



مدیریت آب و آبیاری

دوره ۴ ■ شماره ۲ ■ پاییز و زمستان ۱۳۹۳

صفحه‌های ۱۶۰-۱۴۹

مقایسه مدل رگرسیون درختی، شبکه عصبی مصنوعی و هارگریوز سامانی در برآورد تبخیر تعرق مرجع مناطق خشک

میترا بخشوده^{۱*} و علی رحیمی خوب^۲

۱. کارشناس ارشد گروه مهندسی آبیاری و زهکشی پردیس ابوریحان دانشگاه تهران

۲. استاد گروه مهندسی آبیاری و زهکشی پردیس ابوریحان دانشگاه تهران

تاریخ پذیرش مقاله: ۱۳۹۳/۰۴/۰۷

تاریخ وصول مقاله: ۱۳۹۲/۰۴/۰۷

چکیده

هدف از این تحقیق، ارزیابی سه مدل شبکه عصبی مصنوعی، رگرسیون درختی و مدل هارگریوز سامانی برای برآورد تبخیر تعرق گیاه مرجع بود. بدین منظور از اطلاعات هواشناسی استان‌های سیستان و بلوچستان، کرمان، یزد و خراسان جنوبی در دوره آماری ۱۹۹۸-۲۰۰۸ استفاده شد. با توجه به تأثیر سرعت باد بر میزان تبخیر تعرق منطقه، برآورد تبخیر تعرق براساس تغییرات سرعت باد در قالب سه گروه شامل ایستگاه‌هایی با سرعت باد کمتر از ۲/۴۸ متر بر ثانیه در گروه ۱، ایستگاه‌هایی با سرعت باد بین ۲/۴۸ و ۳/۶۷ متر بر ثانیه در گروه ۲ و ایستگاه‌هایی با سرعت باد بیشتر از ۳/۶۷ متر بر ثانیه در گروه ۳ صورت گرفت. نتایج نشان داد هر سه مدل شبکه عصبی مصنوعی، رگرسیون درختی و هارگریوز سامانی در گروه ۱ بهترین برآورد را داشتند. نتایج آماری میانگین مجذور مربعات خطای مدل شبکه عصبی مصنوعی ۱/۴۱ میلی‌متر در روز با ضریب تبیین ۰/۸۴، مدل رگرسیون درختی ۱/۴۶ میلی‌متر در روز با ضریب تبیین ۰/۸۳ و هارگریوز سامانی کالیبره شده ۲/۰۲ میلی‌متر در روز با ضریب تبیین ۰/۶۹ به دست آمد که بیانگر برتری دو روش شبکه عصبی و رگرسیون درختی بود.

کلیدواژه‌ها: تبخیر تعرق گیاه مرجع، رگرسیون درختی، سرعت باد، شبکه عصبی مصنوعی، هارگریوز سامانی.

۱. مقدمه

تبخیر تعرق^۱ (ET)، شاخص تعیین‌کننده‌ای در فرایند رشد گیاه است که معادل آب مورد نیاز گیاهان زراعی قلمداد می‌شود. به‌همین دلیل تخمین آن با توجه به وضعیت آب‌وهوایی هر منطقه و گیاه صورت می‌گیرد. طی نیم قرن گذشته مدل‌های متعددی به‌صورت روابط ساده تجربی تا معادلات پیچیده ریاضی، به‌منظور برآورد تبخیر تعرق مرجع با استفاده از پارامترهای مختلف هواشناسی ارائه شده است. با توجه به تحقیقات صورت گرفته، روش پنمن مانیتث یکی از دقیق‌ترین روش‌های تخمین تبخیر تعرق گیاه مرجع (ET₀ PM) است که در اقلیم‌های مختلف کاربرد دارد (۲). روش پنمن مانیتث، مبنای فیزیکی دارد که نیاز به پارامترهای ورودی زیادی است. دسترسی به این پارامترها نیازمند صرف وقت و هزینه زیاد است. هرچند انتخاب پارامترهای بیشتر، در راستای افزایش دقت برآوردها است، ذکر این نکته ضروری است که دسترسی به کلیه پارامترها در همه مناطق امکان‌پذیر نیست (۲). بنابراین یافتن مدلی که با استفاده از حداقل داده‌های هواشناسی، مقدار تبخیر تعرق را برآورد کند، بسیار مفید خواهد بود. در چند سال اخیر استفاده از روش‌های حل هوشمند هم به‌عنوان راهکاری مناسب و سریع در بسیاری از زمینه‌ها از جمله برآورد نیاز آبی گیاهان رایج شده است. رهیافت شبکه‌های عصبی مصنوعی به‌دلیل طبیعت غیرخطی و کارآمد بودن در مدل‌سازی سامانه‌های پیچیده مورد توجه بسیاری از محققان است. به‌تازگی مدل رگرسیون درختی^۲ (مدل درختی M5) نیز با موفقیت در چندین برنامه کاربردی منابع آب استفاده شده است. این مدل یک روش کارآمد مبتنی بر مدل‌های رگرسیون‌گیری خطی است. در مطالعه‌ای پتانسیل مدل

رگرسیون درختی برای مدل‌سازی مقادیر روزانه ET₀ با استفاده از داده‌های هواشناسی ایستگاه دیویس کالیفرنیا بررسی شد. بررسی‌های آماری نشان داد که مدل رگرسیون درختی از توانایی بیشتری برای برآورد تبخیر تعرق پتانسیل در مقایسه با روش‌های فائو-پنمن-مانیتث و هارگریوز سامانی برخوردار است (۱۴). در مطالعه‌ای دیگر با استفاده از رگرسیون درختی و به‌کارگیری اطلاعات میدانی باد و موج ثبت‌شده منطقه نکا، ارتفاع امواج ناشی از باد پیش‌بینی و نتایج مدل یادشده با شبکه‌های عصبی مصنوعی مقایسه شد. نتایج حاکی از دقت بیشتر مدل‌های رگرسیون درختی نسبت به شبکه‌های عصبی مصنوعی در منطقه تحقیق بود (۴). در مطالعه دیگری از مدل رگرسیون درختی، شبکه عصبی مصنوعی و ترکیب این دو مدل برای پیش‌بینی سیلاب در مناطق بالادست رودخانه‌های چین استفاده شد. براساس نتایج، هر دو مدل عملکرد خوبی برای پیش‌بینی سیلاب داشتند، اما ترکیب دو مدل رگرسیون درختی و شبکه عصبی مصنوعی بیشترین دقت را برای پیش‌بینی سیلاب در مقایسه هر یک از مدل‌های رگرسیون درختی و شبکه عصبی مصنوعی داشت (۱۷). مقایسه برآوردهای شبکه عصبی مصنوعی و مدل‌های تجربی پنمن، تورک و هارگریوز سامانی بیانگر عملکرد مناسب شبکه عصبی مصنوعی در مدل کردن تبخیر تعرق در کشور ترکیه بود (۱۲). مقایسه کارایی شبکه‌های عصبی مصنوعی و روش هارگریوز سامانی در تخمین تبخیر تعرق گیاه مرجع برای دوازده ایستگاه هواشناسی خوزستان نشان‌دهنده برتری شبکه عصبی مصنوعی در برآورد ET₀ بود (۱۵).

معادله هارگریوز سامانی در نقاطی که داده‌های کافی برای استفاده از روش پنمن مانیتث وجود ندارد به‌خوبی میزان تبخیر تعرق را برآورد می‌کند (۹). این معادله از دقت به‌نسبت خوبی در مناطق خشک و نیمه‌خشک برخوردار بوده و مستلزم دمای بیشینه و کمینه است که در اکثر

1. Evapotranspiration
2. M5' Regression Tree

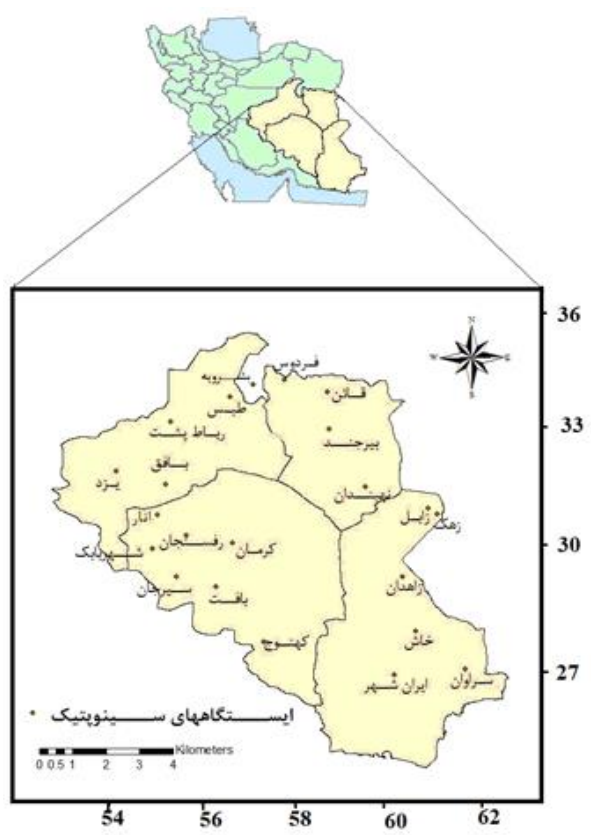
مدیریت آب و آبیاری

دمای روزانه حداکثر (T_x) و حداقل هوا (T_n)، رطوبت نسبی (RH)، ساعات آفتابی واقعی (n) و سرعت باد (U) است. با توجه به اینکه سرعت باد تأثیر زیادی بر میزان تبخیر تعرق دارد (۱) و با در نظر گرفتن این نکته که باد مهم‌ترین شاخص هواشناسی غالب در مناطق شرق و جنوب شرقی کشور است (۳)، در این تحقیق ایستگاه‌های موجود با توجه به سرعت باد در منطقه به سه گروه شامل ایستگاه‌هایی با سرعت باد کمتر از $2/48$ متر بر ثانیه در گروه U_1 ، ایستگاه‌هایی با سرعت باد بین $2/48$ و $3/67$ متر بر ثانیه در گروه U_2 و ایستگاه‌هایی با سرعت باد بیشتر از $3/67$ متر بر ثانیه در گروه U_3 طبقه‌بندی شدند و مدل‌سازی‌ها برای هر گروه به تفکیک انجام گرفت.

ایستگاه‌های هواشناسی به راحتی قابل اندازه‌گیری است. بنابراین هدف این تحقیق، تدوین مدل رگسیون درختی برای برآورد تبخیر تعرق مرجع با استفاده از حداقل داده‌های هواشناسی و مقایسه مدل تدوین شده رگسیون درختی و مدل شبکه عصبی مصنوعی و مدل هارگریوز سامانی برای مناطق خشک مورد تحقیق است.

۲. مواد و روش‌ها

داده‌های مورد نیاز این تحقیق از ۲۲ ایستگاه هواشناسی وابسته به سازمان هواشناسی واقع در استان‌های کرمان، یزد، سیستان و بلوچستان و خراسان جنوبی طی سال‌های ۱۹۹۸ تا ۲۰۰۸ میلادی با موقعیت‌های جغرافیایی نشان داده شده در شکل ۱ جمع‌آوری شد. داده‌ها شامل مقادیر متوسط



شکل ۱. موقعیت ایستگاه‌های هواشناسی در محدوده تحقیق

مدیریت آب و آبیاری

دوره ۴ ■ شماره ۲ ■ پاییز و زمستان ۱۳۹۳

۱.۲. مدل شبکه عصبی مصنوعی

در این تحقیق از شبکه‌های چندلایه پیش‌رونده با الگوریتم آموزشی پس‌انتشار خطا استفاده شد. ساختار شبکه عصبی مصنوعی شامل یک لایه ورودی، یک لایه میانی و یک لایه خروجی است. در هر لایه یک یا چند عنصر پردازشگر (نورون) وجود دارد که با تمامی نورون‌های لایه بعدی با اتصالات وزن‌دار مربوط می‌شوند. بردار داده‌های ورودی مدل به نورون‌های لایه اول وارد می‌شوند و در این لایه هیچ پردازشی انجام نمی‌گیرد و نورون‌های لایه خروجی به بردار خروجی مدل وارد می‌شوند. تعداد نورون‌های لایه‌های ورودی و خروجی، به تعداد متغیرهای ورودی و خروجی مدل بستگی دارد، ولی تعداد نورون‌های لایه میانی با آزمون و خطا تعیین می‌شود. لایه ورودی شبکه شامل دمای کمینه، دمای بیشینه و تابش فرازمینی بود (R_a) که به‌عنوان متغیرهای ورودی برای تخمین تبخیرتعلق مرجع استفاده شد. شایان ذکر است که مقدار تابش فرازمینی براساس رابطه پیشنهادی آلن و همکاران به‌شرح زیر محاسبه شد (۵).

$$R_a = 37.6dr (w_s \sin \phi \cdot \cos \delta + \cos \phi \cdot \cos \delta \cdot \sin w_s) \quad (1)$$

$$W_s = \arccos(-\tan \phi \cdot \tan \delta) \quad (2)$$

$$dr = 1 + 0.033 \cos(0.0172 J) \quad (3)$$

$$\delta = 0.409 \sin(0.0172 J - 1.39) \quad (4)$$

که R_a تابش برون‌زمینی (مگاژول بر متر مربع بر روز)، dr فاصله نسبی زمین تا خورشید، δ زاویه میل خورشید (رادیان)، ϕ عرض جغرافیایی (رادیان)، w_s زاویه ساعتی غروب خورشید (رادیان)، و J شماره روز ژولیوسی از ابتدای سال میلادی است.

برای مدلسازی از داده‌های ۲۲ ایستگاه، به‌منظور اجرای شبکه از شبکه پرسپترون چندلایه (MLP)، تابع انتقال

سیگموئید^۱ (S) و قانون آموزش لونیبرگ مارکواریت^۲ (LM) استفاده شد که با توجه به تحقیقات صورت گرفته، کاربرد و تأثیر بیشتری از خود نشان داده‌اند (۱۱، ۱۸). بعد از آموزش، ارزیابی شبکه با داده‌هایی که در آموزش شرکت نداشتند انجام می‌گیرد و عملکرد آن با استفاده از شاخص‌های آماری بررسی می‌شود. در این پژوهش آموزش شبکه با تعداد نورون‌های مختلف لایه میانی (۱ تا ۱۰ نورون) و استفاده از داده‌های دمای بیشینه و کمینه هوا و تابش بیرون زمینی به‌عنوان نورون‌های لایه ورودی انجام گرفت. اجرای شبکه عصبی مصنوعی به تفکیک هر سه گروه سرعت باد (U_1 ، U_2 و U_3) انجام گرفت. ضمن آنکه از میان داده‌های آموزش و ارزیابی، ۷۰ درصد آن به آموزش و ۳۰ درصد بقیه به ارزیابی شبکه اختصاص داده شد (۲). کلیه عملیات مدلسازی شبکه عصبی مصنوعی در نرم‌افزار NeuroSolutions نسخه ۵ انجام گرفت.

۲.۲. مدل رگرسیون درختی

مدل رگرسیون درختی یک درخت تصمیم دودویی است که در هر گره پایانی (برگ) روابط رگرسیونی خطی دارد که می‌تواند مقادیر عددی را پیش‌بینی کند. تولید مدل درختی در دو مرحله صورت می‌گیرد: مرحله اول شامل تعیین مناسب‌ترین پارامتر ورودی برای انشعاب و نیز ضابطه (معیار) تقسیم برای تولید یک درخت تصمیم است. انتخاب پارامتر تقسیم مناسب و ایجاد نقطه انشعاب در درخت از اهمیت خاصی برخوردار است. روش‌های مختلفی برای انتخاب نقطه شکست وجود دارد. درختان تصمیم، نمونه‌ها را با مرتب کردن آنها در درخت از گره ریشه به سمت گره‌های برگ دسته‌بندی می‌کنند. هر گره

1. SigmoidAxon
2. Levenberg- Marquardt

۳.۲. معادله هارگریوز سامانی (HG)

علاوه بر دو مدل شبکه عصبی مصنوعی و رگرسیون درختی، از روش هارگریوز سامانی برای برآورد تبخیر تعرق مرجع نیز استفاده شد. هارگریوز و سامانی رابطه تجربی زیر را که مبتنی بر درجه حرارت است برای تعیین تبخیر تعرق گیاه مرجع مطرح کردند (۱۲):

(۵)

$$ET_0 = 0.0023 R_a (T_a + 17.8) (T_{max} - T_{min})^{0.5}$$

در معادله بالا، R_a : تابش بیرون زمینی برحسب عمق آب تبخیر شده است (mm d^{-1}); T_{max} و T_{min} حداکثر دمای روزانه هوا ($^{\circ}\text{C}$) هستند. برای واسنجی معادله فوق، آن و همکاران معادله خطی زیر را پیشنهاد کردند (۵):

$$ET_0 \text{ PM} = a + b (ET_0 \text{ HG}) \quad (۶)$$

در معادله بالا $ET_0 \text{ PM}$ و $ET_0 \text{ HG}$ به ترتیب، تبخیر تعرق مرجع محاسبه شده از معادله فائو پنمن مانتیث و هارگریوز سامانی بر حسب میلی متر در روز است. a و b به ترتیب عرض از مبدأ و شیب خط برازش اند. برای واسنجی معادله ۶ از داده های آموزش شبکه های عصبی استفاده شد. پس از واسنجی و تعیین ضرایب a و b ، آزمون معادله واسنجی با استفاده از داده های ارزیابی مدل های شبکه عصبی انجام گرفت.

۴.۲. معادله پنمن مانتیث فائو (PM)

در این تحقیق، نتایج معادله پنمن مانتیث فائو به عنوان مبنا برای واسنجی و تعیین دقت مدل های شبکه عصبی، شبکه درختی و هارگریوز سامانی استفاده شد. معادله پنمن مانتیث فائو به شرح زیر است.

(۷)

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

داخلی در درخت، صفتی از نمونه را آزمایش می کند و هر شاخه ای که از آن گره خارج می شود متناظر یک مقدار ممکن برای آن صفت است. همچنین به هر گره برگ، یک دسته بندی متناسب می شود. هر نمونه، با شروع از گره ریشه درخت و آزمایش صفت مشخص شده به وسیله این گره و حرکت در شاخه متناظر با مقدار صفت داده شده در نمونه، دسته بندی می شود. این فرایند برای هر زیردرختی که گره جدید ریشه آن است تکرار می شود. روش های مختلفی برای انتخاب نقطه شکست وجود دارد که یکی از آنها الگوریتم C4.5 است. در این الگوریتم انشعاب شاخه ها به صورت دوتایی و تنها براساس یک متغیر مستقل است. معیار تقسیم برای الگوریتم مدل درختی بر مبنای در نظر گرفتن انحراف معیار مقادیر دسته منتهی به هر گره، میزان خطا در آن گره به شمار می رود، و محاسبه کاهش مورد انتظار در این خطا، نتیجه آزمون هر صفت در آن گره است. در این تحقیق از نرم افزار WEKA 3.7 برای تولید مدل های درختی استفاده شد.

اجرای مدل رگرسیون درختی نیز مشابه شبکه عصبی مصنوعی برای هر یک از گروه های U_1 ، U_2 و U_3 ، جداگانه براساس داده های آموزش و آزمون انجام گرفت. با توجه به اینکه هدف از این تحقیق پیش بینی تبخیر تعرق با استفاده از حداقل داده های هواشناسی است، داده های دمای کمینه، دمای بیشینه و تابش فرازمینی مربوط به دوره آماری ۱۹۹۸-۲۰۰۸، پارامتر ورودی؛ و مقادیر تبخیر تعرق مرجع حاصل از معادله پنمن مانتیث، پارامتر مطلوب انتخاب شدند. انتخاب ایستگاه ها برای هر یک از مراحل آموزش و آزمون براساس پراکندگی ایستگاه ها برای هر یک از گروه های U_1 ، U_2 و U_3 طوری انجام گرفت که در هر یک از مراحل آموزش و آزمون، اقلیم های مختلف محدوده تحقیق را تحت پوشش قرار دهند.

مربعات خطا به میانگین مقادیر مبنا بر حسب درصد حساب می‌شود.

۳. نتایج و بحث

خلاصه اطلاعات هواشناسی و تبخیرتغرق گیاه مرجع در استان‌های مطالعه شده در این تحقیق در جدول ۱ ارائه شده است. همان‌گونه که ملاحظه می‌شود، بیشترین تبخیرتغرق معادل ۶/۵۳ میلی‌متر در روز به ایستگاه سیستان و بلوچستان، و کمترین آن به مقدار ۴/۴۵ میلی‌متر در روز به ایستگاه خراسان جنوبی تعلق دارد. از آنجا که برآورد تبخیرتغرق با اعمال پارامترهای هواشناسی صورت گرفته است، تغییرات تبخیرتغرق کاملاً هماهنگ با تغییرات دمای هوا است، به طوری که بیشترین دمای هوا (۳۰/۲۶ درجه سانتی‌گراد) به ایستگاه سیستان و بلوچستان، و کمترین دمای روزانه هوا به ایستگاه خراسان جنوبی (۲۵/۴۱ درجه سانتی‌گراد) تعلق دارد. درخصوص ضریب تغییرات تبخیرتغرق نیز بیشترین ضریب تغییرات در ایستگاه خراسان جنوبی و کمترین آن در ایستگاه کرمان اتفاق افتاده است که همسو با تغییرات پارامتر سرعت باد است. تحقیق پارسافر و همکاران نیز در راستای این نتیجه است (۱). به طوری که تحقیق آنان نشان داد تغییرات تبخیرتغرق با تغییرات پارامتر سرعت باد به مراتب بیش از سایر پارامترهای هواشناسی است. مطالعه طالبی و همکاران نیز مؤید تأثیر تغییرات سرعت باد بر نوسانات تبخیرتغرق ایستگاه‌های استان یزد بود (۳).

معادلات واسنجی هارگریوز سامانی با استفاده از داده‌های آموزش برای سه گروه سرعت باد تعیین شد. این معادلات به همراه نتایج آماری در جدول ۲ ارائه شده‌اند. همان‌طور که ملاحظه می‌شود، معادلات واسنجی برای سرعت‌های کمتر، از دقت و ضریب تعیین بهتری برخوردار است.

که در آن، $ET_0 PM$: تبخیرتغرق گیاه مرجع برآورد شده از معادله پنمن مانیتث فائو ($mm d^{-1}$)؛ Δ : شیب منحنی فشار بخار اشباع با دما ($Kpa \ ^\circ C^{-1}$)؛ R_n : تابش خالص رسیده به زمین ($MJ m^{-2} d^{-1}$)؛ G : شار گرمای خاک ($MJ m^{-2} d^{-1}$)؛ γ : ثابت سایکرومتر ($Kpa \ ^\circ C^{-1}$)؛ T_a : متوسط دمای روزانه هوا ($^\circ C$)؛ U_2 : متوسط روزانه سرعت باد در ارتفاع دومتری از سطح زمین ($m s^{-1}$)؛ e_s : فشار بخار اشباع (Kpa)؛ و e_a : فشار بخار هوا (Kpa) است. مؤلفه‌های رابطه بالا با استفاده از روابط ارائه شده در نشریه شماره ۵۶ فائو (۵) و داده‌های هواشناسی تعیین شد.

شاخص‌های آماری

در این تحقیق نتایج ET_0 حاصل از مدل پنمن مانیتث مبنا قرار گرفت و به عنوان مقادیر مشاهده شده (O_i) و نتایج حاصل از مدل‌های شبکه عصبی، شبکه درختی و هارگریوز سامانی معادلات به عنوان مقادیر پیش‌بینی شده (P_i) استفاده شدند. به منظور ارزیابی این مدل‌ها، علاوه بر ترسیم نمودارهای مشاهده شده در مقابل مقادیر پیش‌بینی شده، از شاخص‌های آماری ضریب تبیین (R^2) و ریشه میانگین مربعات خطا ($RMSE$) استفاده شده است. معادلات این شاخص‌ها به شرح زیرند:

$$R^2 = \frac{\left[\sum (P_i - \bar{P})(O_i - \bar{O}) \right]^2}{\sum (P_i - \bar{P})^2 \sum (O_i - \bar{O})^2} \quad (8)$$

$$RMSE = \left[N^{-1} \sum_{i=1}^N (O_i - P_i)^2 \right]^{0.5} \quad (9)$$

که در آن‌ها، P_i : مقادیر برآورد شده مدل‌های شبکه عصبی و معادله هارگریوز سامانی؛ \bar{P} : متوسط مقادیر برآورد شده این مدل‌ها؛ O_i : مقادیر مبنا؛ \bar{O} : متوسط مقادیر مبنا؛ و N : تعداد مشاهدات است. برای مقایسه نتایج این بررسی با سایر تحقیقات، شاخص درصد جذر میانگین مربع خطا نیز برآورد شده است. این شاخص از نسبت ریشه میانگین

مقایسه مدل رگرسیون درختی، شبکه عصبی مصنوعی و هارگریوز سامانی در برآورد تبخیر تعرق مرجع مناطق خشک

جدول ۱. متوسط اطلاعات هواشناسی و تبخیر تعرق مرجع استان‌های تحقیق

استان	پارامتر آماری	ET ₀ PM mm d ⁻¹	T _x °C	T _n °C	RH %	n h	U ₂ m s ⁻¹
کرمان	میانگین	۵/۱۲	۲۵/۶۱	۱۰/۹۰	۳۲/۲۷	۹/۲۷	۳/۰۵
	ضریب تغییرات	۰/۵۴	۰/۳۷	۰/۸۵	۰/۵۳	۰/۳۲	۰/۳۵
یزد	میانگین	۵/۱۲	۲۸/۰۸	۱۴/۱۶	۲۸/۱۶	۹/۳۹	۲/۶۰
	ضریب تغییرات	۰/۵۶	۰/۳۸	۰/۶۵	۰/۶۴	۰/۳۲	۰/۶۴
سیستان و بلوچستان	میانگین	۶/۵۳	۳۰/۲۶	۱۵/۰۸	۲۸/۷۶	۹/۳۵	۴/۲۱
	ضریب تغییرات	۰/۵۸	۰/۳۱	۰/۶۱	۰/۵۴	۰/۲۸	۰/۷۳
خراسان جنوبی	میانگین	۴/۴۵	۲۵/۴۱	۱۰/۳۱	۳۴/۱۸	۹/۱۶	۲/۱۴
	ضریب تغییرات	۰/۶۳	۰/۴۰	۰/۸۹	۰/۵۶	۰/۳۵	۰/۷۹

جدول ۲. معادلات هارگریوز سامانی واسنجی شده

گروه سرعت باد	معادله واسنجی	R ²	RMSE mm d ⁻¹
U ₁	ET ₀ PM= 0.85 ET ₀ HG -0.01	۰/۷۵	۱/۳۰
U ₂	ET ₀ PM= 1.15 ET ₀ HG+0.03	۰/۷۳	۱/۵۱
U ₃	ET ₀ PM= 1.78 ET ₀ HG-1.21	۰/۵۹	۳/۲۶

تبخیر تعرق برآورد شده در سرعت‌های کم باد بهتر از سرعت‌های زیاد است. در مطالعه‌ای دیگر نیز نشان داده شد تغییرات تبخیر تعرق گیاه مرجع نسبت به تغییرات کاهش سرعت باد حساس‌تر است (۱). در خصوص برآوردهای شبکه عصبی مصنوعی، همان‌گونه که جدول ۲ نشان می‌دهد تغییرات شاخص‌های خطاسنجی در این مدل به تفکیک سه گروه سرعت باد مشابه مدل رگرسیون درختی است. به طوری که کمترین خطای شبکه عصبی معادل ۱ میلی‌متر در روز در گروه اول سرعت باد و بیشترین خطا در گروه سوم معادل ۱/۹۹ میلی‌متر در روز اتفاق افتاده است. این نتیجه بیانگر آن است که با وجود تفاوت در الگوریتم‌های اجرای دو روش مدل رگرسیون درختی و شبکه عصبی مصنوعی، نتایج برآورد تبخیر تعرق از هر دو مدل به نسبت یکسان است. با این حال میانگین

با استفاده از داده‌های آموزش، بهترین مدل شبکه عصبی و همچنین شبکه درختی برای هر یک از گروه‌های سرعت باد تعیین و با استفاده از آنها، تبخیر و تعرق مرجع برای داده‌های ارزیابی محاسبه شد. خلاصه نتایج آماری در جدول ۳ ارائه شده است. همان‌طور که مشاهده می‌شود در مدل رگرسیون درختی، بهترین برآورد تبخیر تعرق در گروه اول سرعت باد یعنی بادهایی با سرعت کمتر از ۲/۴۸ متر بر ثانیه اتفاق افتاده که در آن جذر مربع خطا و ضریب تعیین به ترتیب ۱ میلی‌متر در روز و ۰/۸۵ به دست آمده است. با افزایش سرعت باد (گروه‌های دوم و سوم) مقدار خطای برآورد مدل رگرسیون درختی افزایش یافته است، به طوری که میزان جذر مربع خطا از ۱ میلی‌متر در روز به ۲/۱۶ میلی‌متر در روز افزایش و ضریب تعیین از ۰/۸۵ به ۰/۸۳ کاهش یافته است. این بدان معنا است که دقت

مدیریت آب و آبیاری

دوره ۴ ■ شماره ۲ ■ پاییز و زمستان ۱۳۹۳

رگرسیون درختی گروه U_1 برای برآورد تبخیرتغرق دارای کمترین خطا و بیشترین همبستگی با تبخیرتغرق فائو پنمن مانیتث بود، به منظور اطلاع از نحوه اجرای مدل رگرسیون درختی، مدل‌های خطی توسعه یافته در مدل رگرسیونی گروه U_1 به عنوان گروه بهینه در جدول ۳ آورده شد. این جدول برازش مدل‌های مختلف رگرسیون درختی برای برآورد ET_0 را با لحاظ محدوده‌های دمای هوا و تابش خورشیدی نشان می‌دهد. این جدول بیانگر ساختار اجرای مدل رگرسیون درختی است، بدین صورت که پس از آزمون همه تقسیم‌های ممکن، مدل درختی، تقسیمی را که کاهش خطای مورد انتظار را به حداکثر برساند برمی‌گزیند و روابط رگرسیونی خطی بهینه را برای محدوده‌های مختلف داده‌های ورودی ارائه می‌کند. در نهایت مدل رگرسیونی فوق در گروه U_1 با برازش ۲۳ مدل خطی موفق به برآورد ET_0 شده است که در جدول ۴ مشاهده می‌شود.

خطای جذر مربع خطا در مدل رگرسیون درختی $1/46$ میلی‌متر در روز و در شبکه عصبی مصنوعی $1/41$ میلی‌متر در روز است که بیانگر برتری $3/4$ درصدی شبکه عصبی مصنوعی نسبت به رگرسیون درختی است. نتایج تحقیق اعتماد شاهی و محجوبی در پیش‌بینی ارتفاع موج، سولوماتین^۱ و دولال^۲ در مدل‌سازی بارش-رواناب، و باتاچاریا و سولوماتین^۳ در شبیه‌سازی رابطه سطح آب-دبی، بیانگر تشابه نتایج شبکه عصبی مصنوعی با رگرسیون درختی برای برآورد پارامترهای هیدرولوژی است (۱۶،۶). شایان ذکر است که هرچند تا کنون شبکه عصبی مصنوعی به صورت گسترده در مسائل مربوط به آب از قبیل مدل‌سازی بارش-رواناب (۷، ۸) و پیش‌بینی دبی (۱۳) به کار گرفته شده، عملکرد مناسب و قابل فهم مدل درختی سبب شده است که در سال‌های اخیر بیشتر از شبکه عصبی مصنوعی، به کار گرفته شود. از آنجا که خروجی مدل

جدول ۳. خلاصه نتایج آماری مدل‌های شبکه عصبی، رگرسیون درختی و معادله واسنجی شده هارگریوز

مدل	گروه سرعت باد	R^2	RMSE $mm\ d^{-1}$
شبکه عصبی	U_1	۰/۸۵	۱/۰۰
	U_2	۰/۸۲	۱/۲۵
	U_3	۰/۸۵	۱/۹۹
شبکه درختی	U_1	۰/۸۵	۱/۰۰
	U_2	۰/۸۲	۱/۲۳
	U_3	۰/۸۳	۲/۱۶
هارگریوز سامانی واسنجی شده	U_1	۰/۷۵	۱/۳۰
	U_2	۰/۷۳	۱/۵۱
	U_3	۰/۵۹	۳/۲۶

1. Solomatine
2. Dulal
3. Batacharia and Solomatine

مقایسه مدل رگرسیون درختی، شبکه عصبی مصنوعی و هارگریوز سامانی در برآورد تبخیر تعرق مرجع مناطق خشک

جدول ۴. مدل‌های خطی توسعه یافته در مدل رگرسیونی برای گروه سرعت باد یک

LM num: 1[$R_a \leq 7.863$] $ET_0 = 0.0003 * T_{min} + 0.043 * T_{max} + 0.2919 * R_a - 1.5551$	LM num: 13[$15.157 < R_a \leq 15.903$, $T_{min} \leq 16.9$, $T_{max} \leq 30.7$] $ET_0 = 0.0065 * T_{min} + 0.1321 * T_{max} + 0.0555 * R_a + 0.7177$
LM num: 2[$T_{max} \leq 12.3$, $7.863 < R_a \leq 9.252$] $ET_0 = -0.0108 * T_{min} + 0.057 * T_{max} + 0.2934 * R_a - 1.7428$	LM num: 14[$T_{min} \leq 16.9$, $T_{max} \leq 30.7$, $R_a > 15.903$] $ET_0 = 0.1275 * T_{min} + 0.0578 * T_{max} + 0.0617 * R_a + 1.9514$
LM num: 3[$T_{max} > 12.3$, $7.863 < R_a \leq 9.252$] $ET_0 = 0.0219 * T_{min} + 0.0341 * T_{max} + 0.2795 * R_a - 1.1765$	LM num: 15[$R_a > 15.157$, $T_{min} \leq 16.9$, $T_{max} > 30.7$] $ET_0 = 0.16 * T_{min} + 0.002 * T_{max} + 0.1865 * R_a + 1.228$
LM num: 4[$T_{max} \leq 19.9$, $9.252 < R_a \leq 12.433$] $ET_0 = 0.0005 * T_{min} + 0.0846 * T_{max} + 0.2556 * R_a - 1.6403$	LM num: 16[$15.157 < R_a \leq 16.407$, $16.9 < T_{min} \leq 26.7$, $T_{max} \leq 33.45$] $ET_0 = 0.0032 * T_{min} + 0.1973 * T_{max} + 0.0362 * R_a - 0.1341$
LM num: 5[$9.252 < R_a \leq 10.841$, $T_{max} > 19.9$, $T_{min} \leq 10.3$] $ET_0 = 0.0022 * T_{min} + 0.0559 * T_{max} + 0.354 * R_a - 2.1903$	LM num: 17[$15.157 < R_a \leq 15.638$, $16.9 < T_{min} \leq 26.7$, $T_{max} > 33.45$] $ET_0 = 0.0581 * T_{min} - 0.0067 * T_{max} + 0.042 * R_a + 5.2811$
LM num: 6[$9.252 < R_a \leq 10.841$, $T_{max} > 19.9$, $10.3 < T_{min} \leq 18.1$] $ET_0 = 0.0078 * T_{min} - 0.0225 * T_{max} + 0.4022 * R_a - 0.4698$	LM num: 18[$15.638 < R_a \leq 16.407$, $16.9 < T_{min} \leq 26.7$, $T_{max} > 33.45$] $ET_0 = 0.0043 * T_{min} - 0.0987 * T_{max} + 0.0302 * R_a + 10.3466$
LM num: 7[$9.252 < R_a \leq 10.841$, $T_{max} > 19.9$, $T_{min} > 18.1$] $ET_0 = 0.28 * T_{min} - 0.1955 * T_{max} + 1.0003 * R_a - 5.79$	LM num: 19[$16.9 < T_{min} \leq 18.3$, $R_a > 16.407$] $ET_0 = 0.0142 * T_{min} - 0.0027 * T_{max} + 0.032 * R_a + 6.3846$
LM num: 8[$R_a > 9.252$, $T_{max} > 19.9$] $ET_0 = 0.0409 * T_{min} + 0.0287 * T_{max} + 0.3173 * R_a - 1.1517$	LM num: 20[$18.3 < T_{min} \leq 20.9$, $16.407 < R_a \leq 16.771$] $ET_0 = 0.0111 * T_{min} - 0.0027 * T_{max} - 2.3221 * R_a + 45.8609$
LM num: 9[$12.433 < R_a \leq 15.157$, $T_{max} \leq 28.9$] $ET_0 = 0.0008 * T_{min} + 0.1215 * T_{max} + 0.2302 * R_a - 1.8757$	LM num: 21[$18.3 < T_{min} \leq 20.9$, $R_a > 16.771$] $ET_0 = 0.0111 * T_{min} - 0.0027 * T_{max} + 0.032 * R_a + 6.999$
LM num: 10[$12.433 < R_a \leq 13.986$, $T_{max} > 28.9$] $ET_0 = 0.0746 * T_{min} - 0.0637 * T_{max} + 0.6552 * R_a - 2.7131$	LM num: 22[$20.9 < T_{min} \leq 26.7$, $R_a > 16.407$] $ET_0 = 0.0051 * T_{min} - 0.1138 * T_{max} + 1.1425 * R_a - 7.134$
LM num: 11[$13.986 < R_a \leq 15.157$, $28.9 < T_{max} \leq 31.9$] $ET_0 = 0.0058 * T_{min} + 0.2419 * T_{max} + 0.0693 * R_a - 2.9763$	LM num: 23[$R_a > 15.157$, $T_{min} > 26.7$] $ET_0 = 0.42 * T_{min} - 0.3019 * T_{max} + 0.2448 * R_a + 4.5871$
LM num: 12[$13.986 < R_a \leq 15.157$, $T_{max} > 31.9$] $ET_0 = 0.0845 * T_{min} - 0.1188 * T_{max} + 0.7443 * R_a - 1.9534$	

شایان توجه در این نتایج، برتری مدل هارگریوز سامانی گروه اول سرعت باد در تخمین تبخیر تعرق مرجع است. همان گونه که از جدول ۲ استنباط می شود، تبخیر تعرق برآورد شده از مدل کالیبره شده هارگریوز سامانی در سرعت های باد کمتر از ۲/۴۸ متر بر ثانیه (گروه اول) کمترین خطا و بیشترین ضریب تبیین را نسبت به دو گروه دیگر دارد. مقایسه نتایج بهینه هر سه مدل بیانگر برتری دو مدل شبکه عصبی مصنوعی و رگرسیون درختی نسبت به

همان گونه که در بخش مواد و روش ها اشاره شد، در این تحقیق پس از برآزش مدل هارگریوز سامانی به داده های آموزش، ارزیابی این مدل با استفاده از داده های صحت سنجی انجام گرفت. داده های صحت سنجی شامل تمام ایستگاه های محدوده تحقیق بود، به طوری که بتوان از آن برای تخمین ET_0 در مناطق مختلف محدوده تحقیق استفاده کرد. در جدول ۲ مدل های واسنجی شده هارگریوز سامانی برای هر سه گروه سرعت باد مشاهده می شود. نکته

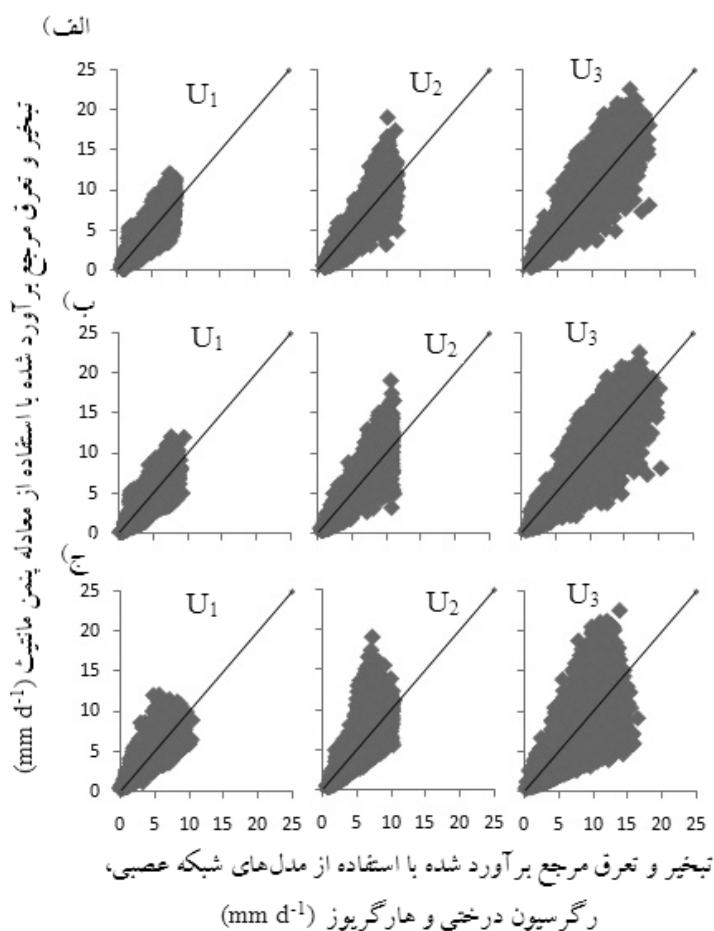
مدیریت آب و آبیاری

دوره ۴ ■ شماره ۲ ■ پاییز و زمستان ۱۳۹۳

پیش‌بینی سیلاب در مناطق بالادست رودخانه هوآی^۱ چین استفاده شد. براساس نتایج، هر دو مدل عملکرد خوبی برای پیش‌بینی سیلاب داشتند، اما مدل هیدرولوژی ساخته شده از ترکیب دو مدل رگرسیون درختی و شبکه عصبی مصنوعی بیشترین دقت را برای پیش‌بینی سیلاب نشان داد (۱۷).

پراکنش نتایج تبخیرتعرق مرجع برآوردشده از هر سه مدل به تفکیک سه گروه سرعت باد در مقابل مقادیر تبخیرتعرق پنمن مانیتث فائو در شکل ۲ ارائه شده است.

هارگریوز سامانی است (جدول ۳)، به‌طوری که خطای مدل هارگریوز سامانی به میزان ۴۳ و ۳۸ درصد به‌ترتیب بیش از شبکه عصبی و رگرسیون درختی است. نتیجه تحقیقی در دانشگاه دیویس کالیفرنیا برای برآورد مقادیر روزانه ET_0 با استفاده از داده‌های هواشناسی نشان داد که مدل رگرسیون درختی از توانایی بیشتری در مقایسه با روش‌های فائو پنمن مانیتث و هارگریوز سامانی برخوردار است (۱۴). در تحقیقی دیگر از مدل درختی، شبکه عصبی مصنوعی و ترکیب این دو مدل برای



شکل ۲. پراکنش نتایج مدل‌های بررسی‌شده این تحقیق در گروه‌های مختلف سرعت باد و معادله پنمن مانیتث فائو. (الف) مدل شبکه عصبی، (ب) مدل رگرسیون درختی، (ج) هارگریوز سامانی واسنجی شده.

میان داده‌ها با حداقل دخالت کاربران استخراج می‌شود. طبقه‌بندی یکی از مسائل مطرح در داده‌کاوی است. از جمله الگوریتم‌هایی که برای طبقه‌بندی در داده‌کاوی استفاده می‌شود، مدل‌های شبکه عصبی مصنوعی و درخت تصمیم‌اند. در تحقیق حاضر، از مدل درختی در محدوده تحقیق (استان‌های کرمان، یزد، سیستان و بلوچستان و خراسان جنوبی) با استفاده از حداقل داده‌های هواشناسی در دوره آماری ۱۹۹۸-۲۰۰۸ به منظور پیش‌بینی تبخیرتعرق مرجع استفاده شد. بدین منظور از داده‌های روزانه دماهای حداکثر و حداقل هوا و تابش برون‌زمینی مربوط به دوره آماری مورد نظر از ۲۲ ایستگاه سینوپتیک در محدوده تحقیق به‌عنوان ورودی مدل‌ها استفاده شد و ارزیابی مدل‌ها بر مبنای رابطه فائو پنمن مانیتث انجام گرفت. بررسی‌ها نشان داد که در گروه U_1 که سرعت باد در آن کمتر از $2/48$ است، هر سه مدل رگسیون درختی، شبکه عصبی مصنوعی و هارگریوز سامانی اصلاح شده بهترین برآورد تبخیرتعرق مرجع را داشتند. از طرفی بین نتایج هر سه مدل، نتایج دو مدل شبکه عصبی و رگسیون درختی به نسبت یکسان بود و شبکه عصبی مصنوعی با اختلاف ناچیز بر رگسیون درختی برتری داشت.

منابع

۱. پارسافرن، سبزی پرورع.ا. و آیینی ع (۱۳۹۱) ارزیابی حساسیت معادله فائو پنمن - مانیتث ۵۶ نسبت به تغییرات سرعت باد در غرب ایران، گزارش کوتاه علمی. پژوهش‌های حفاظت آب و خاک (علوم کشاورزی و منابع طبیعی). ۱۹(۱):۱۹۷-۲۰۷.
۲. زارع ایبانه ح، بیات ورکشی م، معروفی ص. و امیری چایجان ر (۱۳۸۹) ارزیابی سیستم‌های هوشمند مصنوعی در کاهش پارامترهای تخمین تبخیرتعرق

همان‌طور که در شکل ۲ ملاحظه می‌شود، نتایج هر دو مدل شبکه عصبی مصنوعی و رگسیون درختی در هر سه گروه رضایت‌بخش بوده است. پراکندگی نقاط بین خطوط برازشی بیانگر تطابق تبخیرتعرق برآوردی از دو مدل مذکور و تبخیرتعرق مشاهداتی (پنمن مانیتث) است. نتایج شکل ۲ همسو با نتایج جدول ۲ است. نکته شایان توجه در شکل ۲، تغییرات تبخیرتعرق هر سه روش در قالب گروه‌های متفاوت از سرعت باد است. به‌طوری که از شکل ۲ مشاهده می‌شود، با افزایش سرعت باد (از گروه U_1 به گروه U_3) مقدار تبخیرتعرق برآورد شده از هر سه روش شبکه عصبی، هارگریوز سامانی و رگسیون درختی افزایش داشته است. بنابراین می‌توان گفت با افزایش سرعت باد تبخیرتعرق افزایش می‌یابد. این یافته با نتایج جدول ۱ نیز هم‌راستا است. از طرفی با افزایش سرعت باد پراکنش تبخیرتعرق نیز افزایش داشته است. به عبارت دیگر در گروه سرعت کم باد (U_1)، تبخیرتعرق از پراکندگی کمتری برخوردار است، درحالی که در گروه‌های U_2 و U_3 تبخیرتعرق دارای پراکندگی بیشتری است. همچنین با مشاهده شکل ۲ استنباط می‌شود که تبخیرتعرق برآوردی از دو روش شبکه عصبی و رگسیون همسو با تبخیرتعرق حاصل از مدل فائو پنمن مانیتث است. همان‌گونه که شکل ۲ نشان می‌دهد اختلاف بین تبخیرتعرق برآوردی از مدل هارگریوز سامانی و مدل فائو پنمن مانیتث مشهود بوده و توزیع تبخیرتعرق غیریکنواخت است. اما در دو روش شبکه عصبی مصنوعی و رگسیون درختی، تبخیرتعرق هم‌راستا با تبخیرتعرق مدل فائو پنمن مانیتث است که بیانگر دقت زیاد این دو روش در برآورد تبخیرتعرق در مقایسه با روش هارگریوز سامانی است که در جدول ۲ نیز مشهود است.

۴. نتیجه‌گیری

داده‌کاوی روشی است که به‌وسیله آن الگوهای مفید از

- height in Lake Superior. Ocean Engineering. 36(15-16): 1175-1181.
11. Hagan MT and Menhaj M (1994) Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks. 5: 989-993.
12. Hargreaves G.H and Samani, Z.A (1985) Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture. 1(1): 96-99.
13. Kisi O (2007) Evapotranspiration modeling from climatic data using a neural computing technique. Hydrological Processes. 21(6):1925-1934.
14. Muttiah R.S, Srinivasan R and Allen P.M (1997) Prediction of two-year peak stream-discharges using neural networks. the American Water Resources Association. 33(3):513-703.
15. Pal M and Deswal S (2009) M5 model tree based modeling of reference evapotranspiration. Hydrological Processes. 23(2):1437-1443.
16. RahimiKhoob A (2008) Comparative study of Hargreaves's and artificial neural network's Methodologies in estimating reference evapotranspiration in a semiarid environment. Irrigation Science. 26: 253-259.
17. Solomatine D.P and Dulal K.N (2003) Model trees as an alternative to neural networks in rainfall-runoff modeling. Hydrological Science, 48(3): 399-411.
18. Solomatine D.P and Xue Y (2004) M5 model trees and neural networks: application to floodforecasting in the upper reach of the Huai River in China. Hydrologic Engineering. 9(6): 491-501.
- Tan Y and Van Cauwenberghe A (1999) Neural-network-based - step-ahead predictors for nonlinear systems with time delay. Engineering Application of Artificial Intelligence. 12: 21-25.
- گیاه مرجع. آب و خاک (علوم و صنایع کشاورزی). ۲۴ (۲): ۲۹۷-۳۰۵.
۳. طالبی ع، پورمحمدی س. و رحیمیان م. ح (۱۳۸۹) بررسی عوامل مؤثر در تبخیر و تعرق مرجع، با استفاده از آنالیز حساسیت معادله فائو پنمن مانیتث مطالعه موردی ایستگاه‌های یزد، طبس و مروست. پژوهش‌های جغرافیای طبیعی. ۷۳: ۹۷-۱۱۰.
۴. محجوبی ج. و اعتمادشاهی ا (۱۳۸۷) پیش‌بینی پارامترهای امواج ناشی از باد در بندر امیرآباد به کمک درخت‌های تصمیم رگرسیونی. چهارمین کنگره ملی مهندسی عمران: ۱-۶.
5. Allen R.G, Pereira L.S, Raes D and Smith M (1998) Crop evapotranspiration Guidelines for computing crop water requirements. Irrigation and Drainage Paper No.56, FAO, ROME.
6. Bhattacharya B and Solomatine D.P (2005) Neural networks and M5 model trees in modeling water level-discharge relationship. Neurocomputing. 63: 381-396.
7. Dawson C.W and Wilby R (1998) An artificial neural network approach to rainfall-runoff modeling. Hydrological Sciences. 43(1): 47-66.
8. Dibike Y.B and Solomatine D (2001) River Flow Forecasting Using Artificial Neural Networks. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere. 26(1): 1-8.
9. Droogers P and Allen R.G (2002) Estimating reference evapotranspiration under inaccurate data conditions. Irrigation and Drainage Systems. 16:33-45.
10. Etemad Shahidi A and Mahjoobi J (2009) Comparison between M5 model tree and neural networks for prediction of significant wave