

مقاله علمی-پژوهشی

پیش‌بینی تبخیر-تعرق مرجع روزانه با استفاده از روش ترکیبی هوشمند مصنوعی بر پایه الگوریتم پیش‌پردازش‌کننده تجزیه مد تجربی

امین امیر عشایری^۱، جواد بهمنش^{۲*}، وحید رضا وردی‌نژاد^۳، نسرين فتح‌اله زاده عطاری^۴

تاریخ دریافت: ۱۳۹۹/۲/۹ تاریخ پذیرش: ۱۳۹۹/۳/۹

چکیده

تبخیر-تعرق به عنوان یکی از اجزاء مهم چرخه هیدرولوژیک نقش بسیار مهمی در بررسی بیلان آبی حوضه‌های آبریز دارد. در محاسبه نیاز آبی گیاهان، ابتدا مقدار تبخیر-تعرق مرجع محاسبه و سپس با استفاده از آن، تبخیر-تعرق گیاهی محاسبه می‌شود. در این تحقیق جهت برآوردی دقیق از مقدار تبخیر-تعرق مرجع روزانه حوضه آبریز دریاچه ارومیه، ابتدا براساس روش استاندارد فائو-پنمن-مونتیث و داده‌های هواشناسی سه ایستگاه ارومیه، مهاباد و خوی، مقدار تبخیر-تعرق مرجع (ET_0) محاسبه شد. سپس با استفاده از آنالیز ضریب مالو، موثرترین پارامترها جهت استفاده در مدل‌های مورد استفاده مشخص گردید. در این تحقیق از مدل درخت MT که بر اساس الگوریتم استنتاجی کلاس‌بندی توسعه پیدا کرده است، استفاده گردید. برای مقابله با پیچیدگی و ناپایداری داده‌های سری زمانی تبخیر-تعرق از الگوریتم پیش‌پردازش‌کننده تجزیه مد تجربی (EMD) استفاده و نتایج به دست آمده با روابط تجربی تعیین تبخیر-تعرق مرجع شامل روش‌های Romanenko و Schendel مقایسه گردید. نتایج تحقیق نشان داد که عملکرد مدل درخت به شکل منفرد MT، مشابه و گاهی کمتر از عملکرد شبکه عصبی مصنوعی بود. با این حال ترکیب مدل درخت با تکنیک EMD باعث افزایش دقت مدل و کاهش خطا در شبیه‌سازی ET_0 روزانه در ایستگاه‌های ارومیه، مهاباد و خوی گردید. نتایج نشان داد که در مرحله آزمون ترکیب روش MT با EMD باعث ارتقا شاخص همبستگی به میزان ۱/۰۲٪، ۴/۳۹٪ و ۲/۰۴٪ به ترتیب برای ایستگاه‌های ارومیه، مهاباد و خوی گردید. همچنین میان روابط تجربی، رابطه Romanenko نسبت به رابطه تجربی Schendel دارای دقت بالاتری بوده و می‌توان رابطه Romanenko را برای مدل سازی تبخیر-تعرق مرجع برای منطقه مورد مطالعه توصیه نمود.

واژه‌های کلیدی: الگوریتم تجزیه مد تجربی، پیش‌بینی، تبخیر-تعرق مرجع، مدل درخت تصمیم

مقدمه

شدت وابسته به پارامترهای اقلیمی است. در برنامه‌ریزی و مدیریت منابع آب برای آبیاری نیاز به تجزیه و تحلیل متغیرهای هیدرولوژیک نظیر بارش و تبخیر-تعرق وجود دارد (حسن‌پور و همکاران، ۱۳۹۱). استفاده از اطلاعات گذشته در این متغیرها، پیش‌بینی تحولات آینده را که عامل کلیدی در برنامه‌ریزی، طراحی و مدیریت منابع آب است، ممکن می‌سازد. استفاده از مدل‌های ترکیبی (مدل‌های هوشمند مصنوعی با روش‌های پیش‌پردازش‌کننده) روشی پرکاربرد در پیش‌بینی مقادیر آینده تبخیر-تعرق مرجع (ET_0) بوده که می‌تواند موجب کاهش خطای پیش‌بینی گردند. کشاورزان و متصدیان سامانه‌های آبیاری می‌توانند، در مورد آب آبیاری موردنیاز بر اساس پیش‌بینی‌های بارش و آب مورد نیاز محصولات، تصمیم‌گیری‌ها را بر اساس زمان واقعی اعمال کنند. بارش معمولاً از طریق بررسی وضعیت آب و هوا پیش‌بینی شده درحالی که نیاز آبی گیاهان در پیش‌بینی‌ها انجام نمی‌پذیرد (Traore et al., 2017).

ایران کشوری با متوسط ریزش‌های جوی ۲۵۰ میلی‌متر بوده و در زمره مناطق خشک و نیمه خشک جهان طبقه‌بندی می‌گردد (Taheri et al., 2019). محدودیت منابع آب و نبود آب با کیفیت مناسب از عمده‌ترین مشکلات کشاورزی در کشور است؛ بنابراین یافتن راه‌کارهایی برای غلبه بر کمبود آب در این مناطق، همواره از اولویت‌های مطالعاتی و تحقیقاتی می‌باشد (سیفی و همکاران، ۱۳۸۹). تقاضای آب و مصرف آن در اکوسیستم‌های کشاورزی به

۱- دانشجوی کارشناسی ارشد آبیاری زهکشی، گروه مهندسی آب، دانشگاه ارومیه
۲- استاد گروه مهندسی آب، گروه مهندسی آب، دانشگاه ارومیه
۳- دانشیار گروه مهندسی آب، گروه مهندسی آب، دانشگاه ارومیه
۴- دکتری گروه مهندسی آب، گروه مهندسی آب، دانشگاه ارومیه
(* نویسنده مسئول: Email: j.behmanesh@urmia.ac.ir)

(2002). سودهیر و همکاران مقدار ET_0 روزانه گیاه مرجع را با استفاده از روش ANN و با بکارگیری داده‌های هواشناسی روزانه همچون حداقل و حداکثر دما، تشعشع خورشیدی، رطوبت نسبی و سرعت باد تخمین زدند. ایشان به این نتیجه رسیدند که روش شبکه عصبی مصنوعی ANN با دقت بالایی قادر به تخمین تبخیر- تعرق مرجع می‌باشد، آنها برای محاسبه تبخیر- تعرق از داده‌های ساده شده مانند دمای هوا، تابش خورشیدی و ساعات آفتابی استفاده کردند و نتایج حاصل از تحقیقشان را رضایت‌بخش عنوان کردند (Sudheer et al., 2003). در تحقیقی دیگر کیشی و اوزترک جهت تخمین مقدار تبخیر- تعرق به روش فائو- پنمن- مونیتث از تکنیک محاسباتی عصبی به منظور تخمین میزان تبخیر و تعرق در ایالت کالیفرنیا استفاده و نتایج حاصل را با مقادیر محاسباتی توسط معادلات تجربی مقایسه نمودند. حاصل تحقیق حاکی از دقت بالای مدل عصبی در تخمین تبخیر و تعرق بود (Kisi and Ozturk., 2007). همچنین رحیمی و همکاران با استفاده از داده‌های هواشناسی ایستگاه‌های منطقه خوزستان در ایران مقدار تبخیر- تعرق را با روش شبکه‌های عصبی مصنوعی و با دقت بالایی تخمین زدند (Rahimi Khoob., 2011). مقدم‌نیا و همکاران نیز برای مدل‌سازی تبخیر- تعرق روزانه منطقه چاهنیمه، واقع در استان سیستان و بلوچستان از ANN و سیستم استنتاج عصبی فازی استفاده نمودند و در پایان بیان داشتند که نتایج حاصل از ANN برای پیش‌بینی تبخیر- تعرق نسبت به مدل‌های تجربی و سیستم استنتاج عصبی فازی از دقت بالاتری برخوردارند (Mogadamnia et al., 2009). از آنجا که داده‌های هیدرولوژیک نظیر ET_0 به صورت کلی غیرخطی و غیرثابت (Non-stationary) هستند، به‌کارگیری مستقیم آنها برای کالیبره کردن مدل‌های هوشمند ممکن است دقت کافی جهت پیش‌بینی ET_0 نداشته باشد. از این‌رو، استفاده از تکنیک‌های پیش‌پردازش کننده عمدتاً باعث بهبود عملکرد مدل‌ها می‌شوند. محققین زیادی از این تکنیک‌های پیش‌پردازش کننده برای استخراج ویژگی‌های قرار داده شده در داده‌های سری زمانی مانند سطح آب زیرزمینی (Rezaie-Blf et al., 2017a)، تبخیر (Ghaemi et al., 2019)، شاخص تشعشع خورشید (Deo et al., 2017)، سرعت باد (Huang et al., 2019)، شاخص خشکسالی (Ali et al., 2019)، پارامترهای کیفیت آب (Fijani et al., 2019) و ارتفاع موج دریا (Ali and Prasad., 2019) استفاده کردند. این رویکردها به عنوان تکنیک‌های پیش‌پردازنده شامل PCA^o ، CWT^1 ، $MESA^y$ و MA^h

در زنجیره آب، خاک، گیاه، اتمسفر، آب مستقیم از سطح خاک و یا توسط گیاه به داخل اتمسفر وارد می‌شود. انتقال آب از سطح خاک به هوا را تبخیر و خارج شدن آن از گیاه را تعرق گویند. این دو پدیده ماهیت یکسانی داشته و چون تفکیک آنها از یکدیگر امکان پذیر نمی‌باشد، مجموعاً به نام تبخیر- تعرق در نظر گرفته شده و به اختصار به صورت ET^1 نشان داده می‌شود. از آنجا که بدست آوردن تبخیر- تعرق برای هر نوع زراعت یا پوشش گیاهی امری بسیار مشکل است در عمل، ابتدا پارامتری به نام تبخیر- تعرق مرجع (ET_0) محاسبه شده سپس به کمک آن می‌توان تبخیر- تعرق گیاه مورد نظر را بدست آورد (Allen et al., 1998). با توجه به کمبود آب و لزوم استفاده بهینه از آب، تخمین دقیق تبخیر- تعرق مرجع و به تبع آن تبخیر- تعرق واقعی می‌تواند به تصمیم‌گیری مناسب برنامه‌ریزان و مدیران بخش آب در اتخاذ تصمیم‌های درست کمک شایانی نماید (پهمنش و همکاران، ۱۳۹۳).

مقدار ET_0 را می‌توان به صورت مستقیم با استفاده از لایسی متر و یا از روابط تجربی ارائه شده همچون پنمن- مونیتث محاسبه نمود. نسخه اصلاح شده فائو- پنمن- مونیتث هم اینک مبنای نرم‌افزارهایی همچون CropWat بوده که در محاسبات تبخیر- تعرق مورد استفاده قرار می‌گیرد. گرچه لایسی متر تخمین نسبتاً دقیق از تبخیر- تعرق واقعی را به دست می‌دهد ولی با این حال در عمل کارگذاری آن هزینه‌های زیادی به دنبال داشته و صرفه اقتصادی آن کم است. بر این اساس، محققین بسیاری در سراسر دنیا روش فائو- پنمن- مونیتث را به عنوان یک روش استاندارد برای برآورد تبخیر- تعرق مرجع معرفی کرده‌اند که توسط آن و با استفاده از ضریب گیاهی می‌توان تبخیر- تعرق واقعی را برآورد نمود (Chen et al., 2018).

در سال‌های اخیر، در صورت وجود داده‌های دارای کمیت و کیفیت لازم می‌توان از روش‌های هوش مصنوعی نظیر شبکه عصبی مصنوعی ANN^2 ، ماشین‌های بردار پشتیبان SVM^3 ، برنامه‌ریزی بیان ژن GEP^4 برای مدل‌سازی و حل مسائل در مهندسی آب استفاده کرد که به علت مزایای قابل تأمل آن شامل، فراگیری آسان مدل‌ها، شبیه‌سازی پدیده‌ها با کمترین تعداد پارامترها در مقایسه با مدل‌های عددی، مدل‌سازی پدیده‌های پیچیده‌ی غیر خطی فراگیر شده است (Rezaie-Balf et al., 2019). اساساً، تبخیر - تعرق یک فرآیند پیچیده و غیرخطی است و از این رو روش‌های هوشمند بعنوان یک روش جدید که از قابلیت زیادی در مدل‌سازی فرآیندهای پیچیده و غیرخطی برخوردار است، می‌تواند جهت مدل‌سازی تبخیر- تعرق بکار گرفته شود (Kumar et al.,

5- Principal Component Analysis
6- Continuous Wavelet Transform
7- Maximum Entropy Spectral Analysis
8- Moving Average

1- Evapotranspiration
2- Artificial Neural Network
3- Support Vector Machine
4- Gene Expression Programming

گردید. نکته قابل توجه این است که نواقص آماری موجود در داده‌های مربوط، با استفاده از روش‌های همبستگی نظیر میانگین‌گیری تکمیل شد. در جداول ۱ و ۲ مشخصات جغرافیایی و محدوده پارامترهای هیدرولوژیک مربوط به هر یک از ایستگاه‌ها را نشان می‌دهند.

روش فائو- پنمن - مونتیث

محققین بسیاری در سراسر دنیا روش فائو- پنمن- مونتیث را که به صورت تئوری بر اساس توازن انرژی روی یک سطح مرطوب پوشیده از گیاه ارائه گردیده، به عنوان یک روش استاندارد برای برآورد تبخیر- تعرق گیاه مرجع که در حال حاضر اساس محاسبات نیاز آبی می‌باشد، معرفی نموده‌اند. آنچه در معادله فائو- پنمن- مونتیث محاسبه می‌شود، نیاز آبی گیاه مرجع (ET_0) می‌باشد. رابطه مذکور به صورت رابطه (۱) می‌باشد:

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

که در آن ET_0 تبخیر- تعرق گیاه مرجع (mm day^{-1}); R_n تابش خالص در سطح محصول ($\text{Mj m}^{-2} \text{day}^{-1}$); G چگالی شار حرارتی خاک ($\text{Mj m}^{-2} \text{day}^{-1}$); T متوسط درجه حرارت هوای روزانه در ارتفاع ۲ متری ($^{\circ}\text{C}$); U_2 سرعت باد در ارتفاع ۲ متری (m s^{-1}); e_s فشار بخار اشباع (kPa); e_a فشار بخار واقعی (kPa); $e_s - e_a$ کمبود فشار بخار اشباع (kPa); Δ شیب منحنی فشار بخار ($\text{kPa } ^{\circ}\text{C}^{-1}$) و γ ثابت سایکرومتریکی ($\text{kPa } ^{\circ}\text{C}^{-1}$) می‌باشد.

الگوریتم تجزیه مد تجربی (EMD)

تجزیه مد تجربی (EMD) یک روش تجزیه اطلاعات برای استخراج سیگنال‌ها در اطلاعات تولید شده توسط عملیات غیرخطی و غیرایستا بوده که شامل جداسازی محلی و کامل برگرفته از اطلاعات از یک سیگنال در نوسان‌های سریع و آهسته است. EMD یک سری زمانی را به تعداد محدودی از حالت‌های نوسانی به نام توابع مد ذاتی^۲ (IMF) بسط می‌دهد. روش تجزیه مد تجربی توسط نوردنای - هوانگ و همکاران در سال ۱۹۹۸، هنگامی که هوانگ مشغول کار در ناسا بود، ابداع شد (Huang et al., 1998). ابداع روش EMD، در نتیجه مطالعه تبدیل هیلبرت است، این تبدیل به افتخار دیوید هیلبرت تبدیل هیلبرت نامیده شده است. هیلبرت برای اولین بار از این تبدیل برای حل حالت خاصی از مسئله ریمن-هیلبرت استفاده کرد که در پردازش سیگنال از تبدیل هیلبرت برای یافتن سیگنال تحلیلی یک سیگنال استفاده می‌شود. نتایج تجزیه EMD می‌تواند داده‌های مورد

می‌باشند. روش تجزیه مد تجربی^۱ (EMD) به عنوان یک روش تجزیه داده‌ها مبتنی بر نويز توسط هانگ و همکاران ارائه شده که سری ورودی فرکانس بالاتر را به مولفه‌های (سیگنال) حل شده با فرکانس پایین‌تر تفکیک می‌کند (Huang et al., 1998).

هدف از انجام این تحقیق، ارائه مدل پیش‌بینی دقیق، جهت تخمین تبخیر- تعرق مرجع روزانه در سه ایستگاه هواشناسی سینوپتیک واقع در حوضه آبریز دریاچه ارومیه (ارومیه، مهاباد، خوی) با استفاده از مدل‌های هیبریدی می‌باشد. در این تحقیق، الگوریتم پیش‌پردازش‌کننده EMD با روش مدل درخت ترکیب شده تا بتوان از طریق این روش ترکیبی، مدلی دقیق‌تر جهت ET_0 روزانه در سه ایستگاه سینوپتیک واقع در این حوضه شامل ارومیه، خوی و مهاباد ارائه نمود. همچنین از دیگر نوآوری‌های این تحقیق، استفاده از ضریب مالو برای شناسایی بهترین پارامترهای ورودی به مدل‌های هوشمند مصنوعی بوده بطوریکه ضریب مذکور توانایی این را دارد که بهترین و موثرترین پارامترها را از بین تمامی پارامترهای موجود انتخاب و ارائه کند.

مواد و روش‌ها

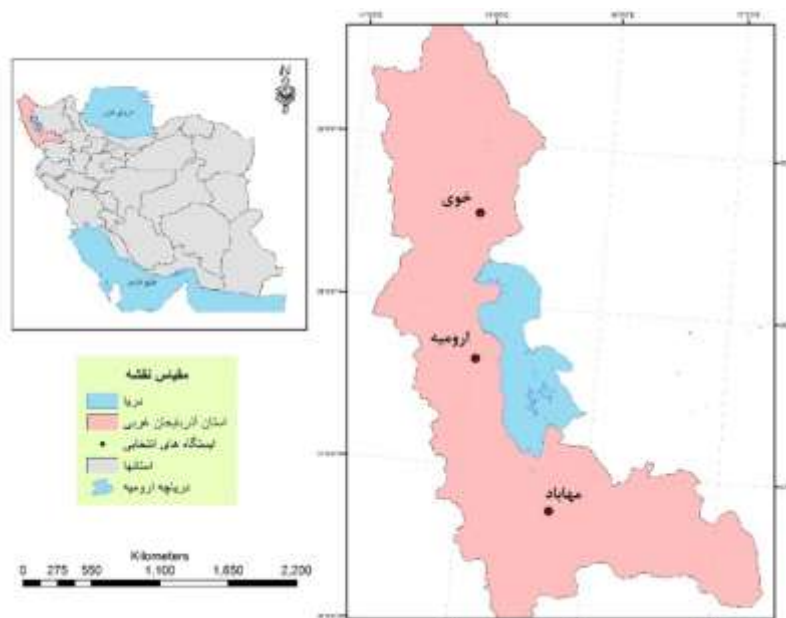
منطقه‌ی مورد مطالعه

دریاچه ارومیه در شمال غرب ایران بیستمین دریاچه جهان از لحاظ وسعت محسوب می‌گردد. وسعت حوضه آبریز دریاچه ۵۲۳۵۵ کیلومتر مربع است که از این مقدار حدود ۵۸۲۲ کیلومتر مربع وسعت خود دریاچه مستقیماً با ارتفاع آب دریاچه ارتباط دارد. تراز کف دریاچه از سطح آزاد دریاها ۱۲۶۸ متر است. متوسط درجه حرارت سالیانه آن ۱۱ درجه سانتیگراد و میانگین سالانه بارندگی در این منطقه بین ۱۸۰ تا ۵۰۰ میلی‌متر متغیر است (گودرزی و همکاران، ۱۳۹۴). رودخانه‌های عمده حوضه، شامل نازلوچای، باراندوزچای، آجی‌چای، زرینه رود و سیمینه رود، مهابادچای و زولاچای و مجموعه‌ای دیگر از رودخانه‌های بخش غربی دریاچه ارومیه می‌باشد (Kakahaji et al., 2013). در این تحقیق، سه ایستگاه سینوپتیک ارومیه، مهاباد و خوی که دارای داده‌های هواشناسی غنی در منطقه می‌باشند، جهت تخمین و مدل‌سازی تبخیر- تعرق مرجع روزانه استفاده گردید. شکل ۱ موقعیت حوضه‌ی دریاچه ارومیه و ایستگاه‌های منتخب برای پیش‌بینی ET_0 روزانه را نشان می‌دهد.

با توجه به اهداف تحقیق، آمار و اطلاعات مورد نیاز متغیرهای سری زمانی در دوره‌ی آماری ۹۷-۱۳۷۸ شامل میانگین دما (T_{mean}), رطوبت نسبی میانگین (RH_{mean}), سرعت باد در ارتفاع ۲ متری (U_2), میانگین فشار بخار (VP_{mean}) از ایستگاه‌های ذکر شده جمع‌آوری و ET_0 متناظر با این پارامترها در دوره‌ی مذکور محاسبه

اساسی هر مجموعه داده پیچیده است، ارائه نمود. آن‌ها سپس روش تازه‌ای به نام تجزیه مد تجربی را ایجاد نموده تا یک سیگنال را به مجموعه‌ای از توابع مد ذاتی تجزیه نماید.

نیاز برای تکمیل تبدیل هیلبرت را فراهم نماید. در موارد زیادی، EMD با تبدیل هیلبرت ترکیب شده و به نام تبدیل هیلبرت- هوانگ (HHT) شناخته می‌شود. پس از مطالعه مفهوم فرکانس آنی در تبدیل هیلبرت، هوانگ مفهوم تازه‌ای از تابع مد ذاتی (IMF) را که مؤلفه



شکل ۱- موقعیت جغرافیایی ایستگاه‌های سینوپتیک مورد مطالعه

جدول ۱- مشخصات جغرافیایی و محدوده پارامترهای هیدرولوژیک ایستگاه‌ها در دوره مورد مطالعه

ایستگاه	موقعیت جغرافیایی		ارتفاع از سطح دریا (m)	بازه تغییرات میانگین دمای هوا (°C)	میانگین رطوبت نسبی	سرعت باد در ارتفاع دو متری (m/s)	میانگین فشار بخار (KPa)	تبخیر- تعرق مرجع (mm/day)
	عرض شمالی	طول شرقی						
ارومیه	۳۷° ۶۵'	۴۵° ۰۵'	۱۳۲۸	-۱۲/۹ - ۳۰	۱۵/۹۹-۳۷	۴۰-۰	-۱۹/۵۸ ۱/۳۷	۰/۲۷ - ۱۴/۴۳
مهاباد	۳۶° ۷۵'	۴۵° ۷۱'	۱۳۵۱	-۱۳/۵ - ۳۴	۱۰-۹۷	۳۷-۰	-۱۶/۷۶ ۱/۱۴	۰/۱۸ - ۱۷/۴۵
خوی	۳۸° ۵۵'	۴۴° ۹۹'	۱۱۰۳	-۱۸/۲ - ۳۳	۱۵/۹۸-۲۵	۳۰-۰	-۱۹/۵۶ ۱/۲۲	۰/۱۳ - ۱۵/۹۷

مقیاس شده یک تابع با طول متناهی و نوسانی شدیداً میرا هستند. الگوریتم EMD تلاش می‌کند تقریباً هر سیگنالی را به مجموعه‌ای متناهی از توابع تجزیه نماید و تبدیل هیلبرت، مقادیر فرکانس لحظه‌ای فیزیکی را می‌دهد. این توابع به نام توابع مد ذاتی نامیده می‌شوند. الگوریتم از یک روند غربال تکرارشونده استفاده می‌کند که با موفقیت می‌تواند میانگین محلی را از یک سیگنال کم نماید. مراحل روند غربال‌گری در ادامه بیان می‌شوند (Cheng et al., 2019):

۱. اکستریمم محلی سیگنال (ماکزیمم، مینیمم) تعیین می‌شود.

برخلاف روش‌های تجزیه سنتی، مانند تبدیل فوریه و موجک، این روش جدید داده- وفقی است. مزیت اصلی EMD نسبت به روش‌های دیگر، امکان تجزیه خودکار به صورت کاملاً داده وفقی است و نیاز به یک مجموعه توابع مفروض، همانند تبدیل موجک ندارد (Huang et al., 1998). موجک (Wavelet) دسته‌ای از توابع ریاضی هستند که برای تجزیه سیگنال پیوسته به مولفه‌های فرکانسی آن بکار می‌رود که رزولوشن هر مولفه برابر با مقیاس آن است. تبدیل موجک تجزیه یک تابع بر مبنای توابع موجک می‌باشد. موجک‌ها که بعنوان موجک‌های دختر شناخته می‌شوند، نمونه‌های انتقال یافته و

دسته‌هایی به نام لایه مرتب می‌شوند. ساختار معمولی شبکه عصبی از سه لایه تشکیل شده است. لایه ورودی، داده‌ها را در شبکه توزیع می‌کند، لایه پنهان داده‌ها را پردازش می‌کند و لایه خروجی نتایج را به ازای ورودی‌های مشخص استخراج می‌کند (Moghaddamnia et al., 2009). در این تحقیق، تعداد لایه‌های پنهان و تعداد نرون‌ها در لایه پنهان، بر مبنای مقایسه عملکرد شبکه‌ها انتخاب شد که برای تشکیل شبکه از تابع آموزش Levenberg-Marquardt و تابع انتقال purelin و logsigmoid استفاده گردید. بهترین معیار برای مقایسه عملکرد شبکه‌های ایجاد شده برای مرحله ارزیابی آموزش، ریشه میانگین مربعات خطا^۱ می‌باشد، که مدل برازش شده خطی بین داده‌های مشاهده شده و داده‌های پیش‌بینی شده انجام گردید (Kumar et al., 2002).

مدل درخت تصمیم

الگوریتم MT رایج‌ترین طبقه‌بندی استفاده شده در خانواده مدل تصمیم‌گیری درختی است که توسط کوینلان در سال ۱۹۹۲ برای پیش‌بینی داده‌های پیوسته ارائه شد (Quinlan., 1992). مدل درختی MT یک الگوریتم پیش‌بینی عددی است و گره‌های درخت با خواص بیشینه خطاهای مورد انتظاری که به عنوان تابعی از انحراف استاندارد پارامترهای خروجی می‌باشد، انتخاب می‌شوند. مدل MT یک مدل درختی برای پیش‌بینی صفات عددی پیوسته است که در آن، توابع رگرسیونی خطی در برگ‌های این درخت ظاهر می‌گردند. این مدل، بر خلاف مدل‌های درخت تصمیم معمول که کلاس یا رده‌های گسسته را به عنوان خروجی ارائه می‌نمایند، یک مدل خطی چند متغیره را برای داده‌ها در هر گره از مدل درختی می‌سازد. ساختار درخت تصمیم‌گیری شبیه یک درخت بوده که از ۴ بخش ریشه، شاخه، گره‌ها و برگ‌ها تشکیل شده است. گره اول در درخت تصمیم به عنوان ریشه درخت در نظر گرفته می‌شود. هر گره مربوط به یک خصوصیت معین می‌باشد و شاخه‌ها به معنای بازه‌ای از مقادیر هستند. این بازه‌های مقادیر باید بخش‌های مختلف مجموعه مقادیر معلوم را برای خصوصیت‌ها به دست دهند.

تشکیل ساختار مدل‌های درخت تصمیم‌گیری شامل مراحل ایجاد درخت و هرس کردن آن است که در مرحله تشکیل درخت، از یک الگوریتم استنتاجی یا معیار تقسیم (انشعاب) برای تولید یک درخت تصمیم استفاده می‌شود (Rezaie-Balf et al., 2017b). معیار تقسیم برای الگوریتم مدل MT، انحراف معیار مقادیر هر کلاسی است که به عنوان کمیتی از خطا به یک گره می‌رسد و کاهش مورد انتظار در این خطا را به عنوان نتیجه آزمون هر صفت در آن گره

۲. مقدار ماکزیمم به یک تابع درون‌یابی متصل شده و پوش بالایی سیگنال را می‌سازد.
 ۳. مقدار مینییمم به یک تابع درون‌یابی متصل شده و پوش پایینی سیگنال را می‌سازد.
 ۴. میانگین محلی به عنوان نصف اختلاف بین پوش بالایی و پایینی محاسبه می‌شود.
 ۵. میانگین محلی از سیگنال کم می‌شود.
 ۶. این روند بر باقیمانده تکرار می‌شود.
- روند غربال‌گری تا زمانی که سیگنال به تعریف یک IMF برسد، تکرار می‌شود. سپس IMF، از سیگنال اصلی، کسر شده و روند غربال بر باقیمانده تکرار می‌شود. این کار تا زمانی که باقیمانده، یک تابع یکنواخت شود، ادامه می‌یابد. آخرین IMF استخراج شده، کم‌ترین مؤلفه فرکانسی است. با اعمال EMD، یک سیگنال می‌تواند به مجموعه‌ای از توابع تک مؤلفه با نام توابع مد ذاتی (IMF_s) تجزیه شود یک تابع تک مؤلفه، تابعی نوسانی نزدیک به رایج‌ترین و ابتدایی‌ترین تابع هارمونیک را نشان می‌دهد. (Al-Musaylh et al., 2018).

ایده اصلی در پس زمینه IMF جداسازی اطلاعات به بخش میانگین محلی با تغییرات آرام و بخش نوسانی متقارن با تغییرات سریع، (بخش آخر IMF)، و میانگین محلی که به عنوان باقیمانده تعریف می‌شود. این باقیمانده به عنوان ورودی برای تجزیه بیشتر، با روندی تا آنجا که هیچ نوسان دیگری نباشد تکرار می‌شود، بکار می‌رود. به عنوان یک حالت نوسانی، دامنه و فرکانس یک IMF با زمان به طریقی تغییر خواهد کرد که بطور محلی متقارن و تک مؤلفه باشد. بنابراین، فرکانس لحظه‌ای می‌تواند به طریقی معنادار محاسبه شود (Zhou et al., 2019).

از این رو، IMF‌ها فرکانس‌هایی در محدوده بالاترین تا پایین‌ترین سیگنال را شامل می‌شوند که به عنوان سیگنال‌های مدوله شده دامنه و فرکانس (AM-FM) موجود هستند، که AM پوش را حمل می‌کند و FM تغییر دامنه ثابت در فرکانس است و با استفاده از یک روش غربالی محاسبه می‌شود. برای انجام آن، یک IMF باید دو شرط را ارضا کند (Huang et al., 1998):

تعداد اکستریم Ne (ماکزیمم و مینییمم محلی) و تعداد عبور از صفرها Nz باید برابر باشند یا حداکثر یکی اختلاف داشته باشند.

$$(N_z - 1) \leq N_e \leq (N_z + 1) \quad (2)$$

در هر نقطه، مقدار میانگین پوشی که به وسیله ماکزیمم محلی و پوشی که به وسیله مینییمم محلی به وجود می‌آید، صفر است.

شبکه عصبی مصنوعی

ساختار شبکه عصبی مصنوعی، به گونه‌ای است که نرون‌ها در

1- Root Mean Square Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^M (P_i - O_i)^2}{M}} \quad (8)$$

$$MAE = \frac{\sum_{i=1}^M |P_i - O_i|}{M} \quad (9)$$

که O_i مقادیرهای مشاهده‌ای تبخیر-تعرق مرجع، P_i مقادیرهای پیش‌بینی شده تبخیر-تعرق مرجع با هر کدام از مدل‌ها، \bar{O} میانگین داده‌های مشاهده‌ای، \bar{P} میانگین داده‌های پیش‌بینی و M تعداد کل داده‌ها می‌باشد. مقدار اندازه‌گیری شده R میزان همبستگی بین دو متغیر وابسته را نشان می‌دهد که مقدار آن بین $+1$ و -1 می‌باشد. اگر مقدار بدست آمده مثبت باشد به معنی این است که تغییرات دو متغیر به طور هم جهت اتفاق می‌افتد، یعنی با افزایش در هر متغیر، متغیر دیگر نیز افزایش می‌یابد و برعکس اگر مقدار R منفی شد یعنی اینکه دو متغیر در جهت عکس هم عمل می‌کنند، یعنی با افزایش مقدار یک متغیر مقادیر متغیر دیگر کاهش می‌یابد و برعکس. اگر مقدار بدست آمده صفر شد، نشان می‌دهد که هیچ رابطه‌ای بین دو متغیر وجود ندارد و اگر $+1$ شد همبستگی مثبت کامل و اگر -1 شد همبستگی کامل و منفی است. R^2 نشانگر ضریب تبیین بوده و مجذور ضریب R می‌باشد. که هر چه این مقدار به 1 نزدیکتر، نشان دهنده مدل صحیح‌تر و منطبق‌تر می‌باشد. همچنین، $RMSE$ برای اندازه‌گیری دقت پیش‌بینی مدل استفاده می‌شود که با جذر گرفتن از خطاها، یک مقدار مثبت را نشان می‌دهد. محدوده $RMSE$ از صفر تا یک عدد مثبت بزرگ ادامه می‌یابد که واحد آن از جنس متغیر خروجی مشاهداتی می‌باشد. مقادیر بالا در پارامتر R^2 و کم در $RMSE$ نشان‌دهنده عملکرد بالای مدل می‌باشد. همین‌طور، MAE مقدار میانگین مثبت خطاها در یک مجموعه از مقادیر پیش‌بینی شده را محاسبه می‌کند و واحد آن از جنس پارامتر خروجی مشاهداتی می‌باشد. مقدار این شاخص آماری معمولاً مشابه مقدار $RMSE$ می‌باشد.

نتایج و بحث

داده‌های ورودی به مدل‌ها

انتخاب متغیرهای ورودی مناسب برای مدل‌های هوشمند مصنوعی و روش‌های داده‌کاوی در عملکرد هر یک از مدل‌ها بسیار اثرگذار می‌باشد. از این رو، بر اساس مشاهدات تجربی و مطالعات سایر محققان از جمله متغیرهای ورودی موثر در پیش‌بینی تبخیر-تعرق مرجع در یک حوضه آبریز می‌توان از پارامترهای هیدرولوژیک نظیر دما، تبخیر، رطوبت، تابش خورشید و وزش باد که اثرگذاری قابل توجهی بر خروجی مدل‌ها دارند، استفاده نمود. از میان روش‌های انتخاب ویژگی، ضریب مالو^۴ در انتخاب تعداد کمتر از متغیرهای پیش‌بینی کننده برای کالیبره کردن مدل برتر،

محاسبه می‌نماید. کاهش معیار (SDR) از رابطه (۳) محاسبه می‌شود (Quinlan., 1992):

$$SDR = sd(T) - \sum_i \frac{T_i}{|T|} \times sd(T_i) \quad (3)$$

$$sd(T) = \sqrt{\frac{1}{N} \left(\sum_{i=1}^N y_i^2 - \frac{1}{N} \left(\sum_{i=1}^N y_i \right)^2 \right)} \quad (4)$$

که در آن T بیانگر مجموعه‌ای از نمونه‌هاست که به گره می‌رسند، Sd بیانگر انحراف معیار، y_i مقدار عددی ویژگی هدف نمونه i ، N شماره داده‌ها و T_i بیانگر زیرمجموعه‌ای از نمونه‌هاست که i امین نتیجه تست پتانسیلی را دارند، می‌باشد. به دلیل فرآیند انشعاب، داده‌های قرار گرفته در گره‌های فرزند، انحراف معیار کمتری نسبت به گره مادر داشته و بنابراین خالص‌تر هستند. پس از حداکثرسازی تمامی انشعاب‌های ممکن، MT صفتی را انتخاب می‌کند که کاهش مورد انتظار را بیشینه نماید. این تقسیم بیشتر ساختار شبه درختی بزرگی را تشکیل می‌دهد که باعث بیش‌برازش می‌گردد. برای غلبه بر مساله بیش‌برازش، درخت تشکیل شده باید هرس شود. این کار با جایگزینی یک درخت فرعی با یک برگ انجام می‌شود. بنابراین، مرحله دوم در طراحی مدل درختی شامل هرس نمودن درخت رشد یافته و جایگزینی درختان فرعی با توابع رگرسیونی خطی است (Solomatine and Dulal., 2003).

در این تحقیق، علاوه بر مدل‌سازی ET_0 ، دو رابطه تجربی (Romanenko (Romanenko., 1961) و Schendel., (Schendel., 1967) برای مقایسه با هر یک از روش‌های پیشنهادی محاسبه و در نظر گرفته شد، که روابط آنها به ترتیب به شرح روابط (۵) و (۶) می‌باشد:

$$ET_0(\text{Romanenko}) = 0.00006(25 + T_{\text{mean}})^2(100 - RH) \quad (5)$$

$$ET_0(\text{RSchendel}) = 16 \frac{T_{\text{mean}}}{RH} \quad (6)$$

که در روابط (۵) و (۶)، T_{mean} و RH به ترتیب متوسط دما بر حسب درجه سلسیوس و درصد رطوبت می‌باشند.

معیار ارزیابی عملکرد مدل‌ها

برای مقایسه عملکرد مدل‌های ارائه شده و روابط تجربی جهت تخمین تبخیر-تعرق مرجع در مراحل آموزش و تست، از سه شاخص آماری شامل ضریب تبیین R^2 (ریشه میانگین مربعات خطا $(RMSE)^2$ و ریشه مربعات خطا^۳ (MAE) استفاده گردید که به صورت زیر و بر طبق معادلات (۷)، (۸) و (۹) محاسبه می‌گردند.

$$R^2 = \left[\frac{\sum_{i=1}^M (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^M (O_i - \bar{O})^2 \sum_{i=1}^M (P_i - \bar{P})^2}} \right]^2 \quad (7)$$

- 1- Standard Deviation Reduction
- 2- Root Mean Square Error
- 3- Mean Absolute Error

4- Mallows's coefficient

در جدول ۲، بهترین پارامترهای ورودی به مدل، پارامترهای میانگین دما، رطوبت نسبی میانگین، سرعت باد در ارتفاع ۲ متری و میانگین فشار بخار بر اساس کمترین مقدار C_p (۸/۱۴) انتخاب شدند. انحراف از معیار (S) و ضریب تبیین (R^2) نیز برای این سناریو به ترتیب ۰/۴۳ و ۰/۸۸ بدست آمد که نسبت به دیگر سناریوها عملکرد بهتری داشته است.

عملکرد موفق داشته است. در این تحقیق، از روش ضریب مالو برای انتخاب ویژگی و کاهش ابعاد پارامترهای ورودی اثرگذار بر خروجی مسئله یعنی پیش‌بینی تبخیر-تعرق مرجع روزانه استفاده گردید. مقدار ضریب مالو کم، نشان‌دهنده یک انتخاب خوب از زیرمجموعه‌های پیش‌بینی کننده و همچنین یک مدل پیش‌بینی کننده قوی است. بر این اساس، نتایج بدست آمده برای انتخاب بهترین پارامترهای ورودی به مدل‌های پیشنهادی، در جدول ۲ آورده شده است.

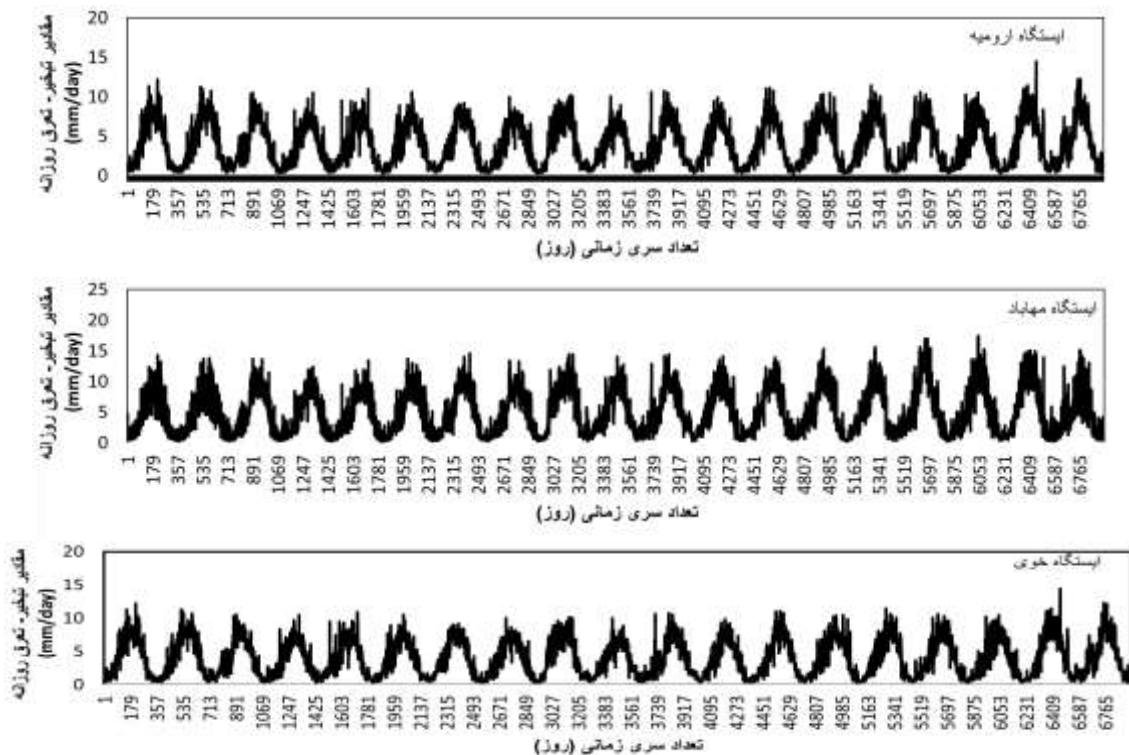
جدول ۲- مقادیر بدست آمده از ضریب مالو برای انتخاب بهترین پارامترهای ورودی به مدل

سناریو ها	R^2	C_p	S
میانگین دما	0/57	325	0/95
میانگین دما، سرعت باد در ارتفاع ۲ متری	0/73	165	0/65
میانگین دما، رطوبت نسبی میانگین، سرعت باد در ارتفاع ۲ متری	0/83	13/45	0/54
میانگین دما، رطوبت نسبی میانگین، سرعت باد در ارتفاع ۲ متری، میانگین فشار بخار	0/88	8/14	0/43
میانگین دما، رطوبت نسبی میانگین، سرعت باد در ارتفاع ۲ متری، میانگین فشار بخار، تشعشع خورشید، ساعات آفتابی	0/86	26/10	0/49
میانگین دما، رطوبت نسبی میانگین، سرعت باد در ارتفاع ۲ متری، میانگین فشار بخار، بارش باران، تشعشع خورشید، ساعات آفتابی	0/88	21/97	48

همانگونه که از مقادیر ET_0 برای هر سه ایستگاه نشان داده شد، مقدار این پارامتر برای هر سه ایستگاه با توجه به موقعیت مکانی جغرافیایی دارای محدوده تقریباً یکسانی می‌باشد.

توسعه مدل‌ها برای پیش‌بینی ET_0

مقدار تبخیر-تعرق مرجع در مقیاس روزانه برای هر یک از سه ایستگاه ارومیه، مهاباد و خوی در شکل ۲ نشان داده شده است.



شکل ۲- سری زمانی تبخیر-تعرق مرجع روزانه در ایستگاه‌های سینوپتیک حوضه دریاچه ارومیه شکل الف) ایستگاه ارومیه، شکل ب) ایستگاه مهاباد، شکل ج) ایستگاه خوی

می‌دهد هر دو مدل، با آموزش مطلوبی مواجه بوده‌اند. از طرف دیگر، شاخص آماری (R^2)، در مرحله تست برای مدل درخت و شبکه عصبی به ترتیب ۰/۹۷ و ۰/۹۶ و RMSE برای مدل‌های مذکور ۰/۴۸ و ۰/۶۰ (میلی‌متر بر روز) تعیین گردید.

همان‌طور که در جدول ۳ برای ایستگاه ارومیه نشان داده شده است، ضریب تبیین در مرحله آموزش برای روش ANN و MT برابر ۰/۹۲ می‌باشد و تفاوت چندانی در پیش‌بینی ET_0 نداشته‌اند. اما، مقدار شاخص آماری RMSE به ترتیب ۰/۵۷ و ۰/۴۷ (میلی‌متر بر روز) بدین ترتیب، شاخص‌های آماری، نشان

جدول ۳- نتایج ارزیابی عملکرد مدل‌های پیشنهادی در مقایسه با روابط تجربی برای ایستگاه ارومیه

مدل‌ها	R2	RMSE (mm/day)	MAE (mm/day)
مرحله آموزش			
ANN	۰/۹۲	۰/۵۷۳	۰/۴۴۷
MT	۰/۹۲	۰/۴۷۹	۰/۳۷
EMD-MT	۰/۹۳	۰/۴۵۵	۰/۳۵۴
مرحله تست			
ANN	۰/۹۲	۰/۶۰۳	۰/۴۶۱
MT	۰/۹۳	۰/۴۸۵	۰/۳۷۳
EMD-MT	۰/۹۳	۰/۴۷۸	۰/۳۶
Romanenko	۰/۸۴	۱/۰۰۲	۴/۶۸۱
Schendel	۰/۷۹	۱/۹۲۱	۱/۳۱

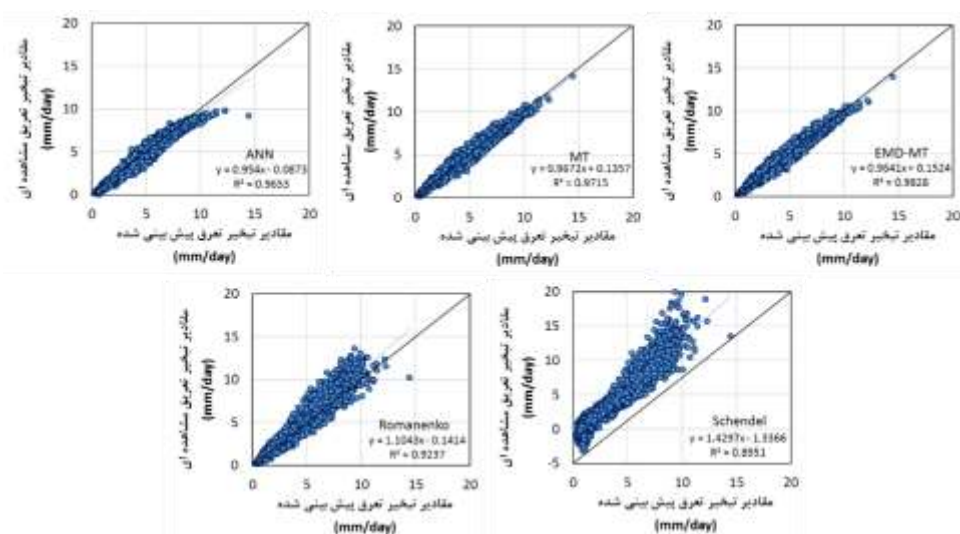
پیشنهادی، ایستگاه مهاباد در نظر گرفته شد، که نتایج همچنین نشان‌گر عملکرد بهتر روش ترکیبی EMD-MT با $R^2 = 0.95$ نسبت به باقی روش‌ها (مدل درخت = ۰/۹۵ و شبکه عصبی مصنوعی = ۰/۹۴) بوده است (جدول ۴). در ترم شاخص خطای RMSE روش مدل درخت مبتنی بر الگوریتم پیش‌پردازش‌کننده در مرحله آموزش، به میزان (mm/day) ۳/۸۷ نسبت به مدل شبکه عصبی مستقل کاهش خطا داشته و در مرحله آزمایش این مقدار برابر ۲۴/۴۷ (mm/day) بوده است. شاخص MAE نیز برای این روش ترکیبی در پیش‌بینی ET_0 روزانه در مراحل آموزش و تست به ترتیب برابر (mm/day) ۰/۳۳۹ و (mm/day) ۰/۸۷۱ بوده که در مقایسه با روابط تجربی Romanenko و Schendel در مرحله تست به ترتیب (mm/day) ۱۵/۷ و (mm/day) ۷۱/۳۲ کاهش خطا داشت. در این ایستگاه روش شبکه عصبی مصنوعی از نظر میزان خطا، نسبت به رابطه تجربی Romanenko عملکرد ضعیف‌تری داشت.

شکل ۵ و ۶ نمودارهای پراکندگی و سری زمانی را برای ایستگاه مهاباد نشان می‌دهد. با توجه به شکل ۵ تخمین برای هر سه روش هوشمند رو به پایین بوده در حالیکه این روند برای روابط تجربی عملکرد عکس داشته است. همچنین مقادیر پیش‌بینی رابطه تجربی Schendel در مقایسه با رابطه Romanenko دچار انحراف بیشتری از مقادیر واقعی شده است. نمودار سری زمانی برای این ایستگاه همانند ایستگاه ارومیه نشان می‌دهد که رابطه تجربی Schendel عملکرد بسیار ضعیفی در پیش‌بینی ET_0 روزانه برای مقادیر بالای (mm/day) ۱۰ دارد.

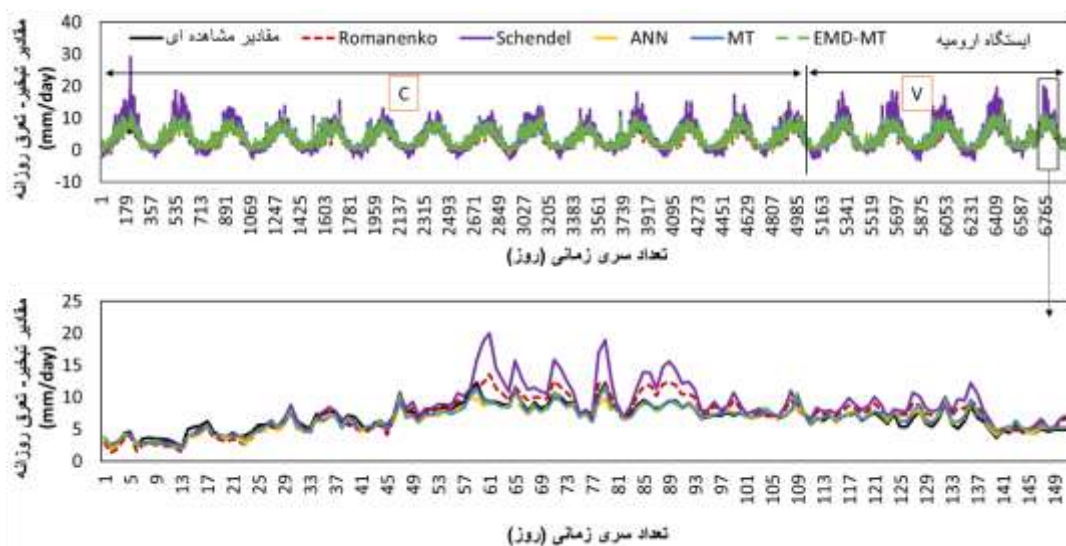
درحقیقت در این ایستگاه، ترکیب کردن روش پیش‌پردازش‌کننده EMD با روش MT باعث می‌شود که ضریب تبیین به میزان ۰/۰۱ و شاخص RMSE نیز ۰/۰۷ میلی‌متر بر روز افزایش یابند. در این ایستگاه روابط تجربی عملکرد نسبتاً خوبی برای پیش‌بینی ET_0 داشته اما ضعیف‌تر از روش‌های هوشمند مصنوعی عمل کرده‌اند.

شکل ۳ و ۴ نیز نمودارهای پراکندگی و سری زمانی مقادیر تبخیر- تعرق مرجع مشاهداتی و پیش‌بینی نظیر آن توسط مدل پیشنهادی ANN، MT و EMD-MT را در مرحله آموزش و آزمون نشان می‌دهد. همان‌طور که در شکل ۳ مشاهده می‌شود، داده‌ها در اطراف خط ۴۵ درجه متمرکز شده که نشان از همبستگی مقادیر واقعی و محاسباتی مدل می‌باشد. یکی از ویژگی‌های اصلی نمودار پراکندگی بررسی پدیده Over-estimation و Under-estimation است که میزان تخمین رو به پایین داده‌های ET_0 و یا تخمین رو به بالای آن می‌باشد. همانگونه که در این شکل مشخص است، تمامی روش‌ها به جز رابطه تجربی Schendel تخمینی مناسب را داشته‌اند.

در شکل ۴ نیز که مقادیر سری زمانی مدل‌های پیشنهادی مستقل و ترکیبی را در پیش‌بینی ET_0 روزانه ایستگاه ارومیه نشان می‌دهد. مشخص است که روابط تجربی در پیش‌بینی مقادیر بیشینه ET_0 دارای خطای بسیاری بوده‌اند که باعث ضعف در عملکرد آن‌ها شده است، همچنین روش ترکیبی EMD-MT با منطبق شدن بر مقادیر ET_0 مشاهداتی (خط مشکی) نشانگر عملکرد قوی این روش برای پیش‌بینی تبخیر- تعرق مرجع روزانه در ایستگاه ارومیه می‌باشد. برای ارزیابی بیشتر هر یک از روش‌های مستقل و ترکیبی



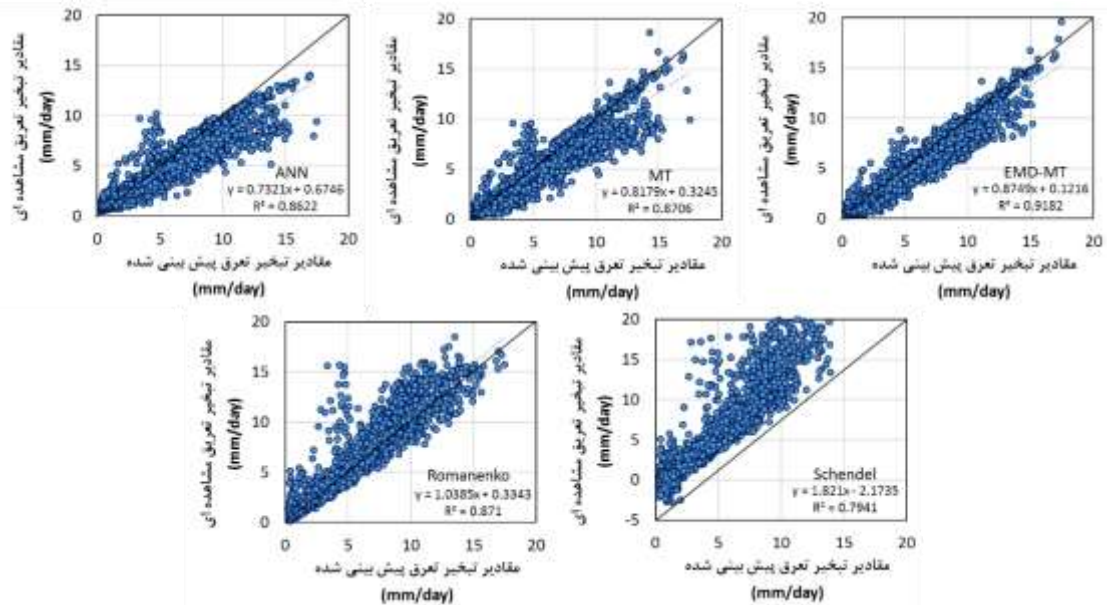
شکل ۳- پراکندگی مدل‌های پیشنهادی در مرحله آزمون برای ایستگاه ارومیه



شکل ۴- سری زمانی پیش‌بینی تبخیر_تعرق مرجع روزانه برای هر یک از مدل‌های پیشنهادی در ایستگاه ارومیه

جدول ۴- نتایج ارزیابی عملکرد مدل‌های پیشنهادی در مقایسه با روابط تجربی برای ایستگاه مهاباد

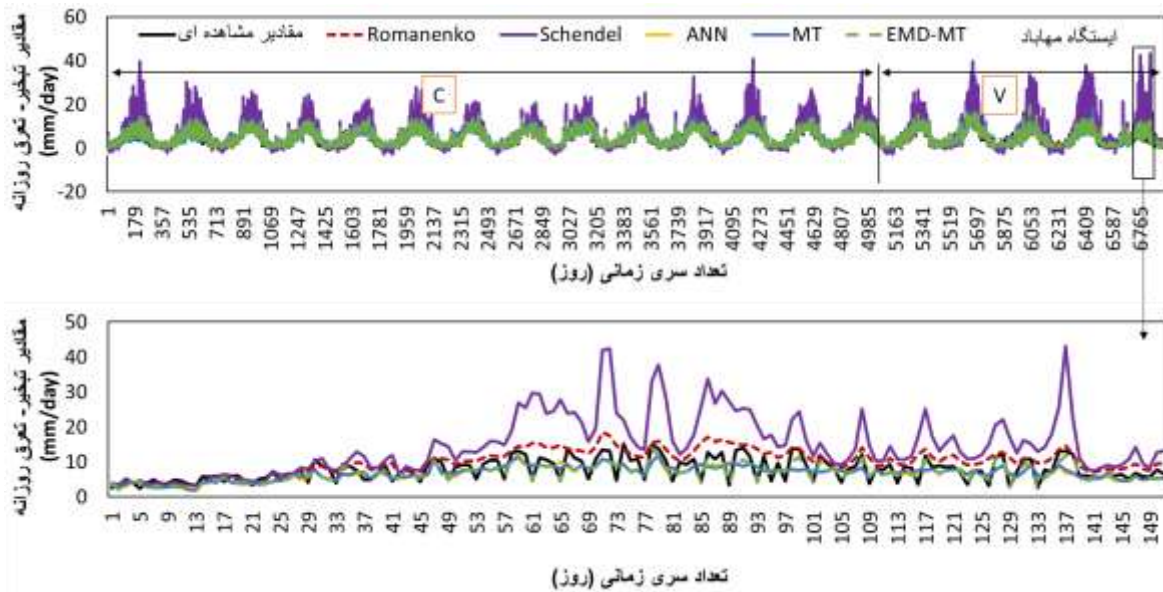
مدل‌ها	R^2	RMSE (mm/day)	MAE (mm/day)
مرحله آموزش			
ANN	۰/۹۴	۰/۶۰۹	۰/۴۷۴
MT	۰/۹۵	۰/۴۵۶	۰/۳۵۱
EMD-MT	۰/۹۵	۰/۴۳۹	۰/۳۳۹
مرحله تست			
ANN	۰/۷۳	۱/۷۵۲	۱/۱۸۴
MT	۰/۷۴	۱/۵۸۷	۱/۰۴۶
EMD-MT	۰/۸۲	۱/۳۷۵	۰/۸۷۱
Romanenko	۰/۷۴	۱/۶۸۵	۱/۰۳۴
Schendel	۰/۶۲	۵/۴۰۶	۳/۰۳۸



شکل ۵- پراکندگی مدل‌های پیشنهادی در مرحله آزمون برای ایستگاه مهاباد

آموزش برای مقادیر پیش‌بینی شده ET_0 برابر ۰/۹۷ گزارش شده که تقریباً برابر با دو روش مستقل MT و ANN می‌باشد (جدول ۵).

در ایستگاه خوی نیز نتایج برای روش ترکیبی EMD-MT تا حدودی مانند دو ایستگاه قبل گزارش شد که ضریب تبیین در مرحله



شکل ۶- سری زمانی پیش‌بینی تبخیر- تعرق مرجع روزانه برای هر یک از مدل‌های پیشنهادی در ایستگاه مهاباد

به مانند دو ایستگاه قبلی عملکرد نسبتاً ضعیف‌تری نسبت به روش‌های هوشمند مصنوعی داشتند. از این رو می‌توان به عنوان ابزاری معتبر در داده کاوی و پژوهش‌های هوش مصنوعی در زمینه هیدرولوژی مورد استفاده قرار گیرند.

در ترم شاخص‌های RMSE و MAE نیز پیش‌پردازش داده‌ها کمکی چندانی به افزایش دقت مدل برای پیش‌بینی ET_0 روزانه در مرحله آموزش نکرد. اما این الگوریتم (EMD) در مرحله تست توانست دقت مدل درخت را از مقدار ۰/۹۱ به میزان ۲/۰۴ افزایش دهد. روابط تجربی نیز

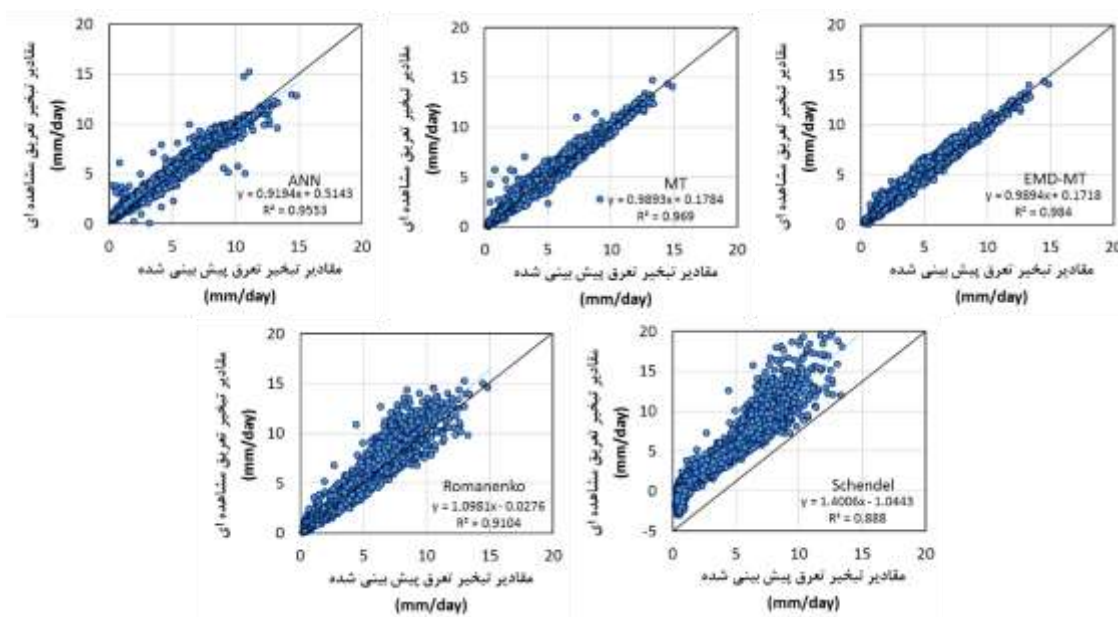
جدول ۵- نتایج ارزیابی عملکرد مدل‌های پیشنهادی در مقایسه با روابط تجربی برای ایستگاه خوی

مدل‌ها	R ²	RMSE (mm/day)	MAE (mm/day)
مرحله آموزش			
ANN	۰/۹۴	۰/۶۱۲	۰/۱۹۵
MT	۰/۹۳	۰/۷۳۸	۰/۲۴۲
EMD-MT	۰/۹۴	۰/۶۰۶	۰/۱۴۱
مرحله تست			
ANN	۰/۹۰	۰/۷۰۴	۰/۵۱
MT	۰/۹۱	۰/۵۸۱	۰/۳۸۱
EMD-MT	۰/۹۶	۰/۴۳۶	۰/۳۱۵
Romanenko	۰/۸۲	۱/۲۱۱	۰/۷۸
Schendel	۰/۷۷	۲/۱۵۳	۱/۳۸۷

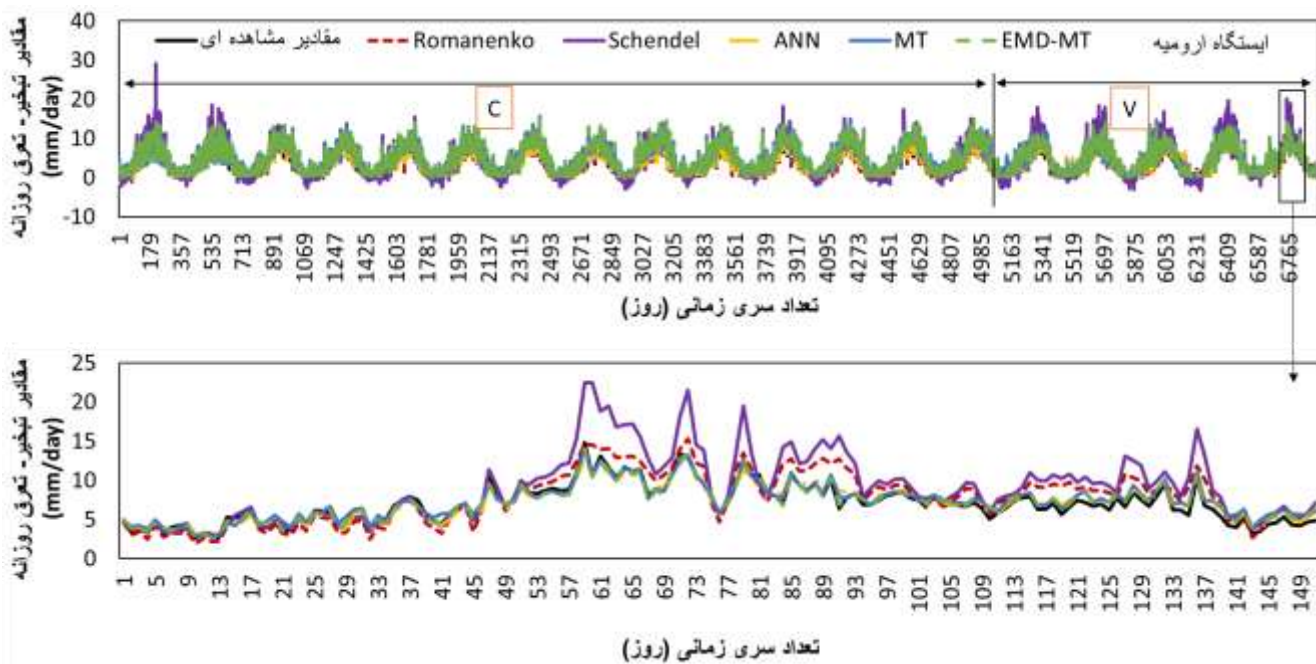
نتیجه‌گیری

برآورد دقیق از فرآیند تبخیر- تعرق مرجع در مدیریت بهینه منابع آب هر کشور مسئله بسیار مهمی به شمار می‌رود. از این رو، محاسبه و پیش‌بینی مقدار دقیق تبخیر و تعرق روزانه و به تبع آن برآورد دقیق نیاز آبی گیاهان زمینه را برای طراحی مناسب تاسیسات آبیاری فراهم نموده و ضمن کاهش هزینه‌های اجرا باعث ارائه برنامه مناسب بهره‌برداری از منابع آب بخش آبیاری و کشاورزی می‌گردد. در این تحقیق از داده‌های هواشناسی سه ایستگاه واقع در حوضه آبریز دریاچه ارومیه شامل ارومیه، مهاباد و خوی برای سال‌های بین ۱۳۷۸ تا ۱۳۹۷ برای پیش‌بینی ET₀ روزانه با استفاده از روش‌های هوشمند مصنوعی استفاده گردید.

نمودار پراکندگی مقادیر پیش‌بینی مدل‌های پیشنهادی و روابط تجربی برای ایستگاه خوی در مرحله آزمایش نیز در شکل ۷ نشان داده شده است. همچنین شکل ۸ مقایسه‌ای بین داده‌های مشاهده‌ای و نتایج محاسباتی مدل‌های هوشمند و روابط تجربی در پیش‌بینی تبخیر- تعرق مرجع را با استفاده از نمودار سری زمانی نشان داده است. مدل EMD-MT در تخمین بیشینه‌ها و کمینه‌های محلی موفقیت بیشتری در مقایسه با شبکه عصبی مصنوعی، هم در مرحله آموزش و هم در مرحله آزمون داشته است.



شکل ۷- پراکندگی مدل‌های پیشنهادی در مرحله آزمون برای ایستگاه خوی



شکل ۸- سری زمانی پیش‌بینی تبخیر-تعرق مرجع روزانه برای هر یک از مدل‌های پیشنهادی در ایستگاه خوی

سیفی ا، میر لطیفی س.م. و ریاحی ح. ۱۳۸۹. توسعه مدل ترکیبی رگرسیون چندگانه-تحلیل مؤلفه‌ها و عامل‌های اصلی (MLR-PCA) در پیش‌بینی تبخیر و تعرق مرجع (مطالعه موردی: ایستگاه کرمان)، نشریه آب و خاک. جلد ۲۴. شماره ۶. گودرزی م، صلاحی ب، حسینی س.ا. ۱۳۹۴. بررسی تغییرات اقلیمی بر تغییرات رواناب سطحی (مطالعه موردی: حوضه آبریز دریاچه ارومیه)، نشریه اکوهیدرولوژی، دوره ۲، شماره ۲.

Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. 1998. Crop evapotranspiration. Guide lines for computing crop evapotranspiration. FAO Irrigation and Drainage Paper No 56. Rome, Italy.

Al-Musaylh, M.S., Deo, R.C., Li, Y., and Adamowski, J. F. 2018. Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting. Applied energy, 217, 422-439.

Ali, M., and Prasad, R. 2019. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. Renewable and Sustainable Energy Reviews, 104, 281-295.

Chen, N., Zhang, Y., Jin, C., Wang, A., Guan, D., and Tian, L. 2018. Intercomparison of three methods to estimate evapotranspiration over temperate meadow in Inner Mongolia: Penman-Monteith, Makkink and

برای شناسایی مهمترین پارامترهای اثرگذار بر ET_0 در این تحقیق از ضریب مالو که کمترین مقدار برای این شاخص نشان‌دهنده پارامترهای اثرگذار بر مدل می‌باشد، استفاده گردید. از این رو، پارامترهای میانگین دما، رطوبت نسبی میانگین، سرعت باد در ارتفاع ۲ متری، میانگین فشار بخار به عنوان متغیرهای ورودی به مدل انتخاب شدند. نتایج نشان داد که با پیش‌پردازش کردن داده‌های سری زمانی با استفاده از الگوریتم تجزیه مد تجزیه دقت مدل درخت را در ترم ضریب همبستگی $0.2/1$ ، $0.39/4$ و $0.4/2$ به ترتیب برای ایستگاه ارومیه، مهاباد و خوی در مرحله تست افزایش داد. همچنین این روش ترکیبی پیشنهادی به عنوان مثال در ایستگاه مهاباد خطای به مراتب کمتری نسبت به دو رابطه تجربی Romanenko و Schendel در پیش‌بینی تبخیر-تعرق مرجع روزانه داشته است.

منابع

بهمنش، ج، آزاد طلائی، ن، منتصری، م، و بشارت، س. ۱۳۹۳. ارزیابی مدل‌های سری زمانی خطی و غیر خطی بیلینیر در پیش‌بینی تبخیر-تعرق گیاه مرجع در ایستگاه سینوپتیک ارومیه. نشریه پژوهش آب در کشاورزی. جلد ۲۸، شماره ۱. حسن پور ب، میرزایی، ف، ارشد، ص. و کوثری، ه. ۱۳۹۱. مقایسه الگوریتم‌های SEBAL و S-SEBI در برآورد تبخیر و تعرق در منطقه کرچ، نشریه آب و خاک، جلد ۲۶، شماره ۶.

- and Han, D. 2009. Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Advances in Water Resources*, 32(1), 88-97.
- Quinlan, J.R. 1992. Learning with continuous classes. In 5th Australian joint conference on artificial intelligence (Vol. 92, pp. 343-348).
- Rahimi, K.A. 2011. A prediction of maximum monthly precipitation recorded at Ilam meteorological station based on Persian Gulf and Red Sea surface temperature through recordings data mining method. *Iranian journal of soil and water research*, 42-1.
- Rezaie-balf, M., Naganna, S.R., Ghaemi, A., and Deka, P.C. 2017a. Wavelet coupled MARS and M5 Model Tree approaches for groundwater level forecasting. *Journal of hydrology*, 553, 356-373.
- Rezaie-Balf, M., Zahmatkesh, Z., and Kim, S. 2017b. Soft computing techniques for rainfall-runoff simulation: local non-parametric paradigm vs. model classification methods. *Water Resources Management*, 31(12), 3843-3865.
- Rezaie-Balf, M., Kisi, O., and Chua, L.H. 2019. Application of ensemble empirical mode decomposition based on machine learning methodologies in forecasting monthly pan evaporation. *Hydrology Research*, 50(2), 498-516.
- Solomatine, D.P., and Dulal, K.N. 2003. Model trees as an alternative to neural networks in rainfall—runoff modelling. *Hydrological Sciences Journal*, 48(3), 399-411.
- Sudheer, K.P., Gosain, A.K., and Ramasastri, K.S. 2003. Estimating actual evapotranspiration from limited climatic data using neural computing technique. *Journal of Irrigation and Drainage Engineering*, 129(3), 214-218.
- Taheri, M., Emadzadeh, M., Gholizadeh, M., Tajrish, M., Ahmadi, M., and Moradi, M. 2019. Investigating the temporal and spatial variations of water consumption in Urmia Lake River Basin considering the climate and anthropogenic effects on the agriculture in the basin. *Agricultural water management*, 213, 782-791.
- Traore, S., Luo, Y., and Fipps, G. 2017. Gene-expression programming for short-term forecasting of daily reference evapotranspiration using public weather forecast information. *Water resources management*, 31(15), 4891-4908.
- Zhou, L., Meng, Y., and Abbaspour, K.C. 2019. A new framework for multi-site stochastic rainfall generator based on empirical orthogonal function analysis and Hilbert-Huang transform. *Journal of Hydrology*, 575, 730-742.
- Priestley-Taylor equation. *Water and environment journal*, 32(4), 500-507.
- Cheng, Y., Wang, Z., Chen, B., Zhang, W., and Huang, G. 2019. An improved complementary ensemble empirical mode decomposition with adaptive noise and its application to rolling element bearing fault diagnosis. *ISA transactions*.
- Deo, R.C., Downs, N., Parisi, A.V., Adamowski, J.F., and Quilty, J.M. 2017. Very short-term reactive forecasting of the solar ultraviolet index using an extreme learning machine integrated with the solar zenith angle. *Environmental research*, 155, 141-166.
- Fijani, E., Barzegar, R., Deo, R., Tziritis, E., and Konstantinos, S. 2019. Design and implementation of a hybrid model based on two-layer decomposition method coupled with extreme learning machines to support real-time environmental monitoring of water quality parameters. *Science of the total environment*, 648, 839-853.
- Ghaemi, A., Rezaie-Balf, M., Adamowski, J., Kisi, O., and Quilty, J. 2019. On the applicability of maximum overlap discrete wavelet transform integrated with MARS and M5 model tree for monthly pan evaporation prediction. *Agricultural and Forest Meteorology*, 278, 107647.
- Huang, Y., Yang, L., Liu, S., and Wang, G. 2019. Multi-Step Wind Speed Forecasting Based On Ensemble Empirical Mode Decomposition, Long Short Term Memory Network and Error Correction Strategy. *Energies*, 12(10), 1822.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., and Liu, H.H. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903-995.
- Kakahaji, H., Banadaki, H.D., Kakahaji, A., and Kakahaji, A. 2013. Prediction of Urmia Lake water-level fluctuations by using analytical, linear statistic and intelligent methods. *Water resources management*, 27(13), 4469-4492.
- Kişi, Ö., and Öztürk, Ö. 2007. Adaptive neurofuzzy computing technique for evapotranspiration estimation. *Journal of Irrigation and Drainage Engineering*, 133(4), 368-379.
- Kumar, M., Raghuvanshi, N. S., Singh, R., Wallender, W.W., and Pruitt, W.O. 2002. Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, 128(4), 224-233.
- Moghaddamnia, A., Gousheh, M.G., Piri, J., Amin, S.,

soil moisture using a universal relationship for a large area. Proceedings. Ukrainian. Hydrometeorol. Res. Inst. 3, 12-25.

Schudel, U. 1967. Vegetations was serverbrauch und-wasserbedarf. Habilitation, Kiel, 137 pp.

Romanenko, V.A. 1961. Computation of the autumn

Prediction of Daily Reference Evapotranspiration Using Hybrid Artificial Intelligence Method Based on Empirical Mode Decomposition

A. Amirashayeri¹, J. Behmanesh^{2*}, V.R. Verdinejad³, N. Fathollahzadeh Attar⁴

Received: Apr.28, 2020

Accepted: may.29, 2020

Abstract

Evapotranspiration (ET) is one of the essential components of the hydrological cycle, which plays a crucial role in the study of a watershed water balance. In calculating the water requirement of plants, it is essential to calculate the reference evapotranspiration, and then, the crop evapotranspiration is estimated using the calculated value. In the present research, for accurate determining of daily reference evapotranspiration of Lake Urmia watershed, three stations of the watershed, including Urmia, Mahabad, and Khoy, were selected and daily reference evapotranspiration values were calculated based on the standard FAO-Penman-Monteith method. The best input parameters for modeling reference evapotranspiration were selected based on Malo's coefficient. The MT model, which is used in the current study, is one of the inference-classification algorithms. To deal with the complexity and instability of time series data, the empirical mode decomposition (EMD) preprocessing algorithm was used. The results of the methods were compared with the empirical relationships of Romanenko and Schendel. The results of this study show that although the tree modeling method performs relatively equal and sometimes weaker than the ANN method, the combination with EMD technique increases the accuracy of the model and reduces the error in daily ET_0 prediction. According to the results, the EMD-MT method in correlation coefficient index for Urmia, Mahabad, and Khoy stations increased 1.02%, 4.39%, and 2.04%, respectively. Also, among the empirical relations, the Romanenko relation is more accurate than the Schendel equation, and it is a reliable empirical model.

Keywords: Empirical mode decomposition, Reference evapotranspiration, Model tree, Prediction

1- M.Sc. Student of Drainage Irrigation, Department of Water Engineering
2- Professor of Water Engineering, Department of Water Engineering, Urmia University
3- Associate Professor, Department of Water Engineering, Urmia University
4- Ph.D. of Water Engineering, Department of Water Engineering, Urmia University
(* - Corresponding Author Email: j.behmanesh@urmia.ac.ir)