

مقاله علمی-پژوهشی

ارزیابی مدل هیبریدی ANN-PSO و روش‌های مختلف داده‌کاوی در برآورد تبخیر و تعرق ماهانه در دو اقلیم مختلف

سید محمدرضا حسینی وردنجانی^۱، مجتبی خوش‌روش^{۲*}، سید ابراهیم حسینی کاکلکی^۳، مسعود پورغلام آمیجی^۴

تاریخ دریافت: ۱۴۰۱/۰۴/۱۴ تاریخ پذیرش: ۱۴۰۱/۰۵/۰۴

چکیده

تبخیر و تعرق یکی از مؤلفه‌های اصلی بیلان آب هر منطقه و برآورد دقیق آن در مطالعات هیدرولوژیکی، طراحی سامانه‌های آبیاری و زهکشی و برنامه‌ریزی سامانه‌های آبیاری بسیار ضروری است. در این تحقیق به ارزیابی مدل درختی M5، M5 Rules، K Star، Rep Tree، مدل شبکه عصبی مصنوعی و مدل هیبرید شبکه عصبی ANN-PSO در تخمین تبخیر و تعرق مرجع در دو نقطه اقلیمی مختلف در استان مرکزی بر مبنای مدل پنمن مونتیت فائو پرداخته شد. داده‌های مورد استفاده شامل دمای حداقل و حداکثر، رطوبت نسبی میانگین، سرعت باد در ارتفاع دو متری و ساعات آفتابی به عنوان ورودی مدل از ایستگاه سینوپتیک دلیجان و تفرش بین سال‌های ۲۰۲۱-۲۰۰۴ بودند و تبخیر و تعرق محاسبه‌شده از روش پنمن مونتیت فائو به عنوان خروجی مدل‌ها استفاده شد. برای ارزیابی مدل‌ها از شاخص‌های جذر میانگین مربعات خطا (RMSE)، میانگین انحراف مطلق (MAE) و ضریب همبستگی (R) استفاده شد. تجزیه و تحلیل نتایج نشان داد که برای هر دو ایستگاه، مدل هیبرید شبکه عصبی ANN-PSO با $RMSE=0.3115$ ، $MAE=0.2441$ و $R=0.9989$ برای ایستگاه دلیجان و $RMSE=0.2915$ ، $MAE=0.2355$ و $R=0.9989$ برای ایستگاه تفرش و مدل درختی M5 با $RMSE=0.2793$ ، $MAE=0.2398$ و $R=0.9967$ برای ایستگاه دلیجان و $RMSE=0.2803$ ، $MAE=0.2306$ و $R=0.9969$ برای ایستگاه تفرش بهترین عملکرد را داشتند. همچنین مدل Rep Tree ضعیف‌ترین عملکرد را در بین مدل‌های مورد بررسی در این پژوهش داشت. با توجه به اینکه مدل درختی علاوه بر دقت مطلوب، روابط ساده، خطی و قابل‌فهم‌تری را برای برآورد تبخیر و تعرق مرجع ارائه می‌نماید، این مدل برای برآورد تبخیر و تعرق در این منطقه قابل توصیه است.

واژه‌های کلیدی: مناطق خشک، تبخیر و تعرق، یادگیری ماشین، محاسبات نرم

مقدمه

چون تفکیک آن‌ها از یکدیگر امکان‌پذیر نیست، مجموعاً به نام تبخیر و تعرق (ET^0) در نظر گرفته می‌شود (Allen et al., 1998; Eichelmann et al., 2022). پیچیدگی سیستم‌های هیدرولوژیکی از نظر تعیین تمامی پارامترهای دخیل و نقص اطلاعات آماری، امکان مدل‌سازی فیزیکی کامل سیستم‌های هیدرولوژیکی را غیرممکن می‌سازد. لذا در این شرایط مدل‌سازی سیستمی که بتواند مبتنی بر روابط ریاضی باشد، مورد توجه قرار می‌گیرد (Khoshravesh et al., 2016; نورانی و سیاح فرد، ۱۳۹۲). روش‌های داده‌محور با توجه به انعطاف‌پذیری بالا در مدل‌سازی شرایط مختلف یک پدیده توانایی خوبی را دارا هستند (بختیاری و همکاران، ۱۳۹۴؛ پورغلام آمیجی و همکاران، ۱۴۰۱). در حال حاضر داده‌کاوی مهم‌ترین فناوری جهت بهره‌برداری از داده‌های حجیم است و اهمیت و کاربرد آن روبه‌فزونی

تبخیرتعرق به انتقال همزمان تبخیر از سطح خاک و تعرق از سطح گیاه به اتمسفر در یک سیستم خاک - گیاه اطلاق می‌شود (Allen et al. 1998). این دو پدیده هر دو ماهیت تبخیری داشته و

- ۱- دانشجوی دکتری گروه مهندسی آب، دانشکده مهندسی زراعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران
- ۲- دانشیار گروه مهندسی آب، دانشکده مهندسی زراعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران
- ۳- دانشجوی دکتری گروه مهندسی برق، دانشکده مهندسی، دانشگاه فردوسی مشهد، مشهد، ایران
- ۴- دانشجوی دکتری گروه مهندسی آبیاری و آبادانی، دانشکده‌گان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران

*- نویسنده مسئول: (Email: m.khoshravesh@sanru.ac.ir)

DOR: 20.1001.1.20087942.1401.16.5.4.9

مدل پنمن مونیتث فائو ۵۶ مقایسه گردید. نتایج حاکی از برتری مدل درختی بود. در تحقیقی در کشور صربستان به تخمین تبخیر و تعرق به کمک شبکه تابع پایه شعاعی با بهینه‌سازی ازدحام ذرات (PSO) و شبکه تابع پایه شعاعی با الگوریتم پس انتشار خطا پرداخته شد. برای ارزیابی مدل‌ها از مدل پنمن مونیتث فائو استفاده گردید. نتایج نشان داد هر دو مدل دقت قابل قبولی در تخمین تبخیر و تعرق مرجع دارند ولی بر طبق شاخص‌های آماری مدل شبکه تابع پایه شعاعی با بهینه‌سازی ازدحام ذرات PSO دقیق‌تر است (Petković et al., 2016). قربانی و همکاران (۱۳۹۵) به ارزیابی عملکرد شبکه‌های عصبی در برآورد تبخیر و تعرق گیاه مرجع در ایستگاه سینوپتیک اهواز پرداختند. آن‌ها از شبکه‌های MNN^3 ، MLP^4 و FF^5 برای تخمین تبخیر و تعرق گیاه مرجع استفاده نمودند. نتایج نشان داد که شبکه‌های MLP و FF عملکرد بهتری نسبت به MNN دارند. ستاری و اسماعیل‌زاده (۱۳۹۵) به ارزیابی مدل درختی M5 و مدل برنامه‌ریزی ژنتیک (GP^6) در برآورد تبخیر تعرق ماهانه در منطقه تبریز آذربایجان آذربایجان شرقی پرداختند. برای ارزیابی مدل‌ها از روش پنمن مونیتث فائو استفاده گردید. نتایج نشان داد اگرچه مدل برنامه‌ریزی ژنتیک دقت بیشتری نسبت به مدل درختی دارد ولی مدل درختی روابط خطی ساده‌تری ارائه می‌کند. پناهی و همکاران (۱۳۹۵) به پیش‌بینی تبخیر و تعرق مرجع به کمک سه نوع شبکه عصبی پرسپترون، شعاعی و ماشین بردار پشتیبان در ایستگاه همدیدی تبریز پرداختند. نتایج نشان داد که شبکه عصبی پرسپترون با $RMSE=3.88$ و $R=0.95$ عملکرد بهتری نسبت به مدل‌های دیگر داشت. در تحقیقی در کشور یونان به تخمین تبخیر و تعرق مرجع روزانه به کمک مدل‌های مختلف شبکه عصبی مصنوعی و مدل‌های تجربی پرداخته شد. نتایج حاکی از دقت بالای مدل شبکه عصبی مصنوعی نسبت به مدل‌های تجربی بود (Antonopoulos and Antonopoulos, 2017). ذرتی پور و همکاران (۱۳۹۸) به محاسبه تبخیر و تعرق پتانسیل به کمک مدل شبکه عصبی مصنوعی، عصبی فازی و مدل درختی M5 پرداختند. نتایج نشان داد هر سه مدل دقت قابل قبولی در محاسبه تبخیر و تعرق دارند. دقت مدل شبکه عصبی مصنوعی و عصبی فازی بیشتر است. حقی‌زاده و همکاران (۱۳۹۸) به مقایسه مدل هیبریدی ANFIS-PSO و مدل تجربی تورک در تخمین تبخیر و تعرق مرجع در پل دختر پرداختند. نتایج نشان داد که مدل هیبریدی ANFIS-PSO با سه پارامتر دمایی میانگین، رطوبت و گرمای نهان تبخیر قادر به تخمین تبخیر و تعرق مرجع روزانه است.

است (Sattari et al, 2013; Rahimikhoob, 2014; Ansari, 2020). شبکه‌های عصبی به عنوان یکی از مؤلفه‌های اصلی هوش محاسباتی^۱ از خواص مهمی برخوردارند که آن‌ها را در علوم و مسائل فنی مهندسی حائز اهمیت می‌نماید. از جمله خواص مهم این شبکه‌ها تقریب زنی توابع، ساختار موازی، قدرت یادگیری و تعمیم و سرعت پردازش بالا است (منهاج، ۱۳۹۳). الگوریتم بهینه‌سازی ازدحام ذرات (PSO^2)، یک الگوریتم بهینه‌سازی تصادفی مبتنی بر حرکت ذرات است (Kennedy Eberhart, 1995) and. این الگوریتم از روی رفتار اجتماعی دسته‌های پرندگان مدل شده است. روش‌های داده‌کاوی برای مجموعه داده‌های بزرگ با متغیرهای زیاد ساخته شده‌اند. بنابراین با روش‌های آماری قدیمی متفاوت هستند. مدل M5 یک مدل درختی برای پیش‌بینی صفات عددی پیوسته است که در آن توابع رگرسیونی خطی در برگ‌های این درخت ظاهر می‌شوند (Pal and Deswal, 2009). مدل درختی M5 اولین بار توسط (Quinlan, 1992) معرفی شده است. از مزایای مدل درختی M5 این است که قابل فهم‌تر بوده و روند آموزش به مدل بسیار ساده است و نیازی به آزمون و خطا ندارد. همچنین در مواجهه با داده‌های گمشده بسیار قدرتمند و قوی عمل کرده، و بر روی داده‌ها و نمونه‌های بزرگ و دارای ابعاد زیاد نیز قابلیت عملکرد مناسبی دارد. از دیگر روش‌های داده‌کاوی می‌توان به روش‌های M5 Rules، K star و Rep Tree اشاره نمود.

در زمینه پیش‌بینی تبخیر و تعرق به کمک روش‌های داده‌محور، پژوهش‌های مختلفی در داخل و خارج کشور انجام شده است. در تحقیقی در کشور ترکیه با به کارگیری داده‌های هواشناسی دریاچه‌ی آغیردر در ترکیه به مدل سازی تبخیر از تشت پرداختند و نشان دادند که پارامترهای دمای هوا، دمای آب، تابش خورشیدی و رطوبت نسبی به عنوان مؤثرترین عوامل بر تبخیر از تشت می‌باشند. در این مطالعه از الگوریتم‌های داده‌کاوی M5، M5Rules، Kstar، شبکه عصبی و روش‌های رگرسیونی استفاده شد. نتایج حاکی از دقت بالای مدل M5 نسبت به سایر روش‌ها بود (Terzi, 2007). پناهی و همکاران (۱۳۹۴) به ارزیابی سه مدل از نرم‌افزار Weka شامل M5p، Rep Tree و K Star در برآورد تبخیر و تعرق پتانسیل و رطوبت نسبی ماه بعد در استان یزد پرداختند. نتایج نشان داد که M5p بهترین عملکرد Rep Tree و Rep Tree ضعیف‌ترین عملکرد را در این منطقه داشتند. وروانی و همکاران (۱۳۹۲) به کمک مدل رگرسیون درختی نرم‌افزار Xlminer در چند ناحیه آب و هوایی مختلف در ایران به تخمین تبخیر و تعرق ماهانه پرداختند. داده‌های ورودی شامل میانگین ماهانه دمای حداقل و حداکثر، رطوبت نسبی تابش خورشیدی و سرعت باد بودند. نتایج با

3- Mobile Neural Network

4- Multi-Layer Perceptron

5- Fuzzy Forests

6- Genetic Programming

7- Adaptive Neuro Fuzzy Inference System

1- Computational Intelligence

2- Particle Swarm Optimization

شمالی می‌باشد. متوسط دمای سالانه ۱۶/۱ درجه سانتی‌گراد و گرم‌ترین و سردترین ماه‌های سال به ترتیب تیر و دی می‌باشند. متوسط سالانه بارندگی زراعی ۱۸۵/۲ میلی‌متر و متوسط رطوبت سالانه ۳۹ درصد می‌باشد. اقلیم این شهرستان به روش نمایه دومارتن خشک و به روش نمایه آمبرژه خشک و سرد است. همین‌طور تفرش یکی دیگر از شهرستان‌های استان مرکزی است که با طول جغرافیایی ۵۰ درجه ۰۱ دقیقه شرقی و عرض جغرافیایی ۳۴ درجه و ۴۱ دقیقه شمالی می‌باشد. متوسط دمای سالانه ۱۳/۳ درجه سانتی‌گراد و گرم‌ترین و سردترین ماه‌های سال به ترتیب مرداد و دی می‌باشند. متوسط سالانه بارندگی زراعی ۳۱۴/۴ میلی‌متر و متوسط رطوبت سالانه ۴۴ درصد می‌باشد. اقلیم این شهرستان به روش نمایه دومارتن نیمه‌خشک و به روش نمایه آمبرژه نیمه‌خشک و سرد است. در این تحقیق از میانگین ماهانه داده‌های هواشناسی ایستگاه سینوپتیک دلجان و تفرش در بازه زمانی ۲۰۰۴ تا ۲۰۲۱ شامل دمای حداقل و حداکثر، رطوبت نسبی میانگین، سرعت باد در ارتفاع ۲ متری و ساعات آفتابی استفاده گردید. محدوده مطالعاتی در شکل (۱) نشان داده شده است.

محاسبه تبخیر و تعرق مرجع

رابطه فائو پنمن مونتیث یکی از معتبرترین روش‌ها برای تخمین ET_0 است که مورد استفاده متخصصان قرار می‌گیرد. در روش مذکور، گیاه مرجع، چمن با ارتفاع ۱۲ سانتی‌متر و ضریب بازتاب تابش آن ۲۳ درصد است (Allen et al., 1998). این مدل به عنوان مدل مرجع تبخیر و تعرق شناخته شده و سایر روش‌ها نسبت به این روش کالیبره می‌شوند. مدل این روش به صورت رابطه (۱) می‌باشد:

$$ET_0 = \frac{0.408 \cdot \Delta (R_n - G) + \gamma \left[\frac{900}{(T + 273)} \right] U_2 (e_a - e_d)}{\Delta + \gamma (1 + 0.34 U_2)} \quad (1)$$

در این رابطه ET_0 تبخیر و تعرق روزانه (mm/day)، R_n تابش خورشیدی روزانه رسیده به سطح گیاه ($Mjm^{-2}d^{-1}$)، G شار گرما به داخل خاک ($Mjm^{-2}d^{-1}$)، γ ثابت سایکرومتری ($Kpac^{-1}$)، T میانگین دمای روزانه ($^{\circ}C$)، u_2 سرعت باد در ارتفاع دو متری (m/s)، e_a فشار بخار اشباع (Kpa)، e_d فشار بخار واقعی (Kpa) و Δ شیب منحنی فشار بخار اشباع ($Kpac^{-1}$) می‌باشد (Allen et al, 1998).

مدل شبکه عصبی مصنوعی

شبکه‌های چندلایه پیش‌خور یکی از مهم‌ترین ساختارهای شبکه عصبی می‌باشد. به طور معمول این شبکه‌ها شامل مجموعه‌ای از واحدهای حسی (نرون) می‌باشند که تشکیل‌دهنده لایه ورودی، یک یا چند لایه پنهان و یک لایه خروجی می‌باشند. سیگنال ورودی در خلأ شبکه و در مسیری رو به جلو به صورت لایه به لایه منتشر می‌شود. این نوع شبکه‌ها معمولاً به عنوان پرسپترون چندلایه نامیده

حسینی وردنجانی و همکاران (۱۳۹۹) به ارائه یک ضریب اصلاحی برای بهبود روش هارگریوز در محاسبه تبخیر و تعرق مرجع روزانه در منطقه خشک سرد شهرکرد به کمک مدل درخت تصمیم و مدل شبکه عصبی مصنوعی پرداختند. نتایج نشان داد که عملکرد مدل شبکه عصبی مصنوعی در تخمین ضریب اصلاحی دقیق‌تر است. سپهری و همکاران (۱۳۹۹) به بررسی و تحلیل حساسیت روش‌های مبتنی بر شبکه عصبی در برآورد تبخیر و تعرق مرجع در استان تهران پرداختند. نتایج نشان داد مدل شبکه عصبی مصنوعی تکنیک بسیار مناسبی برای تحلیل تبخیر و تعرق مرجع می‌باشد. همچنین مشخص گردید شیوه آموزش پرسپترون چندلایه با دو لایه میانی و توابع انتقال تانژانت خطی و تانژانت برای لایه‌های پنهان و خروجی و قاعده آموزشی لوبنبرگ مارکوات برای دو لایه پنهان و خروجی و ساختار ۱-۱۶-۱۶-۶ به عنوان بهترین شبکه برای برآورد تبخیر و تعرق مرجع پیشنهاد می‌گردد.

در تحقیقی در کشور استرالیا در پارکی نزدیک شهر ملبورن به برآورد تبخیر و تعرق مرجع به کمک شبکه عصبی مصنوعی پرداخته شد. برای محاسبه تبخیر و تعرق مرجع از روش پنمن مونتیث فائو استفاده گردید. نتایج نشان داد که مدل شبکه عصبی مصنوعی دقت قابل قبولی در تخمین تبخیر و تعرق مرجع در این منطقه دارد. همچنین مشخص گردید تغییرات آب و هوایی باعث افزایش تبخیر و تعرق مرجع می‌گردد (Algreawee and Alshama, 2021). بیدآبادی و همکاران (۱۴۰۰) به برآورد تبخیر و تعرق مرجع با استفاده از مدل شبکه عصبی مصنوعی و مدل نرو فازی در اقلیم خشک و نیمه خشک (استان کرمان) پرداختند. نتایج نشان داد دو مدل فوق عملکرد بهتری نسبت به مدل‌های تجربی دارند و دقت مدل نرو فازی در مقایسه با سایر مدل‌ها بیشتر است. با توجه به نقش کلیدی تبخیر و تعرق در مدیریت منابع آب و طراحی سامانه‌های آبیاری و زهکش و از آنجا که تحقیقات زیادی در خصوص مدل ترکیبی شبکه عصبی ANN¹-PSO و مقایسه نتایج آن با الگوریتم‌های داده‌کاوی در کشور انجام نشده است، ارزشمندی تحقیق حاضر را بیان می‌کند. هدف این تحقیق ارزیابی مدل درختی M5 Rules, M5, K Star, Rep Tree, مدل شبکه عصبی مصنوعی و مدل هیبرید شبکه عصبی ANN-PSO در تخمین تبخیر و تعرق مرجع در دو نقطه اقلیمی مختلف در استان مرکزی، منطقه تفرش و دلجان می‌باشد.

مواد و روش‌ها

منطقه مورد مطالعه

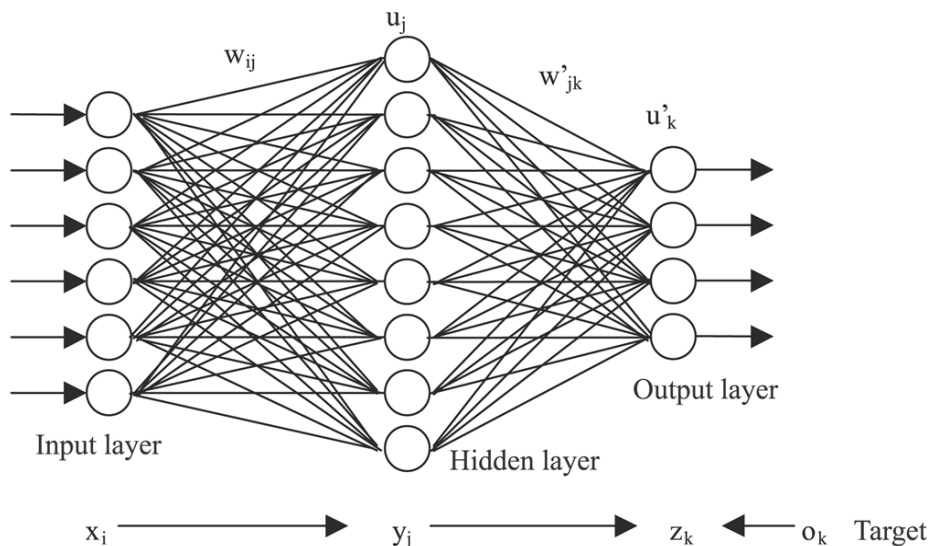
دلجان یکی از شهرستان‌های استان مرکزی با طول جغرافیایی ۵۰ درجه ۴۲ دقیقه شرقی و عرض جغرافیایی ۳۳ درجه و ۵۹ دقیقه

منظور دستیابی به پاسخی پایدارتر و غلبه بر مشکل برآزش بیش از حد در آموزش مدل ANN، از الگوریتم تنظیم بیزی استفاده شد. الگوریتم تنظیم بیزی خطاها را در خطاهای شبکه و وزن شبکه و بایاس به حداقل می‌رساند (Mackay, 1992). شکل (۲)، شماتیک یک شبکه عصبی مصنوعی را نشان می‌دهد.

می‌شوند (Haykin, 1994). برای پیش‌بینی تبخیر و تعرق مرجع از شبکه پرسپترون چندلایه پیش‌خور با یک قانون یادگیری پس انتشار و با یک لایه پنهان استفاده شد. این نوع شبکه قادر به مدل‌سازی و تشخیص هر نوع رابطه بین متغیرهای ورودی و خروجی است (Minasny and McBratney, 2002a, 2002b;) به (Khalilmoghdam et al., 2009; Mirzaee et al., 2017).



شکل ۱- موقعیت جغرافیایی مناطق مورد مطالعه



شکل ۲- شماتیک یک شبکه عصبی مصنوعی (ANN)

که در این رابطه x_n مقدار نرمالیز شده، x مقدار هر متغیر ورودی، x_{min} مینیمم داده‌ی ورودی، x_{max} ماکزیمم داده‌ی ورودی می‌باشد.

الگوریتم PSO

مراحل اجرای الگوریتم PSO شامل مراحل ذیل است. هر فرد در این توده به عنوان یک ذره در فضای n بعدی در نظر گرفته می‌شود

نخستین گام در ایجاد شبکه عصبی مصنوعی انتخاب داده‌های ورودی و نرمال‌سازی داده‌ها قبل از ورود به شبکه می‌باشد. نرمال‌سازی داده‌ها به منظور افزایش دقت و سرعت شبکه در پاسخ به پیام‌های ورودی صورت می‌گیرد. با استفاده از رابطه (۲) اقدام به نرمال‌سازی داده‌ها شد (احمدی و همکاران، ۱۳۹۴).

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

به صفر می‌رسد. با این کار درخت بزرگی توسعه می‌یابد. بنابراین برای رسیدن به یک درخت بهینه و کارآمد باید شاخه‌های اضافی درخت هرس شوند. دومین مرحله در طراحی مدل، هرس کردن درخت بیش از حد رشد کرده، و جایگزینی روابط رگرسیونی خطی به جای شاخه‌ها خواهد بود. این تکنیک تولید مدل درختی، فضای پارامترهای ورودی را به زیرحوزه‌هایی تقسیم نموده و برای هر کدام از آن‌ها یک رابطه خطی ارائه می‌کند (Kisi et al., 2021; Genolini, 2022).

مدل M5 Rules

از مدل درختی ساخته شده توسط مدل M5p Tree، قانون رگرسیونی تولید می‌نماید.

مدل Rep Tree

این مدل با استفاده از مقادیر واریانس درخت تصمیم یا درخت رگرسیونی می‌سازد و با هرس خطای کاهش یافته، آن را هرس میکند. برای بهبود سرعت، این روش تنها یک بار مقادیر را برای صفات عددی مرتبط می‌نماید (Belouch et al., 2017). می‌توان حداقل تعداد نمونه‌ها را در هر برگ، ماکزیمم عمق درخت حداقل نسبت واریانس سری آموزشی برای یک انشعاب (تنها کلاس عددی) و تعداد فولدها برای هرس نمودن را تنظیم نمود (Ngo et al., 2022).

مدل K star

الگوریتم مبتنی بر فاصله که در آن تابع آنتروپی استفاده می‌شود. در واقع این الگوریتم شباهت داده با داده‌های آموزشی را در نظر می‌گیرد (فضل مقصودی و مومنی، ۱۳۹۲).

معیارهای ارزیابی

در این تحقیق به منظور ارزیابی روش‌ها، شاخص‌های آماری میانگین انحراف مطلق^۱، جذر میانگین مربعات خطا^۲ و ضریب همبستگی^۳ بر اساس رابطه‌های (۶، ۷ و ۸) محاسبه شدند:

$$MAE = \frac{\sum_{i=1}^N (|O_i - P_i|)}{N} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_i - P_i)^2}{n}} \quad (7)$$

$$R = \frac{\sum_{i=1}^N O_i P_i}{\sqrt{\sum_{i=1}^N O_i^2 \sum_{i=1}^N P_i^2}} \quad (8)$$

در این روابط: O_i تبخیر و تعرق به روش پنمن موتنیث فائو؛ P_i تبخیر و تعرق محاسبه شده به هر روش؛ MAE میانگین مطلق خطا

(Hou et al., 2022). سرعت و موقعیت ذره i به ترتیب به صورت رابطه (۳) تعریف می‌شوند:

$$V_i(t) = [V_{i1}(t), V_{i2}(t), \dots, V_{in}(t)] \quad \text{and} \quad X_i(t) = [X_{i1}(t), X_{i2}(t), \dots, X_{in}(t)] \quad (3)$$

حرکت ذرات به صورت رابطه (۴ و ۵) محاسبه می‌شود:

$$V_{i,d}(t+1) = W(t)V_{i,d}(t) + c_1 r_1 (P_{i,d}(t) - X_{i,d}(t)) + c_2 r_2 (P_{g,d}(t) - X_{i,d}(t)) \quad (4)$$

$$X_{i,d}(t+1) = X_{i,d}(t) + V_{i,d}(t+1), i = 1, 2, \dots, M; d = 1, 2, \dots, n \quad (5)$$

که در این رابطه: $V_{i,d}$ سرعت ذره i در تکرار $t+1$ است که در بازه $[-V_{max}, V_{max}]$ محدود شده است، $X_{i,d}$ مقدار موقعیت ذره i است، c_1 و c_2 ضرایب شتاب هستند. بردار $\{P_{i,1}, P_{i,2}, \dots, P_{i,n}\}$ نشان دهنده قرارگیری است که بهترین ارزش تناسب (بهترین جایگذاری قبلی) ذره را می‌دهد. بردار $P_g = \{P_{g,1}, P_{g,2}, \dots, P_{g,n}\}$ قرارگیری بهترین ذره در مقایسه با سایر ذرات جمعیت است و به عنوان بهترین موقعیت جهانی (gbest) شناخته می‌شود. متغیرهای r_1 و r_2 دو عدد تصادفی هستند که به طور یکنواخت در بازه $[0, 1]$ توزیع شده‌اند، و $W(t)$ وزن اینرسی است که توسط (Shi and Eberhart, 1998) برای سرعت بخشیدن به همگرایی الگوریتم PSO معرفی شد (Petković et al., 2016).

مدل درختی M5

تولید مدل درختی در دو مرحله صورت می‌گیرد. مرحله اول شامل تعیین مناسب‌ترین پارامتر ورودی برای انشعاب و نیز معیار تقسیم برای تولید یک درخت تصمیم است. انتخاب پارامتر تقسیم مناسب و ایجاد نقطه انشعاب در درخت از اهمیت خاصی برخوردار است. هر گره بر اساس شباهت بین اعضای آن، به دو زیرگره تفکیک می‌شود. معیار تقسیم برای الگوریتم مدل درختی M5 بر مبنای در نظر گرفتن انحراف معیار مقادیر دسته‌ی منتهی به هر گره به عنوان میزان خطا در آن گره و محاسبه‌ی کاهش مورد انتظار^۱ در این خطا به عنوان نتیجه آزمون هر صفت در آن گره می‌باشد (Binkert et al., 2006). کاهش انحراف معیار از رابطه (۶) محاسبه می‌شود:

$$SDR = sd(t) - \sum \frac{|T_i|}{|T|} sd(T_i) \quad (6)$$

که در آن، T دسته‌ای از نمونه‌ها که به گره می‌رسند؛ T_i زیرمجموعه‌ی نمونه‌هایی که آمین خروجی از دسته‌ی پتانسیل را دارد و Sd انحراف از معیار است. چون فرآیند تقسیم داده‌ها در گره‌های کوچک انحراف استاندارد کمتری نسبت به گره‌های بزرگ دارد بنابراین نتایج خالص‌تر است. فرآیند انشعاب در هر گره بارها و بارها تکرار می‌شود تا به گره پایانی برسد که مجذور انحراف از میانگین داده‌ها تقریباً

2 Root Mean Square of Error
3 Correlation Coefficient
4 Expected reduction

1- Mean Absolute Error

حداکثر، رطوبت نسبی میانگین، ساعات آفتابی و سرعت باد در ارتفاع دو متری به عنوان ورودی مدل استفاده گردید. داده‌ها به دو قسمت ۷۵ درصد آموزش و اعتبارسنجی و ۲۵ درصد آزمایش تقسیم شدند و مدل اجرا گردید.

آنالیز آماری با رتبه‌بندی بین شاخص‌ها انجام گردید به نحوی هر شاخص در ایده‌آل‌ترین حالت بیشترین امتیاز و در ضعیف‌ترین حالت ممکن کم‌ترین امتیاز به آن شاخص تعلق یابد. نتایج مدل‌سازی تبخیر و تعرق با مدل شبکه عصبی مصنوعی و مدل هیبرید شبکه عصبی ANN-PSO (جدول ۱) نشان داد که هر دو مدل دقت قابل قبولی در تخمین تبخیر و تعرق مرجع دارند. نتایج نشان داد که مدل هیبرید شبکه عصبی ANN-PSO با $MAE=0.17$, $RMSE=0.244$ و $R=0.9992$ برای ایستگاه دلیجان و $MAE=0.2$, $RMSE=0.2331$ و $R=0.9993$ برای ایستگاه تفرش و مدل شبکه عصبی مصنوعی با $MAE=0.2441$, $RMSE=0.3115$ و $R=0.9989$ برای ایستگاه دلیجان و $MAE=0.2355$, $RMSE=0.2915$ و $R=0.9989$ دقت قابل قبولی در تخمین تبخیر و تعرق مرجع در این مناطق دارد. نتایج نشان می‌دهد که مدل هیبرید شبکه عصبی ANN-PSO ضمن افزایش دقت شبکه عصبی باعث کم شدن زمان آموزش شبکه می‌گردد.

نتایج مدل‌سازی تبخیر و تعرق با مدل درختی M5 نشان داد که مدل درختی دقت بالایی در برآورد تبخیر و تعرق در این منطقه دارد (جدول ۲).

مدل درختی با $MAE=0.2398$, $RMSE=0.2793$ و $R=0.9967$ برای ایستگاه دلیجان و $MAE=0.2306$, $RMSE=0.2803$ و $R=0.9969$ برای ایستگاه تفرش محاسبه گردید که عملکرد مطلوبی را در تخمین تبخیر و تعرق مرجع در این ایستگاه‌ها را دارد. مدل درختی روابط خطی و ساده‌ای برای برآورد تبخیر و تعرق مرجع ارائه می‌کند، هر کدام از این روابط برای شرایط خاصی از دما، رطوبت، ساعات آفتابی، سرعت باد و غیره صادق هستند. شکل (۴) و (۵) درخت بهینه تولیدشده را نشان می‌دهد که در این شکل‌ها دایره‌ها نماینده گره‌های تقسیم، خطوط رابط نشانگر شرایط تقسیم، و مستطیل‌ها گره‌های نهایی یا همان برگ‌ها هستند که عبارت LM نشانگر شماره رابطه‌ی رگرسیونی است که در ذیل درخت آمده است و شماره درون پرانتزها نیز نشانگر تعداد داده‌های آموزشی است که با هر برگ پوشش داده و درصد خطای هر رابطه می‌باشد. با توجه به شکل مهم‌ترین پارامتر در ایستگاه دلیجان دمای ماکزیمم T_{max} مهم‌ترین پارامتر است و درخت را به دو شاخه $T_{max} \leq 21.597$ و $T_{max} > 21.597$ تقسیم می‌کند. بر اساس معیار انشعاب $T_{max} > 21.597$ دمای مینیمم بیش‌ترین اهمیت را دارد و بر اساس معیار انشعاب $T_{max} \leq 21.597$ دمای ماکزیمم بیش‌ترین اهمیت را دارد. برای ایستگاه تفرش نیز دمای ماکزیمم

که هر چه این شاخص به صفر نزدیک‌تر باشد، نشان‌دهنده انحراف کمتر و دقت بالاتر مدل می‌باشد. $RMSE$ جذر میانگین مربعات خطا که هر چه این شاخص به صفر نزدیک‌تر باشد نشان‌دهنده خطای کمتر و دقت بالای مدل می‌باشد و R ضریب همبستگی که هر چه این شاخص به یک نزدیک‌تر باشد، نشان‌دهنده همبستگی بالای مدل می‌باشد.

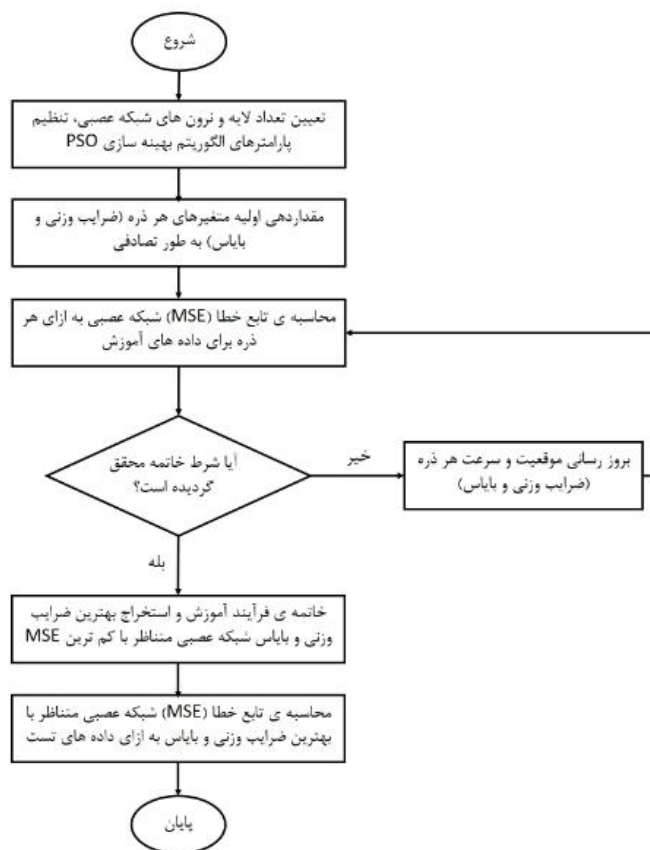
نتایج و بحث

در این تحقیق به ارزیابی مدل درختی $M5$ ، $M5rules$ ، K star، Rep Tree و مدل شبکه عصبی مصنوعی با و مدل هیبرید شبکه عصبی ANN-PSO در تخمین تبخیر و تعرق مرجع در دو نقطه اقلیمی مختلف در استان مرکزی، منطقه تفرش و دلیجان پرداخته شد. داده‌های ورودی به مدل شبکه عصبی و مدل هیبرید شبکه عصبی ANN-PSO عبارت از دمای حداکثر و دمای حداقل، ساعات آفتابی، سرعت باد در ارتفاع دو متری و رطوبت نسبی میانگین بودند. به دلیل عدم دسترسی به داده‌های لایسیمتری، مدل پنمن مونتیت فائو به عنوان خروجی مدل‌ها در نظر گرفته شد (Kumar et al, 2011). داده‌ها به دو گروه ۷۵ درصد آموزش^۱ و اعتبار^۲ و ۲۵ درصد آزمایش^۳ تقسیم شدند. در این تحقیق از شبکه‌های پیش‌خور استفاده شده است. بهینه‌سازی لایه پنهان به روش سعی و خطا انجام گردید (Jain et al., 2008). نرون‌های مختلفی در این لایه مورد آزمون قرار گرفتند. افزایش نرون‌ها اثر چشمگیری بر کاهش خطای شبکه نداشت. از قاعده‌ی آموزش تنظیم بیز^۴ برای همگرایی شبکه استفاده گردید (Mackay, 1992). متداول‌ترین توابع فعال^۵ توابع سیگموئیدی و خطی می‌باشند (منهاج، ۱۳۹۳). در این تحقیق تابع لوگ سیگموئید^۶ در لایه پنهان و تابع خطی در لایه خروجی استفاده شده است. مراحل کدنویسی مدل هیبرید شبکه عصبی ANN-PSO در شکل (۳) نشان داده شده است. تمام این مراحل به کمک نرم‌افزار MATLAB(R2017b) انجام شده است. برای اجرای الگوریتم‌های داده‌کاوی داده‌های ورودی و خروجی وارد نرم‌افزار وکا^۷ (Weka3.9.2) گردید. ابزار وکا نرم‌افزاری است که توسط دانشگاه Waikato در کشور نیوزلند ارائه شده است. از ویژگی‌های بارز این نرم‌افزار می‌توان به پوشش تقریباً کامل الگوریتم‌های داده‌کاوی در آن اشاره کرد. مقادیر تبخیر و تعرق محاسبه شده به کمک مدل پنمن مونتیت فائو به عنوان خروجی الگوریتم و متغیرهای دمای حداقل و

- 1- Training
- 2- Validation
- 3- Testing
- 4- Bayesian-Regulation
- 5- Active function
- 6- logsig
- 7- Waikato Environment for knowledge Analysis (Weka)

ترین پارامترهایی که در درخت تصمیم باعث انشعاب می‌شوند کدام متغیرها هستند. از مزیت‌های دیگر این مدل در صورتی که داده‌ها روزانه، ماهانه و یا هر مقیاس دیگری گم‌شده باشند، می‌توان به کمک الگوریتم‌هایی امکان جایگزینی داده مفقودشده را فراهم کرد.

T_{max} مهم‌ترین پارامتر است و درخت را به دو شاخه $T_{max} \leq 21.597$ و $T_{max} > 21.597$ تقسیم می‌کند. بر اساس معیار انشعاب $T_{max} \leq 21.597$ دمای ماکزیمم بیش‌ترین اهمیت را دارد. به عبارت دیگر در شرایط مختلف داده‌های آب و هوایی، دما، رطوبت، ساعات آفتابی و سرعت باد چه رابطه‌ای حاکم است و مهم-



شکل ۳- مدل هیبرید شبکه عصبی ANN-PSO

جدول ۱- نتایج مدل شبکه عصبی مصنوعی و مدل ترکیبی شبکه عصبی ANN-PSO

داده	مدل	منطقه مطالعاتی	RMSE*	MAE**	R***
$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$	شبکه عصبی ANN	دلیجان	۰/۳۱۱۵	۰/۲۴۴۱	۰/۹۸۹۹
$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$	شبکه عصبی ANN	تفرش	۰/۲۹۱۵	۰/۲۳۵۵	۰/۹۹۹۱
$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$	شبکه عصبی ANN-PSO	دلیجان	۰/۲۴۴۰	۰/۱۷۰۰	۰/۹۹۹۲
$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$	شبکه عصبی ANN-PSO	تفرش	۰/۲۳۳۱	۰/۲۰۰۰	۰/۹۹۹۳

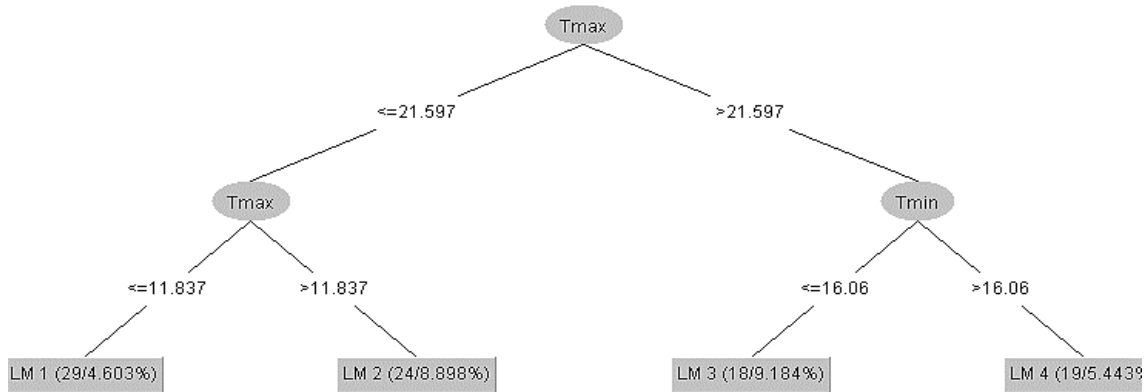
* جذر میانگین مربعات خطا، ** میانگین مطلق خطا و *** ضریب همبستگی

شبکه عصبی مصنوعی و مدل هیبرید شبکه عصبی ANN-PSO دقت کمتری دارد. برای مدل K star شاخص‌های آماری $RMSE=0.463$ ، $MAE=0.2831$ و $R=0.9896$ برای ایستگاه دلیجان و $RMSE=0.32$ ، $MAE=0.207$ و $R=0.995$ برای ایستگاه تفرش محاسبه گردید که نسبت به مدل درختی M5، مدل

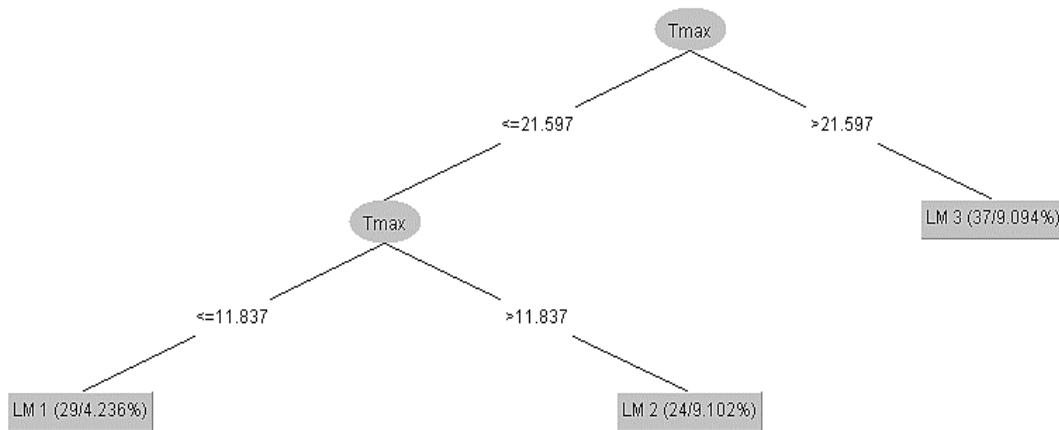
نتایج آنالیز آماری مدل‌های M5 Rules و K star در جدول (۲) ذکر گردیده است. برای مدل M5 Rules شاخص‌های آماری $RMSE=0.3385$ ، $MAE=0.28$ و $R=0.9951$ برای ایستگاه دلیجان و $RMSE=0.293$ ، $MAE=0.2258$ و $R=0.9965$ برای ایستگاه تفرش محاسبه گردید که نسبت به مدل درختی M5 و مدل

شبکه عصبی مصنوعی، مدل هیبرید شبکه عصبی ANN-PSO و مدل M5 Rules عملکرد ضعیف‌تری دارد. نتایج اجرای مدل Rep Tree نشان داد که این مدل با $MAE=0.4628$, $RMSE=0.6335$ و $R=0.982$ برای ایستگاه دلیجان و $MAE=0.618$ و $R=0.9697$ برای ایستگاه تفرش ضعیف‌ترین عملکرد را در بین مدل‌های مورد بررسی در این تحقیق دارد (جدول ۲).

شکل ۴- مدل درختی M5 برای ایستگاه دلیجان



شکل ۴- مدل درختی M5 برای ایستگاه دلیجان



شکل ۵- مدل درختی M5 برای ایستگاه تفرش

جدول ۲- نتایج الگوریتم‌های داده‌کاوی

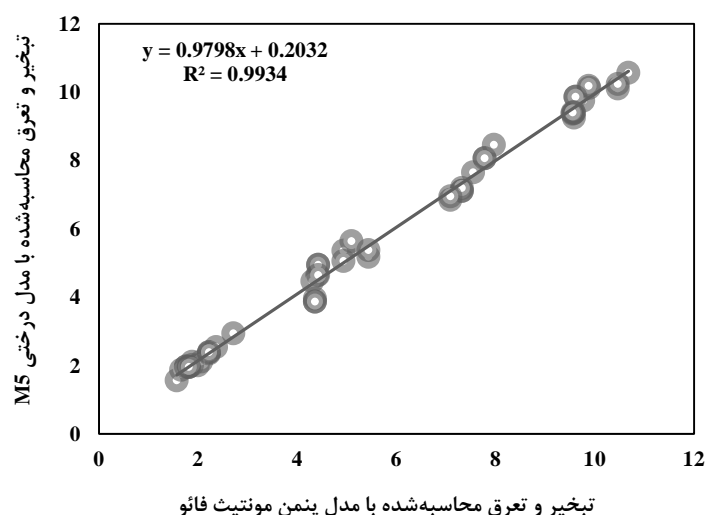
عملکرد مدل‌ها	R***	MAE**	RMSE*	منطقه مطالعاتی	الگوریتم	داده
بهترین عملکرد	۰/۹۹۶۷	۰/۲۳۹۸	۰/۲۷۹۳	دلیجان	M5	$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$
بهترین عملکرد	۰/۹۹۶۹	۰/۲۳۰۶	۰/۲۸۰۳	تفرش	M5	$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$
-	۰/۹۹۵۱	۰/۲۸۰۰	۰/۳۳۸۵	دلیجان	M5 Rules	$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$
-	۰/۹۹۶۵	۰/۲۲۵۸	۰/۲۹۳۰	تفرش	M5 Rules	$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$
-	۰/۹۸۹۶	۰/۲۹۳۱	۰/۴۶۳۰	دلیجان	K star	$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$
-	۰/۹۹۵۰	۰/۲۰۷۰	۰/۳۲۰۰	تفرش	K star	$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$
-	۰/۹۸۲۰	۰/۴۶۲۸	۰/۶۳۳۵	دلیجان	Rep Tree	$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$
-	۰/۹۶۹۷	۰/۶۱۸۰	۰/۸۳۳۷	تفرش	Rep Tree	$ET_0 = F(T_{max}, T_{min}, RH_{mean}, n, U_2)$

* جذر میانگین مربعات خطا، ** میانگین مطلق خطا و *** ضریب همبستگی

منطقه مختلف از کشور آمریکا پرداختند نتایج نشان داد که عملکرد مدل M5 نسبت به مدل شبکه عصبی دقیق تر بود (Kisi and Kilic, 2016).

اگرچه دقت مدل درختی M5 از مدل شبکه عصبی بیشتر و از مدل هیبرید شبکه عصبی ANN-PSO کمتر است ولی مدل درختی علاوه بر دقت مطلوب روابط ساده، خطی و قابل فهم ارائه می کند. لذا این تحقیق مدل درختی را برای برآورد تبخیر و تعرق در این منطقه توصیه می کند. پراکنش تبخیر و تعرق به کمک مدل درختی با روش پنمن مونتیث فائو در (شکل ۶ و ۷) برای هر دو ایستگاه نشان داده شده است. پیشنهاد می شود در تحقیقات آینده به ارزیابی مدل ها در اقلیم های مختلف با تعداد ایستگاه های بیشتر با الگوریتم های مختلف و جدید هوشمند و داده کاوی و مقایسه نتایج آن ها با هم و با ایستگاه های مجاور پرداخته شود. در صورت دسترسی به داده های لایسیمیتری معتبر برای ارزیابی صحت نتایج استفاده شود.

با توجه به عدم دسترسی به آمار و داده های مورد نیاز در برآورد تبخیر و تعرق و عدم اندازه گیری بعضی از این داده های مورد نیاز در بیشتر ایستگاه ها مدل شبکه عصبی، مدل هیبرید شبکه عصبی ANN-PSO و الگوریتم های داده کاوی راهکار مناسب تری نسبت به روش های تجربی در شرایط عدم وجود اطلاعات کافی می باشند. ستاری و همکاران (۱۳۹۲) پیش بینی تبخیر تعرق مرجع روزانه با استفاده از مدل شبکه عصبی مصنوعی و مدل درختی M5 پرداختند نتایج نشان داد که عملکرد مدل شبکه عصبی مصنوعی در تخمین تبخیر و تعرق دقیق تر است. قهرمان و صامتی به برآورد تبخیر و تعرق مرجع به کمک مدل درختی M5 و شبکه عصبی مصنوعی در ایستگاه های شیراز و کرمانشاه پرداختند. نتایج نشان داد که مدل درختی M5 عملکرد بهتری دارد (Ghahreman and Sameti, 2014). کیسی و کیلیچ به بررسی قابلیت تعمیم شبکه های عصبی مصنوعی و درخت مدل M5 در مدل سازی تبخیر و تعرق مرجع در دو



شکل ۶- پراکنش تبخیر و تعرق مدل درختی M5 با مدل پنمن مونتیث فائو در ایستگاه دلیجان

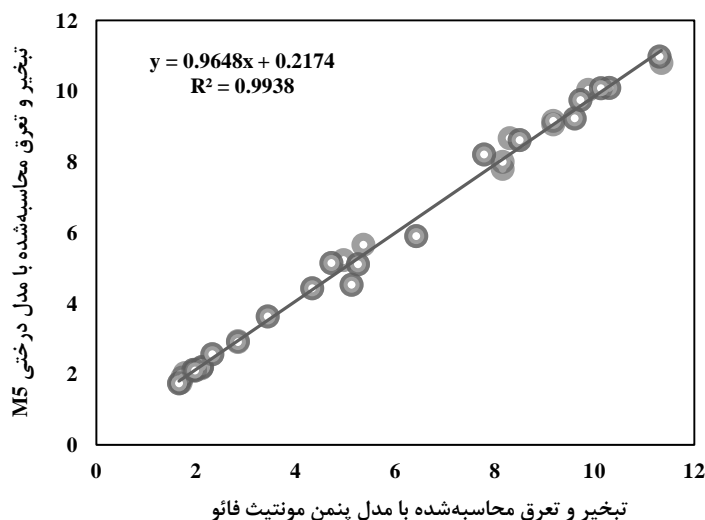
فائو پرداخته شد. ابتدا داده های هواشناسی دو ایستگاه سینوپتیک دلیجان و تفرش از سازمان هواشناسی اخذ گردید. مقادیر تبخیر و تعرق با روش پنمن مونتیث فائو که طبق تحقیقات انجام شده در نبود داده های لایسیمیتری معتبرترین روش محاسبه تبخیر و تعرق مرجع می باشد، محاسبه گردید و به عنوان مقادیر هدف در مدل های مورد بررسی استفاده شد. سپس با استفاده از نرم افزارهای مربوطه و کدنویسی مدل ها اجرا گردید. نتایج نشان داد که با استفاده از متغیرهای در دسترس دمای حداقل و حداکثر، رطوبت نسبی میانگین، ساعات آفتابی و سرعت باد در ارتفاع دو متری، مدل های هوشمند و داده کاوی دقت خوبی در برآورد تبخیر و تعرق مرجع را دارند. نتایج

نتیجه گیری

کاهش بارندگی و خشک سالی های متناوب اخیر، بخش کشاورزی را به عنوان پرمصرف ترین بخش استفاده کننده آب در کشور با چالش جدی مواجه ساخته است. یکی از پارامترهای کلیدی در طراحی سامانه های آبیاری (آبیاری سطحی و آبیاری نوین) تخمین دقیق نیاز آبی است که منجر به صرفه جویی در مصرف آب می گردد. لذا در این تحقیق به ارزیابی مدل درختی M5، M5rule، K star، Rep Tree، مدل شبکه عصبی مصنوعی و مدل هیبرید شبکه عصبی ANN-PSO در تخمین تبخیر و تعرق مرجع در دو نقطه اقلیمی مختلف در استان مرکزی، منطقه تفرش و دلیجان بر مبنای مدل پنمن مونتیث

ساده‌تر نتایج آن، ارائه روابط خطی و ساده، عدم نیاز به آزمون و خطا برای فرآیند آموزش و دقت بالای این مدل اشاره نمود. نتایج نهایی این پژوهش مدل درختی برای برآورد تبخیر و تعرق در این منطقه قابل توصیه می‌کند.

آنالیزهای آماری نشان داد که برای هر دو ایستگاه مدل درختی M5 علاوه بر ارائه روابط ساده، خطی و قابل فهم تر نسبت به سایر مدل‌های مورد بررسی در این تحقیق دقت و عملکرد بالاتری را دارد. از جمله ویژگی‌های بارز مدل درختی می‌توان به درک این مدل و تحلیل



شکل ۷- پراکنش تبخیر و تعرق مدل درختی M5 با مدل پنمن مونیتث فائو در ایستگاه تفرش

پژوهشی توسعه و ترویج علوم کشاورزی، منابع طبیعی و محیط‌زیست ایران، تهران. ۲۸ و ۲۹ شهریورماه.

پورغلام آمیجی، م.، خوش‌روش، م.، دیوبند هفشجانی، ل. و قدمی فیروزآبادی، ع. ۱۴۰۱. اثر آبیاری با پساب مغناطیسی تصفیه شده بر بهره‌وری آب ذرت. نشریه آبیاری و زهکشی ایران. ۱۶ (۱): ۲۴۳-۲۴۳.

حسینی وردنجانی، س. م. ر.، فتاحی، ر. و محترمی، ا. ۱۳۹۹. ارائه یک ضریب اصلاحی برای بهبود دقت روش هارگریوز در برآورد تبخیر و تعرق مرجع. نیوار. ۴۴ (۱۱۰-۱۱۱): ۴۳-۵۲.

حق‌زاده، ع.، یوسفی، ح.، یاراحمدی، ی. و ابراهیمیان، ط. ۱۳۹۸. مقایسه مدل هیبریدی PSO-ANFIS و مدل تجربی تورک در تخمین تبخیر و تعرق مرجع (مطالعه موردی: پل دختر- لرستان). اکوهیدرولوژی. ۶ (۳): ۶۸۵-۶۹۴.

ذرتی پور، ا.، نیسی، ل.، گلابی، م.، بزاز، ا. و ذرتی پور، ا. ۱۳۹۸. شبیه‌سازی و مقایسه‌ی تبخیر و تعرق پتانسیل به روش‌های شبکه عصبی مصنوعی، نروفازی و درخت تصمیم‌گیری M5 (مطالعه موردی: ایستگاه سینوپتیک شیراز). تحقیقات منابع آب ایران. ۱۵ (۱): ۳۶۵-۳۷۱.

سپهری، س.، عباسی، ف.، زارعی، ق. و نخجوانی مقدم، م. ۱۳۹۹.

منابع

احمدی، ف.، آیشم، س.، خلیلی، ک. و بهمنش، ج. ۱۳۹۴. کاربرد سیستم‌های استنتاج عصبی-فازی تطبیقی و برنامه‌ریزی ژنتیک برای برآورد تبخیر و تعرق ماهانه در شمال غرب ایران. پژوهش‌های آب در کشاورزی. ۲۹ (۲): ۲۳۵-۲۴۷.

بختیاری، ب.، محبی دهقانی، ع. و قادری، ک. ۱۳۹۴. برآورد تبخیر تعرق مرجع روزانه با حداقل داده‌های هواشناسی در اقلیم‌های نیمه‌خشک منتخب ایران. تحقیقات منابع آب ایران. ۱۱ (۳): ۱۳۱-۱۴۴.

بیدآبادی، م.، بابازاده، ح.، شیری، ج. و صارمی، ع. ۱۴۰۰. برآورد تبخیر و تعرق گیاه مرجع با استفاده از ANN و ANFIS در اقلیم نیمه‌خشک و خشک. آبیاری و زهکشی ایران. ۱۵ (۶): ۱۴۲۰-۱۴۱۲.

پناهی، س.، کرباسی، م. و نیکبخت، ج. ۱۳۹۵. پیش‌بینی تبخیر-تعرق مرجع با استفاده از شبکه‌های عصبی مصنوعی MLP, RBF و SVM. محیط‌زیست و مهندسی آب. ۲ (۱): ۶۳-۵۱.

پناهی، م. و میرهاشمی، س. ح. ۱۳۹۴. ارزیابی سه مدل نرم‌افزار weka در پیش‌بینی تبخیر تعرق پتانسیل و رطوبت نسبی ماهانه برای ماه بعد (مطالعه موردی استان یزد). اولین کنگره علمی

- Iranian Journal of Soil and Water Research. 51(8): 2051-2063.
- Antonopoulos, V. and Antonopoulos, A. 2017. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables, *Computers and Electronics in Agriculture*. 132: 86-96.
- Belouch, M., El Hadaj, S. and Idhammad, M. 2017. A two-stage classifier approach using reptree algorithm for network intrusion detection. *International Journal of Advanced Computer Science and Applications*. 8(6): 389-394.
- Binkert, N. L., Dreslinski, R. G., Hsu, L. R., Lim, K. T., Saidi, A. G. and Reinhardt, S. K. 2006. The M5 simulator: Modeling networked systems. *Ieee micro*. 26(4): 52-60.
- Eberhart, R. and Kennedy, J. 1995. A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the sixth international symposium on micro machine and human science* (pp. 39-43). Ieee.
- Eichelmann, E., Mantoani, M. C., Chamberlain, S. D., Hemes, K. S., Oikawa, P. Y., Szutu, D. and Baldocchi, D. D. 2022. A novel approach to partitioning evapotranspiration into evaporation and transpiration in flooded ecosystems. *Global Change Biology*. 28(3): 990-1007.
- Genolini, P. B. 2022. Wrapped M5-branes and complex saddle points. *Journal of High Energy Physics*. 2022(1): 1-45.
- Ghahreman, N. and Sameti, M. 2014. Comparison of M5 Model Tree and Artificial Neural Network for Estimating Potential Evapotranspiration in Semi-arid. *Desert*. 19(1): 75-81.
- Haykin, S. 1994. *Neural Networks: A Comprehensive Foundation*, MacMillan College Publishing Co. New York.
- Hou, Y., Hao, G., Zhang, Y., Gu, F., and Xu, W. 2022. A multi-objective discrete particle swarm optimization method for particle routing in distributed particle filters. *Knowledge-Based Systems*. 240: 108068.
- Jain, S. K., Nayak, P. C. and Sudheer, K. P. 2008. Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. *Hydrological Processes: An International Journal*. 22(13): 2225-2234.
- Khalilmoghadam, B., Afyuni, M., Abbaspour, K. C., Jalalian, A., Dehghani, A. A. and Schulin, R. 2009. Estimation of surface shear strength in Zagros region of Iran—a comparison of artificial neural networks and multiple-linear regression models. *Geoderma*. 153(1-2): 29-36.
- Kumar, M., Raghuwanshi, N. S. and Singh, R. 2011. Artificial neural networks approach in بررسی و تحلیل حساسیت روش‌های مبتنی بر شبکه‌های عصبی مصنوعی در برآورد تبخیر و تعرق گیاه مرجع. آبیاری و زهکشی ایران. ۱۴ (۶): ۲۰۹۹-۲۰۸۹.
- ستاری، م. ت. و اسماعیل‌زاده، ب. ۱۳۹۵. مقایسه نتایج مدل درختی M5 و برنامه‌ریزی ژنتیک با روش پنمن مونتیث فائو برای تخمین تبخیر و تعرق مرجع. مهندسی منابع آب. ۹ (۳۱): ۲۰-۱۱.
- ستاری، م. ت.، نهرین، ف. و عظیمی، و. ۱۳۹۲. پیش‌بینی تبخیر-تعرق مرجع روزانه با استفاده از مدل شبکه عصبی مصنوعی و مدل درختی M5 (مطالعه موردی: ایستگاه بناب). آبیاری و زهکشی ایران. ۷ (۱): ۱۱۳-۱۰۴.
- فضلی مقصودی، حسن و مومنی، حسین. ۱۳۹۲. مقایسه و بررسی الگوریتم‌های داده‌کاوی کاهل و ماشین بردار پشتیبان برای تشخیص نفوذ، همایش مهندسی کامپیوتر و توسعه پایدار با محوریت شبکه‌های کامپیوتری، مدل‌سازی و امنیت سیستم‌ها، مشهد.
- قربانی، م.، شکری، س. و برومندنسب، س. ۱۳۹۵. بررسی عملکرد شبکه‌های عصبی در برآورد تبخیر و تعرق گیاه مرجع (مطالعه موردی: ایستگاه سینوپتیک اهواز). اکوبیولوژی تالاب. ۸ (۲۸): ۳۴-۲۳.
- منهاج، م. ب. ۱۳۹۳. مبانی شبکه‌های عصبی و هوش محاسباتی. جلد اول، چاپ نهم، انتشارات دانشگاه صنعتی امیر کبیر، تهران، ۷۱۸ ص.
- نورانی، و. و سیاح فرد، م. ۱۳۹۲. آنالیز حساسیت داده‌های ورودی به شبکه عصبی مصنوعی به منظور برآورد مقدار تبخیر روزانه. آب و فاضلاب. ۲۴ (۳): ۱۰۰-۸۸.
- وروانی، ه.، مرادی، م. و وروانی، ج. ۱۳۹۲. برآورد تبخیر و تعرق ماهانه گیاه مرجع توسط مدل رگرسیون درختی در نواحی مختلف آب و هوایی ایران. پژوهش آب در کشاورزی. ۲۷ (۴): ۵۳۴-۵۲۳.
- Algetawee, H. and Alshama, G. 2021. Modeling of Evapotranspiration (ET_c) in a Medium Urban Park within a Megacity by Using Artificial Neural Network (ANN) Model. *Periodica Polytechnica Civil Engineering*. 65(4): 1260-1268.
- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. 1998. *Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56*. Fao, Rome. 300(9): D05109.
- Ansari Ghojghar, M., Pourgholam-Amiji, M., Bazrafshan, J., Liaghat, A. and Araghinejad, S. 2020. Performance Comparison of Statistical, Fuzzy and Perceptron Neural Network Models in Forecasting Dust Storms in Critical Regions in Iran.

- algorithm and machine learning model for energy use forecast in non-residential buildings. *Scientific Reports*. 12(1): 1-18.
- Pal, M. and Deswal, S. 2009. M5 model tree based modelling of reference evapotranspiration. *Hydrological Processes: An International Journal*. 23(10): 1437-1443.
- Petković, D., Gocic, M., Shamsirband, S., Qasem, S. N. and Trajkovic, S. 2016. Particle swarm optimization-based radial basis function network for estimation of reference evapotranspiration. *Theoretical and applied climatology*. 125(3): 555-563.
- Quinlan, J. R. 1992. Learning with continuous classes. In 5th Australian joint conference on artificial intelligence. 92: 343-348.
- Rahimikhoob, A. 2014. Comparison between M5 model tree and neural networks for estimating reference evapotranspiration in an arid environment. *Water resources management*. 28(3): 657-669.
- Sattari, M. T, Pal, M., Apaydin, H. and Ozturk, F. 2013. M5 model tree application in daily river flow forecasting in Sohu Stream, Turkey. *Water Resources*. 40(3). 233-242.
- Shi, Y. and Eberhart, R. 1998. A modified particle swarm optimizer. In 1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360). 69-73.
- Terzi, O. 2007. Data mining approach for estimation evaporation from free water surface. *Journal of Applied Sciences*. 7(4). 593-596.
- evapotranspiration modeling: a review. *Irrigation Science*. 29:11-25.
- Khoshravesh, M., Abedi-Koupai, J. and Nikzad-Tehrani, E. 2016. Detection of trends in hydroclimatological variables using parametric and non-parametric tests in Neka basin. *Journal of Water and Soil Science*. 19(74): 1-14.
- Kisi, O. and Kilic, Y. 2016. An investigation on generalization ability of artificial neural networks and M5 model tree in modeling reference evapotranspiration. *Theoretical and applied climatology*. 126(3): 413-425.
- Kisi, O., Keshtegar, B., Zounemat-Kermani, M., Heddami, S. and Trung, N. T. 2021. Modeling reference evapotranspiration using a novel regression-based method: radial basis M5 model tree. *Theoretical and Applied Climatology*. 145(1): 639-659.
- MacKay, D. J. 1992. Bayesian interpolation. *Neural computation*. 4(3): 415-447.
- Minasny, B. and McBratney, A. B. 2002b. Evaluation and development of hydraulic conductivity pedotransfer functions for Australian soil. *Soil Research*. 38(4): 905-926.
- Minasny, B. and McBratney, A. B. 2002a. The neuro-m method for fitting neural network parametric pedotransfer functions. *Soil Science Society of America Journal*. 66(2): 352-361.
- Mirzaee, S., Ghorbani-Dashtaki, S., Mohammadi, J., Asadzadeh, F. and Kerry, R. 2017. Modeling WEPP erodibility parameters in calcareous soils in northwest Iran. *Ecological Indicators*. 74: 302-310.
- Ngo, N. T., Truong, T. T. H., Truong, N. S., Pham, A. D., Huynh, N. T., Pham, T. M. and Pham, V. H. S. 2022. Proposing a hybrid metaheuristic optimization

Evaluation of Hybrid Model ANN-PSO and Different Data Mining Methods in Estimating Monthly Evapotranspiration in Two Different Climates

S.M.R. Hosseini-Vardanjani¹, M. Khoshravesh^{2*}, S.E. Hosseini-Kakolaki³, M. Pourgholam-Amiji⁴

Received: Jul.26, 2021

Accepted: Jul.05, 2021

Abstract

Evapotranspiration is one of the main components of any region's water balance. Its accurate estimation is very necessary for hydrological studies, designing irrigation and drainage systems, and planning irrigation systems. In this research, the M5 tree model, M5 Rules, K Star, Rep Tree, artificial neural network model, and ANN-PSO neural network hybrid model in the estimation of reference evapotranspiration in two different climatic regions in Markazi Province based on the FAO Penman-Monteith model were evaluated. The data used included minimum and maximum temperature, average relative humidity, and wind speed at a height of two meters and sunny hours from the synoptic stations of Delijan and Tafresh between 2004-2021. To evaluate the models, the Root Mean Square of Error (RMSE), Mean Absolute Error (MAE) and Correlation Coefficient (R) indices were used. The analysis of the results showed that for both stations, ANN-PSO neural network hybrid model had the best performance with RMSE=0.3115, MAE=0.2441 and R=0.9989 for the Delijan station and RMSE=0.2915, MAE=0.2355 and R=0.9989 for Tafresh station and tree model M5 with RMSE=0.2793, MAE=0.2398 and R=0.9967 for Delijan station and RMSE=0.2803, MAE=0.2306 and R=0.9969 for Tafresh station. Also, the Rep Tree model had the weakest performance among the models examined in this research. Considering that the tree model provides simple, linear and more comprehensible relationships for estimating reference evapotranspiration in addition to optimal accuracy, this model is recommended for estimating evapotranspiration in this region.

Keywords: Dry Regions, Evaporation and Transpiration, Machine Learning, Soft Computing.

1- Ph.D. Student, Department of Water Engineering, Faculty of Agricultural Engineering, Sari Agricultural Sciences and Natural Resources University, Sari, Iran

2- Associate Professor, Department of Water Engineering, Faculty of Agricultural Engineering, Sari Agricultural Sciences and Natural Resources University, Sari, Iran

3- Ph.D. Student, Department of Electrical Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

4- Ph.D. Candidate, Department of Irrigation and Reclamation Engineering, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

(*- Corresponding Author Email: Khoshravesh_m24@yahoo.com, m.khoshravesh@sanru.ac.ir)