دانشگاه علوم کشاورزی و منابع طبیعی ساری
پژوهشنامه مدیریت حوزه آبخیز

# برآورد تبخیر- تعرق مرجع روزانه با استفاده از روش‌های داده‌کاوی رگرسیون بردار پشتیبان و مدل درختی M5

**سعید صمدیان فرد[1] و سولماز پناهی[2]**

۱- استادیار، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، (نویسنده مسوول: s.samadian@tabrizu.ac.ir)
۲- دانشجوی کارشناسی ارشد آبیاری و زهکشی، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز
تاریخ دریافت: ۹۶/۸/۱۶ تاریخ پذیرش: ۹۷/۴/۱۱

**چکیده**

**تبخیر و تعرق یکی از اجزای اصلی چرخه هیدرولوژی است و تعیین صحیح آن برای بسیاری از مطالعات از قبیل توازن هیدرولوژیکی آب، طراحی و مدیریت سیستم‌های آبیاری، شبیه‌سازی میزان محصول و طراحی و مدیریت منابع آب از درجه اهمیت بالایی برخوردار است. خاصیت غیرخطی، عدم قطعیت ذاتی و نیاز به اطلاعات متنوع اقلیمی در برآورد تبخیر و تعرق از دلایلی بوده‌اند که باعث شده پژوهشگران به سوی روش‌های داده کاوی همچون درخت تصمیم (M5) و رگرسیون بردار پشتیبان (SVR) روی آورند. در تحقیق حاضر دقت روش های مذکور در تخمین تبخیر و تعرق گیاه مرجع در مقایسه با روش‌های تجربی هارگریوز و تورنت وایت مورد مطالعه قرار گرفته است. بدین منظور، با استفاده از سری داده‌های هواشناسی سال‌های ۱۳۹٤- ۱۳۷۱ ایستگاه سینوپتیک تبریز، ابتدا مقادیر تبخیر و تعرق روزانه گیاه مرجع از روش استاندارد فائو- پنمن- مونتیث محاسبه گردید. سپس با استفاده از این مقادیر به عنوان خروجی‌های هدف، ۱۷ سناریو مختلف از داده‌های هواشناسی که ترکیبی از حداقل داده شامل یک پارامتر تا حداکثر شش پارامتر هواشناسی بودند، توسط روش‌های مذکور ارزیابی گردید. در نهایت قابلیت روش‌های SVR و M5 در برآورد تبخیر و تعرق گیاه مرجع با استفاده از داده‌های صحت‌سنجی مورد بررسی قرار گرفت. نتایج بررسی‌های آماری و دیاگرام تیلور نشان دادند که روش‌های SVR و M5 در حالت در نظر گرفتن تمامی پارامترهای هواشناسی به ترتیب با دارا بودن جذر میانگین مربعات خطای ۰/۳۹۸ و ۰/٤۱۱ در مقایسه با روش‌های تجربی هارگریوز و تورنت وایت، نتایج بهتری در برآورد تبخیر و تعرق گیاه داشته‌اند.**

**واژه‌های کلیدی: تبخیر و تعرق گیاه مرجع، روش‌های تجربی، روش‌های یادگیری ماشینی، فائو- پنمن- مونتیث**

## مقدمه

به دلیل اینکه برآورد میزان آبی که در طرح‌های آبیاری به مصرف کشاورزی می‌رسد، بر اساس تبخیر و تعرق انجام می‌پذیرد، تخمین و تعیین این پارامتر در مطالعات آبیاری و زهکشی بسیار ضروری می‌باشد. تعیین آب مصرفی گیاهان باغی و زراعی در مناطق خشک و نیمه‌خشک، برنامه‌ریزی آبیاری، تقسیم آب، حقآبه‌ها و تهیه و تنظیم اسناد ملی آبیاری همگی بر اساس تبخیر و تعرق می‌باشد. همچنین در طراحی انواع سازه‌های انتقال آب، کنترل سیلاب در طرح‌های آبیاری و زهکشی و حتی مصارف آب شهری، نیاز مبرمی به آگاهی از مقدار صحیح تبخیر و تعرق احساس می‌گردد (۴). یکی از متداول‌ترین روش‌های تخمین تبخیر و تعرق برای گیاهان زراعی، محاسبه تبخیر و تعرق گیاه مرجع نظیر تبخیر و تعرق چمن و یا تبخیر و تعرق یونجه از یک سطح استاندارد و سپس استفاده از ضرایب گیاهی می‌باشد.

مقدار تبخیر و تعرق گیاه مرجع را می‌توان به‌صورت مستقیم، به وسیله لایسیمتر و غیرمستقیم (روش‌های آیرودینامیک، توازن انرژی و ترکیبی) و یا از طریق روش‌های تجربی به دست آورد. به علت مشکلات احداث، نیاز به دقت بالا، انجام آزمایش‌های متعدد و هزینه نسبتاً زیاد، استفاده از روش‌های غیرمستقیم بر اساس اطلاعات هواشناسی ضرورت پیدا می‌کند. روش فائو- پنمن- مونتیث یکی از روش‌های ترکیبی برای تخمین تبخیر و تعرق گیاه مرجع می‌باشد که در آن گیاه مرجع، گیاه فرضی (معمولاً چمن) با ارتفاع ۱۲ سانتی‌متر، ضریب بازتاب ۲۳ درصد و مقاومت روزنه ۷۰ ثانیه بر متر می‌باشد. این روش از آن جهت حائز اهمیت است که بیش از آن که بر اساس تجربه استوار باشد، مبنای فیزیکی داشته و اکثر عواملی که در تبخیر و تعرق موثرند در آن دخالت داده شده است. کارشناسان سازمان فائو به منظور استاندارد کردن روش محاسبه تبخیر و تعرق، روش فائو- پنمن- مونتیث را به عنوان مبنا انتخاب کرده و بر اساس آن تبخیر و تعرق گیاه مرجع را برای گیاه فرضی محاسبه نمودند (۳). با توجه به تحقیقات صورت گرفته، روش فائو- پنمن - مونتیث یکی از دقیق‌ترین روش‌های تخمین تبخیر و تعرق می‌باشد و می‌توان از آن برای اقلیم‌های مختلف استفاده نمود (۱۰). در زمینه محاسبه تبخیر و تعرق مطالعات زیادی در سطح جهان و ایران انجام گرفته است که در زیر به چند نمونه از این مطالعات اشاره می‌گردد:

جنسن و همکاران (۱۴) برای ۱۱ شرایط مختلف اقلیمی، ۱۹ روش مختلف را برای محاسبه تبخیر و تعرق پتانسیل گیاه مرجع ارزیابی نموده و معادله فائو- پنمن- مونتیث را دقیق‌ترین روش معرفی نمودند. آلن وهمکاران (۵) روش هارگریوز را درشرایطی که داده‌های هواشناسی در محدودیت باشند، برای تعیین تبخیر و تعرق پتانسیل توصیه نمودند. این روش فقط نیاز به اندازه‌گیری دمای حداقل و حداکثر هوا دارد و تابش برون‌زمینی (تابش بالای جو) مورد استفاده در آن به صورت تابعی از عرض جغرافیایی و شماره روز سال محاسبه می‌شود. تراجکوویچ (۳۰) به علت اینکه روش فائو- پنمن- مونتیث نیازمند برخی پارامترهای هواشناسی است که در تمامی مناطق و ایستگاه‌ها در دسترس نمی‌باشد، استفاده از

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]
[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]
[ DOI: 10.29252/jwmr.9.18.157 ]

روش تجربی هارگریوز به جای روش مذکور را در مناطق مرطوب امکان‌سنجی نمود و به این نتیجه رسید که استفاده از روش هارگریوز با ضرایب تعدیل شده مقادیر تبخیر و تعرق پتانسیل را فقط ۱٪ بیشتر از روش فائو- پنمن- مونتیث برآورد می‌کند و این روش به خاطر اینکه فقط از داده دما استفاده می‌کند، قابل استفاده بوده و توصیه می‌گردد. آرون‌راج و جیراج (۶) کاربرد رگرسیون بردار پشتیبان (SVR)[1] را در مدل‌سازی تبخیر و تعرق مرجع با پارامترهای مختلف هواشناسی از قبیل دما، تابش خورشیدی، رطوبت نسبی و سرعت باد بررسی نمودند. آن‌ها با بررسی عملکرد مدل SVR، مشاهده نمودند که مدل مذکور با ارائه تخمین‌های مناسب از مقادیر تبخیر و تعرق، روش سودمندی برای برآورد تبخیر تعرق می‌باشد. میرهاشمی و پناهی (۱۸) توانایی مدل درختی M5[2] را جهت برآورد متوسط درجه حرارت و تبخیر و تعرق ماهانه ماه‌های بعد ارزیابی نمودند. آن‌ها از داده‌های متوسط ماهانه ایستگاه هواشناسی یزد از جمله متوسط درجه حرارت، ساعت آفتابی، دمای نقطه شبنم، رطوبت نسبی، متوسط سرعت باد و کمبود فشار بخار اشباع استفاده نموده و بیان داشتند که مدل درختی M5 عملکرد بهتری در برآورد مقادیر ماه‌های بعد دارد. ماهش پال و دسوال (۲۱) پتانسیل مدل درختی M5 را براساس روش رگرسیون جهت مدل سازی تبخیر و تعرق مرجع روزانه با استفاده از داده‌های اقلیمی بررسی کردند. آن‌ها برای مقایسه عملکرد مدل درختی M5 در پیش بینی تبخیر و تعرق مرجع از معادله فائو- پنمن- مونتیث و رابطه هارگریوز - سامانی استفاده نموده و نشان دادند که مدل درختی M5 در مقایسه با هر دو روش مورد مطالعه بهتر عمل کرده و می‌تواند در مدل‌سازی تبخیر و تعرق مرجع به کار برده شود. مارینیکومای و وینودهینی (۱۷) نیز تبخیر و تعرق مرجع را بوسیله مدل رگرسیون SVR و رگرسیون خطی چندگانه مدل سازی کرده و نتیجه گرفتند که مدل SVR برآورد مناسب‌تری از تبخیر تعرق مرجع ارائه می‌دهد. تراجکوویچ و کولاکوویچ (۳۱) استفاده از روش فائو- پنمن- مونتیث با داده‌های محدود را در مناطق مرطوب صربستان در شرق اروپا مورد بررسی قرار داده و به این نتیجه رسیدند که استفاده از روش فائو- پنمن- مونتیث با داده‌های محدود در مقایسه با روش‌های تجربی تورک و هارگریوز تطابق بهتری با روش فائو- پنمن- مونتیث با داده‌های کامل دارد. آن‌ها اظهار نمودند که استفاده از داده‌های هواشناسی دمای حداقل، دمای حداکثر و سرعت باد برای برآورد دقیق تبخیر و تعرق پتانسیل به عنوان حداقل داده ضروری می‌باشند. سلطانی و همکاران (۲۸) دقت معادله فائو- پنمن- مونتیث برای برآورد تبخیر تعرق مرجع در ۹ ایستگاه هواشناسی مختلف در اقلیم‌های مختلف ایران برای داده‌های هواشناسی محدود را بررسی نموده و نشان دادند که مدل‌های کاهش داده فائو- پنمن- مونتیث اغلب نتایج بهتری نسبت به مدل‌های تجربی ارائه دادند. آنها همچنین نشان دادند که حداقل داده لازم برای برآورد دقیق تبخیر تعرق مرجع به روش فائو- پنمن- مونتیث داده‌های حداقل دما، حداکثر دما و سرعت باد می‌باشند.

ستاری و همکاران (۲۶) در منطقه بناب استان آذربایجان شرقی پژوهشی جهت برآورد دقیق تبخیر و تعرق مرجع روزانه انجام دادند و با ترکیب‌هایی مختلف از پارامترهای هواشناسی ایستگاه بناب در قالب سناریوهای مختلف به عنوان ورودی مدل درختی M5 و شبکه‌های عصبی مصنوعی پرسپترون چندلایه انتشار خطا به عقب تبخیر تعرق را محاسبه نمودند. آن‌ها به این نتیجه رسیدند که مدل درختی M5 نتایج قابل قبول و روابط ساده خطی قابل فهم و کاربردی‌تری ارائه می‌دهد. ستاری و همکاران (۲۵) تبخیر از سطح آزاد آب سد علویان در مراغه را مورد بررسی قرار داده و با اعمال پارامترهای مختلف اقلیمی در چندین سناریو مختلف به عنوان ورودی مدل M5، مدل مذکور را روش بهتری جهت برآورد تبخبر و تعرق معرفی کردند.

پژوهشگران بسیاری در سراسر دنیا روش فائو- پنمن- مونتیث را در مقایسه با لایسیمتر به عنوان روشی دقیق، معتبر و استاندارد برای محاسبه تبخیر و تعرق گیاه مرجع معرفی کرده‌اند که در این پژوهش نیز به عنوان روش مبنا مورد استفاده قرار گرفته است (۱۹،۱۶). تبخیر و تعرق فرآیندی پیچیده و غیرخطی است و به عوامل متعدد اقلیمی نظیر دما، رطوبت، سرعت باد، تابش، نوع و مرحله رشد گیاه و غیره وابسته است (۲۷). بنابراین، در تحقیق حاضر با در نظر گرفتن ۱۷ سناریو مختلف از داده‌های هواشناسی روزانه مربوط به منطقه تبریز، مقادیر تبخیر و تعرق متوسط روزانه بر اساس نتایج حاصل از روش فائو- پنمن- مونتیث مدل‌سازی شده و دقت روش‌های درخت تصمیم و رگرسیون بردار پشتیبان با روش‌های تجربی هارگریوز و تورنت وایت مقایسه شده است.

## مواد و روش‌ها

### منطقه مورد مطالعه

در پژوهش حاضر، داده‌های مربوط به ایستگاه سینوپتیک تبریز در دوره آماری ۱۳۷۱ تا ۱۳۹۴ برای محاسبه تبخیر و تعرق متوسط ماهیانه گیاه مرجع مورد استفاده قرار گرفته است. ایستگاه سینوپتیک مذکور (شکل ۱) در عرض جغرافیایی ′۰۵ °۳۸ و طول جغرافیایی ′۱۷ °۴۶ و در ارتفاع ۱۳۶۱ متری از سطح دریا واقع شده است. میانگین حداکثر و حداقل دما در ایستگاه به ترتیب ۱۸ و ۷ درجه سانتی‌گراد و میانگین سرعت باد در منطقه ۶ متر بر ثانیه می‌باشد. به دلیل عدم دسترسی به داده‌های لایسیمتری، از روش فائو- پنمن- مونتیث برای تخمین تبخیر و تعرق گیاه مرجع استفاده شده است. روش فائو- پنمن- مونتیث در میان روش‌های مختلف تبخیر و تعرق گیاه مرجع همواره بیشترین دقت را دارا بوده است. از این رو مقادیر محاسبه شده از روش فائو- پنمن- مونتیث به عنوان مقادیر استاندارد و هدف برای واسنجی و صحت‌سنجی سناریوهای متفاوت روش‌های درخت تصمیم و رگرسیون بردار پشتیبان و روش‌های تجربی هارگریوز و تورنت وایت مورد استفاده قرار گرفت.

1- Support Vector Regression

2- M5 Model Tree

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]

[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]

[ DOI: 10.29252/jwmr.9.18.157 ]

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]

[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]

[ DOI: 10.29252/jwmr.9.18.157 ]

شکل ۱- موقعیت جغرافیایی ایستگاه سینوپتیک تبریز
Figure 1. Geographical location of Tabriz synoptic station

## روش‌های محاسبه تبخیر و تعرق مرجع

### روش فائو- پنمن- مونتیث

آلن و همکاران (۵) رابطه ۱ را که به نام رابطه فائو- پنمن- مونتیث معروف است، به عنوان دقیق‌ترین روش برای برآورد نیاز آبی گیاه ارئه نموده اند:

$$ETo = \frac{0.408\Delta(R_n - G) + ¥\frac{900}{T+273}u_2(e_s - e_a)}{\Delta + ¥(1 + 0.34u_2)} \quad (۱)$$

که در آن $ETo$ تبخیر و تعرق گیاه مرجع $(mm/day)$، $R_n$ تابش خالص در سطح پوشش گیاهی $(MJ/m.day)$، $G$ شار گرما به داخل خاک $(MJ/m.day)$، $T$ متوسط دمای هوا در ارتفاع دو متری از سطح زمین $(°C)$، $u_2$ سرعت باد در ارتفاع دو متری از سطح زمین $(m/s)$، $e_a$ فشار بخار واقعی $(KPa)$، $e_s$ فشار بخار واقعی $(KPa)$، $e_s - e_a$ کمبود فشار بخار اشباع $(KPa)$، $\Delta$ شیب منحنی فشار بخار $(KPa/°C)$ و ¥ ضریب رطوبتی $(KPa/°C)$ می‌باشد.

### روش تورنت وایت

روش تورنت وایت بر اساس یک رابطه تجربی بین مقدار تبخیر و تعرق مرجع و میانگین درجه حرارت هوا ارائه شده است. معادلات این روش به صورت زیر است.

$$ETo = 16N_m\left(\frac{10\times T_i}{I}\right)^{6.75\times10^{-3}I^3 - 7.71\times10^{-5}I^2 + 0.0179I + 0.492} \quad (۲)$$

$$I = \sum_{i=1}^{12}\left(\frac{T_i}{5}\right)^{1.51} \quad (۳)$$

که در آن، $T_i$ متوسط دمای هوا، $N_m$ ضریب اصلاحی برای عرض‌های جغرافیایی و دوره‌های زمانی مختلف، $I$ شاخص حرارتی سالانه می‌باشد.

### روش هارگریوز

در روش هارگریوز نیز از معادله مشابه با روش جنسن- هیز استفاده می‌شود. این معادله به صورت زیر ارائه شده است.

$$ETo = 0.0135(K_T).R_a.TD^{0.5}(T + 17.8) \quad (۴)$$

$$K_T = 0.00185(TD)^2 - 0.0433TD + 0.4023 \quad (۵)$$

در روابط ۴ و ۵، $ETo$ تبخیر و تعرق گیاه مرجع $(mm/day)$، $T$ متوسط دمای هوا $(°C)$، $TD$ اختلاف دمای حداکثر و حداقل $(°C)$ و $R_a$ تابش برون زمینی $(mm/day)$ می‌باشد (۱۱).

### رگرسیون بردار پشتیبان (SVR)

ماشین بردار پشتیبان یکی از روش‌های یادگیری است که در سال ۱۹۹۲ میلادی توسط بوسر و همکاران (۹) بر پایه تئوری یادگیری آماری معرفی شد. آن‌ها طی سال‌های بعد تئوری ابرصفحه بهینه را به عنوان طبقه‌بندی‌کننده خطی ارائه داده و طبقه‌بندی‌کننده‌های غیرخطی را با کمک توابع کرنل معرفی نمودند. اصول اساسی آن چه که امروز به عنوان ماشین بردار پشتیبان شناخته می‌شود، نتایج کارهای بوسر و همکاران (۹) است و در نهایت گسترش ماشین بردار پشتیبان بر اساس رگرسیون نیز در سال ۱۹۹۵ توسط وپنیک به نتیجه رسید (۳۲). ماشین بردار پشتیبان مبتنی بر کمینه کردن ساختاری ریسک می‌باشد که از نظریه آموزش آماری گرفته شده است (۳۳). مدل‌های ماشین‌های بردار پشتیبان به دو گروه عمده الف) مدل طبقه‌بندی ماشین بردار پشتیبان و

ب) مدل رگرسیون بردار پشتیبان تقسیم‌بندی می‌شوند. از مدل طبقه‌بندی ماشین بردار پشتیبان جهت حل مسائل طبقه‌بندی داده‌هایی که در کلاس‌های مختلف قرار می‌گیرند، استفاده می‌گردد و مدل رگرسیون بردار پشتیبان در حل مسائل پیش‌بینی کاربرد دارد (۱۲).

یکی از روش‌های متداول برای حل مسائل غیرخطی، استفاده از توابع کرنل می‌باشد که بر اساس ضرب داخلی داده‌ها تعریف می‌شود. در واقع، با یک تبدیل غیرخطی از فضای ورودی به فضایی با ابعاد بیشتر (حتی نامتناهی) می‌توان مسائل را به صورت خطی، تفکیک‌پذیر ساخت. در پژوهش حاضر از سه تابع کرنل، که عموماً در کاربردهای مهندسی متداول هستند (جدول ۱)، استفاده شده است (۲۰).

جدول ۱- توابع کرنل مورد استفاده در پژوهش حاضر

Table 1. Used kernel functions in the current research

| توابع کرنل | رابطه ریاضی |
|---|---|
| تابع کرنل چند جمله‌ای | $k(x_i,x_j) = (x_i^T x_j + 1)^p$ |
| تابع کرنل با پایه شعاعی | $k(x_i,x_j) = e^{-\gamma\|x_i - x_j\|^2}$ |
| تابع کرنل پیرسون | $k(x_i,x_j) = \dfrac{1}{\left[1 + \left(2\sqrt{\|x_i - x_j\|^2}.\sqrt{2^{1/\omega}}.\dfrac{1}{\sigma}\right)^2\right]^{\omega}}$ |

**مدل درختی M5**

مدل درختی M5 (۲۴) زیر مجموعه‌ای از روش‌های یادگیری ماشینی و داده‌کاوی است. داده‌کاوی به فرآیند جست و جو و کشف مدل‌های گوناگون، مختصرسازی‌ها و اخذ مقادیر از مجموعه‌ای از مقادیر معلوم اطلاق می‌گردد (۲). روش‌های داده‌کاوی برای مجموعه داده‌های بزرگ با متغیرهای زیاد ساخته شده‌اند، بنابراین متفاوت از روش‌های آماری قدیمی هستند که برای مجموعه داده‌های کوچک با متغیرهای اندک طراحی شده‌اند. روش‌های مبتنی بر درخت تصمیم به عنوان یکی از شناخته‌شده‌ترین تکنیک‌های داده کاوی، ویژگی هدف را به عنوان خروجی به صورت یک مدل با سازه درختی با استفاده از داده‌های ورودی، پیش‌بینی و یا طبقه‌بندی می‌نمایند. مدل M5 یک مدل درختی برای پیش بینی صفات عددی پیوسته است که در آن توابع رگرسیونی خطی در برگ‌های این درخت متظاهر می‌شوند (۲۰) که در سال‌های اخیر تحول قابل توجهی در مسایل طبقه‌بندی و پیش‌بینی ایجاد نموده است. نتایج مدل درختی برای فهم و شبیه‌سازی آسان هستند و خروجی مدل از دقت بالایی برخوردار است که می‌توان آن را با سایر مدل‌ها مقایسه کرد، مدل درختی می‌تواند در مسایل گوناگون مورد استفاده پژوهشگران قرار گیرد (۸ و ۱۲). یک درخت تصمیم معمولاً از چهار بخش ریشه، شاخه، گره‌ها و برگ‌ها تشکیل شده است که گره‌ها با دایره نشان داده می‌شوند و شاخه‌ها نشان‌دهنده اتصال بین گره‌ها می‌باشند. درخت تصمیم به منظور سادگی در رسم معمولاً از چپ به راست و یا از بالا به پایین کشیده می‌شود، به طوری که ریشه (گره اول) در بالا قرار می‌گیرد. انتهای یک زنجیره را برگ می‌نامند. هر گره مربوط به یک خصوصیت معین است و شاخه‌ها به معنای بازه‌ای از مقادیر هستند، این بازه‌های مقادیر بخش‌های مختلف مجموعه مقادیر معلوم را برای خصوصیت‌ها به دست می دهند. عمل انشعاب توسط یکی از متغیرهای پیش‌بینی‌کننده انجام می‌پذیرد، بازه‌های انشعاب طوری انتخاب می‌شوند که مجموع مجذور انحراف از میانگین داده‌های هر گره را حداقل کنند (۱۲). اولین مرحله برای ایجاد یک مدل درختی، استفاده از یک معیار انشعاب است. معیار انشعاب برای الگوریتم M5 بر اساس عملکرد انحراف استاندارد مقادیر هر کلاس و یا طبقه است که در هر گره به دست آمده است. این روش اساس روش‌های طبقه‌بندی است که آنتروپی نامیده می‌شود. آنتروپی می‌تواند به عنوان معیار میزان آشفتگی و بی‌نظمی یک سیستم تفسیر شود. معیار انشعاب بیانگر میزان خطا در آن گره می‌باشد و مدل حداقل خطای مورد انتظار را به عنوان نتیجه‌ی آزمایش هر صفت در آن گره محاسبه می‌کند. خطای مدل عموماً با اندازه‌گیری دقت پیش‌بینی مقادیر هدف موارد دیده نشده سنجش می‌شود. فرمول محاسبه کاهش انحراف استاندارد (SDR) به صورت رابطه ۶ می‌باشد.

$$SDR = Sd(T) - \sum \frac{|T_i|}{|T|} Sd(T_i) \qquad (۶)$$

که در آن، T مجموعه‌ای از نمونه‌ها است که به هر گره وارد می‌شود، $T_i$ زیر مجموعه‌ای از نمونه‌هاست که iامین نتیجه آزمون بالقوه را دارند، $S_d$ انحراف معیار می‌باشند (۱).

**توابع عملکرد**

عملکرد سناریوهای مختلف M5، SVR و روش‌های تجربی هارگریوز و تورنت وایت در تخمین مقادیر تبخیر و تعرق مرجع با استفاده از پارمترهای جذر میانگین مربعات خطا (RMSE) ، میانگین خطای مطلق (MAE)، نش ساتکلیف (NS) و ویلموت (WI) سنجیده شده است. فرمول محاسبه این توابع به صورت روابط ۷ الی ۱۰ می‌باشند.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2} \qquad (۷)$$

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]
[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]
[ DOI: 10.29252/jwmr.9.18.157 ]

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|P_i - O_i\right| \quad (۸)$$

$$NS = 1 - \left[\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O}_i)^2}\right], 0 \le NS \le 1 \quad (۹)$$

$$WI = \left|1 - \left[\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(\left|P_i - \bar{O}_i\right| + \left|O_i - \bar{O}_i\right|)^2}\right]\right|, 0 \le WI \le 1 \quad (۱۰)$$

در روابط ۷ تا ۱۰، N تعداد دادها، $O_i$ مقادیر محاسبه شده تبخیر و تعرق از روش فائو- پنمن- مونتیث و $P_i$ مقادیر تخمینی از روش‌های مورد مطالعه می‌باشند.

علاوه بر این، دیاگرام تیلور (۲۹) برای تحلیل دقت روش‌های مورد استفاده در پیش‌بینی تبخیر و تعرق مرجع روزانه به کار گرفته شد. دیاگرام تیلور، راه‌حلی گرافیکی برای ارزیابی دقت داده‌های پیش‌بینی شده با به تصویر کشیدن همزمان پارامترهای آماری می‌باشد. در دیاگرام مذکور، هر نقطه بیانگر عملکرد روش متناظر بوده و هرچه نقاط متناظر روش‌ها به نقطه داده‌های مشاهداتی در مختصات قطبی نزدیک‌تر باشد، نشان دهنده دقت بالاتر و خطای کمتر آن روش می‌باشد (۱۳).

## نتایج و بحث

به منظور بررسی امکان استفاده از ترکیب‌های متفاوتی از داده‌های هواشناسی برای پیش‌بینی دقیق‌تر تبخیر و تعرق مرجع، ۱۷ سناریو مختلف، متشکل از داده‌های متفاوت هواشناسی، تعریف گردید. سپس، با استفاده از نرم‌افزار Weka، که در دانشگاه وایکاتو نیوزلند توسعه داده شده، محاسبات واسنجی بر روی دوسوم از داده‌ها و محاسبات صحت‌سنجی بر روی یک‌سوم داده‌های باقیمانده انجام پذیرفت. جدول ۲ پارامترهای دخیل در هر سناریو را نمایش می‌دهد.

جدول ۲- پارامترهای دخیل در هر سناریو در روش‌های M5 و SVR

Table 2. Implemented parameters in each scenario of M5 and SVR methods

| شماره سناریو | پارامترهای ورودی | شماره سناریو | پارامترهای ورودی |
|---|---|---|---|
| ۱ | دمای میانگین، دمای حداقل، دمای حداکثر، رطوبت نسبی، سرعت باد، نسبت ساعات آفتابی | ۱۰ | دمای حداقل، دمای حداکثر |
| ۲ | دمای حداقل، دمای حداکثر، رطوبت نسبی، سرعت باد، نسبت ساعات آفتابی | ۱۱ | دمای میانگین، سرعت باد |
| ۳ | دمای حداقل، دمای حداکثر، رطوبت نسبی، نسبت ساعات آفتابی | ۱۲ | دمای حداکثر، سرعت باد |
| ۴ | دمای حداقل، دمای حداکثر، سرعت باد، نسبت ساعات آفتابی | ۱۳ | دمای میانگین، سرعت باد |
| ۵ | دمای میانگین، رطوبت نسبی، سرعت باد، نسبت ساعات آفتابی | ۱۴ | دمای حداقل، نسبت ساعات آفتابی |
| ۶ | دمای حداقل، دمای حداکثر، رطوبت نسبی، نسبت ساعات آفتابی | ۱۵ | دمای حداکثر، نسبت ساعات آفتابی |
| ۷ | دمای حداقل، دمای حداکثر، نسبت ساعات آفتابی | ۱۶ | دمای میانگین، نسبت ساعات آفتابی |
| ۸ | دمای میانگین، سرعت باد، نسبت ساعات آفتابی | ۱۷ | سرعت باد، نسبت ساعات آفتابی |
| ۹ | رطوبت نسبی، سرعت باد، نسبت ساعات آفتابی | | |

با تعریف ترکیب‌های متفاوت از پارامترهای هواشناسی موثر بر پدیده تبخیر و تعرق، محاسبات مربوط به پیش‌بینی مقدار تبخیر و تعرق گیاه مرجع با روش‌های M5 و SVR، با استفاده از نرم‌افزار داده کاوی Weka انجام پذیرفته و دقت روش‌های مذکور و همچنین روابط تجربی هارگریوز و تورنت وایت در مرحله صحت‌سنجی، بر اساس معیارهای آماری (روابط ۷ الی ۱۰) محاسبه گردیده و در جدول‌های ۳ و ۴ ارائه شده‌اند.

جدول ۳- نتایج کلی محاسبات صورت گرفته برای سناریوهای تعریف شده برای روش‌های M5 و SVR

Table 3. Overall results of performed computations for all defined scenarios of M5 and SVR methods

| شماره | M5 | | | | SVR | | | |
|---|---|---|---|---|---|---|---|---|
| | RMSE | MAE | NS | WI | RMSE | MAE | NS | WI |
| ۱ | ۰/۴۱۱ | ۰/۳۰۹ | ۰/۹۹۹ | ۰/۹۹۲ | ۰/۳۹۸ | ۰/۲۷۹ | ۰/۹۹۹ | ۰/۹۹۳ |
| ۲ | ۰/۵۶۲ | ۰/۴۱۹ | ۰/۹۹۸ | ۰/۹۸۴ | ۰/۵۴۷ | ۰/۳۹۴ | ۰/۹۹۹ | ۰/۹۸۶ |
| ۳ | ۰/۶۳۸ | ۰/۴۷۹ | ۰/۹۹۸ | ۰/۹۷۹ | ۰/۶۶۲ | ۰/۴۷۵ | ۰/۹۹۸ | ۰/۹۷۸ |
| ۴ | ۰/۵۶۹ | ۰/۴۲۸ | ۰/۹۹۸ | ۰/۹۸۴ | ۰/۷۵۹ | ۰/۶۲۱ | ۰/۹۹۷ | ۰/۹۶۹ |
| ۵ | ۰/۵۶۳ | ۰/۴۳۰ | ۰/۹۹۸ | ۰/۹۸۴ | ۰/۵۶۴ | ۰/۴۲۱ | ۰/۹۹۸ | ۰/۹۸۴ |
| ۶ | ۰/۵۹۹ | ۰/۴۴۹ | ۰/۹۹۸ | ۰/۹۸۲ | ۰/۵۶۹ | ۰/۴۱۸ | ۰/۹۹۸ | ۰/۹۸۴ |
| ۷ | ۰/۶۴۳ | ۰/۴۸۹ | ۰/۹۹۸ | ۰/۹۷۹ | ۰/۶۶۴ | ۰/۴۸۲ | ۰/۹۹۸ | ۰/۹۷۸ |
| ۸ | ۰/۵۶۸ | ۰/۴۳۸ | ۰/۹۹۸ | ۰/۹۸۴ | ۰/۵۷۳ | ۰/۴۳۲ | ۰/۹۹۸ | ۰/۹۸۴ |
| ۹ | ۱/۰۹۱ | ۰/۸۲۸ | ۰/۹۹۳ | ۰/۹۳۳ | ۱/۰۸۵ | ۰/۷۹۲ | ۰/۹۹۳ | ۰/۹۳۹ |
| ۱۰ | ۰/۶۵۸ | ۰/۵۰۵ | ۰/۹۹۸ | ۰/۹۷۸ | ۰/۶۶۱ | ۰/۴۹۸ | ۰/۹۹۸ | ۰/۹۷۸ |
| ۱۱ | ۰/۹۹۰ | ۰/۷۵۶ | ۰/۹۹۳ | ۰/۹۴۶ | ۰/۹۹۳ | ۰/۷۵۳ | ۰/۹۹۵ | ۰/۹۴۷ |
| ۱۲ | ۰/۶۴۲ | ۰/۴۹۴ | ۰/۹۹۸ | ۰/۹۷۹ | ۰/۶۳۳ | ۰/۴۷۸ | ۰/۹۹۸ | ۰/۹۸۰ |
| ۱۳ | ۰/۶۵۱ | ۰/۵۰۰ | ۰/۹۹۸ | ۰/۹۷۹ | ۰/۶۴۵ | ۰/۴۹۰ | ۰/۹۹۸ | ۰/۹۷۹ |
| ۱۴ | ۰/۸۱۹ | ۰/۶۳۰ | ۰/۹۹۵ | ۰/۹۶۴ | ۰/۸۳۰ | ۰/۶۲۵ | ۰/۹۹۶ | ۰/۹۶۴ |
| ۱۵ | ۰/۶۴۱ | ۰/۴۹۲ | ۰/۹۹۸ | ۰/۹۷۹ | ۰/۶۷۱ | ۰/۴۹۷ | ۰/۹۹۸ | ۰/۹۷۷ |
| ۱۶ | ۰/۶۰۰ | ۰/۴۶۲ | ۰/۹۹۸ | ۰/۹۸۲ | ۰/۶۳۰ | ۰/۴۶۹ | ۰/۹۹۸ | ۰/۹۸۰ |
| ۱۷ | ۱/۵۲۲ | ۱/۲۱۸ | ۰/۹۸۴ | ۰/۸۳۷ | ۱/۵۴۳ | ۱/۱۸۸ | ۰/۹۸۵ | ۰/۸۵۱ |

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]

[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]

[ DOI: 10.29252/jwmr.9.18.157 ]

جدول ۴- نتایج کلی محاسبات صورت گرفته برای روش‌های تجربی

Table 4. Overall results of performed computations for empirical methods

| روش | RMSE | MAE | NS | WI |
|---|---|---|---|---|
| هارگریوز | ۰/۵۸۴ | ۰/۴۱۶ | ۰/۹۹۸ | ۰/۹۸۳ |
| تورنت وایت | ۱/۴۹۷ | ۱/۲۸۴ | ۰/۹۸۸ | ۰/۸۸۱ |

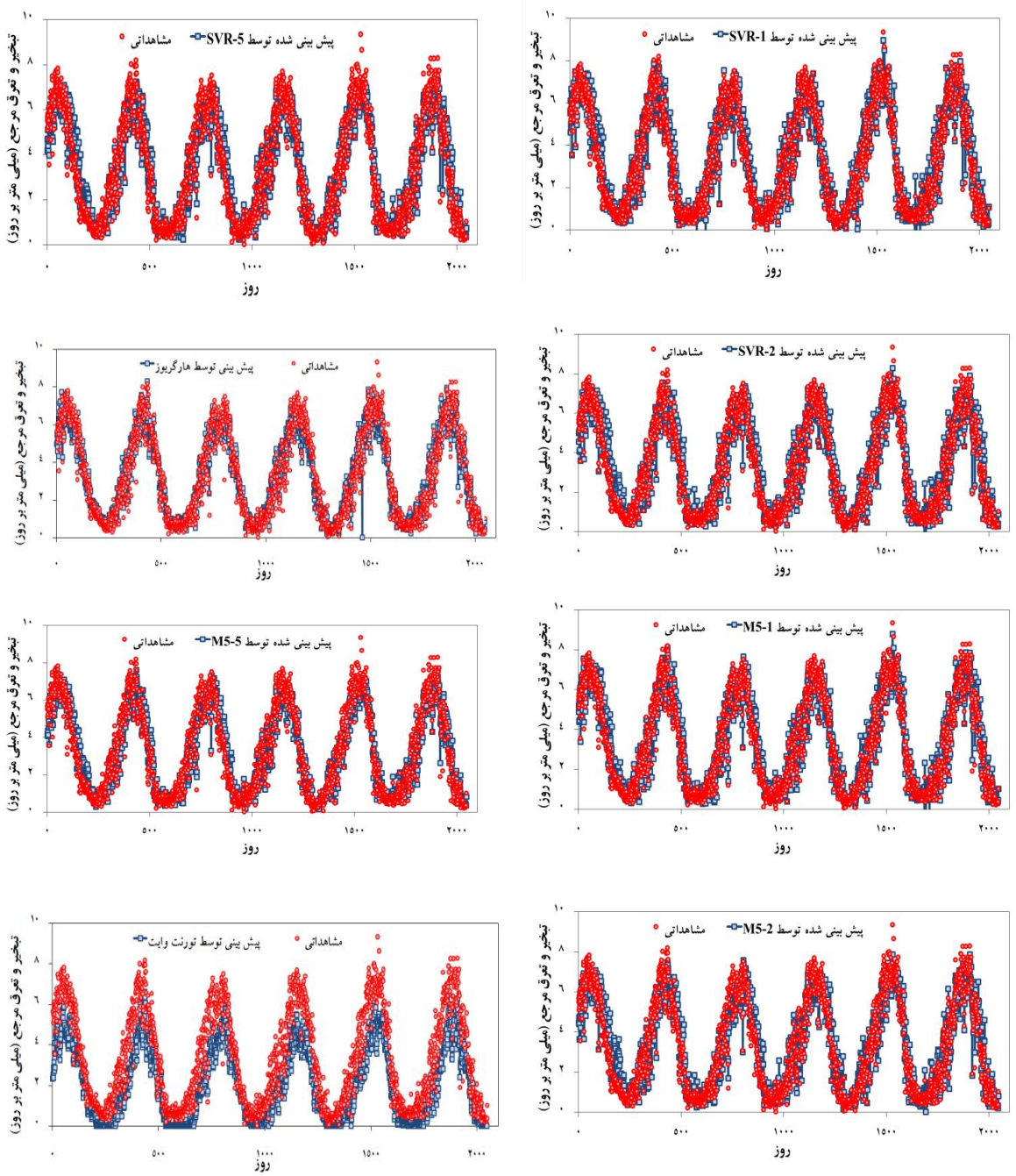
چنانچه از جدول‌های ۳ و ۴ مشخص است، مدل SVR-1 با داشتن کمترین خطا و دارا بودن پارامترهای آماری جذر میانگین مربعات خطای ۰/۳۹۸، میانگین خطای مطلق ۰/۲۷۹، ضریب نش ساتکلیف ۰/۹۹۹ و ضریب ویلموت ۰/۹۹۳ بهترین عملکرد را در مقایسه با سناریوهای مختلف تعرف شده برای مدل‌های M5 و SVR و همچنین در مقایسه با روش‌های تجربی هارگریوز و تورنت وایت داشته است. در مرتبه دوم نیز، مدل M5-1 با داشتن جذر میانگین مربعات خطای ۰/۴۱۱، میانگین خطای مطلق ۰/۳۰۹، ضریب نش ساتکلیف ۰/۹۹۹ و ضریب ویلموت ۰/۹۹۲ قرار گرفته که دقت به مراتب بیشتری نسبت به روش‌های تجربی مورد مطالعه دارد. همچنین، با بررسی‌های به عمل آمده در مقایسه عملکرد سناریوهای مختلف روش‌های یادگیری ماشینی با عملکرد روابط تجربی، مشاهده می‌گردد که ۵ سناریو از تعداد ۱۷ سناریو روش M5 (M5-1, M5-2, M5-4, M5-5, M5-8) و ۵ سناریو روش SVR ( SVR-1, SVR-2, SVR-5, SVR-6, ,SVR-8) دقت بیشتر و خطای کمتری نسبت به روش‌های تجربی هارگریوز در پیش‌بینی تبخیر و تعرق مرجع دارند. بر خلاف دقت نسبتا بالا و مناسب روش هارگریوز در پیش‌بینی تبخیر و تعرق گیاه مرجع، روش تورنت وایت دقت قابل قبولی ارایه نکرده و فقط از یک سناریو روش M5 (M5-17) و یک سناریو روش SVR (SVR-17) خطای کمتری داشته است. در نهایت سناریوهای ۱، ۲ و ۵ برای هر دو روش M5 و SVR به عنوان سه سناریو برتر در پیش‌بینی تبخیر و تعرق گیاه مرجع انتخاب شدند. از طرفی دیگر، سناریو شماره ۱۶ نیز با توجه به اینکه تنها از دو پارامتر دمای میانگین و نسبت ساعات آفتابی استفاده می‌نماید، با داشتن جذر میانگین مربعات خطای ۰/۶۰۰، میانگین خطای مطلق ۰/۴۶۲، ضریب نش ساتکلیف ۰/۹۹۸ و ضریب ویلموت ۰/۹۸۲ برای روش M5(M5-16) و جذر میانگین مربعات خطای ۰/۶۳۰، میانگین خطای مطلق ۰/۴۶۹، ضریب نش ساتکلیف ۰/۹۹۸ و ضریب ویلموت ۰/۹۸۰ برای روش SVR(SVR-16) دقت قابل قبولی در مقایسه با روش تجربی هارگریوز و دقت بیشتری نسبت به روش تجربی تورنت وایت ارائه کرده است. مقایسه‌ی نتایج بدست آمده با نتایج حاصل از پژوهش بابا میری و همکاران (۷) نشان داد که دقت سناریوهای انتخابی با استفاده از روش‌های M5 و SVR به مراتب بیشتر از مدل‌های تصادفی سری زمانی در برآورد تبخیر و تعرق در ایستگاه تبریز بود. در پژوهشی دیگر پناهی و همکاران (۲۲) تبخیر و تعرق مرجع در ایستگاه تبریز را با استفاده از شبکه‌های عصبی مصنوعی و ماشین بردار پشتیبان پیش‌بینی نمودند. مقایسه نتایج حاصله نشان داد که سناریو SVR-1، به عنوان دقیق‌ترین سناریو پژوهش حاضر، با داشتن جذر میانگین مربعات خطای ۰/۳۹۸ پیش‌بینی‌های دقیق‌تری از تبخیر و تعرق مرجع در مقایسه با بهترین الگوی ورودی به شبکه‌ی SVM-RBF با جذر میانگین مربعات خطای ۰/۵۰۹ گزارش شده در پژوهش پناهی و همکاران(۲۲) ارائه نمود.

همچنین نمودارهای تغییرات مقادیر تبخیر و تعرق مرجع پیش‌بینی‌شده برای سه سناریو برتر روش‌های یادگیری ماشینی و روش‌های تجربی هارگریوز و تورنت وایت در مقایسه با نتایج حاصل از روش فائو- پنمن- مونتیث در مرحله صحت‌سنجی در شکل ۲ و نمودارهای پراکنش روش‌های مذکور در شکل ۳ نشان داده شده‌اند.

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]

[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]

[ DOI: 10.29252/jwmr.9.18.157 ]



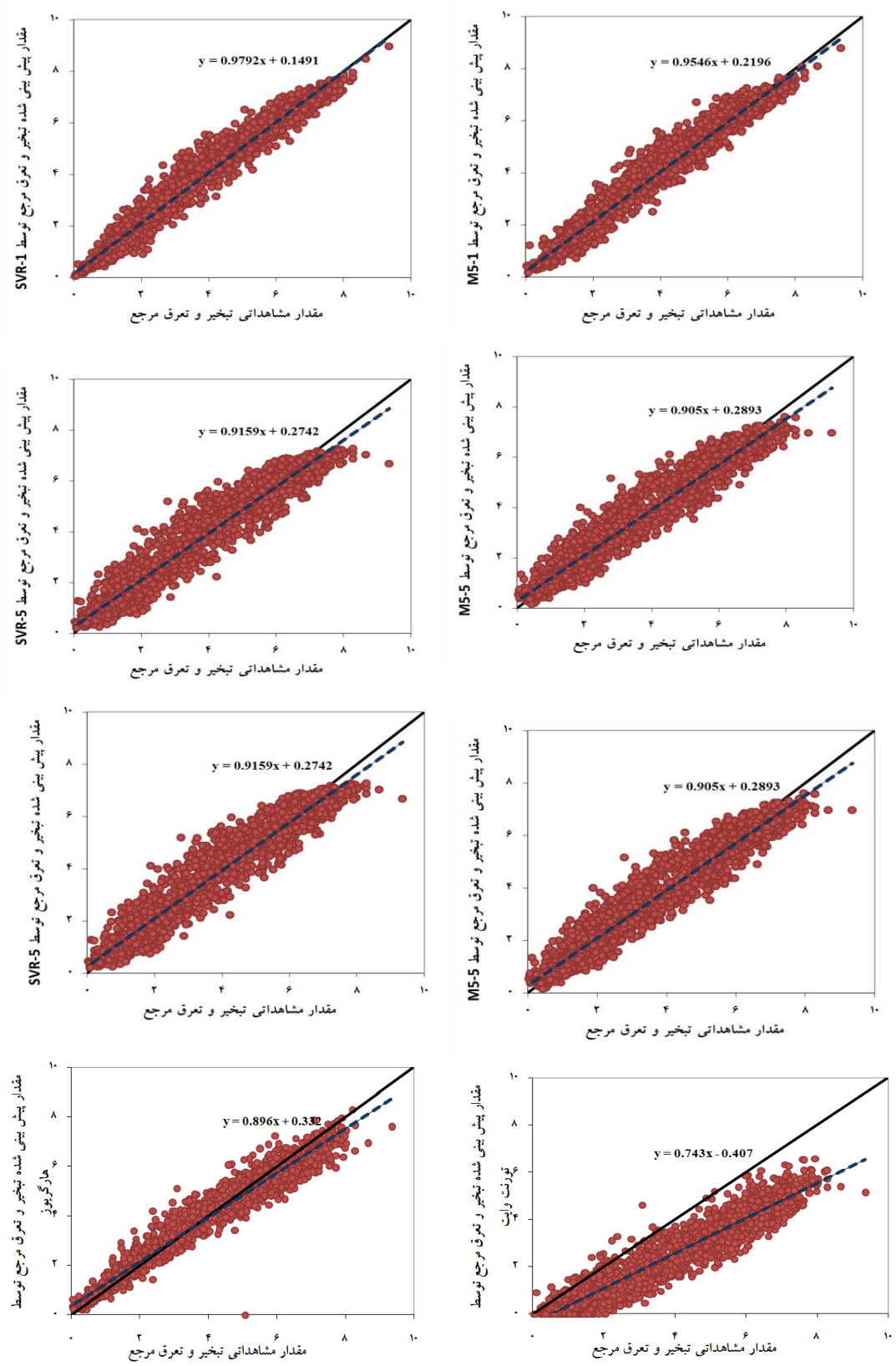

شکل ۲- نمودارهای تغییرات زمانی مقادیر تبخیر و تعرق مرجع با استفاده از روش فائو- پنمن- مونتیث و روش‌های SVR، M5 و روش‌های تجربی تورنت وایت و هارگریوز

Figure 2. Time series plots of reference evapotranspiration using FAO-Penman Monteith, SVR, M5 and empirical methods of Hargreaves and thornthwaite

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]

[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]

[ DOI: 10.29252/jwmr.9.18.157 ]

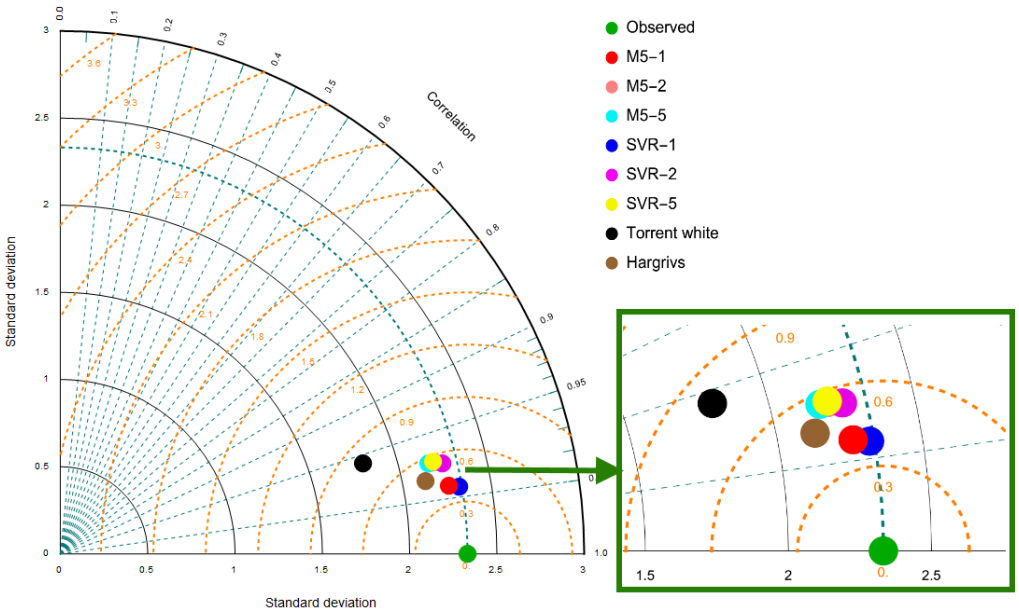



شکل ۳- نمودارهای پراکنش تبخیر و تعرق مرجع با استفاده از فائو- پنمن- مونتیث و روش‌های M5 و SVR و روش‌های تجربی

Figure 3. Scatterplots of reference evapotranspiration using FAO-Penman Monteith, SVR, M5 and empirical methods

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]
[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]
[ DOI: 10.29252/jwmr.9.18.157 ]

روند اشاره شده قبلی در مورد بالابودن دقت سناریوهای ۱، ۲ و ۵ برای هر دو روش M5 و SVR از شکل‌های ۲ و ۳ نیز قابل نتیجه‌گیری است. از شکل ۲ می توان چنین استنباط نمود که پیش‌بینی‌های سناریو شماره ۱ برای هر دو روش M5 و SVR تطابق بیشتری با مقادیر حاصل از روش فائو- پنمن- مونتیث دارد. همچنین از شکل ۳ نیز به وضوح مشاهده می‌گردد که پراکنش نقاط رسم شده حول خط یک به یک (خط نیم ساز) برای هر دو روش M5-1 و SVR-1 کمتر بوده و مدل‌های مذکور دقت به مراتب بالاتری را از خود نشان داده‌اند. علاوه بر این، دیاگرام تیلور به منظور بررسی و تحلیل مقادیر انحراف معیار و ضرایب همبستگی بین داده‌های مشاهداتی و پیش‌بینی شده توسط سناریوهای برتر SVR و M5 و روش‌های تجربی رسم گردید (شکل ۴). لازم به ذکر است که در دیاگرام تیلور، فاصله شعاعی از نقطه مشاهداتی (نقطه سبزرنگ)، نشان‌دهنده مقدار جذر میانگین مربعات خطای روش‌های مورد مطالعه می‌باشد.




شکل ۴- دیاگرام تیلور روش‌های مورد مطالعه در پیش‌بینی تبخیر و تعرق مرجع روزانه
Figure 4. Taylor diagram of studied methods in predicting daily reference evapotranspiration

همانطوری که از شکل ۴ مشاهده می‌گردد، مدل SVR-1 (نقطه آبی رنگ) با در نظر گرفتن تمامی پارامترهای ورودی، دقت بالاتری را در پیش‌بینی تبخیر و تعرق مرجع روزانه در ایستگه تبریز داشته است. بعد از این روش نیز، مدل M5-1 (نقطه قرمز رنگ) فاصله شعاعی کمتری با داده‌های مشاهداتی (نقطه سبز رنگ) داشته و به عنوان دومین روش برتر شناخته گردید.

در تحقیق حاضر، توانایی روش‌های یادگیری ماشینی شامل M5 و SVR برای تخمین تبخیر و تعرق گیاه مرجع مورد بررسی و کاوش قرار گرفت و نتایج به دست آمده از این دو روش با نتایج روش‌های تجربی هارگریوز و تورنت وایت مقایسه گردید. بدین منظور، ایستگاه سینوپتیک تبریز، به عوان یکی از قدیمی‌ترین ایستگاه‌های شمال غرب کشور، انتخاب گردید و داده‌های هواشناسی آن در طی دوره زمانی سال‌های ۹۴-۱۳۷۱ از سازمان هواشناسی اخذ شد. سپس، با محاسبه تبخیر و تعرق گیاه مرجع با روش فائو- پنمن- مونتیث و در نظر گرفتن این مقادیر به عنوان مقادیر هدف در مرحله واسنجی، ۱۷ سناریو مختلف متشکل از ترکیب‌های متفاوتی از پارامترهای هواشناسی برای روش‌های یادگیری ماشینی اشاره شده تعریف گردید و با در نظر گرفتن روش‌های تجربی هارگریوز و تورنت وایت، عملکرد آنها توسط داده‌های صحت سنجی مورد بررسی قرار گرفت. با مقایسه مقادیر پیش‌بینی شده توسط روش‌های مورد مطالعه و مقادیر حاصله از روش فائو- پنمن- مونتیث توسط پارامترهای آماری، مشخص شد که مدل‌های SVR-1 و M5-1 با داشتن کمترین مقدار خطا در مقایسه با دیگر سناریوهای تعریف شده و روش‌های تجربی مورد مطالعه، دقت به مراتب بالاتری را از خود نشان داده‌اند. به عبارت دیگر، می‌توان با استفاده از پارامترهای هواشناسی دمای میانگین، دمای حداقل، دمای حداکثر، رطوبت نسبی ،سرعت باد و نسبت ساعات آفتابی و استفاده از روش‌هایSVR و M5، مقدار تبخیر و تعرق گیاه مرجع را با دقت بالایی پیش‌بینی نمود. در نهایت پیشنهاد می‌گردد توانایی روش‌های مورد مطالعه در تخمین تبخیر و تعرق مرجع در اقلیم های مختلف توسط دیگر پژوهشگران مورد بحث و بررسی بیشتری قرار گیرد.

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]
[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]
[ DOI: 10.29252/jwmr.9.18.157 ]

## منابع


1. Alberg, D., M. Last and A. Kandel. 2012. Knowledge discovery in data streams with regression tree methods. Wiley Interdisciplinary Reviews Data Mining Knowledge Discovery, 2: 69-78.
2. Alikhanzadeh, A. 2006. Data minig. Olome Rayaneh, 340 pp (In Persian).
3. Alizadeh, A. 2002. Irrigation System Design. Ferdowsi University of Mashhad, 450 pp (In Persian).
4. Alizadeh, A. and Gh. Kamali 2008. Crops Water Requirements. Emam reza University of Mashhad, Iran, 227 pp (In Persian).
5. Allen, R.G., L.S. Pereira, D. Raes and M. Smith. 1998. Crop Evapotranspiration. Guidelines for Computing Crop Water Requirements. Irrigation and Drainage Paper No. 56, FAO, Rome, Italy, 300 pp.
6. Arun Raj, V.E. and P.G. Jairaj. 2014. Reference evapotranspiration modelliing using support vector regression. International Journal of Scientific & Engineering Research, 5(7): 2229-5518.
7. Babamiri, O., H. Nowzari and S. Maroofi. 2017. Potential Evapotranspiration Estimation using Stochastic Time Series Models (Case Study: Tabriz), Journal of Watershed Management Research, 8(15): 137-146.
8. Bhattacharya, B. and D.P. Solomatine. 2004. Neural networks and M5 model trees in modeling water level-discharge relationship. Department of Hydroinformatics and Knowledge Management, NESCO-IHE Institute for Water Education, P.O. Box 3015, 2601 DA Delft, The Netherlands.
9. Boser, B.E., Guyon, I.M. and V.N. Vapnik. 1992. A training algorithm for optimal margin classifiers. In D.Haussler, editor, 5th Annual ACM Workshop on COLT, pages 144-152, Pittsburgh, PA.
10. Chiew, F.H.S., N.N. Kamaladassa, H.M. Malano and T.A. MacMahon. 1995. Penman-Monteith, FAO-24 reference crop evapotranspiration and class-A pan data in Australia. Agricultural Water management, 28: 9-21.
11. Dastorani, M.T., S. Poormohammadi, A.R. Massah Bavani and M.H. Rahimian. Evapotranspiration Condition in Yazd Station under Uncertainties of Different GHG Emission Scenarios and ET Estimation Models, Journal of Watershed Management Research, 1(2): 1-20
12. Fallahi, M.R., H. Varvani and S. Goliyan. 2012. Precipitation forecasting using regression tree model to flood control. 5th International conference on watershed & soil and water management, Kerman, Iran (In Persian).
13. Gleckler, P.J., K.E. Taylor and C. Doutriaux. 2008. Performance metrics for climate models. Journal of Geophysical Research: Atmospheres, 113(D6): 1-20.
14. Jensen, M.E., R.D. Burman and R.G. Allen. 1990. Evapotranspiration and irrigation water requirements. ASCE Manuals and Report on Engineering Practices No. 70. American Society of Civil Engineers, New York, 360 pp.
15. Kisi, O. and M. Cimen. 2011. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. Journal of Hydrology, 399: 132-140.
16. Kumar, M., N.S. Raghuwanshi, R. Singh, W.W. Wallender and W.O. Pruitt. 2002. Estimating evapotranspiration using artificial neural network. Journal of Irrigation and Drainage Engineering-ASCE, 128(4): 224-233.
17. Manikumari, N. and G. Vinodhini. 2016. Regression models for predicting reference evapotranspiration. International Journal of Engineering Trends and Technology (IJETT), 38(3): 134-139.
18. Mirhashemi, H. and M. Panahi. 2015. Evaluation of a data mining model in predicting of average temperature and potential evapotranspiration month for the next month in the synoptic weather station Yazd. Biological Forum- An International Journal, 7(1): 1469-1473
19. Naderi, N. and A. Alizadeh. 1998. Determining reference crop evapotranspiration in Mashad and comparing with empirical methods, MSc Thesis, Ferdowsi University of Mashhad. 110 pp (In Persian).
20. Pal, M. 2006. M5 model tree for land cover classification. International Journal of Remote Sensing, 27(4): 825-831.
21. Pal, M. and S. Deswal. 2009. M5 model tree based modeling of reference evapotranspiration. Hydrological Processes. 23:1437-1443.
22. Panahi S., M. Karbasi and J. Nikbakht. 2016. Forecasting of reference evapotranspiration using MLP, RBF and SVM neural networks. Journal of Environment and Water Engineering, 2(1): 51-63
23. Quinlan, J.R. 1992. Learning with continuous classes. In: Proceedings of Australian Joint Conference on Artificial Intelligence (Singapore: World Scientific Press), pp: 343-348.
24. Samadianfard, S. and E. Asadi. 2017. Prediction of SPI drought index using support vector and multiple linear regressions. Journal of Water and Soil Resources Conservation, 6(4): 1-16 (In Persian).
25. Sattari, M.T., V. Ahmadifar and R. Pashapourar. 2014. M5 tree model based modeling of evaporation losses in Eleviyan reservoir and comparison with empirical equations. Irrigation & Water Engineering. 5(17): 110-122 (In Persian).
26. Sattari, M.T., F. Nahrein and V. Azimi. 2014. M5 model trees and neural networks based prediction of daily ET0 (Case Study: Bonab Station). Irrigation and drainage, 1: 104-113 (In Persian).
27. Sayyadi, H., A. Oladghaffari, A. Faalian and A.A. Sadraddini. 2009. Comparison of RBF and MLP neural networks performance of reference crop evapotranspiration, Water and Soil Science, 19(1): 1-12 (In Persian).
28. Soltani, A., S.M. Mirlatifi and H. Dehghanisanij. 2012. Estimating Reference Evapotranspiration Using Limited Weather Data under Different Climatic Conditions. Journal of Water and Soil, 26(1): 139-149 (In Persian).
29. Taylor, K.E. 2001. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research: Atmospheres, 106: 7183-7192.
30. Trajkovic, S. 2007. Hargreaves versus Penman-Monteith under humid conditions. Journal of Irrigation and Drainage Engineering-ASCE, 133(1): 38-42.
31. Trajkovic, S. and S. Kolakovic. 2009. Estimating reference evapotranspiration using limited weather data. Journal of Irrigation and Drainage Engineering-ASCE, 135(4): 443-449.
32. Vapnik, V.N. 1995. The Nature of Statistical Learning Theory. Springer, New York. 314 pp.
33. Vapnik, V.N. 1998. Statistical Learning Theory. Wiley, New York. 736 pp.


[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]

[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]

[ DOI: 10.29252/jwmr.9.18.157 ]

# Estimating Daily Reference Evapotranspiration using Data Mining Methods of Support Vector Regression and M5 Model Tree

**Saeed Samadianfard[1] and Solmaz Panahi[2]**

1- Assistant Professor, Department of Water Engineering, Faculty of Agriculture, Tabriz University, (Corresponding author: s.samadian@tabrizu.ac.ir)
2- M.Sc. Student, of Irrigation and Drainage, Department of Water Engineering, Faculty of Agriculture, Tabriz University
Received: November 7, 2017 Accepted: July 2, 2018

## Abstract

Evapotranspiration is one of the most important components of the hydrological circle and its proper determination is highly important in most researches such as water hydrological balance, design and management of irrigation systems, simulation of crop production and design and management of water resources. Nonlinear characteristic, uncertainty and needing for different climatological data in simulating evapotranspiration are the reasons that motivate researchers to investigate data mining methods such as M5 model trees and support vector regression. In the present study, the precision of mentioned methods in estimation of reference crop evapotranspiration in comparison with empirical methods such as Hargrivs and Torrent white equations was studied. For that purpose, using meteorological dataset of 1371-1394 years of Tabriz synoptic station, the daily values of reference crop evapotranspiration were computed by FAO-Penman-Monteith method. Then, using these computed values as target outputs, 17 various scenarios combining at last one to up to six meteorological parameters have been considered using mentioned methods. Finally, the capability of support vector regression and M5 model trees for estimation of evapotranspiration was analyzed using test data set. Results of statistical analysis and Taylor diagram showed that support vector regression and M5 model trees in a case of considering all meteorological parameters with root mean square of 0.398 and 0.44, respectively, provided precise results comparing with empirical methods such as Hargrivs and Torrent white.

**Keywords:** Empirical Methods, FAO-Penman-Monteith Method, Machine Mearning Algorithms, Reference Evapotranspiration

[ Downloaded from jwmr.sanru.ac.ir on 2025-12-23 ]
[ DOR: 20.1001.1.22516174.1397.9.18.21.5 ]
[ DOI: 10.29252/jwmr.9.18.157 ]