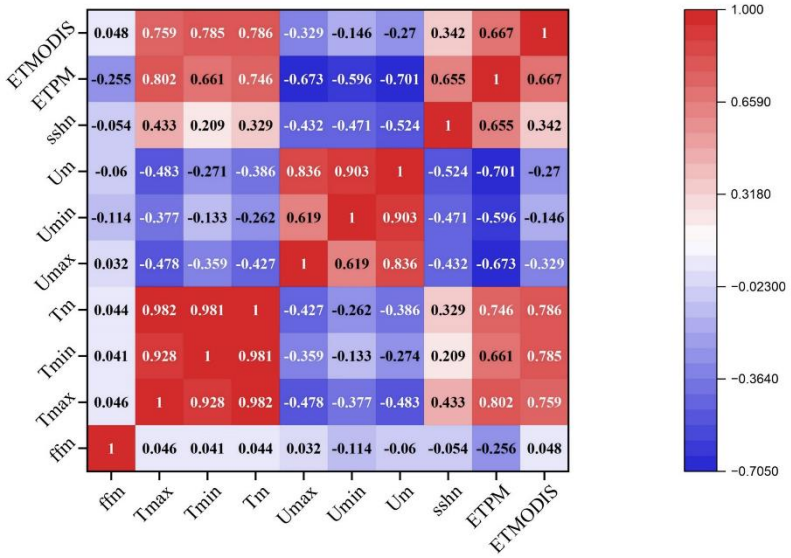
$$WI = 1 - \left[\frac{\sum_{i=1}^{n}(P_i-O_i)^2}{\sum_{i=1}^{n}(|P_i-\bar{O}_i|+|O_i-\bar{O}|)^2}\right] \quad (۷)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i-P_i)^2}{\sum_{i=1}^{n}(O_i-\bar{O}_i)^2} \quad (۸)$$

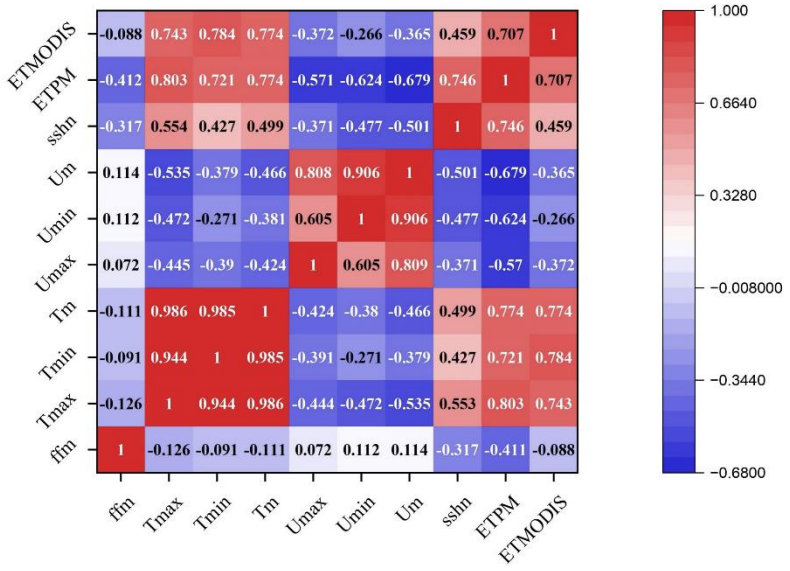
$$NRMSE = \frac{RMSE}{\bar{O}} \times 100 \quad (۹)$$

نتایج و بحث

در این پژوهش، با هدف ارتقای دقت پیش‌بینی تبخیرتعرق مرجع، از ترکیب داده‌های سنجش از دور حاصل از سنجنده MODIS و داده‌های هواشناسی روزانه دو ایستگاه رامسر و بندرانزلی در بازه زمانی ۲۰۰۱ تا ۲۰۲۳ استفاده شده است. برای مدل‌سازی، دو الگوریتم یادگیری ماشین شامل XGB و نسخه بهبودیافته آن HXGB به‌کار گرفته شده‌اند. مدل‌ها با بهره‌گیری از پارامترهای دما، رطوبت نسبی، ساعات آفتابی، سرعت باد و ET_{MODIS} آموزش دیده‌اند تا عملکرد آن‌ها در شرایط اقلیمی مرطوب و بسیار مرطوب ارزیابی شود. شکل‌های ۴ و ۵ نمودار گرمایشی ایستگاه‌های رامسر و بندرانزلی را نشان می‌دهد. همان‌طور که درنمودار گرمایشی هردو شهر مشاهده می‌شود، پارامترهایی مانند ساعات آفتابی و رطوبت نسبی بیشترین ارتباط را با تبخیر و تعرق دارند. دلیل این امر آن است که ساعات آفتابی مستقیماً بر میزان انرژی تابشی مؤثر بوده و رطوبت نسبی ظرفیت جو برای پذیرش بخار آب را تعیین می‌کند. بنابراین این متغیرها نقش کلیدی در فرآیند تبادل انرژی و رطوبت ایفا می‌کنند. در شکل ۴، مربوط به نمودار گرمایی ایستگاه رامسر، مشاهده می‌شود که در آن ET_{MODIS} و تبخیرتعرق فائو-پنمن-مانتیث با پارامترهای دمایی همبستگی بالایی دارد که نشان دهنده تاثیر زیاد داده‌های دمایی بر تبخیرتعرق است. همبستگی ET_{MODIS} با ساعات آفتابی (SSHN) برابر ۰/۳۴۲ است که نشان‌دهنده تأثیر نسبتی ساعات آفتابی بر تبخیرتعرق MODIS است. اما همبستگی ET_{PM} با SSHN برابر ۰/۶۵۵ است که نشان‌دهنده تأثیر نسبتا زیاد ساعات آفتابی بر تبخیرتعرق فائو-پنمن- مانتیث است. U_{min} و U_{max} با تبخیرتعرق همبستگی قوی و منفی دارد که نشان‌دهنده این است که با بالا رفتن رطوبت تبخیرتعرق کاهش می‌یابد. ffm نیز با تبخیرتعرق، همبستگی نسبتا کم اما تاثیرگذاری دارد، که بیانگر تأثیر باد بر تبخیرتعرق است که با برداشتن هوای اشباع در سطح تبخیر سرعت تبخیرتعرق را افزایش می‌دهد. این ماتریس به‌طور عددی ساختار هم‌وابستگی متغیرها را آشکار می‌سازد و در انتخاب ورودی‌های مؤثر برای مدل‌سازی اقلیمی و هیدرولوژیکی نقش کلیدی دارد. در شکل ۵، مربوط به نمودار گرمایی ایستگاه بندرانزلی، مشاهده می‌شود که همبستگی بالای بین ET_{MODIS} و ET_{PM} (ضریب ۰/۷۰۷) نشان‌دهنده هم‌راستایی بالای بین تبخیرتعرق ماهواره‌ای و مدل فائو-پنمن-مانتیث است، که اعتبار مدل را تأیید می‌کند.



شکل ٤- نمودار گرمایشی همبستگی بین پارامترهای هواشناسی و تبخیرتعرق MODIS ایستگاه رامسر



شکل ۵- نمودار گرمایشی همبستگی بین پارامترهای هواشناسی و تبخیرتعرق MODIS ایستگاه بندرانزلی

ffm با ET_{PM} همبستگی ۰/۴۱۲- دارد، که نشان‌دهنده وابستگی نسبتا کم بین سرعت باد و تبخیرتعرق فائو-پنمن-مانتیث است. که منفی بودن آن بعلت اقلیم مرطوب شهر بندرانزلی است. پارامترهای دمایی همبستگی قوی با تبخیرتعرق دارند، که بیانگر افزایش تبخیرتعرق با افزایش دما است. U_m همبستگی بالای ۰/۶۹۷ با تبخیرتعرق فائو-پنمن-مانتیث دارد که بعد از پارامترهای دمایی بالاترین هبستگی را با تبخیرتعرق دارد. همبستگی بالای ساعات آفتابی نیز نشان دهنده این است که با افزایش ساعات آفتابی تبخیرتعرق نیز افزایش می‌یابد.

نتایج نشان داد که بین سرعت باد و تبخیرتعرق همبستگی منفی وجود دارد. این امر می‌تواند ناشی از شرایط اقلیمی منطقه مطالعه باشد. به‌گونه‌ای که در دوره‌های افزایش سرعت باد، رطوبت نسبی نیز افزایش یافته و یا تابش خورشیدی کاهش یافته است، در نتیجه اثر خنک‌کنندگی باد بر سطح گیاه و خاک غالب شده و میزان تبخیرتعرق کاهش یافته است. بنابراین همبستگی منفی مشاهده‌شده بازتابی از تعامل همزمان متغیرهای اقلیمی در منطقه است. همانطور که سو و همکاران (۲۰۲۳) عنوان کردند که بر اساس معادله فائو-پنمن-مانتیث، تبخیرتعرق تحت تأثیر همزمان تابش خورشیدی، دما، رطوبت نسبی و سرعت باد قرار دارد. در منطقه مورد مطالعه، افزایش سرعت باد غالباً با افزایش رطوبت نسبی یا کاهش تابش همراه بوده است. در نتیجه اثر خنک کنندگی باد بر سطح گیاه و خاک غالب شده و مقدار تبخیرتعرق کاهش یافته است. بنابراین همبستگی منفی مشاهده شده میان سرعت باد و تبخیرتعرق در نقشه همبستگی، بازتابی از تعامل همزمان این متغیرها در چارچوب معادله فائو-پنمن-مانتیث است (Su et al., 2023).

جدول ۳ ترکیب‌های متفاوت متغیرهای ورودی شامل پارامترهای هواشناسی و سنجش از دور را برای تخمین تبخیرتعرق مرجع ارائه می‌دهد. این ترکیب‌ها بر اساس ضریب همبستگی پیرسون مشخص شده‌اند. هر ترکیب، مجموعه‌ای منحصربه‌فرد از پارامترها را منعکس می‌کند و فرآیند تلفیق داده‌های چندمنبعی را برای توسعه مدل‌های قوی‌تر برجسته می‌سازد. به‌عنوان نمونه، استفاده هم‌زمان از دمای میانگین، دمای کمینه و دمای بیشینه درک دقیق‌تری از دینامیک تبخیرتعرق فراهم می‌کند. دما به‌طور مستقیم بر نرخ تبخیرتعرق تأثیر می‌گذارد و از این‌رو یکی از عناصر کلیدی در هر مدل تبخیرتعرق محسوب می‌شود. انتخاب هدفمند این پارامترها به پژوهشگر امکان می‌دهد تا اثرات تغییرات اقلیمی را بر برآورد ET_0 با دقت بیشتری تحلیل کند. علاوه بر دما، افزودن پارامترهای مکمل مانند رطوبت نسبی بعد جدیدی به ساختار مدل می‌بخشد. رطوبت نسبی یکی از محرک‌های اصلی تبخیر است، زیرا انتقال بخار آب از سطح به جو را تنظیم می‌کند. این تلفیق دقیق داده‌ها موجب شناخت عمیق‌تری از عوامل مؤثر بر تبخیرتعرق شده و دقت تخمین‌ها را به‌طور قابل‌توجهی افزایش می‌دهد. از سوی دیگر، ساعات آفتابی که با پارامتر SSHN سنجیده می‌شود، نقش محوری در فرآیند تبخیرتعرق دارد. تابش نه‌تنها منبع انرژی برای تبخیر از سطح زمین است، بلکه فعالیت‌های زیستی گیاهان از جمله فتوسنتز را نیز تحت تأثیر قرار می‌دهد. بنابراین، گنجاندن این پارامتر در مدل موجب تقویت تحلیل‌ها و افزایش دقت برآورد ET_0 می‌شود. نکته مهم دیگر، استفاده از پارامتر ET_{MODIS} به‌عنوان داده سنجش از دور در کنار سایر ویژگی‌هاست که نشان‌دهنده ادغام اطلاعات ماهواره‌ای در فرآیند مدل‌سازی است. سامانه‌های سنجش از دور بینش منحصربه‌فردی از وضعیت سطح

زمین ارائه می‌دهند و اطلاعات ضروری درباره توزیع مکانی ET_0 در محیط‌های مختلف فراهم می‌سازند. استفاده از ET_{MODIS} به مدل‌سازان این امکان را می‌دهد تا با بهره‌گیری از داده‌های وسیع و به‌روز، دقت مدل را افزایش دهند. داده‌هایی که ممکن است از طریق مشاهدات زمینی به‌طور کامل قابل دستیابی نباشند. این رویکرد نه‌تنها موجب افزایش پایداری مدل‌ها می‌شود، بلکه دامنه کاربرد نتایج پژوهش را نیز در شرایط اقلیمی متنوع گسترش می‌دهد.

بررسی هریک از هشت ترکیب ارائه‌شده در جدول، امکان ارزیابی سیستماتیک تأثیر متغیرهای ورودی بر خروجی مدل را فراهم می‌سازد. به‌ویژه ترکیب‌هایی که شامل پارامترهای هواشناسی متعدد در کنار ET_{MODIS} هستند، عملکرد مدل را به‌طور چشمگیری بهبود می‌بخشند، زیرا نمایی جامع از شرایط جوی مؤثر بر تبخیرتعرق ارائه می‌دهند.

جدول ۳- ترکیب‌های ورودی مدل برای تخمین ET_0

شماره	سناریو
۱	T_m, T_{min}, T_{max}
۲	SSHN, T_m, T_{min}, T_{max}
۳	U_{max}, U_m, U_{min}, SSHN, T_m, T_{min}, T_{max}
۴	ffm, U_{max}, U_m, U_{min}, SSHN, T_m, T_{min}, T_{max}
۵	ET_{MODIS}, T_m, T_{min}, T_{max}
۶	ET_{MODIS}, SSHN, T_m, T_{min}, T_{max}
۷	ET_{MODIS}, U_{max}, U_m, U_{min}, SSHN, T_m, T_{min}, T_{max}
۸	ET_{MODIS}, ffm, U_{max}, U_m, U_{min}, SSHN, T_m, T_{min}, T_{max}

در جدول ۴ نتایج ارزیابی مدل XGB و HXGB در ایستگاه بندرانزلی نشان داده شده است. در ایستگاه بندرانزلی، مدل XGB در سناریو اول عملکرد متوسطی از خود نشان داده است. مقدار MSE برابر با ۱/۴۷ و RMSE برابر با ۱/۲۱ میلی‌متر در روز بوده که نسبت به سایر سناریوها بالاتر است و نشان‌دهنده خطای بیشتر در پیش‌بینی است. مقدار NRMSE برابر با ۰/۳۱ و ضریب همبستگی برابر با ۰/۷۷ است که نشان‌دهنده همبستگی قابل قبول اما نه مطلوب بین مقادیر پیش‌بینی‌شده و مشاهداتی است. شاخص نش-ساتکلیف نیز با مقدار ۰/۷۷ بیانگر عملکرد قابل قبول مدل است، اما شاخص ویلموت با مقدار ۰/۹۳ نشان می‌دهد که تطابق کلی بین داده‌های مشاهده‌شده و شبیه‌سازی‌شده نسبتاً خوب بوده است. در سناریو چهارم، مدل XGB در بندرانزلی بهبود چشمگیری در تمامی معیارها نشان داده است. مقدار MSE به ۰/۱۶ کاهش یافته و RMSE به ۰/۴۰ میلی‌متر در روز رسیده که نشان‌دهنده کاهش قابل توجه خطاهاست. مقدار NRMSE برابر با ۰/۱۰ و ضریب همبستگی برابر با ۰/۹۷ است که بیانگر همبستگی بسیار قوی بین داده‌های پیش‌بینی‌شده و واقعی است. شاخص NSE نیز با مقدار ۰/۹۷ عملکرد بسیار مطلوب مدل را تأیید می‌کند و شاخص WI با مقدار ۰/۹۹ نشان‌دهنده تطابق عالی بین داده‌هاست. این سناریو با ترکیب کامل پارامترهای هواشناسی و ET_{MODIS}، جامع‌ترین ورودی‌ها را داشته و همین امر موجب افزایش دقت مدل شده است. در سناریو پنجم، مدل XGB در بندرانزلی با استفاده از داده‌های دمایی و ET_{MODIS} عملکردی نسبتاً ضعیف‌تر نسبت به سناریو چهارم دارد. مقدار MSE برابر با ۱/۲۸ و RMSE برابر با ۱/۱۳ میلی‌متر در روز است که نسبت به سناریوهای ترکیبی‌تر، خطای بیشتری دارد. مقدار NRMSE برابر با ۰/۲۹ و R^2 برابر با ۰/۸۰ است که همبستگی متوسطی را نشان می‌دهد. شاخص NSE با مقدار ۰/۸۰ و WI با مقدار ۰/۹۴ نشان‌دهنده عملکرد قابل قبول اما نه بهینه مدل در این ترکیب ورودی است. در سناریو هشتم، مدل XGB در بندرانزلی بهترین عملکرد را در میان تمامی سناریوها ارائه داده است. مقدار MSE به ۰/۰۸ و RMSE به ۰/۲۹ میلی‌متر در روز کاهش یافته که نشان‌دهنده دقت بسیار بالا در پیش‌بینی است. مقدار NRMSE برابر با ۰/۰۷ و R^2 برابر با ۰/۹۸ است که بیانگر همبستگی بسیار قوی بین داده‌هاست. شاخص NSE نیز با مقدار ۰/۹۸ و WI با مقدار ۰/۹۹ نشان‌دهنده تطابق تقریباً کامل بین داده‌های مشاهده‌شده و شبیه‌سازی‌شده است. این سناریو با ترکیب کامل‌ترین مجموعه ورودی‌ها شامل پارامترهای دمایی، رطوبت نسبی، سرعت باد، ساعات آفتابی و ET_{MODIS} جامع‌ترین مدل‌سازی را انجام داده و بهینه‌ترین نتایج را ارائه کرده است. در جدول ۵ نتایج ارزیابی مدل XGB و HXGB در ایستگاه رامسر نشان داده شده است. در ایستگاه رامسر، مدل XGB در سناریو اول عملکردی مشابه بندرانزلی دارد. مقدار MSE برابر با ۱/۲۴ و RMSE برابر با ۱/۱۱ میلی‌متر در روز بوده و NRMSE برابر با ۰/۲۵ است. ضریب همبستگی برابر با ۰/۷۷ و شاخص NSE نیز برابر با ۰/۷۷ است که نشان‌دهنده همبستگی متوسط و خطای نسبتاً بالا در پیش‌بینی است. شاخص WI با مقدار ۰/۹۳ تطابق قابل قبولی را نشان می‌دهد، اما نسبت به سناریوهای پیشرفته‌تر ضعیف‌تر است. در سناریو چهارم، مدل XGB

در رامسر نیز عملکرد بسیار مطلوبی داشته است. مقدار MSE برابر با ۰/۱۶ و RMSE برابر با ۰/۴۰ میلی‌متر در روز بوده و NRMSE برابر با ۰/۰۹ است. ضریب همبستگی برابر با ۰/۹۷ و شاخص NSE نیز برابر با ۰/۹۷ است که نشان‌دهنده دقت بالا و همبستگی قوی بین داده‌هاست. شاخص WI با مقدار ۰/۹۹ نیز تطابق بسیار خوب بین داده‌های مشاهده‌شده و پیش‌بینی‌شده را تأیید می‌کند. این سناریو نیز مانند بندرانزلی از ترکیب کامل پارامترهای هواشناسی بهره برده و نتایج بسیار دقیقی ارائه داده است. در سناریو پنجم، مدل XGB در رامسر عملکردی مشابه بندرانزلی دارد. مقدار MSE برابر با ۱/۱۴ و RMSE برابر با ۱/۰۷ میلی‌متر در روز بوده و NRMSE برابر با ۰/۲۴ است. ضریب همبستگی برابر با ۰/۷۹ و شاخص NSE نیز برابر با ۰/۷۹ است که نشان‌دهنده عملکرد متوسط مدل در این ترکیب ورودی است. شاخص WI با مقدار ۰/۹۳ نیز تطابق قابل قبولی را نشان می‌دهد، اما نسبت به سناریوهای جامع‌تر ضعیف‌تر است. در سناریو هشتم، مدل XGB در رامسر نیز بهترین عملکرد را در میان تمامی سناریوها داشته است. مقدار MSE برابر با ۰/۰۸ و RMSE برابر با ۰/۲۹ میلی‌متر در روز بوده و NRMSE برابر با ۰/۰۶ است. ضریب همبستگی برابر با ۰/۹۸ و شاخص NSE نیز برابر با ۰/۹۸ است که نشان‌دهنده دقت بسیار بالا و همبستگی قوی بین داده‌هاست. شاخص WI با مقدار ۰/۹۹ نیز تطابق عالی بین داده‌های مشاهده‌شده و شبیه‌سازی‌شده را تأیید می‌کند. این سناریو با بهره‌گیری از کامل‌ترین مجموعه ورودی‌ها، دقیق‌ترین نتایج را در ایستگاه رامسر ارائه داده است.

در مقایسه بین دو ایستگاه، عملکرد مدل‌ها در سناریوهای اول، چهارم، پنجم و هشتم نشان‌دهنده الگوی مشابهی است. در هر دو ایستگاه، سناریو هشتم با ترکیب جامع‌ترین ورودی‌ها شامل پارامترهای دمایی، رطوبت نسبی، سرعت باد، ساعات آفتابی و ET_{MODIS}، بهترین نتایج را ارائه داده است. مقدار RMSE در بندرانزلی برابر با ۰/۲۹ و در رامسر برابر با ۰/۲۹ میلی‌متر در روز بوده، مقدار R^2 در هر دو ایستگاه برابر با ۰/۹۸ و شاخص NSE نیز برابر با ۰/۹۸ بوده که بیانگر همبستگی بسیار قوی و عملکرد عالی مدل‌هاست. شاخص WI نیز در هر دو منطقه برابر با ۰/۹۹ بوده که تطابق کامل بین داده‌های مشاهده‌شده و پیش‌بینی‌شده را نشان می‌دهد. در سناریو چهارم نیز عملکرد مدل‌ها در هر دو ایستگاه بسیار مطلوب بوده و تفاوت عددی بین آن‌ها ناچیز است. در سناریو پنجم، که تنها از داده‌های دمایی و ET_{MODIS} استفاده شده، عملکرد مدل‌ها نسبت به سناریوهای جامع‌تر ضعیف‌تر بوده و RMSE در بندرانزلی برابر با ۱/۱۳ و در رامسر برابر با ۱/۰۷ میلی‌متر در روز بوده است. در سناریو اول نیز که تنها از داده‌های دمایی استفاده شده، مدل‌ها عملکرد متوسطی داشته‌اند و RMSE در بندرانزلی برابر با ۱/۲۱ و در رامسر برابر با ۱/۱۱ میلی‌متر در روز بوده است.

جدول ٤- نتایج ارزیابی مدل XGB و HXGB در ایستگاه بندرانزلی

روش	پارامتر					
	R^2	RMSE (mm day^{-1})	WI	NSE	NRMSE	MSE
XGB 1	۰/۷۷	۱/۲۱	۰/۹۳	۰/۷۷	۰/۳۱	۱/۴۷
XGB 2	۰/۸۴	۰/۹۹	۰/۹۵	۰/۸۴	۰/۲۵	۰/۹۹
XGB 3	۰/۸۹	۰/۸۲	۰/۹۷	۰/۸۹	۰/۲۱	۰/۶۸
XGB 4	۰/۹۷	۰/۴۰	۰/۹۹	۰/۹۷	۰/۱۰	۰/۱۶
XGB 5	۰/۸۰	۱/۱۳	۰/۹۴	۰/۸۰	۰/۲۹	۱/۲۸
XGB 6	۰/۸۶	۰/۹۵	۰/۹۶	۰/۸۶	۰/۲۴	۰/۹۱
XGB 7	۰/۹۰	۰/۷۹	۰/۹۷	۰/۹۰	۰/۲۰	۰/۶۳
XGB 8	۰/۹۸	۰/۲۹	۰/۹۹	۰/۹۸	۰/۰۷	۰/۰۸
HXGB 1	۰/۷۸	۱/۲۰	۰/۹۳	۰/۷۸	۰/۳۰	۱/۴۵
HXGB 2	۰/۸۴	۱/۰۰	۰/۹۵	۰/۸۴	۰/۲۵	۱/۰۰
HXGB 3	۰/۹۰	۰/۷۸	۰/۹۷	۰/۹۰	۰/۱۹	۰/۶۰
HXGB 4	۰/۹۹	۰/۲۳	۰/۹۹	۰/۹۹	۰/۰۵	۰/۰۵
HXGB 5	۰/۸۲	۱/۰۷	۰/۹۴	۰/۸۲	۰/۲۷	۱/۱۶
HXGB 6	۰/۸۷	۰/۹۱	۰/۹۶	۰/۸۷	۰/۲۳	۰/۸۳
HXGB 7	۰/۹۲	۰/۶۹	۰/۹۸	۰/۹۲	۰/۱۷	۰/۴۸
HXGB 8	۰/۹۹	۰/۱۱	۰/۹۹	۰/۹۹	۰/۰۲	۰/۰۱

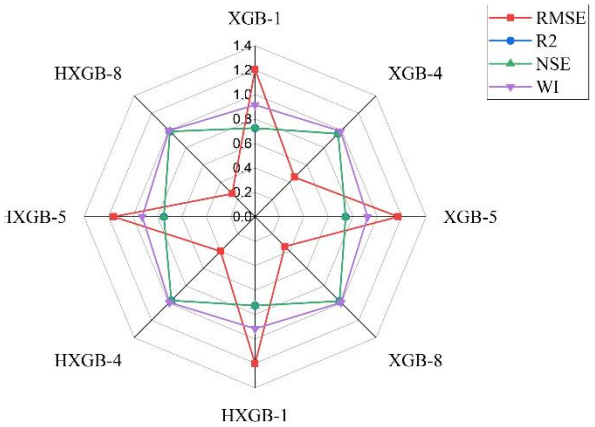
جدول ۵- نتایج ارزیابی مدل XGB و HXGB در ایستگاه رامسر

روش	پارامتر					
	R^2	RMSE ($mm\ day^{-1}$)	WI	NSE	NRMSE	MSE
XGB 1	۰/۷۷	۱/۱۱	۰/۹۳	۰/۷۷	۰/۲۵	۱/۲۴
XGB 2	۰/۸۲	۰/۹۸	۰/۹۵	۰/۸۲	۰/۲۲	۰/۹۷
XGB 3	۰/۸۸	۰/۸۰	۰/۹۶	۰/۸۸	۰/۱۸	۰/۶۴
XGB 4	۰/۹۷	۰/۴۰	۰/۹۹	۰/۹۷	۰/۰۹	۰/۱۶
XGB 5	۰/۷۹	۱/۰۷	۰/۹۳	۰/۷۹	۰/۲۴	۱/۱۴
XGB 6	۰/۸۳	۰/۹۶	۰/۹۵	۰/۸۳	۰/۲۱	۰/۹۲
XGB 7	۰/۸۹	۰/۷۶	۰/۹۷	۰/۸۹	۰/۱۷	۰/۵۹
XGB 8	۰/۹۸	۰/۲۹	۰/۹۹	۰/۹۸	۰/۰۶	۰/۰۸
HXGB 1	۰/۷۸	۱/۰۹	۰/۹۳	۰/۷۸	۰/۲۴	۰/۱۹
HXGB 2	۰/۸۲	۰/۹۸	۰/۹۴	۰/۸۲	۰/۲۲	۰/۹۷
HXGB 3	۰/۸۹	۰/۷۶	۰/۹۷	۰/۸۹	۰/۱۷	۰/۵۸
HXGB 4	۰/۹۸	۰/۲۶	۰/۹۹	۰/۹۸	۰/۰۶	۰/۰۷
HXGB 5	۰/۸۲	۰/۹۹	۰/۹۴	۰/۸۲	۰/۲۲	۰/۹۸
HXGB 6	۰/۸۴	۰/۹۳	۰/۹۵	۰/۸۴	۰/۲۰	۰/۸۶
HXGB 7	۰/۹۰	۰/۷۳	۰/۹۷	۰/۹۰	۰/۱۶	۰/۵۳
HXGB 8	۰/۹۹	۰/۱۹	۰/۹۹	۰/۹۹	۰/۰۴	۰/۰۳

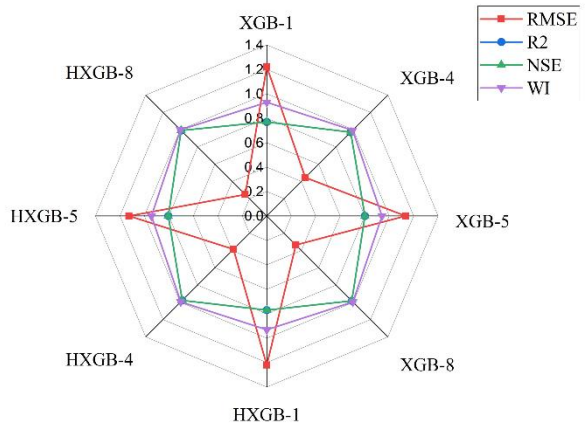
در مجموع نتایج نشان می‌دهد که ترکیب داده‌های ماهواره‌ای با پارامترهای هواشناسی به‌ویژه در سناریو هشتم موجب افزایش قابل‌توجه دقت مدل‌ها در هر دو ایستگاه شده است. همچنین، تفاوت عملکرد بین بندرانزلی و رامسر در سناریوهای مختلف بسیار اندک بوده و مدل HXGB در تمامی موارد نسبت به مدل پایه XGB عملکرد بهتری از خود نشان داده است. این موضوع نشان‌دهنده پایداری و قابلیت تعمیم بالای مدل ترکیبی HXGB در شرایط اقلیمی مرطوب و بسیار مرطوب شمال کشور است.

نمودار راداری ارائه‌شده در شکل‌های ۶ و ۷، شاخص‌های آماری مربوط به بهترین مدل‌ها در مرحله اعتبارسنجی برای ایستگاه‌های رامسر و بندرانزلی را به‌تصویر می‌کشد و نمایی بصری از عملکرد مقایسه‌ای آن‌ها در پیش‌بینی مقادیر ET_0 فراهم می‌سازد. شکل ۶، برای ایستگاه بندرانزلی نشان می‌دهد که مدل HXGB در سناریو هشتم (HXGB-8) به‌وضوح عملکرد برتری نسبت به سایر مدل‌ها دارد. در این سناریو، تمامی معیارهای ارزیابی در بالاترین سطح قرار گرفته‌اند. مقدار RMSE به حداقل رسیده و شاخص‌های همبستگی و توافق ویلموت تقریباً به مقدار کامل نزدیک شده‌اند. این نشان می‌دهد که ترکیب کامل‌ترین مجموعه ورودی‌ها موجب دستیابی به بالاترین دقت در پیش‌بینی ET_0 شده است. در سناریو چهارم نیز مدل HXGB عملکرد بسیار مطلوبی دارد. شاخص‌های R^2 و NSE در سطح بالایی قرار دارند و RMSE نیز به‌طور قابل توجهی کاهش یافته است. این سناریو نیز از ترکیب جامع پارامترهای هواشناسی استفاده کرده و بدون داده‌های ماهواره‌ای نیز توانسته دقت بالایی ارائه دهد، هرچند نسبت به HXGB-8 اندکی ضعیف‌تر است. مدل XGBدر سناریوهای مشابه نیز عملکرد خوبی دارند، اما در مقایسه با نسخه‌های ترکیبی HXGB، در تمامی معیارها اندکی پایین‌تر هستند. به‌ویژه در سناریو XGB-1 و XGB-5 که ورودی‌ها محدودترند، افت قابل توجهی در دقت مدل مشاهده می‌شود. RMSE در این موارد بالاتر است و شاخص‌های R^2 و NSE نیز در سطح متوسط قرار دارند. به‌طور کلی، نمودار راداری بندرانزلی نشان می‌دهد که هرچه ترکیب ورودی‌ها کامل‌تر و متنوع‌تر باشد، به‌ویژه با حضور داده‌های ماهواره‌ای ET_{MODIS}، عملکرد مدل به‌طور چشمگیری بهبود می‌یابد. همچنین استفاده از مدل ترکیبی HXGB نسبت به مدل پایه XGB در تمامی سناریوها منجر به افزایش دقت و پایداری شده است. این موضوع اهمیت تلفیق داده‌های چندمنبعی و بهینه‌سازی ساختار مدل را در مناطق مرطوب با نوسانات اقلیمی بالا به‌خوبی اثبات می‌کند. شکل ۷، برای ایستگاه رامسر نشان می‌دهد که مدل HXGB در سناریو هشتم (HXGB-8) عملکردی ممتاز و کاملاً بهینه از خود نشان داده است. در این سناریو، مقدار RMSE به حداقل رسیده و شاخص‌های R^2 ، NSE و WI همگی در بالاترین سطح قرار دارند. این موضوع نشان‌دهنده دقت بسیار بالا، همبستگی قوی و تطابق کامل بین داده‌های مشاهده‌شده و پیش‌بینی‌شده است. ترکیب جامع پارامترهای هواشناسی و داده‌های ماهواره‌ای ET_{MODIS} در این سناریو موجب شده تا مدل بتواند به‌خوبی دینامیک تبخیرتعرق را در منطقه رامسر

بازنمایی کند .در سناریو چهارم نیز مدل HXGB عملکرد بسیار مطلوبی دارد. شاخص‌های R^2 و NSE در سطح بالا قرار گرفته‌اند و RMSE نیز به‌طور قابل توجهی کاهش یافته است. این سناریو با استفاده از داده‌های زمینی و بدون ET_{MODIS} توانسته دقت بالایی ارائه دهد، که نشان‌دهنده قدرت مدل در استخراج وابستگی‌های اقلیمی از داده‌های هواشناسی است. با این حال، همچنان نسبت به سناریو هشتم اندکی ضعیف‌تر است. مدل‌های پایه XGB در سناریوهای مشابه نیز عملکرد خوبی دارند، اما در مقایسه با نسخه‌های ترکیبی HXGB، در تمامی معیارها اندکی پایین‌تر هستند. به‌ویژه در سناریوهای محدودتر مانند XGB-1 و XGB-5، افت قابل توجهی در دقت مدل مشاهده می‌شود. RMSE در این موارد بالاتر است و شاخص‌های R^2 و NSE نیز در سطح متوسط قرار دارند، که نشان‌دهنده محدودیت مدل در شرایط ورودی ناقص است. در مجموع، نمودار راداری رامسر نیز مانند بندرانزلی تأیید می‌کند که استفاده از مدل ترکیبی HXGB در کنار ورودی‌های جامع، به‌ویژه داده‌های ماهواره‌ای ET_{MODIS}، موجب افزایش چشمگیر دقت و پایداری مدل در پیش‌بینی تبخیرتعرق مرجع شده است. همچنین، تطابق الگوهای عملکرد در هر دو ایستگاه نشان می‌دهد که مدل HXGB قابلیت تعمیم بالایی در مناطق مرطوب شمال کشور دارد و می‌تواند به‌عنوان ابزار قابل اعتماد در مطالعات منابع آب و اقلیم مورد استفاده قرار گیرد.



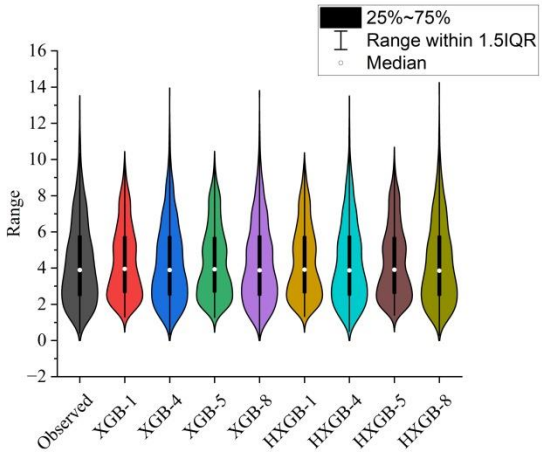
شکل ٦- نمودار راداری ایستگاه بندرانزلی



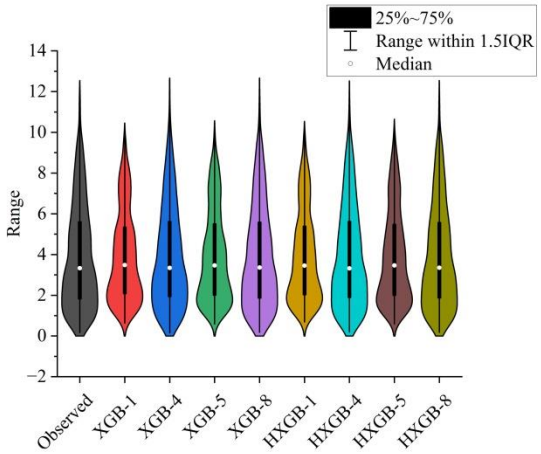
شکل ۷- نمودار راداری ایستگاه رامسر

شکل‌های ۸ و ۹ توزیع مقادیر پیش‌بینی‌شده تبخیرتعرق مرجع توسط مدل‌های مختلف XGB و HXGB را در مقایسه با داده‌های مشاهده‌شده در ایستگاه بندرانزلی و رامسر نشان می‌دهند. شکل ۸ نمودار ویولون برای ایستگاه بندرانزلی نشان می‌دهد که مدل‌های اولیه مانند XGB-1 و HXGB-1 دارای پراکندگی وسیع‌تری هستند و میانه مقادیر آن‌ها فاصله بیشتری از داده‌های مشاهده‌شده دارد. این موضوع نشان‌دهنده خطای بالاتر و همبستگی ضعیف‌تر در شرایطی است که ورودی‌ها محدود بوده‌اند. در مقابل در سناریوهای پیشرفته‌تر مانند XGB-4 و HXGB-4 توزیع‌ها باریک‌تر شده و میانه مقادیر به داده‌های واقعی نزدیک‌تر است که بیانگر بهبود دقت مدل‌ها با افزایش پارامترهای ورودی است. در سناریو پنجم، که ترکیب داده‌های دمایی با ET_{MODIS} به‌کار رفته، توزیع‌ها همچنان به داده‌های واقعی نزدیک هستند اما پهنای ویولون‌ها نسبت به سناریو چهارم و هشتم بیشتر است. این موضوع نشان می‌دهد که استفاده صرف از داده‌های دمایی و ET_{MODIS} به‌تنهایی کافی نیست و برای دستیابی به دقت بالاتر نیاز به ترکیب جامع‌تر پارامترها وجود دارد. در نهایت، در سناریو هشتم، به‌ویژه در مدل HXGB-8 توزیع مقادیر تقریباً به‌طور کامل با داده‌های مشاهده‌شده هم‌پوشانی دارد. ویولون‌ها باریک‌تر و متقارن‌تر هستند و میانه مقادیر دقیقاً در محدوده داده‌های واقعی قرار گرفته است. این موضوع نشان‌دهنده دقت بسیار بالای مدل در این سناریو است و با نتایج عددی معیارهای ارزیابی مانند RMSE پایین و R^2 و NSE نزدیک به یک کاملاً همخوانی دارد. شکل ۹ نمودار ویولون برای ایستگاه رامسر نشان می‌دهد که در سناریوهای ابتدایی مانند XGB-1 و HXGB-1 پهنای ویولون‌ها بیشتر است و میانه مقادیر فاصله محسوسی از داده‌های واقعی دارد. این موضوع بیانگر خطای بالاتر و پراکندگی زیادتر پیش‌بینی‌ها در شرایطی است که ورودی‌ها محدود بوده‌اند. در سناریو چهارم، توزیع‌ها باریک‌تر شده و میانه مقادیر به داده‌های مشاهده‌شده نزدیک‌تر است. این نشان می‌دهد که افزودن پارامترهای بیشتر از جمله رطوبت نسبی و سرعت باد توانسته دقت پیش‌بینی را بهبود دهد و خطاها را کاهش دهد. در سناریو پنجم، که ترکیب داده‌های دمایی با ET_{MODIS} به‌کار رفته، توزیع‌ها همچنان به داده‌های واقعی نزدیک هستند اما پهنای ویولون‌ها نسبت به سناریو

چهارم و هشتم بیشتر است. در سناریو هشتم به‌ویژه در مدل HXGB-8 توزیع مقادیر تقریباً به‌طور کامل با داده‌های مشاهده‌شده هم‌پوشانی دارد. ویولون‌ها باریک‌تر و متقارن‌تر هستند و میانه مقادیر دقیقاً در محدوده داده‌های واقعی قرار گرفته است. این موضوع نشان‌دهنده دقت بسیار بالای مدل در این سناریو است و با نتایج عددی معیارهای ارزیابی مانند RMSE پایین و R^2 و NSE نزدیک به یک کاملاً همخوانی دارد. به‌طور کلی، نمودار ویولون رامسر همانند بندرانزلی تأیید می‌کند که هرچه ورودی‌ها کامل‌تر باشند و مدل ترکیبی HXGB به‌کار گرفته شود، دقت پیش‌بینی ET_0 به سطحی نزدیک به داده‌های واقعی ارتقا می‌یابد.



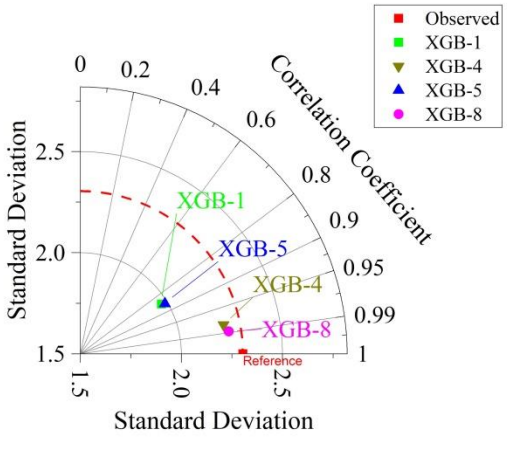
شکل ۸- نمودار ویولون ایستگاه بندرانزلی



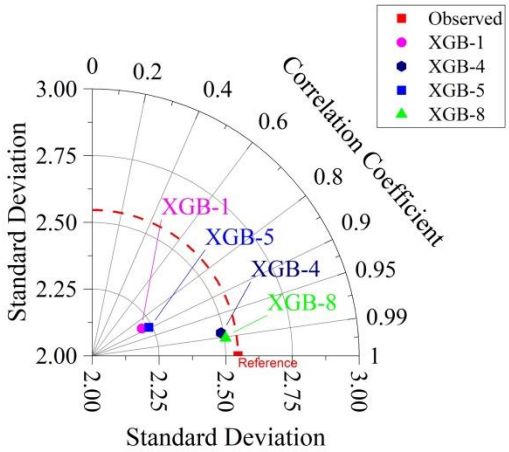
شکل ۹- نمودار ویولون ایستگاه رامسر

شکل‌های ۱۰ و ۱۱ به ترتیب نمودارهای تیلور برای ایستگاه‌های بندرانزلی و رامسر را نشان می‌دهند. در بررسی شکل ۱۰ مربوط به نمودار تیلور ایستگاه بندرانزلی نتایج نشان می‌دهد که مدل‌ها در چه فاصله‌ای از داده‌های مشاهده‌ای قرار دارند و همبستگی، انحراف معیار و RMSE آن‌ها چگونه تغییر می‌کند. در این نمودار، مدل XGB-8 نزدیک‌ترین موقعیت را به نقطه مرجع داده‌های مشاهده‌ای دارد. این یعنی همبستگی آن با داده‌های واقعی بسیار بالا بوده و انحراف معیار آن تقریباً با داده‌های مشاهده‌ای هم‌خوان است، در حالی که RMSE نیز در پایین‌ترین سطح قرار دارد. مدل XGB-4 نیز عملکرد مطلوبی دارد و نسبت به داده‌های واقعی نزدیک است، اما اندکی از XGB-8 ضعیف‌تر است. در مقابل مدل‌های XGB-1 و XGB-5 فاصله بیشتری از نقطه مرجع دارند. این فاصله نشان‌دهنده خطای بالاتر و همبستگی ضعیف‌تر است، به‌ویژه در XGB-1 که ورودی‌های محدودی داشته و نتوانسته تغییرات واقعی را به‌خوبی بازتاب دهد. به‌طور کلی، نمودار تیلور بندرانزلی تأیید می‌کند که هرچه ورودی‌ها کامل‌تر باشند (مانند سناریو هشتم)، مدل به داده‌های واقعی نزدیک‌تر می‌شود و دقت آن به‌طور چشمگیری افزایش می‌یابد. همچنین، سناریوهای میانی مانند چهارم نیز عملکرد بسیار خوبی دارند و نشان می‌دهند که حتی بدون داده‌های ماهواره‌ای، ترکیب جامع پارامترهای هواشناسی می‌تواند نتایج دقیقی ارائه دهد. شکل ۱۱ که نمودار تیلور ایستگاه رامسر است، همانند بندرانزلی امکان مقایسه همزمان سه معیار کلیدی یعنی ضریب همبستگی، انحراف معیار و RMSE را فراهم می‌کند. در این نمودار، مدل XGB-8 نزدیک‌ترین موقعیت را به نقطه مرجع داده‌های مشاهده‌ای دارد. این یعنی همبستگی آن با داده‌های واقعی بسیار بالا بوده و انحراف معیار آن تقریباً با داده‌های مشاهده‌ای هم‌خوان است، در حالی که RMSE نیز در پایین‌ترین سطح قرار دارد. مدل XGB-4 نیز عملکرد مطلوبی دارد و نسبت به داده‌های واقعی نزدیک است، اما اندکی از XGB-8 ضعیف‌تر است. در مقابل مدل‌های XGB-1 و XGB-5 فاصله بیشتری از نقطه مرجع دارند. این فاصله نشان‌دهنده خطای بالاتر و همبستگی ضعیف‌تر است، به‌ویژه در XGB-1 که ورودی‌های محدودی داشته و نتوانسته تغییرات واقعی را به‌خوبی بازتاب دهد. به‌طور کلی، نمودار تیلور رامسر همان الگوی بندرانزلی را تأیید می‌کند. شکل ۱۲ و ۱۳ نمودار تیلور ایستگاه بندرانزلی و رامسر را برای سناریوهای برگزیده هردومدل HXGB و XGB نشان می‌دهند.

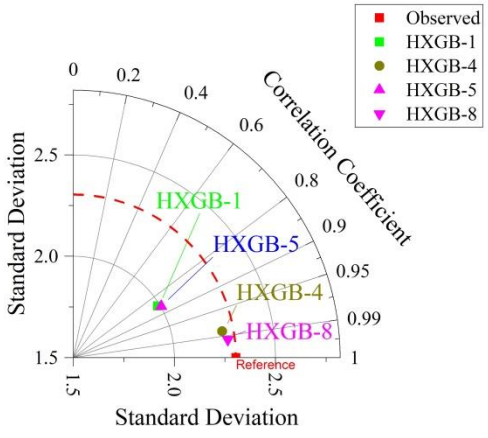
نتایج مشابهی توسط سایر محققان گزارش شده است. خدامرادپور و همکاران (۱۴۰۰) عنوان کردند که در تمام تیپ‌های اقلیمی تبخیرتعرق MODIS نسبت به روش فائو-پنمن-مانتیث دارای واریانس بیشتری در داده‌ها هست و همچنین رابطه خطی قوی بین داده MODIS و روش فائو-پنمن-مانتیث در اکثر تیپ‌های اقلیمی به دلیل ضریب همبستگی بالا (بیشتر از ۰/۸۵) وجود دارد (خدامرادپور و همکاران، ۱۴۰۰). همچنین منور سابق و همکاران (۲۰۲۳) عنوان کردند که از مدل‌های یادگیری ماشین می‌توان برای پیش‌بینی ET_0 به منظور رفع محدودیت‌های روش فائو-پنمن-مانتیث که به چندین پارامتر هواشناسی وابسته است استفاده نمود.



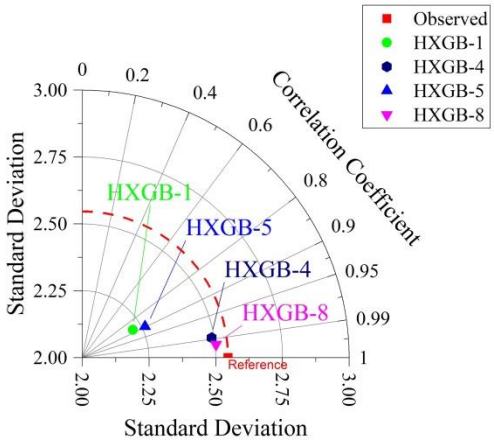
شکل ۱۰- نمودار تیلور ایستگاه بندرانزلی مدل XGB



شکل ۱۱- نمودار تیلور ایستگاه رامسر مدل XGB



شکل ۱۲- نمودار تیلور ایستگاه بندرانزلی مدل HXGB



شکل ۱۳- نمودار تیلور ایستگاه رامسر مدل HXGB

داده‌های هواشناسی روزانه ایستگاه‌های تبریز، سراب و مراغه (دوره ۲۰۰۰-۲۰۲۰) با بکارگیری مدل‌های جنگل تصادفی (RF) و RF بهینه‌شده با الگوریتم ژنتیک (GA-RF) به همراه رویه‌های تجربی مانند بلانی-کریدل، هارگریوز مورد ارزیابی قرار گرفتند. سناریوی ششم که شامل متغیرهای دما، رطوبت نسبی، ساعات آفتاب و سرعت باد بود در تمامی ایستگاه‌ها عملکرد برتری از خود نشان داد. به طور مداوم مدل GA-RF نسبت به روش‌های تجربی عملکرد بهتری داشته و دقت آماری را افزایش داده است (MonavvarSabegh et al., 2023). در مجموع، یافته‌های این پژوهش در کنار مطالعات پیشین، گواه روشنی بر توانمندی مدل‌های یادگیری ماشین در ارتقای دقت پیش‌بینی تبخیرتعرق مرجع و فراهم‌سازی بستری قابل اعتماد برای مدیریت هوشمند منابع آب در شرایط اقلیمی متنوع ایران هستند.

نتیجه‌گیری

نتایج این پژوهش نشان داد که مدل‌های مبتنی بر یادگیری ماشین، به‌ویژه الگوریتم‌های XGB و نسخه ترکیبی HXGB، توانایی بالایی در برآورد تبخیرتعرق مرجع دارند. در میان سناریوهای مختلف، مدل HXGB-8 با ترکیب جامع داده‌های هواشناسی و اطلاعات ماهواره‌ای ET_{MODIS} بهترین عملکرد را ارائه داد و شاخص‌های آماری آن (R^2، NSE، WI بالا و RMSE پایین) نزدیک‌ترین تطابق را با داده‌های مشاهده‌ای نشان داد. همچنین بررسی سناریوهای مختلف نشان داد که هرچه ورودی‌های مدل کامل‌تر باشند، دقت پیش‌بینی به‌طور چشمگیری افزایش می‌یابد. سناریوهای محدود مانند XGB-1 و HXGB-1 به دلیل کمبود داده ورودی، خطای بالاتری داشتند، در حالی که سناریوهای میانی (مانند سناریوهای ۴ و ۵) عملکرد قابل قبولی ارائه کردند اما همچنان نسبت به سناریوی هشتم ضعیف‌تر بودند. تحلیل نمودارهای راداری، ویولون و تیلور به‌طور

هم‌زمان نیز تأیید کرد که مدل HXGB در تمامی شاخص‌ها برتری دارد. این نمودارها نشان دادند که توزیع مقادیر پیش‌بینی‌شده در سناریو هشتم تقریباً به‌طور کامل با داده‌های واقعی هم‌پوشانی دارد و پراکندگی خطاها به حداقل رسیده است. نتایج به‌دست‌آمده در دو ایستگاه بندرانزلی و رامسر، که هر دو در اقلیم مرطوب شمال ایران قرار دارند، الگوی مشابهی را نشان دادند. این موضوع بیانگر آن است که مدل HXGB نه‌تنها در یک ایستگاه بلکه در مناطق مشابه نیز قابلیت تعمیم و اعتمادپذیری بالایی دارد و می‌تواند به‌عنوان یک چارچوب منطقه‌ای مورد استفاده قرار گیرد. همچنین یافته‌های این پژوهش می‌تواند به‌طور مستقیم در مدیریت منابع آب، برنامه‌ریزی کشاورزی و پایش تغییرات اقلیمی مورد استفاده قرار گیرد. دقت بالای مدل HXGB در برآورد ET_0 امکان تصمیم‌گیری بهتر در زمینه تخصیص آب، طراحی سیستم‌های آبیاری و ارزیابی اثرات تغییر اقلیم بر نیاز آبی گیاهان را فراهم می‌سازد.

تشکر و قدردانی

نویسندگان این مقاله بر خود لازم می‌دانند از تمامی افرادی که در مراحل مختلف این پژوهش، از جمله گردآوری داده‌ها، تحلیل نتایج، و ارائه بازخوردهای علمی، همکاری و همراهی داشته‌اند، صمیمانه قدردانی نمایند. همچنین از کارشناسان محترم هواشناسی و منابع آب که با ارائه اطلاعات تخصصی، نقش مؤثری در ارتقای کیفیت این مطالعه ایفا نمودند، سپاسگزاریم.

تعارض منافع

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تعارض منافعی در ارتباط با انجام این پژوهش، نگارش مقاله، یا انتشار آن وجود ندارد.

منابع

جوزی، م.، ابراهیمی پاک، ن.ع. و تافته، آ. ۱۴۰۲. برآورد تبخیرتعرق گیاه مرجع با روش‌های هارگریوز سامانی و هارگریوز سامانی اصلاح‌شده و ارزیابی آن‌ها با استفاده از سامانه نیازآبی گیاهان در استان کرمانشاه. نشریه هیدروژئولوژی. ۱ (۸):۱–۱۶.

خدامرادپور، م. و قوی، ل. ۱۴۰۰. ارزیابی تبخیرتعرق سنجنده MODIS با تبخیرتعرق فائو–پنمن–مانتیث و پریستلی–تیلور در اقلیم‌های متفاوت ایران. نشریه پژوهش‌های حفاظت آب و خاک. ۱ (۲۸):۲۰۱–۲۱۸.

صداقت، آزاده.، ابراهیمی پاک، ن.ع.، تافته، آ. و حسینی، س.ن. ۱۴۰۱. ارزیابی تبخیرتعرق مرجع با استفاده از روش‌های داده‌کاوی و مقایسه آن با نتایج سامانه نیازآب در استان قزوین. نشریه آب و خاک. ۶ (۳۶):۶۱۱–۶۲۷.

صادقی، ع. و احمدی، ح. ۱۴۰۱. ارزیابی تبخیرتعرق مرجع ماهانه در ایران براساس برونداد مدل‌های دینامیکی ریزمقیاس‌شده پروژه CORDEX-MNA. نشریه پژوهش‌های جغرافیایی طبیعی. ۲ (۵۴):۱۸۵–۲۰۲.

یحیوی دیزج، آ.، اکبری ازیرانی، ط.، خالدی، ش. و جوان، خ. ۱۴۰۲. بررسی فصلی تبخیرتعرق مرجع و تحلیل حساسیت آن به عناصر هواشناسی در ایران. نشریه آب و خاک. ۴ (۳۷):۶۴۳–۶۵۷.

Afzaal, H., Farooque, A.A., Abbas, F., Acharya, B. and Esau, T. 2020. Computation of Evapotranspiration with Artificial Intelligence for Precision Water Resource Management. Applied Sciences. 10(5):1621.

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. 301 p.

Dadashi Roudbari, A.A., Ahmad, M. and Shakiba, A. 2020. Seasonal Study of Dust Deposition and Fine Particles (PM 2.5) in Iran Using MERRA-2 Data. Iranian Journal of Geophysics. 4 (13): 43-59.

Gelcer, E.M., Fraisse, C.W. and Sentelhas, P.C. 2010. Evaluation of Methodologies to Estimate Reference Evapotranspiration in Florida. Environmental Science. 123:189–195.

Gonenc, A., Ozerdem, M.S. and Acar, M. 2019. Comparison of NDVI and RVI Vegetation Indices Using Satellite Images. 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics). Istanbul. Turkey.

Guo, D., Olesen, J.E., Manevski, K., Pullens, W.M., Li, A. and Liu, E. 2023. Change Trend and Attribution Analysis of Reference Evapotranspiration under Climate Change in the Northern China. Agronomy .13 (12):3036.

Hargreaves, G.H. and Samani, Z. 1985. Reference Crop Evapotranspiration from Temperature. Applied Engineering in Agriculture. 1 (2):96–99.

Hargreaves, G.H. and Allen, R.G. 2003. History and Evaluation of Hargreaves Evapotranspiration Equation. Journal of Irrigation and Drainage Engineering. 129 (1):53–63.

Huang, B., Wang, J., Song, H., Fu, D. and Wang K. 2013. Generating High Spatiotemporal Resolution Land Surface Temperature for Urban Heat Island Monitoring. IEEE Geoscience and Remote Sensing Letters. 5 (10): 1011–1015.

Hussain, S., Mubeen, M., Nasim, W., Fahad, Sh., Ali, M., Ehsan, M.A. and Reza, A. 2023. Investigation of

Comparative Evaluation of Different Potential Evapotranspiration Estimation Approaches. International Journal of Research in Engineering and Technology. 3 (6):544–552.

Su, W., Shao, H., Xian, W., Xie, Z., Zhang, C. and Yang, H. 2023. Quantification of Spatiotemporal Variability of Evapotranspiration (ET) and the Contribution of Influencing Factors for Different Land Cover Types in the Yunnan Province. Water. 15(18): 3309. MDPI.

Trajkovic, S. 2007. Hargreaves versus Penman-Monteith under Humid Conditions. Journal of Irrigation and Drainage Engineering. 133 (1):38–42.

Wiens, M., VeroneBoyle, A., Henscheid, N., Podichetty, J. and Burton, J. 2025. A Tutorial and Use Case Example of the eXtreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications. Clinical and Translational Science. 3 (18):70172.

Yoder, R.E., Odhiambo, L.O. and Wright, W.C. 2005. Evaluation of Methods for Estimating Daily Reference Crop Evapotranspiration at a Site in the Humid Southeast United States. Applied Engineering in Agriculture. 21 (2):197–202.

Zhou, T., Wu, P., Sun, Sh., Li, X., Wang, T. and Luan, X. 2017. Impact of Future Climate Change on Regional Crop Water Requirement—A Case Study of Hetao Irrigation District, China. Water. 9 (6):429.

Irrigation Water Requirement and Evapotranspiration for Water Resource Management in Southern Punjab, Pakistan. Sustainability. 15 (3):1768.

Liu, Z., Jing, D., Han, Y., Yu, J., Lu, T. and Zhangzhong, L. 2022. Spatiotemporal Distribution Characteristics and Influencing Factors Analysis of Reference Evapotranspiration in Beijing–Tianjin–Hebei Region from 1990 to 2019 under Climate Change. Sustainability. 14 (10):6277.

Luo, Y., Gao, P. and Mu, X. 2021. Influence of Meteorological Factors on the Potential Evapotranspiration in Yanhe River Basin, China. Water. 13 (9):1222.

Monavvar Sabegh, S., Zare Haghi, D., Samadianfard, S. and Neishabouri, M.R. 2023. Estimation of Daily Reference Evapotranspiration Using Random Forest Optimized by Genetic Algorithm. Water and Soil Science. 33 :33–53.

Najafi, M.S. and Alizadeh, O. 2023. Climate zones in Iran. Meteological Applications. 5 (30):2147.

Nie, T., Yuan, R., Liao, S., Zhang, Zh., Gong, Zh., Zhao, X., Chen, P., Li, T., Lin, Y., Du, Ch., Dai, Ch. and Jiang, H. 2022. Characteristics of Potential Evapotranspiration Changes and Its Climatic Causes in Heilongjiang Province from 1960 to 2019. Agriculture. 12 (12):2017.

Nikam, B., Kumar, P., Garg, V. and Thakur, P. 2014.

Iranian Journal of Irrigation and Drainage
No.4, Vol. 20, Oct.-Nov. 2026, p. 547-565

نشریه آبیاری و زهکشی ایران
شماره۴، جلد ۲۰، مهر - آبان ۱۴۰۵، ص. ۵۶۵-۵۴۷

Improving Reference Evapotranspiration Prediction Using Remote Sensing Data and Machine Learning Algorithms in Northern Iran

Z. Rousta[1], S. Samadianfard[2*], R. Delirhassannia[3], S. Karimzadeh [4]

Recived: Nov.13, 2025 Accepted: Jun.08, 2026

Abstract

Faced with growing challenges from climate change and rising water demand, accurate prediction of reference evapotranspiration (ET_0) has become a key component of sustainable water resource management. Traditional models like FAO-Penman-Monteith, although scientifically reliable, often lose accuracy in areas with limited meteorological data. This study aims to improve ET_0 prediction by combining MODIS remote sensing data with advanced machine learning algorithms. Two models were developed: the basic Extreme Gradient Boosting (XGB) and an enhanced version called Hybrid XGB (HXGB), which offers better generalization. Meteorological and satellite data from Ramsar and Bandar Anzali stations (2001–2023) were used for training and evaluation. Results showed that HXGB outperformed the base model at both stations. In Bandar Anzali, Scenario 8 (using both meteorological data and satellite-based ET_{MODIS}) reduced RMSE to 0.11 mm/day. In Ramsar, the same scenario achieved an RMSE of 0.19 mm/day. This data fusion approach increased the models sensitivity to spatial and temporal variations, significantly improving prediction accuracy.

Keywords: Bandar Anzali, MODIS sensor, Ramsar, Satellite imagery, Scenario development

1- PhD student, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
2- Associate Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
3- Associate Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
4- Associate Professor, Department of Remote Sensing and Geographic Information Systems, Faculty of Planning and Environmental Sciences, University of Tabriz, Tabriz, Iran
(*- Corresponding Author E-mail: s.samadian@tabrizu.ac.ir)