

## مقایسه نتایج مدل درختی M5 و برنامه‌ریزی ژنتیک با روش پنمن-مونتیث-فائو برای

### تخمین تبخیر-تعرق مرجع

محمد تقی ستاری<sup>۱\*</sup>، بهرام اسمعیل زاده<sup>۲</sup>

#### چکیده

تبخیر-تعرق به عنوان یکی از اجزاء مهم چرخه هیدرولوژیک، نقش مهمی در بیلان حوضه‌های آبریز دارد. برای محاسبه حجم تبخیر-تعرق و نیاز آبی در گیاهان، لازم است ابتدا مقدار تبخیر-تعرق مرجع ( $ET_0$ ) محاسبه و اعمال ضرائبی، تبخیر-تعرق برای هر نوع گیاهی محاسبه شود. در این تحقیق، ابتدا براساس روش استاندارد و متداول فائو-پنمن-مونتیث و داده‌های هواشناسی ایستگاه تبریز، مقدار تبخیر-تعرق مرجع ماهانه در منطقه تبریز استان آذربایجان شرقی محاسبه شد. سپس با استفاده از پارامترهای هواشناسی شامل میانگین، حداقل و حداکثر دما و رطوبت نسبی، مقدار بارش، سرعت باد و ساعات آفتابی به عنوان ورودی برنامه‌ریزی ژنتیک و مدل درختی M5 در مقیاس زمانی ماهانه، میزان تبخیر-تعرق مرجع به عنوان خروجی مدل برآورد گردید. نتایج نشان داد که هر دو روش مذکور نتایج دقیقی (با ضریب تبیین ۰/۹۹ برای مدل درختی M5 و ۰/۹۶ برای برنامه‌ریزی ژنتیک) را جهت پیش‌بینی تبخیر-تعرق مرجع ماهانه در منطقه تبریز ارائه می‌دهند، ولی مدل درختی M5 روابط خطی ساده، قابل فهم و کاربردی‌تری برای پیش‌بینی تبخیر-تعرق مرجع ارائه می‌دهد.

**واژه‌های کلیدی:** تبخیر-تعرق مرجع، فائو-پنمن-مونتیث، برنامه‌ریزی ژنتیک، مدل درختی M5.

<sup>۱</sup> عضو هیئت علمی گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز

ایمیل: mtsattar@gmail.com تلفن: ۰۹۱۴۴۰۱۵۸۰۲

<sup>۲</sup> عضو هیئت علمی گروه مهندسی عمران، دانشگاه آزاد اسلامی، واحد تبریز

ایمیل: bahram1411@gmail.com تلفن: ۰۹۱۴۴۱۱۴۷۶۴

## مقدمه

یکی از راه‌های هدر رفت آب در مناطق مختلف آب و هوایی تبخیر و تعرق می‌باشد. تبخیر و تعرق مرجع یک پدیده چند متغیره و پیچیده است که به عوامل متعدد اقلیمی بستگی دارد. مقدار تبخیر-تعرق مرجع  $ET_0$  را می‌توان بصورت مستقیم با استفاده از لایسی‌متر اندازه‌گیری نمود و یا با استفاده از روابط تجربی ارائه شده همچون پنمن-مونتیث تخمین زد. نسخه اصلاح شده پنمن-مونتیث-فائو هم اینک مبنای نرم‌افزارهایی همچون CropWat بوده که در محاسبات تبخیر-تعرق مورد استفاده قرار می‌گیرد. گرچه لایسی‌متر اندازه‌گیری نسبتاً دقیقی از میزان تبخیر-تعرق ارائه می‌دهد، اما در عمل نصب و بهره‌برداری از آن زمان‌بر بوده و هزینه‌های زیادی به دنبال دارد و فاقد صرفه اقتصادی است. تا به حال روش‌های زیادی مبتنی بر داده‌های هواشناسی برای شرایط مختلف جغرافیایی و اقلیمی با هدف محاسبه  $ET_0$  تدوین شده است. این روش‌ها گستره‌ای از معادلات پیچیده چون روش پنمن-مونتیث که به داده‌های هواشناسی زیادی نیاز دارند (آلن و همکاران ۱۹۹۸) تا معادلات ساده‌تر چون روش‌های بلانی-کریدل (۱۹۵۰) و هارگریوز-سامانی (۱۹۸۲) که به داده‌های کمتری نیاز دارند را دربر می‌گیرد. در این تحقیق از روش فائو-پنمن-مونتیث به عنوان روش مبنا استفاده گردیده است. امروزه سیستم‌های هوشمند همچون برنامه‌ریزی ژنتیک<sup>۱</sup> و مدل‌های درختی در مدل‌سازی فرایندهای هیدرولوژیکی و مهندسی آب مورد استفاده قرار می‌گیرد. اساساً تبخیر-تعرق یک فرایند پیچیده بوده و روش‌هایی نوین که از قابلیت بسیار زیادی در مدل‌سازی فرایندهای پیچیده برخوردارند، می‌توانند جهت مدل‌سازی فرایند تبخیر-تعرق بکار گرفته شوند. گوون و همکاران (۲۰۰۸) برای داده‌های منطقه کالیفرنیا در آمریکا مقدار تبخیر-تعرق مرجع را با استفاده از روش برنامه‌ریزی ژنتیک برآورد نمودند و با مقایسه نتایج بدست آمده از برنامه‌ریزی ژنتیک با داده‌های مشاهداتی نشان دادند که روش مذکور از دقت بالایی در تخمین  $ET_0$  برخوردار است. کیشی و گوون (۲۰۱۰) با بکارگیری ترکیب‌های مختلفی از

پارامترهای هواشناسی برای داده‌های کالیفرنیا و با استفاده از دو روش شبکه عصبی مصنوعی و برنامه‌ریزی ژنتیک خطی اقدام به مدل‌سازی تبخیر-تعرق مرجع نمودند. نتایج نشان داد برنامه‌ریزی ژنتیک خطی دقت بالاتری نسبت به شبکه عصبی مصنوعی دارد. تریزی (۲۰۱۰) با استفاده از برنامه‌ریزی ژنتیک مقدار تبخیر روزانه از سطح آزاد آب دریاچه ایردیر<sup>۲</sup> در ترکیه را مدل نمود. از جمله تحقیقات انجام یافته با مدل درختی M5 می‌توان به مطالعات زیر اشاره نمود. ماهس پال و دسوال (۲۰۰۹) با استفاده از داده‌های هواشناسی سیستم اطلاعات و مدیریت آبیاری دیویس کالیفرنیا اقدام به مدل‌سازی تبخیر-تعرق روزانه با استفاده از مدل درختی M5 نمودند. ایشان جهت مقایسه نتایج از روابط هاریگریوز-سامانی<sup>۳</sup> و فائو<sup>۴</sup>-۵۶ استفاده نموده و دریافتند که مدل درختی M5 بخوبی قادر به مدل‌سازی تبخیر-تعرق روزانه می‌باشد. دیتهاکیت و چیناراسری (۲۰۱۲) با استفاده از مدل M5 اقدام به توسعه ضریب جدید Kp مربوط به معادلات تشتک‌های تبخیر کلاس A<sup>۵</sup> و سانکن کلرادو<sup>۶</sup> تحت شرایط خشک و مرطوب نمودند، ستاری و همکاران (۲۰۱۳) با استفاده از مدل‌های M5 و شبکه‌های عصبی مصنوعی اقدام به مدل‌سازی تبخیر و تعرق ماهانه گیاه مرجع در آنکارا نمودند. ایشان نتیجه گرفتند گرچه شبکه‌های عصبی مصنوعی نسبت به مدل درختی نتایج نسبتاً دقیقی ارائه می‌دهد، اما خروجی مدل درختی ساده‌تر، قابل فهم‌تر و کاربردی‌تر است. طباطبائی و همکاران (۲۰۱۴) با استفاده از روش‌های هوش مصنوعی مانند برنامه‌ریزی ژنتیک و انفیس<sup>۷</sup> به مدل‌سازی تبخیر و تعرق مرجع پرداختند. نتایج نشان داد روش انفیس نتایج دقیق‌تری نسبت به روابط تجربی و برنامه‌ریزی ژنتیک ارائه می‌کند. اسماعیل‌زاده (۲۰۱۵) با استفاده از شبکه عصبی مصنوعی و برنامه‌ریزی ژنتیک به مدل‌سازی تبخیر و تعرق مرجع پرداختند. نتایج نشان داد برنامه‌ریزی ژنتیک نتایج دقیق‌تر و کاربردی‌تری ارائه می‌کند. مرادی و رحیمی خوب (۱۳۹۱) به کمک تصاویر ماهواره‌ای نوا<sup>۸</sup> و مدل درختی M5 اقدام به برآورد تبخیر و تعرق مرجع در شبکه آبیاری قزوین نمودند و نتیجه گرفتند مدل درختی M5 با دقت نسبتاً خوبی قادر به برآورد تبخیر-تعرق مرجع می‌باشد.

<sup>5</sup> - Class A Pan

<sup>6</sup> - Colorado Sunken Pan

<sup>7</sup> - ANFIS

<sup>8</sup> - NOAA

<sup>1</sup> - Lysimeter

<sup>2</sup> - Genetic Programming

<sup>3</sup> - Egirdir

<sup>4</sup> - Hargreaves-Samani

داده‌های آموزشی<sup>۱</sup> و ۷ سال، از نوامبر سال ۲۰۰۰ تا دسامبر سال ۲۰۰۶ (۲۵ درصد داده‌ها) به عنوان داده‌های آزمایشی<sup>۲</sup> انتخاب شدند. عملکرد مدل‌های مورد استفاده در این تحقیق بر پایه محاسبه جذر میانگین مربعات خطا<sup>۳</sup> و ضریب همبستگی<sup>۴</sup> مورد ارزیابی قرار گرفت. فرمول‌های محاسبه دو آماره فوق در معادلات ۱ و ۲ آمده است. خصوصیات آماری برای داده‌های مورد استفاده در این تحقیق نیز در جدول ۱ ارائه شده است.

$$R^2 = \left( \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2 * \sum_{i=1}^N (Y_i - \bar{Y})^2}} \right)^2 \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - x_i)^2}{N}} \quad (2)$$

### برنامه‌ریزی ژنتیک

برنامه‌ریزی ژنتیک تعمیم یافته روش الگوریتم ژنتیک می‌باشد که برای اولین بار توسط جان کزا (۱۹۹۲) بر اساس تئوری داروین ارائه شد. برنامه‌ریزی ژنتیک یک تکنیک برنامه‌ریزی خودکار می‌باشد که راه حل مسئله را با استفاده از برنامه کامپیوتری ارائه می‌کند. در این روش در ابتدای فرآیند، هیچگونه رابطه تابعی در نظر گرفته نشده و این روش قادر به بهینه‌سازی ساختار مدل و مؤلفه‌های آن می‌باشد. برنامه‌ریزی ژنتیک بر خلاف الگوریتم ژنتیک روی ساختار درختی فرمول‌ها به جای سلسله ارقام دودویی عمل می‌کند. ساختارهای درختی از مجموعه توابع (عملگرهای ریاضی مورد استفاده در فرمول‌ها) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) تشکیل می‌شوند. قبل از مراحل اجرایی برنامه‌ریزی ژنتیک گام‌های مقدماتی زیر باید توسط کاربر تعیین شوند (صمدیان فرد و دلیر حسن نیا، ۱۳۹۴). ۱. مجموعه ترمینال‌ها (متغیرهای مسئله، اعداد ثابت تصادفی) ۲. مجموعه عملگرهای ریاضی مورد استفاده در فرمول‌ها ۳. انتخاب تابع برازش مناسب (RMSE, MSE, R, ...), برای سنجش برازش روابط

ستاری و همکاران (۱۳۹۲) اقدام به پیش‌بینی تبخیر-تعرق روزانه گیاه مرجع در منطقه بناب با دو روش شبکه عصبی مصنوعی پرسپترون چندلایه پس انتشار خطا و مدل درختی M5 نمودند و نتیجه گرفتند گرچه روش شبکه عصبی مصنوعی با اختلاف خیلی کمی در مقایسه با M5 پیش‌بینی نسبتاً دقیق‌تری از تبخیر-تعرق مرجع ارائه می‌دهد، ولی مدل درختی M5 روابط ساده خطی و قابل فهم‌تری توسعه می‌دهد. سامتی و همکاران (۱۳۹۲) از مدل M5 جهت برآورد تبخیر-تعرق مرجع در ایستگاه‌های شیراز و کرمانشاه استفاده نموده و نتیجه گرفتند که در هر دو ایستگاه، مدل درختی M5 در تخمین تبخیر-تعرق پتانسیل نسبت به هر دو روش پنمن-مانتیث و هارگریوز-سامانی دارای عملکرد مناسبی است، اما توانایی آن در تخمین تبخیر-تعرق به روش پنمن-مونتیث نسبت به روش هارگریوز-سامانی بالاتر می‌باشد. نتایج تحقیقات فوق نشان داده‌اند که مدل درختی M5 می‌تواند به طور موفقیت‌آمیزی در مسائل مختلف مربوط به مهندسی آب به کار رود. هدف از این تحقیق، مدل‌سازی تبخیر-تعرق ماهانه شهر تبریز با استفاده از برنامه‌ریزی ژنتیک و مقایسه آن با مدل درختی M5 می‌باشد. در این مدل‌سازی پارامترهای هواشناسی موثر شامل: دما، بارش، باد، رطوبت نسبی، ساعات آفتابی و ... در محاسبه تبخیر-تعرق بعنوان ورودی مدل‌ها استفاده شده است.

### مواد و روش‌ها

**معرفی منطقه مورد مطالعه و داده‌های مورد استفاده**  
شهر تبریز به عنوان مرکز استان آذربایجان شرقی در غرب این استان و در منتهی‌الیه شرق و جنوب شرق جلگه تبریز واقع شده است. این شهر در طول جغرافیایی ۴۶ درجه و ۲۶ دقیقه و عرض جغرافیایی ۳۸ درجه و ۶ دقیقه و با ارتفاع ۱۳۵۰ متر تا ۱۵۵۰ متر از سطح دریا قرار گرفته است. شهر تبریز از سمت شمال، جنوب و جنوب شرق به کوهستان و از سمت غرب به زمین‌های هموار و شوره زارهای تلخه رود محدود شده است. موقعیت جغرافیایی شهرستان تبریز در شکل ۱ نشان داده شده است.

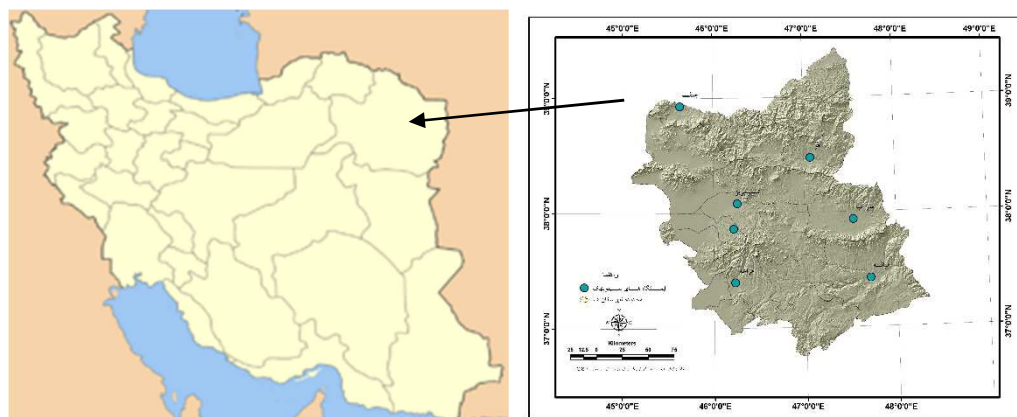
در این تحقیق داده‌های ۱۹ سال، از ژانویه سال ۱۹۸۲ تا اکتبر سال ۲۰۰۰ (۷۵ درصد از داده‌ها) به عنوان

<sup>3</sup> - Root Mean Square Errors

<sup>4</sup> - Correlation Coefficient

<sup>1</sup> - Training Data

<sup>2</sup> - Testing Data



شکل ۱- موقعیت جغرافیایی شهرستان تبریز.

جدول ۱- خصوصیات آماری داده‌های مورد استفاده در بازه زمانی سال‌های ۱۹۸۲ تا ۲۰۰۶.

آماره	تبخیر و تغرق (FPN) (میلی متر)	بارش (میلی متر)	حداکثر رطوبت نسبی (درصد)	حداکثر رطوبت نسبی (درصد)	سرعت باد (متر بر ثانیه)	ساعات آفتابی (ساعت)	حداکثر دما (درجه سانتی گراد)	حداقل دما (درجه سانتی گراد)
میانگین	۲۵/۱۱	۲۰/۹۸	۶۹/۵	۳۴/۵۶	۳/۰۰	۲۳۶/۰۶	۱۸/۶۳	۷/۷۵
انحراف استاندارد	۱۸/۳۲	۲۱/۳۳	۱۳/۷۳	۱۴/۵۶	۱/۹۱	۸۳/۹۳	۱۰/۹۱	۸/۹۸
حداکثر	۶۵/۸۵	۱۲۳/۴	۹۴/۱۶	۷۵/۴۸	۳۱/۰۰	۳۷۹/۵۰	۳۵/۷۶	۳۱/۰۰
حداقل	۲/۰۳	۰/۰	۳۱/۰۰	۱۲/۱۷	۰/۷۸	۵۱/۸۰	-۳/۵۵	-۱۲/۱۶

مهمترین عمل‌های ژنتیکی مورد استفاده در برنامه‌ریزی ژنتیک می‌باشند. عمل‌های دیگری مثل اصلاح ساختار و ... نیز با احتمال کمتر مورد استفاده قرار می‌گیرند. (ب) تعداد مناسبی از افراد جمعیت حاضر انتخاب می‌شوند (انتخاب فرد یا افرادی از جمعیت مذکور به صورت احتمالاتی می‌باشد که در این انتخاب احتمالاتی منفردهای با برآزش بهتر به منفردهای نامرغوب ترجیح داده می‌شوند و این بدان معنی نیست که حتماً منفردهای نامرغوب حذف می‌شوند). (ج) از عمل ژنتیکی انتخاب شده برای تولید فرزند (رابطه جدید) استفاده می‌شود. (د) فرزند (فرمول جدید) تولید شده در یک جمعیت جدید وارد می‌شود. (ه) مدل مورد نظر با استفاده از تابع برآزش مورد ارزیابی واقع می‌شود. ۴. گام سوم تا نیل به حداکثر تعداد تولید، تکرار خواهد شد. در این تحقیق برای اعمال روش برنامه‌ریزی ژنتیک روی داده‌ها از نرم‌افزار GeneXproTools 4.0 استفاده شده است.

۴. تعیین پارامترهای کنترل‌کننده اجرای برنامه (اندازه جمعیت، احتمال مربوط به بکارگیری اعمال ژنتیکی و جزئیات دیگر مربوط به اجرای برنامه) ۵. معیار پایان و ارائه نتایج اجرای برنامه (مثل تعداد تولید جمعیت جدید، تعیین یک مقدار مشخص برای برآزش روابط که اگر میزان برآزش برابر یا بیشتر از آن مقدار شد، اجرا متوقف شود). فرآیند اجرایی گام به گام برنامه‌ریزی ژنتیک به صورت مراحل زیر است (صمدیان فرد و دلیر حسن نیا، ۱۳۹۴): ۱. تولید یک جمعیت اولیه از روابط که این روابط از ترکیب تصادفی مجموعه توابع (عملگرهای ریاضی مورد استفاده در روابط) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) ایجاد می‌شوند. ۲. هر یک از افراد جمعیت مذکور با استفاده از توابع برآزش مورد ارزیابی قرار می‌گیرند. ۳. تولید یک جمعیت جدید از روابط، بدین منظور مراحل زیر برای تولید یک جمعیت جدید دنبال می‌شود (صمدیان فرد و دلیر حسن نیا، ۱۳۹۴): (الف) یکی از عمل‌های ژنتیکی تلاقی، جهش و تولیدمثل انتخاب می‌شود (این سه عمل ژنتیکی،

## مدل درختی M5

مدل درختی M5 (کوئینلن ۱۹۹۲) زیر مجموعه‌ای از روش‌های یادگیری ماشینی و داده‌کاوی است. روش‌های یادگیری ماشینی و داده‌کاوی قابلیت کشف نیمه اتوماتیک الگوهای ارزشمندی از داده‌ها را دارند. در سال‌های اخیر مدل درختی M5 تحول قابل توجهی در مسایل طبقه‌بندی و پیش‌بینی ایجاد نموده است. از جمله دلایل استفاده از مدل درختی می‌توان به این موارد اشاره کرد: (۱) مدل درختی به طور مستقیم با متغیرهای پیش‌بینی‌کننده مرتبط می‌باشد، بنابراین نتایج مدل برای فهم و شبیه‌سازی آسان هستند. (۲) درخت‌های تصمیم‌گیری غیر پارامتریک بوده و هیچ دخالتی از سوی کاربر بر روی آن‌ها صورت نمی‌گیرد. (۳) خروجی مدل از دقت بالایی برخوردار است که می‌توان آن را با سایر مدل‌ها مقایسه کرد.

یک درخت تصمیم معمولاً از چهار بخش ریشه، شاخه، گره‌ها و برگ‌ها تشکیل شده است. گره اول در درخت تصمیم به‌عنوان ریشه‌ی درخت در نظر گرفته می‌شود. هر گره مربوط به یک خصوصیت معین است و شاخه‌ها به معنای بازه‌ای از مقادیر هستند. این بازه‌های مقادیر باید بخش‌های مختلف مجموعه‌ی مقادیر معلوم را برای خصوصیت‌ها به‌دست دهند. عمل انشعاب توسط یکی از متغیرهای پیش‌بینی‌کننده انجام می‌پذیرد (آلبرگ و همکاران ۲۰۱۲). اولین مرحله برای ایجاد یک مدل درختی، استفاده از یک معیار انشعاب است. معیار انشعاب برای الگوریتم M5 بر اساس عملکرد انحراف استاندارد مقادیر هر کلاس و یا طبقه است که در هر گره به‌دست آمده است. این روش اساس روش‌های طبقه‌بندی است که آنتروپی<sup>۱</sup> نامیده می‌شود. آنتروپی می‌تواند به عنوان معیار میزان آشفتگی و بی‌نظمی یک سیستم تفسیر شود. معیار انشعاب بیانگر میزان خطا در آن گره می‌باشد و مدل حداقل خطای مورد انتظار را به عنوان نتیجه آزمایش هر صفت در آن گره محاسبه می‌کند. خطای مدل عموماً با اندازه‌گیری دقت پیش‌بینی مقادیر هدف موارد دیده نشده سنجش می‌شود. معادله محاسبه کاهش انحراف استاندارد (SDR) به صورت معادله ۳ و ۴ می‌باشند.

در این معادله T مجموعه‌ای از نمونه‌هاست (موارد)

$$SDR = Sd(T) - \sum_{i=1}^N \frac{|T_i|}{|T|} Sd(T_i) \quad (3)$$

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

که به هر گره وارد می‌شوند،  $T_i$  زیرمجموعه‌ای از نمونه‌هاست که در نتیجه تقسیم گره براساس ویژگی منتخب حاصل می‌شوند،  $Sd$  بیانگر انحراف معیار،  $y_i$  مقدار عددی ویژگی هدف نمونه  $i$  و  $N$  شماره داده‌ها را نشان می‌دهد (آلبرگ و همکاران ۲۰۱۲). مدل‌های تصمیم درختی انواع مختلف دارد. هنگامی که خروجی یک درخت، یک مجموعه‌ی گسسته از یک مجموعه مقادیر ممکن است، به آن طبقه‌بندی درختی گفته می‌شود. هنگامی که بتوان خروجی درخت را یک عدد حقیقی در نظر گرفت، آن را رگرسیون درختی می‌نامند. به عبارت دیگر، اگر متغیرهای ورودی به سیستم، عددی باشند، از رگرسیون درختی و اگر مطلق و قیاسی باشند، از طبقه‌بندی درختی استفاده می‌کنیم (کوئینلن، ۱۹۹۲). مدل M5 یک مدل درختی برای پیش‌بینی صفات عددی پیوسته است که در آن توابع رگرسیونی خطی در برگ‌های این درخت ظاهر می‌شوند (آلبرگ و همکاران ۲۰۱۲). فرآیند انشعاب در هر گره بارها تکرار می‌شود تا به گره پایانی (برگ) برسد که در برگ، مجموع مجذور انحراف از میانگین داده‌ها حدوداً به صفر می‌رسد. با این کار درخت بزرگی توسعه پیدا خواهد کرد. کار با این درخت بزرگ که شاخه‌ها و گره‌های زیادی دارد سخت می‌باشد، بنابراین برای رسیدن به یک درخت بهینه و کارآمد باید شاخه‌های اضافی درخت هرس شود. دو روش برای هرس کردن درخت وجود دارد (کوئینلن، ۱۹۹۲): (۱) هرس قبل از شکل‌گیری درخت حداکثر (۲) هرس بعد از شکل‌گیری درخت حداکثر. در روش اول فرآیند هرس اجازه نمی‌دهد شاخه‌های اضافی تولید شوند، ولی در روش دوم ابتدا درخت حداکثر تشکیل می‌شود، سپس فرآیند هرس انجام می‌گیرد. درخت بهینه بر اساس حداقل کردن خطای پیش‌بینی انتخاب می‌شود که دو روش برای محاسبه‌ی خطای پیش‌بینی وجود دارد: (۱) آزمون دستگاه مستقل (۲) آزمون صحت‌سنجی. روش اول

<sup>۱</sup>- Entropy

مرجع به‌دست آمده از برنامه‌ریزی ژنتیک و مدل درختی M5، در مقابل مقادیر محاسبه شده به وسیله‌ی رابطه‌ی پنمن-مونتیث-فائو برای پارامترهای ورودی در شکل‌های ۲ و ۳ نشان داده شده است.

شکل ۴ نیز جهت نشان دادن تشابه بالای نتایج حاصل از دو روش GP و مدل درختی M5 با نتایج عاده پنمن-مونتیث-فائو برای داده‌های آزمایش نشان داده شده است.

مقدار بالای  $R^2$  در شکل ۲ و ۳، نشان‌گر قدرت بالای مدل‌ها در پیش‌بینی تبخیر-تعرق مرجع در منطقه مورد مطالعه می‌باشد. پارامترهای آماری متعددی جهت مقایسه‌ی عملکرد روش‌های برنامه‌ریزی ژنتیک و مدل درختی M5 استفاده شده است. جدول ۲ مقادیر این پارامترها را برای سه روش فائو-پنمن-مونتیث، برنامه‌ریزی ژنتیک و مدل درختی M5 با استفاده از داده‌های آزمایشی را نشان می‌دهد.

همچنان‌که از جدول ۲ پیداست نتایج بدست آمده از هر دو مدل با ارائه مقادیر آماره‌های ضریب همبستگی بالا و ریشه میانگین مربعات خطای پایین دارای دقت مناسبی در زمینه پیش‌بینی تبخیر و تعرق دارا می‌باشند. نتایج حاصل از کاربرد مدل درختی M5 در پیش‌بینی تبخیر-تعرق مرجع در ناحیه‌ی مذکور، که منجر به ایجاد ۶ معادله خطی شده است، در شکل ۵ نشان داده شده است. معادلات خطی مذکور در بازه‌های مشخصی از داده‌های ورودی ارائه شده است. اعداد داخل پرانتز در تصویر مدل درختی M5 به ترتیب تعداد مواردی از داده‌ها را که در شرایط هر رابطه خطی صدق می‌کند نشان می‌دهد. بعنوان نمونه در مدل درختی ارائه شده در شکل ۵ براساس شاخه سمت راست چنانچه مقدار میانگین دما بزرگتر از ۱۹/۱ باشد در آنصورت مقدار تبخیر-تعرق مرجع با استفاده از رابطه خطی LM6 قابل محاسبه خواهد بود. در غیر اینصورت براساس شاخه سمت چپ درخت اقدام خواهد شد و براساس بازه‌های ارائه شده به سایر پارامترها و شرایط مرزی معادلات خطی LM1 الی LM5 مقادیر تبخیر-تعرق محاسبه خواهد گردید.

معادله و تابع حاصل از کاربرد روش برنامه‌ریزی ژنتیک نیز در رابطه ۶ نشان داده شده است.

هنگامی مورد استفاده قرار می‌گیرد که تعداد داده‌های زیادی داشته باشیم، در غیر این صورت از روش دوم استفاده می‌شود (کوئینلن، ۱۹۹۲). پس از هرس کردن، پروسه صاف کردن برای جبران گسیختگی‌های تند که به طور اجتناب ناپذیر بین مدل‌های خطی هم جوار در برگ‌های درخت هرس شده، مخصوصاً برای مدل‌هایی که از مقادیر کمتری از نمونه‌ها ساخته شده‌اند، انجام می‌شود (آلبرگ و همکاران ۲۰۱۲). در این تحقیق جهت مدل‌سازی روش M5 از نرم افزار WEKA که در دانشگاه Waikato نیوزلند توسعه داده شده، استفاده گردیده است.

### روش پنمن-مونتیث-فائو

محققین بسیاری در سراسر دنیا روش پنمن-مونتیث-فائو را به عنوان دقیق ترین روش برای برآورد تبخیر و تعرق گیاه مرجع که در حال حاضر اساس محاسبات نیازآبی می باشد، معرفی نموده اند. آنچه در معادله پنمن-مونتیث-فائو محاسبه می‌شود نیاز آبی مرجع و یا نیاز آبی یک گیاه فرضی است که به نام  $ET_0$  معروف است. هر چند این گیاهان فرضی با چمن شباهت زیادی دارد اما نباید آن‌را چمن در نظر گرفت. معادله پنمن-مونتیث-فائو در معادله ۵ نشان داده شده است (آلن و همکاران، ۱۹۹۸).

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{890}{T+273} U_2 (e_s - e_a)}{\Delta + \gamma(1+0.34U_2)} \quad (5)$$

در این معادله

$ET_0$  تبخیر-تعرق مرجع (mm/day)

$R_n$  تابش خالص در سطح پوشش گیاهی ( $MJm^{-2}d^{-1}$ )

T متوسط دمای هوا ( $^{\circ}C$ )

$U_2$  متوسط سرعت باد در ارتفاع ۲ متری ( $ms^{-1}$ )

es-ea کمبود فشار بخار در ارتفاع دو متری (KPa)

$\Delta$  شیب منحنی فشار بخار ( $KPa^{\circ}C^{-1}$ )

G شار گرما به داخل خاک ( $MJm^{-2}d^{-1}$ )

Y ضریب رطوبتی ( $KPa^{\circ}C^{-1}$ ) می‌باشد.

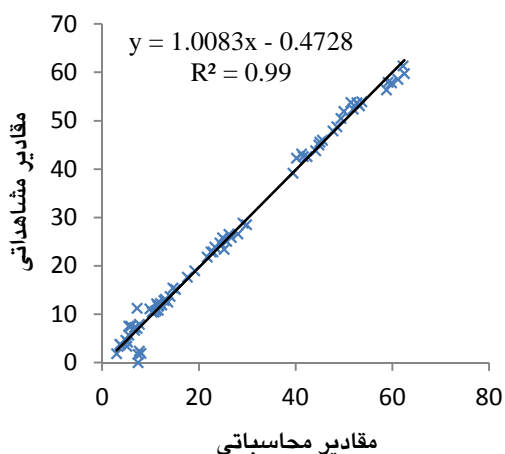
### نتایج و بحث

در این تحقیق پارامترهای ورودی شامل شاخص ماهانه، میانگین دما، رطوبت حداکثر، رطوبت حداقل، سرعت باد، ساعات آفتابی در مقیاس زمانی ماهانه جهت معرفی به مدل‌های M5 و برنامه‌ریزی ژنتیک در نظر گرفته شده است. نتایج مربوط به مقادیر مشاهداتی تبخیر-تعرق

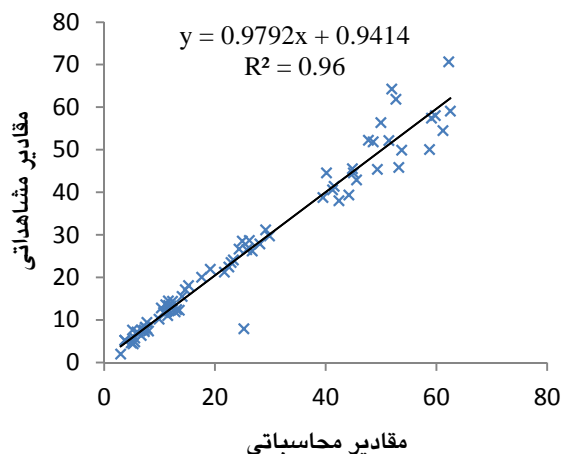
می‌باشد. مقادیر  $R^2$  و RMSE جهت مقایسه‌ی دو روش برنامه‌ریزی ژنتیک و M5 در جدول ۳ نشان داده شده است.

با توجه به این نتایج مشاهده می‌شود که به ترتیب روش‌های مدل درختی M5 و برنامه‌ریزی ژنتیک با دقت بالا و مطلوبی قادر به پیش‌بینی تبخیر-تعرق پتانسیل در ایستگاه هواشناسی تبریز می‌باشند.

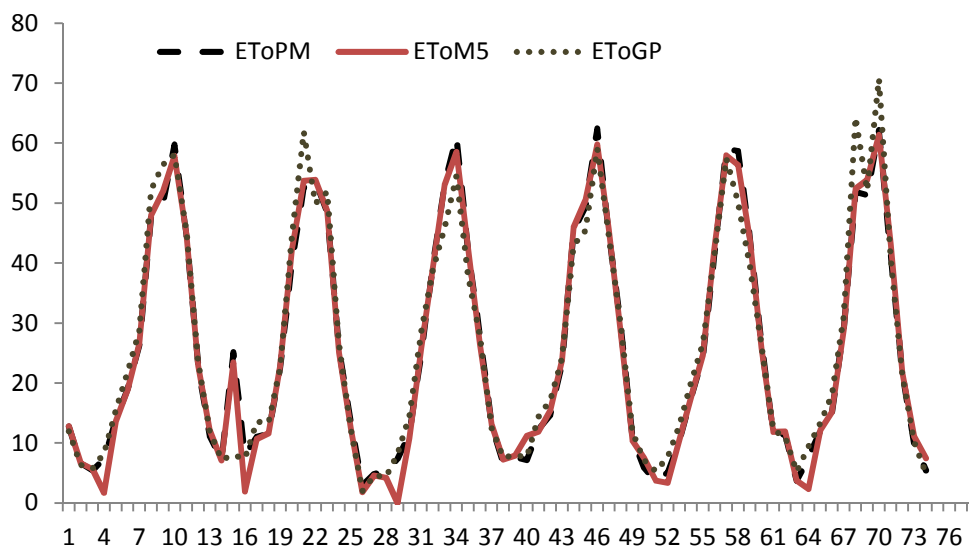
در روابط ارائه شده توسط برنامه‌ریزی ژنتیک و مدل درختی M5، پارامتر MinT حداقل دما (سانتی‌گراد)، MaxT حداکثر دما (سانتی‌گراد) و SunH ساعات آفتابی (ساعت)، MaxH رطوبت نسبی حداکثر (درصد)، MinH رطوبت نسبی حداقل (درصد)، AvT میانگین دما (سانتی‌گراد) و سرعت باد در ارتفاع ۲ متری از زمین (متر بر ثانیه)



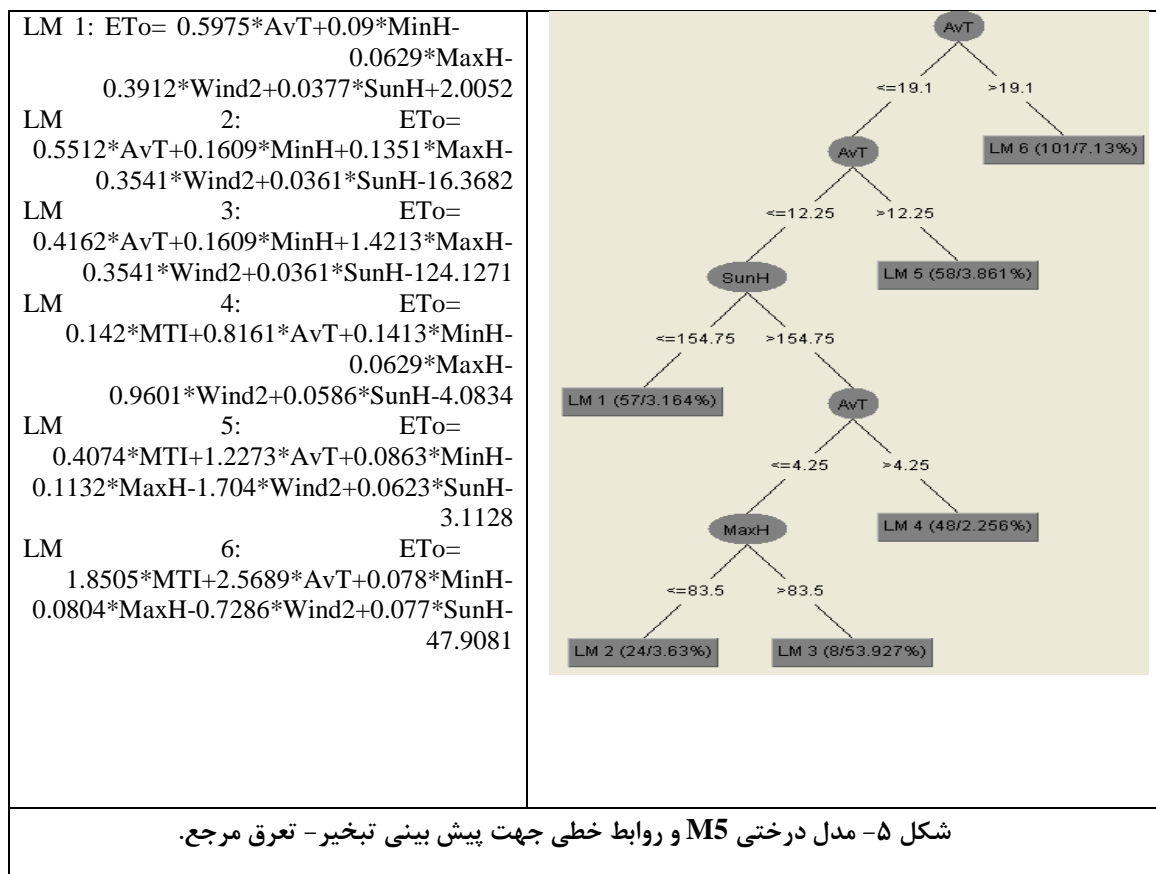
شکل ۳- مقادیر محاسباتی و مشاهداتی تبخیر-تعرق مرجع در مدل درختی M5.



شکل ۲- مقادیر محاسباتی و مشاهداتی تبخیر-تعرق مرجع در مدل GP.



شکل ۴- مقایسه نتایج مربوط به روش‌های مختلف در محاسبه تبخیر-تعرق مرجع.



جدول ۲- ویژگی های آماری نتایج به دست آمده از روش های مختلف محاسبه تبخیر-تعرق مرجع (برحسب میلی متر) برای داده های آزمون

آماره	FPM	GP	M5
میانگین	۲۵/۵	۲۵/۹۱	۲۵/۲۴
انحراف استاندارد	۱۸/۸۹	۱۸/۸۹	۱۹/۱۴
دامنه	۳۵۶/۹۹	۳۵۶/۹	۳۶۶/۲
کمینه	۲/۹۳	۱/۹۹	۰/۰۰
بیشینه	۶۲/۴۸	۷۰/۷۱	۶۱/۳۶
همبستگی با تاخیر یک روزه	۰/۷۹۳	۰/۷۸۵	۰/۸

جدول ۳- نتایج حاصل از روش های برنامه ریزی ژنتیک و مدل درختی M5 برای محاسبه تبخیر-تعرق مرجع

روش	R <sup>2</sup>	RMSE (میلی متر)
M5 مدل درختی	۰/۹۹	۱/۸۱
برنامه ریزی ژنتیک	۰/۹۶	۳/۸۴

$$ET = \frac{SunH}{MaxH - MinH - 7.306} + \frac{AvT - \frac{AvT}{SunH - AvT + MinH}}{MinH} + \frac{AvT^3}{6.077 Wind2. MaxH} \quad (۶)$$



(۲) 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. مجله آبیاری و زهکشی ایران. ۷(۱): ۱۰۴-۱۱۳.

(۳) صمدیان فرد س، دلیر حسن نیار. ۱۳۹۴. پیش بینی چریان رودخانه شهرچای در حوضه آبریز دریاچه اورمیه با استفاده از برنامه ریزی ژنتیک و مدل درختی M5. آب و خاک. ۲۹(۵): ۱۲۰۶-۱۱۹۰.

(۴) مرادی م الف، رحیمی خوب ع. ۱۳۹۱. برآورد تبخیر و تعرق مرجع با استفاده از تصاویر ماهواره نوا و مدل درختی M5 برای شبکه‌های آبیاری - مطالعه موردی شبکه آبیاری قزوین. ۱۶(۶۲): ۱۲۳-۱۳۵.

5) Alberg D, Last M and Kandel A, 2012. Knowledge discovery in data streams with regression tree methods. WIREs Data Mining Knowledge Discovery. (2): 69-78.

6) 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, Italy, 300 pp.

7) Blaney H.F, and Criddle W.D, 1950. Determining water requirements in irrigated areas from climatological and irrigation data. Soil Conservation Service Technical Paper 96. 44 pp.

8) Dittthakit P, and Chinnarasri CH, 2012. Estimation of pan coefficient using M5 model tree. American journal of environmental sciences 8 (2): 95-103

9) Esmailzadeh B, and Sattari MT, 2015. Monthly Evapotranspiration Modeling Using Intelligent Systems in Tabriz, Iran. Agricultural Science Developments. 4(3): 35-40.

10) Guven K, Aytok A, Yuce MI and Aksoy H, 2008. Genetic Programming – based empirical model for daily reference evapotranspiration estimation. Clean, 36 (10-11): 905-912.

11) Hargreaves G.H, and Samani Z.A, 1982. Estimating potential evapotranspiration. J. Irrigation and Drainage Engineering. ASCE, 108(3): 223-230.

12) Kisi O, and Guven A, 2010. Evapotranspiration modeling using linear genetic programming technique. Journal of Irrigation and Drainage. 136(10): 715-723.

13) Koza John R, 1992. Genetic Programming: On the Programming of Computers by Means of Natural Selection. Cambridge, MA: The MIT Press.

14) Pal M, and Deswal S, 2009. M5 model tree based modeling of reference evapotranspiration. Hydrological Process. 23(10): 1437-1443

15) Quinlan J.R, 1992. Learning with continuous classes. In proceedings AI,92 (Adams & Sterling, Eds), 343-348, Singapore: World Scientific

لازم به ذکر است که معادله پنمن مونتیث فائو علاوه بر دما و سرعت باد از پارامترهای فیزیکی دیگری نظیر تابش خالص در سطح پوشش گیاهی، کمبود فشار بخار در ارتفاع دو متری، شیب منحنی فشار، شار گرما به داخل خاک و ضریب رطوبتی که دسترسی به مقادیر آنها نسبتاً سخت است، بهره می برد. این درحالیست که معادلات ارائه شده در این تحقیق صرفاً از داده‌های هواشناسی قابل دسترس در ایستگاه‌های هواشناسی جهت تخمین استفاده می کند.

## نتیجه گیری

محاسبه و پیش‌بینی مقدار دقیق تبخیر-تعرق و به تبع آن برآورد دقیق نیازآبی 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 ایستگاه هواشناسی تبریز شامل میانگین دما، حداقل و حداکثر رطوبت نسبی، سرعت باد، ساعات آفتابی و ماه مورد نظر به عنوان ورودی جهت پیش‌بینی تبخیر-تعرق مرجع ماهانه با استفاده از دو روش برنامه‌ریزی ژنتیک (GP) و مدل درختی M5 استفاده گردید. با توجه به مقدار بالای ضریب همبستگی نتایج حاصل از پیش‌بینی تبخیر-تعرق نشان داد هر دو مدل از دقت قابل قبول و بالایی برخوردار است. همچنین مدل درختی M5 به لحاظ سادگی، قابل فهم بودن و ارائه روابط خطی ساده در محدوده خاصی از داده‌های ورودی که بصورت روابط خطی ارائه می‌گردد، می‌تواند به‌عنوان یک روش جایگزین قابل رقابت با سایر روش‌ها طرح شود. پال و دسوال (۲۰۰۹)، میزان تبخیر-تعرق مرجع را فقط با استفاده از مدل درختی M5 و به کارگیری چهار پارامتر هواشناسی (میانگین دما، رطوبت نسبی متوسط، ساعات آفتابی و سرعت باد) به عنوان ورودی مدل درختی M5 تخمین زدند. نتایج تحقیق آنان نشان داد روش مدل درختی M5 با توجه به ارائه روابط خطی ساده و قابل فهم دقت بالایی داشته و بسیار کاربردی می باشد، که این نتیجه با نتایج تحقیق حاضر همخوانی دارد.

## مراجع

(۱) سامتی م، قهرمان ن و قربانی خ. ۱۳۹۲. کاربرد مدل M5 برآورد تبخیر-تعرق مرجع در ایستگاه های شیراز و کرمانشاه. پژوهش آب در کشاورزی. ۲۷(۳): ۲۸۹-۲۹۸.

- 16) Sattari, M.T, Pal, M., Yürekli, K., Ünlükara, A, 2013. M5 model trees and neural network based modelling of ETO in Ankara, Turkey. Turkish Journal of Engineering and Environmental Science. 37: 211-219.
- 17) Tabatabaee M, Mirshekari M, Sheikh Alipour Z, Tahmasebizade R, 2014. Comparing The Results Of An Empirical Equation And Artificial Intelligence To Calculate Evapotranspiration And Compare Them With The Actual Results (Case Study: Kavar And Doroudzan Weather Stations). Journal of current research in science. 2(3): 340-345.
- 18) Terzi O, 2010. Modeling of daily pan evaporation of Lake Egirdir using data-driven techniques. International Symposium on Innovations in Intelligent Systems and Applications. 320-324. Istanbul. Turkey.