

ارزیابی الگوریتم M5P در بر آورد تبخیر تعرق پتانسیل، دمای کمینه و بیشینه (مطالعه موردی: ایستگاه هواشناسی ساری)

سید حسن میرهاشمی^۱، مهدی پناهی^{۲*}، لیلا زارعی^۲

^۱ دانش آموخته دکترای آبیاری و زهکشی دانشکده آب و خاک دانشگاه زابل، ایران

^۲ گروه مهندسی آب، دانشکده کشاورزی، دانشگاه زنجان، ایران

^۲ دانشجوی کارشناسی ارشد مهندسی منابع آب دانشگاه بین المللی امام خمینی(ره)، ایران

تاریخ دریافت: ۹۷/۴/۲۸، تاریخ پذیرش: ۹۸/۷/۲۱

چکیده

در این پژوهش ارزیابی عملکرد مدل درختی M5P در تخمین سه پارامتر تبخیر تعرق پتانسیل، دمای کمینه و بیشینه انجام شد. داده‌های مورد استفاده از اطلاعات هواشناسی ایستگاه سینوپتیک ساری طی دوره ۳۰ ساله از سال ۱۳۶۷ تا ۱۳۹۷ به صورت ماهانه مورد استفاده قرار گرفت. با توجه به نتایج مدل M5P با ضریب همبستگی برابر ۰/۹۲، ریشه میانگین خطا برابر ۰/۲۵ و متوسط مطلق خطا برابر ۰/۱۲ دارای عملکرد بهتری در پیشبینی تبخیر تعرق پتانسیل نسبت به پیشبینی دمای کمینه و بیشینه بوده است. با توجه به آنالیز حساسیت، الگوریتم M5P با متغیرهای ورودی شامل دمای متوسط، ساعات آفتابی، درصد رطوبت هوا، دمای نقطه شبنم و کمبود فشار بخار اشباع دارای بیشترین و با متغیرهای ورودی ساعات آفتابی و سرعت باد دارای کمترین عملکرد درست در پیشبینی تبخیر تعرق پتانسیل بوده است. همچنین مشخص شد که پارامتر سرعت باد تأثیر منفی در عملکرد پیشبینی مدل M5P در مقدار تبخیر تعرق پتانسیل داشته است.

کلمات کلیدی: آنالیز حساسیت، سرعت باد، ضریب همبستگی، پارامتر هواشناسی

علیرغم کاربرد وسیع مفهوم تبخیر تعرق، ابهاماتی در استفاده از چنین عباراتی مانند تبخیر تعرق پتانسیل و تبخیر تعرق گیاه مرجع وجود دارد. برای رفع این، سازمان خوار و بار کشاورزی ملل متحد^۱ نشریه شماره ۵۶ را منتشر نمود (آلن^۲ و همکاران، ۱۹۹۸). تبخیر و تعرق بعد از بارش مهمترین پارامتر سیکل هیدرولوژیک می باشد و همچنین مطالعه تبخیر و تعرق در علوم اقلیم شناسی، هیدرولوژی، جغرافیا، کشاورزی و اکولوژی دارای اهمیت است (هان^۳ و همکاران، ۲۰۱۵).

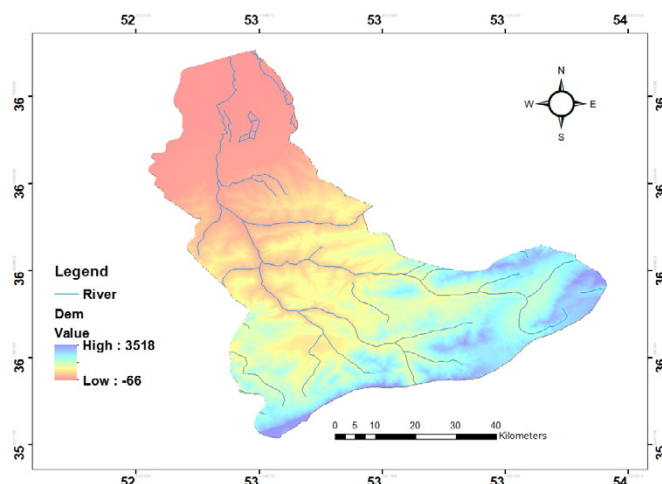
آگاهی یافتن از وضعیت آینده تبخیر - تعرق مرجع می تواند به تصمیم گیری های آتی به منظور روش های مدیریتی و برنامه ریزی منابع آب مناسب کمک کند (به منش و همکاران، ۱۳۹۳). با توجه به وقوع پدیده تغییر اقلیم استفاده از شاخص هایی که علاوه بر بارش، تبخیر و تعرق پتانسیل را نیز در نظر بگیرد، اجتناب ناپذیر است (اسدی زارچ^۴ و همکاران، ۲۰۱۵). تاکنون مطالعات بسیاری برای تخمین درست تبخیر و تعرق صورت گرفته که منجر به ارائه روش های متنوعی برای برآورد تبخیر و تعرق شده است (مارتی^۵ و همکاران، ۲۰۱۵). روش فاوو-پنمن-مانتیت از دقت بیشتری نسبت به سایر روش ها برخوردار است (دینگ^۶ و همکاران، ۲۰۱۵). روش پنمن مانتیت یک مدل فیزیکی است که جنبه های آئرو دینامیک و ترمودینامیک را در بر می گیرد (اسدی زارچ و همکاران، ۲۰۱۷). در ناحیه شمال- شرقی هند به دنبال مقایسه روش های تبخیر تعرق مرجع با بررسی معادله های مختلف می توان انطباق نتایج حاصل از مدل های با پایه تابش خورشید و دمای هوا را با نتایج حاصل از مدل فائو پنمن مانتیت بدست آورد (پندی^۷ و همکاران، ۲۰۱۶). در ارزیابی عملکرد ۵۲ معادله تبخیر تعرق مرجع با معادله فائو پنمن مانتیت به این نتیجه شد که خروجی معادلات بر پایه دمای هوا در مقایسه با خروجی

معادله فائو پنمن مانتیت کمترین درصد خطا را نشان می دهد (دجمن^۸ و همکاران، ۲۰۱۵).

داده کاوی یکی از روش های است که در تحقیقات زیادی مانند هواشناسی و هیدرولوژی استفاده شده است. فرآیند انتخاب، کاوش و مدل بندی داده های حجیم، جهت کشف روابط نهفته با هدف بدست آوردن نتایج واضح و مفید برای مالک پایگاه داده ها را، داده کاوی می گویند (مشکانی، ۱۳۸۸). داده کاوی فرآیندی است که ابزارهای مختلف تحلیل داده را به کار می گیرد تا الگوها و روابط فیزیکی متغیرها را در مجموعه داده های مختلف کشف کند (توکرس^۹، ۱۹۹۹). یکی از الگوریتم های پیش بینی کننده داده کاوی الگوریتم M5P^{۱۰} می باشد. ستاری و همکاران (۱۳۹۲) تبخیر- تعرق مرجع روزانه ایستگاه سینوپتیک بناب را با استفاده از مدل درختی M5 و شبکه عصبی مصنوعی پیش بینی کردند. با توجه به نتایج مشخص شد که گرچه روش شبکه عصبی پیش بینی مصنوعی با اختلاف بسیار کمی در مقایسه با روش M5 به نسبت دقیقتری از میزان تبخیر - تعرق مرجع دارد، اما روش M5 روابط ساده خطی، قابل فهمتر و کاربردی تری را برای پیش بینی تبخیر- تعرق عرضه می کند. میرهاشمی و پناهی (۱۳۹۲) در پژوهشی توانایی مدل های درختی CART برای تخمین متوسط تبخیر تعرق پتانسیل و درصد رطوبت نسبی ماه آینده به صورت ماهانه برای ایستگاه سینوپتیک اراک مورد ارزیابی قرار دادند. الگوریتم CART با استفاده از ضریب همبستگی و متوسط مطلق خطا مورد ارزیابی قرار گرفت. عملکرد مدل درختی CART با مقدار ضریب همبستگی برابر ۰/۸۴ و MAE برابر ۰/۶۷ مربوط به تخمین متوسط تبخیر تعرق پتانسیل ماهانه و با مقدار ضریب همبستگی ۰/۸۱ و MAE برابر ۷/۹ مربوط به تخمین درصد رطوبت نسبی ماهانه، بدست آمد. با توجه به نتایج مدل درختی CART برای تخمین متوسط تبخیر تعرق پتانسیل نسبت به تخمین درصد رطوبت نسبی هوا، عملکرد بهتری داشته است.

8- Djaman
9- Two Crows
10- M5 model trees

1- FAO
2- Allen
3- Han
4- Asadi Zarch
5- Marti
6- Ding
7- Pandey



شکل ۱: توپوگرافی و موقعیت رودخانه‌ها در شهرستان ساری

مواد و روش‌ها

منطقه مطالعاتی

شهر ساری واقع در کوهپایه‌های رشته کوه البرز دارای دو بخش کوهستانی و دشت و همچنین در طول جغرافیایی $26^{\circ} 37' 53''$ تا $25^{\circ} 44' 54''$ و عرض‌های جغرافیایی $18^{\circ} 28' 36''$ تا $9^{\circ} 41' 36''$ درجه واقع شده است. آب و هوای این شهر در تابستان‌ها معتدل و نمناک و در زمستان‌ها نسبتاً سرد و خشک است. بادهایی که از نواحی غربی می‌وزد باعث برودت و سردی هوا گشته و گاهی موجب ریزش برف می‌گردد (شاهی نژاد و دهقانی، ۱۳۹۵) وجود رشته کوه البرز در قسمت جنوبی مازندران یک قوس عظیمی به وجود آورده است که مانع عبور رطوبت دریای مازندران به مرکز ایران می‌شود. جلگه پست دریای مازندران را از فلات مرکزی جدا ساخته و حد فاصل بین این دو آب و هوا متضاد می‌باشد. این شرایط سبب شده است که بارش منطقه الگوی پیچیده و متفاوتی با سایر نقاط کشور داشته باشد. میانگین بارش سالانه مازندران طی دوره مورد مطالعه $681/5$ میلی‌متر می‌باشد. میانگین کمترین بارش $194/7$ میلی‌متر می‌باشد که منطبق با ارتفاعات و در شرق استان می‌باشد (محمدپورزیدی و همکاران، ۱۳۹۸). در شکل (۱) توپوگرافی و موقعیت رودخانه‌ها منطقه مطالعاتی مشخص شده است.

به بررسی اثر تغییر اقلیم بر تبخیر و تعرق پتانسیل در دشت مشهد طی دوره ۲۰۲۱-۲۰۷۰ تحت سناریوهای RCP و با بکارگیری یک مدل دینامیکی جهت ریزمقیاس‌نمایی پرداخته شد و نتیجه شد که میزان تبخیر و تعرق پتانسیل در دوره‌های آتی نسبت به دوره پایه در ماه‌های گرم سال کاهش و در ماه‌های سرد سال افزایش خواهد یافت (قهرمان و همکاران ۱۳۹۵).

برآورد صحیح دمای هوا کاربردهای متعددی در تعیین تابش طول موج بلند و کوتاه (Stanelle et al., 2010) و مدیریت کشاورزی (Venkadesh et al., 2013) دارد. پناهی و میرهاشمی (۱۳۹۲) دو الگوریتم داده کاوی CART و CHAID در پیش‌بینی دمای هوای ایستگاه سینوپتیک اراک مورد ارزیابی قرار دادند. بر طبق نتایج مدل درختی CHAID با مقدار ضریب همبستگی برابر $0/91$ و متوسط مطلق خطا برابر $2/7$ عملکرد بهتری در پیش‌بینی دمای متوسط ماه بعد نسبت به الگوریتم CART داشته است.

با توجه به تحقیقات گذشته توانایی و عملکرد مدل‌های داده کاوی مانند M5P در پیش‌بینی در حد قابل قبول می‌باشد. همچنین نتایج مدل درختی M5P به صورت ساده و قابل فهم برای تصمیم‌گیران و مدیران است. بنابراین هدف از این پژوهش ارزیابی عملکرد مدل M5P در پیش‌بینی سه پارامتر تبخیر تعرق، دمای کمینه و بیشینه ایستگاه هواشناسی ساری می‌باشد.

تبخیر تعرق پتانسیل

فائو پیشنهاد نموده است که سطح مرجع فرضی، دقیقاً مانند سطح وسیعی از چمن سبز با ارتفاع یکنواخت، رشد فعال، سایه اندازی کامل زمین و با آب کافی (آلن و همکاران، ۱۹۹۸) به عنوان سطح مرجع در نظر گرفته شود. در رهیافت فائو، مشخصات سطح که بر تبخیر تعرق مؤثرند بصورت واضح کمی شده اند (ایتنفسو و همکاران، ۲۰۰۳).

نرخ تبخیر تعرق از سطح مرجع، بدون کمبود آب تبخیر تعرق گیاه مرجع یا تبخیر تعرق مرجع نامیده شده و به صورت ET_0 نامیده می شود (آلن و همکاران، ۱۹۹۸). معادلات تبخیر تعرق مرجع فراوانی بسته به قابل دسترس بودن داده های هواشناسی، توسعه و استفاده شده اند. این معادلات از معادلات پیچیده مانند معادلات تجربی بر مبنای تابش خورشید یا دما تا معادلات پیچیده بر اساس فرآیندهای فیزیکی مانند روش ترکیبی پنمن^{۱۱} (۱۹۶۵) وجود دارند. جنسن^{۱۲} و همکاران (۱۹۹۰) عملکرد ۲۰ روش مختلف را در مقابل تبخیر تعرق اندازه گیری شده برای ۱۱ ایستگاه واقع در مناطق اقلیمی مختلف جهان تحلیل نمودند. روش پنمن - مونیت به عنوان بهترین روش برای تمامی شرایط اقلیمی رتبه بندی شد. پنمن مانیت تعیین شده به معادله وسیله FAO-56 (آلن و همکاران، ۱۹۹۸):

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

ET = تبخیر تعرق مرجع (میلی متر بر روز)،
 R = تابش خالص ورودی به سطح گیاه (مگاژول بر مترمربع بر روز)،
 G = شار گرمای خاک (مگاژول بر مترمربع بر روز)،
 T = میانگین روزانه دمای هوا در ارتفاع دو متری (درجه سلسیوس)،
 U_2 = میانگین روزانه سرعت باد در ارتفاع دو متری (متر بر ثانیه) = e_s = فشار بخار اشباع (کیلو پاسکال)،
 e_a = فشار بخار واقعی (کیلو پاسکال)

داده های که در ایستگاه های مختلف هواشناسی اندازه گیری و بایگانی می شوند حجم زیادی از اطلاعات را شامل می شود و به مرور زمان بر حجم آن ها افزوده می شود. بر این اساس لزوم استفاده از روش های جدید استخراج اطلاعات از آن ها بیشتر احساس می شود. تبخیر تعرق یکی از متغیرهایی است که معمولاً بر اساس روابط پیشنهادی متغیرهای هواشناسی که در ایستگاه های هواشناسی اندازه گیری می شوند تخمین زده می شوند اما در برخی از این روابط پیشنهادی متغیرهای زیادی مورد استفاده قرار می گیرند که ممکن است برخی از آن ها در تمام ایستگاه های هواشناسی اندازه گیری نشوند بنابراین ضرورت ایجاد می کند تا از روش های نوین مدل سازی مانند داده کاوی استفاده کرد.

تعاریف داده کاوی

داده کاوی دارای تعاریف متنوعی است. داده کاوی به فرآیند استخراج اطلاعات نهفته، قابل فهم، قابل پیگیری از پایگاه داده های بزرگ و استفاده از آن در تصمیم گیری های تجاری مهم، اطلاق می شود. داده کاوی تکنیکی است که آزمون های فرض را ترکیب می کند و داده های مشتقی را کشف می کند. در آزمون های فرض، محقق ایده را در برابر داده آزمون می کند تا اعتبار آن را تایید یا رد کند. واندربرگ^{۱۳} و همکاران (۱۹۹۹) توضیح می دهند که در کشف، محقق نتیجه را از داده ها ترسیم می کند و به داده ها اجازه می دهد تا نتیجه را بپذیرد. اغلب مسائل داده کاوی با استفاده از یک ترکیب از هر دو روش حل می شود. به طور مثال از نتیجه ممکن است فرضیه جدیدی برخیزد که می تواند آزمون شود و این آزمون تایید یا رد شود. داده کاوی فرآیند انتخاب، شناسایی و مدل سازی از مقادیر زیاد داده ها برای کشف نظم یا رابطه ای که در ابتدا ناشناخته هستند با هدف بدست آوردن نتایج مفید و شفاف برای مالک پایگاه می باشد (گیودسی، ۲۰۰۳).

تفاوت اصلی که بین داده کاوی و آمار وجود دارد این است که داده کاوی یک رهیافت بدون پیش فرض

11- Penman

12- Jensen

13- Vanderberg

مدل‌های بر مبنای درخت به روش تقسیم و نتیجه‌گیری ساخته می‌شوند. تولید مدل درختی به دو مرحله نیاز دارد. اولین مرحله شامل استفاده از یک معیار انشعاب برای تولید یک تصمیم درختی است. معیار انشعاب برای الگوریتم مدل درختی M5P براساس رفتار انحراف معیار مقادیر کلاسی است که به عنوان کمیتی از خطا به یک گره می‌رسد و کاهش مورد انتظار در این خطا را به عنوان نتیجه‌ی آزمون هر صفت در آن گره محاسبه می‌نماید. فرمول محاسبه کاهش انحراف معیار^۴ بصورت رابطه زیر است

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

که T بیانگر یکسری نمونه‌هایی است که به گره می‌رسد؛ T_i بیانگر یکسری نمونه‌هایی است که i امین برآمد تست پتانسیلی را دارند، sd بیانگر انحراف معیار است. بدلیل فرآیند انشعاب، داده در گره‌های فرزند انحراف معیار کمتری از گره مادر داشته و بنابراین خالص‌تر است. پس از بیشینه نمودن تمامی انشعاب‌های ممکن، M5P صفتی را انتخاب می‌کند که کاهش مورد انتظار را بیشینه می‌نماید. این تقسیم اغلب ساختار شبه درختی بزرگی می‌سازد که باعث بیش‌برازش می‌گردد. برای غلبه بر مسئله‌ی بیش‌برازش، درخت بایستی به عقب هرس شود. بنابراین، مرحله دوم در طراحی مدل درختی شامل هرس نمودن درخت رشد یافته و جایگزینی درختان فرعی با توابع رگرسیونی خطی است. این تکنیک تولید مدل درختی، فضای پارامتر را به زیرفضاهایی منشعب نموده و در هر کدام از آن‌ها یک مدل رگرسیونی خطی می‌سازد (Quinlan, 1992).

داده‌های مورد استفاده

در این مطالعه از داده‌های ماهانه ایستگاه هواشناسی سینوپتیک ساری که شامل متوسط دمای هوا (C)، ساعات آفتابی (ساعت)، دمای نقطه شبنم (C)، متوسط رطوبت

است در حالیکه بیشتر تکنیک‌های آماری معمول نیاز به پیش فرض دارند و تحلیل گران داده‌ها در جستجوی معادلاتی برای مطابقت دادن با پیش فرض‌ها هستند. در مقابل الگوریتم‌های داده‌کاوی می‌توانند این معادلات را به طور اتوماتیک از اطلاعات موجود در مجموعه داده‌ها توسعه دهند (کبنا و همکاران، ۱۹۹۸).

درخت تصمیم‌گیری

درخت تصمیم‌گیری به شکل وارونه می‌باشد. به طوریکه یک ریشه در بالای آن و در پایین برگ‌های آن وجود دارند. زمانی که یک داده وارد ریشه می‌شود بر روی آن یک آزمون انجام می‌گیرد تا مشخص شود که این داده باید به کدامیک از شاخه‌های پایین‌تر وارد شود. با وجود اینکه روش‌های مختلفی برای تعیین آزمون اولیه وجود دارد اما در نهایت هدف تمامی آزمون‌ها انجام بهترین جداسازی ممکن می‌باشد. این عمل انقدر تکرار می‌شود تا یک رکورد به یک گره برگ برسد. تمامی رکوردهایی که به یک گره برگ می‌رسند در یک کلاس قرار می‌گیرند. برای رسیدن از گره ریشه تا برگ تنها یک راه وجود دارد و آن راه در واقع یکی از قوانین تولید شده توسط درخت می‌باشد. در واقع الگوریتم درخت تصمیم‌گیری با انتخاب آزمونی آغاز می‌شود که بهترین جداسازی را برای دسته‌ها انجام دهد. در مراحل بعدی همین کار برای گره‌های بعدی با داده‌های کمتر صورت می‌گیرد تا بهترین قانون حاصل شود. درخت تصمیم‌قادر به تولید توصیفات قابل درک برای انسان، از روابط موجود در یک مجموعه داده‌ای است و می‌تواند به شکل تکنیک‌های ریاضی و محاسباتی که به توصیف و سته‌بندی یک مجموعه از داده‌ها کمک می‌کنند، نیز معرفی شوند (Kantardzic, 2003).

الگوریتم M5P

الگوریتم M5P یک درخت رگرسیونی دودویی است که در گره‌های نهایی خود (برگ‌ها) توابع رگرسیونی خطی دارند که می‌تواند صفات عددی پیوسته تولید نمایند.

14- standard deviation

نسبی (٪)، متوسط سرعت باد (متر بر ثانیه)، کمبود فشار بخار اشباع (میلی بار) و تبخیر تعرق (میلی متر بر روز) محاسبه شده از رابطه ی پنمن مانیتث طی دوره چهل ساله از سال ۱۳۶۷ تا ۱۳۹۷ میلادی بوده است. از ۷۵ درصد داده ها به عنوان تولید مدل با استفاده از مدل M5P و از ۲۵ درصد آن به عنوان آزمون مدل استفاده گردید.

به منظور اجرای الگوریتم M5P باید دو دسته داده (داده های مستقل و وابسته) به الگوریتم معرفی کرد. برای پیش بینی تبخیر تعرق پتانسیل توسط الگوریتم M5P، از شش پارامتر هواشناسی ماهانه (ساعات آفتابی (h)، دمای نقطه شبنم (°C)، متوسط رطوبت نسبی (٪)، متوسط سرعت باد (m/s)، کمبود فشار بخار اشباع (mb)) به عنوان داده های مستقل و داده های تبخیر تعرق پتانسیل ماهانه (mm/day) به عنوان داده وابسته (هدف) به الگوریتم معرفی شد.

همچنین به منظور پیش بینی کمینه و بیشینه دمای هوا توسط الگوریتم M5P، از پنج پارامتر هواشناسی ماهانه (ساعات آفتابی (h)، دمای نقطه شبنم (°C)، متوسط رطوبت نسبی (٪)، متوسط سرعت باد (m/s)، کمبود فشار بخار اشباع (mb)) به عنوان داده های مستقل و داده های کمینه و بیشینه دمای هوا (°C) به عنوان داده وابسته (هدف) به الگوریتم معرفی شد.

نتایج

نتایج پیش بینی الگوریتم M5P توسط شاخص های ضریب همبستگی (R)، ریشه متوسط خطا (RMSE) و متوسط مطلق خطا MAE ارزیابی شد. نتایج ارزیابی عملکرد مدل درختی M5P در جدول ۱ نشان داده شده

است. با توجه به جدول (۱) عملکرد الگوریتم M5P در پیش بینی تبخیر تعرق پتانسیل نسبت دمای بیشینه و کمینه بهتر بوده است.

با توجه به عملکرد بهتر الگوریتم M5P در برآورد تبخیر تعرق پتانسیل، نمودار درختی و رگرسیون خطی مربوط به آن در شکل (۱) نشان داده شده است. یکی از مزیت های مدل درختی M5P در دسترس قراردادن یک ترکیب رابطه خطی ساده به صورت مدل درختی (شکل ۱) است که بتواند برای پیش بینی تبخیر تعرق پتانسیل استفاده شود. تولید مدل درختی در مرحله اول شامل تعیین مناسبترین پارامتر ورودی برای انشعاب و نیز ضابطه تقسیم برای تولید یک درخت تصمیم است. با توجه به اینکه کمبود فشار بخار اشباع (میلی بار) به عنوان اولین پارامتر ورودی به مدل درختی است، بنابراین به عنوان مناسب ترین پارامتر ورودی به تشخیص الگوریتم M5P می باشد. در ادامه نیز در شاخه های مدل درختی، پارامتر کمبود فشار بخار اشباع (میلی بار) استفاده شده است که نشان دهنده تاثیر زیاد این پارامتر در پیش بینی تبخیر تعرق پتانسیل توسط الگوریتم M5P می باشد. در شکل زیر RH: متوسل رطوبت نسبی (درصد)، T: متوسط دمای روزانه (سانتیگراد)، W: سرعت باد (متر بر ثانیه)، e: کمبود فشار بخار اشباع (mb)، DEW: دمای نقطه شبنم (°C) و n: ساعات آفتابی (ساعت) می باشد.

در داخل هر برگ انتهایی شکل (۲) یک رگرسیون خطی (LM) مربوط به همان برگ می باشد. به عنوان مثال برای رگرسیون خطی L1 می توان گفت در ابتدای نمودار درختی داده های مربوط به کمبود فشار بخار اشباع به دو شاخه کمتر-مساوی و بیشتر از ۱۸/۷ (mb) تقسیم شده است. اگر داده های کمبود فشار بخار اشباع کمتر یا مساوی

جدول ۱: مقایسه عملکرد مدل M5P با سه شاخص آماری در تخمین تبخیر تعرق پتانسیل، دمای بیشینه و کمینه ماه آبی در ایستگاه هواشناسی سینوپتیک ساری

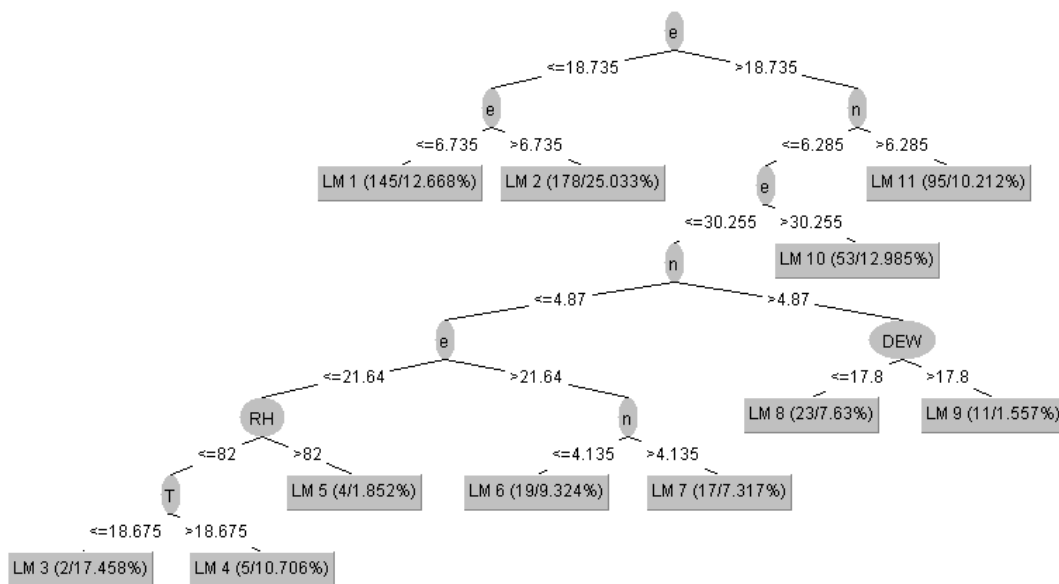
MAE	RMSE	R	
۰/۱۲	۰/۲۵	۰/۹۲	تبخیر تعرق پتانسیل
۱/۳۶	۲/۱۲	۰/۷۸	دمای کمینه
۱/۱۲	۱/۸۲	۰/۸۱	دمای بیشینه

۱۸/۷ (mb) باشند، به شاخه دیگر که با مقدار کمبود فشار بخار اشباع برابر ۶/۷ (mb) تقسیم شده است، هدایت می شوند. اگر داده های کمبود فشار بخار اشباع کمتر یا مساوی ۶/۷ باشد به سمت برگ انتهایی (L₁) منتقل می شوند. در برگ L₁ نیز رگرسیون خطی مورد نظر آورده شده است. با قرار دادن پارامترهای مورد نظر در معادله L₁ می توان مقدار تبخیر تعرق پتانسیل را پیش بینی کرد.

آنالیز حساسیت

برای تعیین مهمترین عوامل موثر در پیش بینی

تبخیر تعرق پتانسیل ماهانه توسط مدل درختی M5P با تغییر دادن متغیرهای ورودی صورت گرفت. نتایج مدل M5P با توجه به ترکیب پارامترهای ورودی توسط سه شاخص آماری ضریب رگرسیون، متوسط مطلق خطا و ریشه میانگین خطا مورد ارزیابی قرار گرفت. طبق جدول (۲) ردیف پنجم که شامل پنج پارامتر دمای متوسط، ساعات آفتابی، درصد رطوبت هوا، دمای نقطه شبنم و کمبود فشار بخار اشباع با مقدار ضریب همبستگی برابر ۰/۹۶، متوسط مطلق خطا برابر ۰/۱۰ و ریشه میانگین خطا برابر ۰/۱۵ بیشترین عملکرد صحیح مدل M5P در



شکل ۲: نمودار درختی الگوریتم M5P در پیش‌بینی تبخیر تعرق پتانسیل

جدول ۲: معادلات خروجی از مدل درختی M5P به منظور پیش‌بینی تبخیر تعرق پتانسیل ماهانه

معادلات	
LM 1:	$ET = (0.0063 * T) - (0.0087 * RH) + (0.0094 * W) + (0.0167 * n) - (0.0031 * DEW) + (0.0857 * e) + 1.0577$
LM 2:	$ET = (0.0951 * T) - (0.0301 * RH) + (0.0094 * W) + (0.1154 * n) - (0.065 * DEW) + (0.0475 * e) + 2.4679$
LM 3:	$ET = (0.0619 * T) - (0.0141 * RH) + (0.2756 * W) + (0.0905 * n) - (0.0548 * DEW) + (0.0276 * e) + 2.1807$
LM 4:	$ET = 0.0619 * T - 0.0141 * RH + 0.2756 * W + 0.0905 * n - 0.0548 * DEW + (0.0276 * e) + 2.1899$
LM 5:	$ET = (0.065 * T) - (0.0151 * RH) + (0.2756 * W) + (0.0905 * n) - (0.0548 * DEW) + (0.0276 * e) + 2.1576$
LM 6:	$ET = (0.1 * T) - (0.005 * RH) + (0.4425 * W) + (0.1321 * n) - (0.0711 * DEW) + (0.0218 * e) + 0.8771$
LM 7:	$ET = (0.083 * T) - (0.005 * RH) + (0.3346 * W) + (0.1136 * n) - (0.0721 * DEW) + (0.0218 * e) + 1.4021$
LM 8:	$ET = (0.0791 * T) - (0.0035 * RH) + (0.3163 * W) + (0.198 * n) - (0.0765 * DEW) + (0.0169 * e) + 1.2517$
LM 9:	$ET = (0.0977 * T) - (0.0043 * RH) + (0.3163 * W) + (0.1962 * n) - (0.0831 * DEW) + (0.0169 * e) + 0.9716$
LM 10:	$ET = (0.1182 * T) - (0.0023 * RH) + (0.1494 * W) + (0.2664 * n) - (0.0778 * DEW) + (0.0136 * e) + 0.368$
LM 11:	$ET = (0.0846 * T) - (0.001 * RH) + (0.2738 * W) + (0.2285 * n) - (0.0344 * DEW) + (0.0285 * e) - 0.1164$

جدول ۳: تحلیل حساسیت مدل درختی M5P برای ایستگاه سینوپتیک ساری

RMSE	MAE	R	ترکیب پارامترهای هواشناسی
0.25	0.12	0.92	T,n,w,RH,dwe,e
0.35	0.22	0.89	T,n,w,RH,dwe
0.36	0.32	0.88	T,n,w,RH,e
0.38	0.35	0.88	T,n,w,dwe,e
0.15	0.10	0.96	*T,n,RH,dwe,e
0.41	0.42	0.8	T,n,w,dwe
0.48	0.45	0.79	T,n,w,e
0.49	0.43	0.81	T,n,w,RH
0.57	0.53	0.75	T,n,w
0.59	0.51	0.72	T,n
0.57	0.53	0.75	T,w
0.75	0.61	0.70	n,w

و بیشینه ایستگاه هواشناسی ساری می باشد. با استفاده از الگوریتم درختی M5P و فرمول پنمن مانیتث، می توان به عنوان الگوریتم پیش بینی کننده تبخیر و تعرق پتانسیل با دقت مناسب، برای ماه های آتی استفاده کرد. یکی از مزیت های الگوریتم درختی M5P استفاده از داده های کمتری نسبت به روابط معرفی شده توسط محققان مختلف، در تعیین مقدار تبخیر تعرق می باشد. در تحقیق که توسط میرهاشمی و پناهی (۱۳۹۲) به منظور عملکرد سه مدل M5P, REPTree و Kstar در تخمین تبخیر و تعرق پتانسیل و رطوبت نسبی ماهانه ماه بعد، مربوط به ایستگاه هواشناسی سینوپتیک یزد مورد ارزیابی قرار گرفت، مشخص شد که مدل درختی M5P مدل مناسب تری برای تخمین تبخیر تعرق پتانسیل ماهانه ماه بعد می باشد. از بین عوامل هواشناسی مور مطالعه در این تحقیق، کمبود فشار بخار اشباع (میلی بار) بیشترین تأثیر در برآورد تبخیر تعرق پتانسیل توسط الگوریتم M5P داشته است. در آنالیز حساسیت مشخص شد که پنج پارامتر ورودی شامل دمای متوسط، ساعات آفتابی، درصد رطوبت هوا، دمای نقطه شبنم و کمبود فشار بخار اشباع بیشترین و دو پارامتر ورودی شامل ساعات آفتابی و سرعت باد کمترین عملکرد صحیح مدل M5P در پیش بینی تبخیر تعرق پتانسیل داشته اند. در آخر مشخص شد که سرعت

پیش بینی نسبت به ترکیب پارامترهای دیگر هواشناسی داشته است. کمترین عملکرد درست پیش بینی M5P مربوط به پارامترهای ورودی ساعات آفتابی و سرعت باد در ردیف آخر با مقدار ضریب همبستگی برابر ۰/۷۰، متوسط مطلق خطا برابر ۰/۴۳ و ریشه میانگین خطا برابر ۰/۱۵ است.

با توجه به جدول ۳ بهترین عملکرد در زمانی بوده است که تمام عوامل هواشناسی مورد مطالعه به غیر از سرعت باد به عنوان پارامترهای ورودی استفاده شدند. در آخرین ردیف جدول ۳ نیز هنگامی که پارامترهای ورودی شامل ساعات آفتابی و سرعت باد بوده است پیش بینی الگوریتم M5P ضعیف ترین عملکرد را داشته است. با توجه به اینکه در رابطه فائو- پنمن- مونیتث، اثر سرعت باد بر روی مقدار تبخیر و تعرق، به صورت غیرخطی هست (حامی کوچه باغی و همکاران، ۱۳۹۵) و اینکه مدل M5P یک مدل درختی رگرسیونی است، سرعت باد تأثیر منفی در عملکرد صحیح الگوریتم M5P در پیش بینی تبخیر تعرق گذاشته است.

نتیجه گیری

با توجه به نتایج مشخص شد عملکرد الگوریتم M5P در پیش بینی تبخیر تعرق پتانسیل بهتر از دمای کمینه

باد دارای بیشترین تأثیر منفی در برآورد تبخیر تعرق پتانسیل توسط الگوریتم M5P داشته است.

منابع

- به منش، ج.، آزاد طلائی، ن.، منتصری، م.، بشارت، س. ۱۳۹۳. ارزیابی مدل های سری زمانی خطی و غیر خطی بیلینیر در پیش بینی تبخیر- تعرق گیاه مرجع در ایستگاه سینوپتیک ارومیه. نشریه پژوهش آب در کشاورزی، دوره ۲۸، ۱، شماره ۱ ص ۸۵-۹۶
- پناهی، م. و میرهاشمی، س.ج.، ۱۳۹۴. ارزیابی دو الگوریتم داده کاوی CART و CHAID در پیش بینی دمای هوای ایستگاه سینوپتیک اراک. فصلنامه علوم محیطی، دوره سیزدهم، شماره ۴. حامی کوچه باغی، م.، ناظمی، ام.، صدرالدینی، ا.ع.، دلیرحسن نیا، ر. ۱۳۹۵. نگرش ریاضی به نقش مؤلفه سرعت باد در رابطه فائو- پنمن - موتنتیت برای محاسبه تبخیر- تعرق مرجع. نشریه دانش آب و خاک. ۲۶ (۲): ۱-۱۴
- سازمان آمار ایران، ۱۳۸۹. سالنامه آماری استان مازندران ستاری، م.ت.، نهرین، ف.، عظیمی، و. ۱۳۹۲. پیش بینی تبخیر- تعرق مرجع روزانه با استفاده از مدل شبکه عصبی مصنوعی و مدل درختی M5 مطالعه موردی: ایستگاه بناب. (نشریه آبیاری و زهکشی ایران). ۱۰۷. ۱۰۴-۱۱۳
- قهرمان، نوذر؛ بابائیان، ایمان؛ طباطبائی، سیدمحمد رضا. ۱۳۹۵. بررسی اثرات تغییر اقلیم بر نیاز آبی و طول دوره رشد گیاه نیشکر تحت سناریوهای واداشت تابشی، نشریه حفاظت منابع آب و خاک، ۶۳-۷۵. (۶)
- محمدپورزیدی، ع.، علیجانی، ب.، سلیمه، م.، گرامی، م. ص. ۱۳۹۸. تحلیل فضایی بارش استان مازندران، نشریه تحقیقات کاربردی علوم جغرافیایی سال نوزدهم، شماره ۲۵
- مشکانی، علی و ناظمی، عبدالرضا. ۱۳۸۸، مقدمه ای بر داده کاوی، انتشارات دانشگاه آزاد اسلامی واحد نیشابور، ص ۴۵۶.
- میرهاشمی، س.ج. و پناهی، م. ۱۳۹۲. ارزیابی سه مدل از نرم افزار weka در پیش بینی تبخیر تعرق پتانسیل و رطوبت نسبی ماهانه برای ماه بعد (مطالعه موردی: استان یزد). اولین کنگره علمی پژوهشی توسعه و ترویج علوم کشاورزی، منابع طبیعی و محیط زیست ایران، ۲۸ و ۲۹ شهریور ۱۳۹۲
- میرهاشمی، س.ج. و پناهی، م. ۱۳۹۲. ارزیابی مدل CART در تخمین تبخیر تعرق پتانسیل و درصد رطوبت نسبی هوای ماه بعد (مطالعه موردی ایستگاه اراک). اولین همایش سراسری کشاورزی و منابع طبیعی پایدار، ۱۰ بهمن ۱۳۹۲.
- Allen RG., Pereira LS., Raes D. and Smith M. 1998. Crop evapotranspiration, Irrigation and Drainage Paper No. 56. Food and Agriculture Organization: Rome, Italy.
- Asadi Zarch, M. A., Sivakumar, B., Sharma .A. 2015 . Droughts in a warming climate: a global assessment of Standardized precipitation index (SPI) and Reconnaissance drought index (RDI). Journal of Hydrology, 526, 183-195.
- Asadi Zarch, M.A., Sivakumar, B., Malekinezhad, H., Sharma, A. 2017 . Future aridity under conditions of global climate change. Journal of Hydrology, 554, 451-469.
- Cabena PH., Stadler R., Verhees J., and Zanasi .: 1998, Discovering data mining: From concept to implementation, IMB, New Jersey, 195 pp.
- Ding, R., Kang, S., Zhang, Y., Hao, X., Tong, L., Li, S. 2015. A dynamic surface conductance to predict crop water use from partial to full canopy cover. Agricultural Water Management, 150, 1-8.
- Djaman, k., Balde, A., Sow, A., Muller, B., Irmak, S., N'Diaye, M., Manneh, B., D.Moukoumbi, Y., Futakuchi, K., Saito, K. 2015. Evaluation of sixteen reference evapotranspiration methods under sahelian conditions in the Senegal River Valey. Journal of Hydrology: Regional Studies 3, pp:139-159
- Giudici P. 2003. Applied data Mining: statistical methods for business and industry. Wily, London . pp. 364
- Han, X., Liu, W., Lin, W. 2015 . Spatiotemporal analysis of potential evapotranspiration in the Changwu tableland from 1957 to 2012. Meteorological Applications, 22, 586-591.
- Itenfisu D., Elliott, RL., Allen RG. and Walter IA. 2003. Comparison of reference evapotranspiration calculations as part of the ASCE standardization effort. Journal of Irrigation and Drainage Engineering, ASCE 129: 440-448
- Jensen ME., Burman RD. and Allen RG. 1990. Evapotranspiration and irrigation water requirements. ASCE Manual and Reports on Engineering Practice No. 70. ASCE: New York.
- Kantardzic, M. 2003. Advances in Data Mining.” Data Mining: Concepts, Models, Methods, and Algorithms, Second Edition. 328-384.
- Martí, P., González-Altozano, P., López- Urrea, R., Mancha, L.A., Shiri, J. 2015 . Modeling reference evapotranspiration with calculated targets. Assessment and implications. Agricultural Water Management, 149, 81-90.
- Monteith JL. 1965. The state and movement of water in living organisms. Proceedeings of Evaporation and Enviroment, XIX Symposium, Society for Experimental Biology, Swansea. Cambridge University Press: New York: 205-234.
- Pandey, P., P.Dabral, P., Pandey, V. 2016. Evaluation of reference evapotranspiration methods for the northeastern region of India. International Soil and Water Conservation Research 4. pp: 52-63
- Quinlan, John R. 1992. “Learning with continuous classes.” 5th Australian joint conference on artificial intelligence. Vol. 92.
- Stanelle, T., Vogel, B., Vogel, H., Baumer, D., Kottmeier, C. 2010. Feedback between dust particles and atmospheric processes over West Africa during dust episodes in March 2006 and June 2007. Atmos. Chem. Phys., 10: 10771-10788.
- Two Crows Corporation, 1999. Introduction to data mining and knowledge discovery, third ed., Postmac, MD. Available at: www.twocrows.com, (April 29, 2000)
- Vanderberg H., Sogard P. and Motoroni S. 1999. MineSetTM 3.0 Enterprise Edition Tutorial for Windows, Doc. No. 007-4006-001, Silicon Graphics Inc.

Evaluation of M5P Algorithm for Estimation of Potential Evapotranspiration, Minimum and Maximum Temperature (Case study: Sari Weather Station)

Seyed Hassan Mirhashemi¹, Mehdi Panahi^{2,*}, Leila Zareei³

¹ Graduated Ph.d, Dept. of Water Engineering, Faculty of Water and Soil, University of Zabol, Iran

² Assistant Professor, Dept. of Water Engineering, Faculty of Agriculture, University of Zanjan, Iran

³ Master student of water resources engineering, Imam khomeini international university, Iran

*Corresponding Author Email: panahi40@yahoo.com

Received: 19 July 2019, accepted: 13 October 2019

ABSTRACT

In this study, we evaluated the performance of M5P tree model in estimating potential evapotranspiration, minimum and maximum air temperature. The monthly mean data of Sari synoptic weather station during the 30 year period from 1989 to 2019 were analyzed. According to the results, M5P model with correlation coefficient of 0.92, RMSE of 0.25 and MAE of 0.12, has better performance in predicting potential evapotranspiration compare with prediction of minimum and maximum air temperature. According to sensitivity analysis, M5P algorithm with inputs such as mean air temperature, sunshine hours, air humidity, dew point temperature and saturation vapor pressure deficit had the best and with inputs such as sunshine hours and wind speed had the worst performance in predicting potential evapotranspiration. It was also found that the wind speed had a negative effect on the prediction of evapotranspiration by the M5P model.

Keywords: Sensitivity analysis, wind speed, correlation coefficient, meteorological parameter

HOW TO CITE THIS ARTICLE

Mirhashemi, S.H.; Panahi, M.; Zareei L. (2020). Evaluation of M5P Algorithm for Estimation of Potential Evapotranspiration, Minimum and Maximum Temperature (Case study: Sari Weather Station). *J. Meteorol. Atmos. Sci.*, 2(4): 287-295.

COPYRIGHTS

Copyright for this article is retained by the author(s), with publication rights granted to the JMAs Journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

