

ارزیابی مدل‌های تجربی و هوشمند در برآورد تبخیر و تعرق مرجع

یاسر سبزواری^۱، مرادعلی قنبرپوری^{۲*}

تاریخ پذیرش: ۱۴۰۱/۷/۱۴

تاریخ دریافت: ۱۴۰۰/۵/۱۹

چکیده

تبخیر و تعرق مرجع که مهم‌ترین جزء چرخه هیدرولوژی است، با روش‌های مختلفی محاسبه می‌شود و هرکدام به اطلاعات ورودی مختلف و بعضاً زیادی نیاز دارد که در بسیاری شرایط دسترسی به همه اطلاعات امکان‌پذیر نیست. این امر انتخاب روش‌های با داده‌های ورودی کمتر و دقت مناسب را ضروری می‌سازد. در این پژوهش هفت مدل برآورد تبخیر و تعرق مرجع و مدل برنامه‌ریزی بیان ژن (GEP) نسبت به فائوپنمن‌مانتیت ارزیابی شدند. بدین منظور از اطلاعات روزانه بیشینه و کمینه درجه حرارت، بیشینه و کمینه رطوبت نسبی، سرعت باد و ساعات آفتابی ۳۵ ساله (۱۹۸۳-۲۰۱۷) ایستگاه سینوپتیک الیگودرز استفاده شد. ۷۰٪ داده‌ها برای آموزش و ۳۰٪ داده‌ها برای آزمون مدل به کار گرفته شدند. نتایج نشان داد مدل‌های کیمبرلی پنمن و تشعشعی فائو دارای دقت بیشتری نسبت به سایر مدل‌های تجربی‌اند. تحلیل ضرایب مدل رگرسیون چندمتغیره حاکی از بیشترین تأثیر بیشینه درجه حرارت با ضریب ۰/۵۸ بر تبخیر و تعرق مرجع بود. در GEP، الگوی شماره ۲ با عملگرهای پیش فرض مدل با $RMSE=0/843$ و $R^2=0/932$ در مرحله آموزش و $RMSE=0/76$ و $R^2=0/941$ در مرحله آزمون عملکرد بهتری را داشته است. مقایسه مدل‌های برآورد تبخیر و تعرق مرجع حاکی از برتری مدل GEP نسبت به سایر مدل‌ها بود. نتایج این پژوهش نشان داد که مدل GEP دارای توانایی قابل قبولی در تخمین تبخیر و تعرق مرجع تحت شرایط آب‌وهوایی الیگودرز بوده و به عنوان مدل قابل استفاده در این زمینه معرفی کرد.

کلیدواژه‌ها: مدل تشعشعی، مدل دمایی، مدل تشعشعی-دمایی، رگرسیون، GEP.

۱. دانشجوی دکتری آبیاری و زهکشی، گروه مهندسی آب، دانشگاه صنعتی اصفهان، Y.sabzevari@ag.iut.ac.ir

۲. محقق بخش تحقیقات خاک و آب، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی لرستان، نویسنده مسئول، moradghanbari@gmail.com

* این مقاله برگرفته از طرح پژوهشی است.

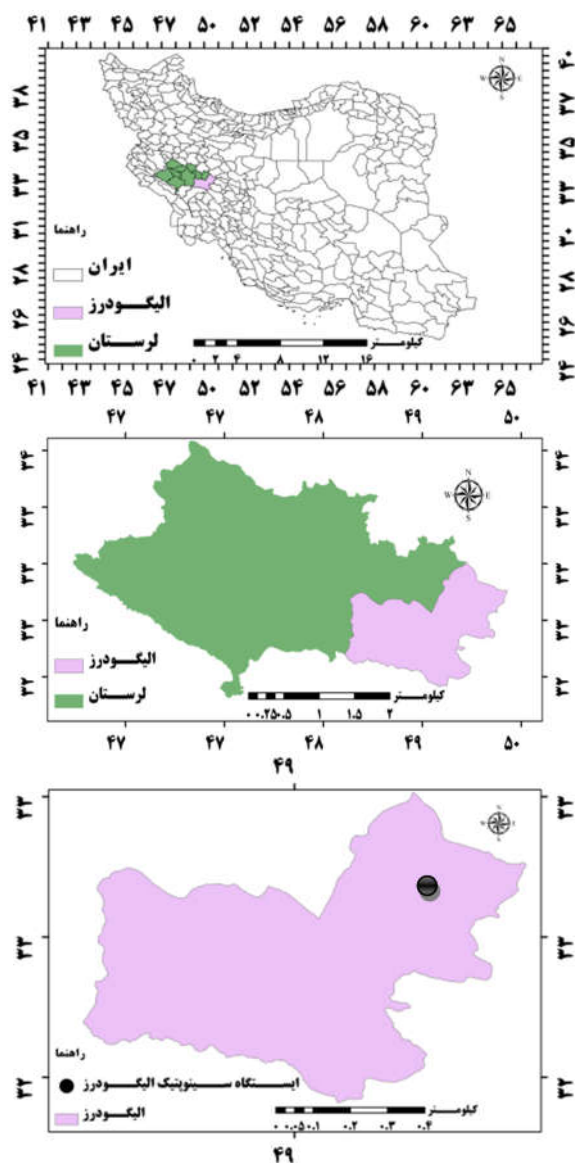
مقدمه

برآورد دقیق نیاز آبی گیاهان، عاملی اصلی در کنترل برنامه‌ریزی و مدیریت منابع آب (طبری و همکاران، ۲۰۱۲؛ لابودی^۱ و همکاران، ۲۰۱۲؛ ون^۲ و همکاران، ۲۰۱۵)، تعیین قیمت آب (پورعلی‌بابا و همکاران، ۲۰۱۳) و نیاز آب آبیاری (یاسین و همکاران، ۲۰۱۶) است. یکی از مهم‌ترین اجزای نیاز آبی، تبخیر و تعرق مرجع (ET_0) بوده که تابع متغیرهای اقلیمی مختلفی است. روش‌های مختلفی برای محاسبه تبخیر و تعرق مرجع وجود دارد که هرکدام با توجه به فرضیات و داده‌های هواشناسی مختلفی که در نظر می‌گیرند، اغلب نتایج متفاوتی به دست می‌دهند (گریسمر^۳ و همکاران، ۲۰۰۲). به‌طور معمول مدل‌های تبخیر و تعرق با استفاده از لایسی‌متر یا با یک مدل مرجع واسنجی می‌شوند که معمولاً مدل فانوینمن‌مانتیت (FAO56) به‌عنوان مدل مرجع در نظر گرفته می‌شود (طبری و همکاران، ۲۰۱۳؛ نیاقی^۴ و همکاران، ۲۰۱۳؛ لیو^۵ و همکاران، ۲۰۱۷). سازمان خواربار جهانی (FAO^۶) این روش را به‌عنوان معادله استاندارد توصیه کرده است (آلن^۷ و همکاران، ۱۹۹۸) و مطالعات زیادی در سراسر دنیا صحت نتایج مربوط به مدل فانوینمن‌مانتیت در شرایط مختلف آب‌وهوایی را نشان می‌دهد (برتی^۸ و همکاران، ۲۰۱۴؛ ژامان^۹ و همکاران، ۲۰۱۵). این روش به تعدادی از متغیرهای اقلیمی نظیر حداقل و حداکثر دمای هوا، تابش خورشیدی، رطوبت نسبی و ساعات آفتابی نیاز دارد. با وجود این، این متغیرهای هواشناسی به‌ویژه در کشورهای در حال توسعه، یا در دسترس نبوده یا معمولاً قابل اعتماد نیستند (آلموروکس^{۱۰} و همکاران، ۲۰۱۵)؛ بنابراین ارائه روشی که بتواند با داده‌های ورودی کمتر، دقت مناسبی را به همراه داشته باشد، به‌خصوص در این مناطق، الزامی است. در سال‌های اخیر، روش‌های هوش مصنوعی (AI^{11}) با داده‌های

هواشناسی محدود، به‌طور موفقیت‌آمیزی در شبیه‌سازی ET_0 مورد استفاده قرار گرفته‌اند. از جمله روش‌های هوش مصنوعی به‌کارگرفته‌شده برای شبیه‌سازی ET_0 می‌توان به برنامه‌ریزی بیان ژن (شیری و همکاران، ۲۰۱۴؛ مارتی^{۱۲} و همکاران، ۲۰۱۵) اشاره کرد. در همین راستا مطالعات مختلفی برای مقایسه نتایج مدل‌های ساده‌تر برآورد تبخیر و تعرق با مدل مرجع و مدل‌سازی آن برای ارائه روابط با ورودی‌های کمتر انجام شده است. لیو و همکاران (۲۰۱۷) به ارزیابی ۱۶ مدل برآورد تبخیر و تعرق با لایسی‌متر وزنی در پکن چین پرداختند؛ نتایج حاکی از بیش‌برآوردی همه ۱۶ مدل نسبت به نتایج لایسی‌متر بود. هژبر و همکاران (۲۰۱۷) به مقایسه ۱۲ مدل محاسبه تبخیر و تعرق و مدل‌سازی شبکه عصبی با نتایج لایسی‌متر در ایستگاه کهرئز ارومیه پرداختند؛ نتایج بیانگر دقت بالاتر مدل‌سازی شبکه عصبی نسبت به سایر مدل‌ها بود و از بین مدل‌ها، مدل تورک مناسب‌ترین مدل شناخته شد. خوشحال و همکاران (۲۰۱۶) چند مدل تبخیر و تعرق را با نتایج تشت تبخیر در حوزه آبریز شرق و جنوب شرق کشور ارزیابی کردند و به این نتیجه رسیدند که مدل هارگریوز سامانی، بلانی کریدل و تورک مناسب‌ترین مدل‌ها در منطقه مطالعاتی است. با توجه به نتایج پژوهش‌های مورد اشاره، هر مدل برآوردکننده تبخیر و تعرق مرجع در منطقه‌ای خاص با شرایط اقلیمی مربوط به آن محل استخراج شده است. حسینی و همکاران (۲۰۱۵) به ارزیابی و تحلیل حساسیت روش‌های مختلف تخمین تبخیر و تعرق مرجع روزانه در یک اقلیم خشک سرد پرداختند. نتایج مشخص کرد که شبکه عصبی مصنوعی عملکرد مناسب‌تری نسبت به مدل‌های تجربی و مدل رگرسیون چندمتغیره دارد. سناری و اسماعیل‌زاده (۲۰۱۶) به مقایسه نتایج مدل درختی M5 و برنامه‌ریزی ژنتیک با روش فانوینمن‌مانتیت برای تخمین تبخیر و تعرق مرجع پرداختند؛ نتایج نشان داد که هر دو روش مذکور نتایج دقیقی را برای پیش‌بینی تبخیر و تعرق مرجع ماهانه در منطقه تبریز ارائه می‌دهند. فیض‌الله‌پور و همکاران (۲۰۱۷) به ارزیابی و تحلیل عدم قطعیت برآورد تبخیر و تعرق گیاه مرجع با استفاده از

1. Laaboudi
2. Wen
3. Grismer
4. Niaghi
5. Liu
6. Food and Agriculture Organization
7. Allen
8. Berti
9. Djaman
10. Almorox
11. Artificial Intelligence

درجه و ۲۴ دقیقه عرض شمالی با وسعت ۱۳۰ هکتار در شرق استان لرستان واقع شده است. این منطقه دارای ارتفاع ۲۰۲۲ متری از سطح دریا و بارندگی سالانه بین ۳۰۰ تا ۹۰۰ میلی‌متر (فلاح و طهماسبی‌پور، ۲۰۱۶) و تبخیر سالانه ۲۰۴۸/۲ میلی‌متر است. طبق روش‌های طبقه‌بندی اقلیمی دمارتن و آمبرژه دارای اقلیم نیمه‌خشک و خشک سرد، با تابستان معتدل و زمستان بسیار سرد است. شکل (۱) نشان‌دهنده موقعیت جغرافیایی الیگودرز در لرستان و کشور است. این شهرستان دارای یک ایستگاه سینوپتیک با عرض جغرافیایی ۳۳° ۲۴' و طول جغرافیایی ۴۹° ۴۲' و ارتفاع از سطح دریا ۲۰۲۲ متر بوده که در این مطالعه از اطلاعات این ایستگاه استفاده شده است.



شکل (۱): موقعیت جغرافیایی الیگودرز در لرستان و کشور
 Figure (1): Geographical position of Aligoodarz plain in Lorestan and Iran

برنامه‌ریزی ژنتیک پرداختند؛ نتایج نشان داد که مدل‌های برنامه‌ریزی ژنتیک دارای دقت بیشتری نسبت به روش‌های تجربی‌اند. لدلنی^۱ و همکاران (۲۰۱۴) تبخیر و تعرق روزانه در منطقه مدیریتانه‌ای کشور الجزایر را توسط روش‌های سیستم فازی^۲ و رگرسیون چندمتغیره خطی مدل‌سازی کردند؛ نتایج به‌دست‌آمده نشان از عملکرد قابل قبول هر دو مدل در پیش‌بینی تبخیر و تعرق داشت. یاسین و همکاران (۲۰۱۶) به ارزیابی شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن برای شبیه‌سازی تبخیر و تعرق مرجع در اقلیم خشک پرداختند؛ نتایج نشان داد که ۸ مدل برآورد تبخیر و تعرق مرجع با استفاده از شبکه عصبی مصنوعی نسبت به همان مدل‌ها با برنامه‌ریزی بیان ژن، دقت بیشتری دارند. شیری (۲۰۱۷) به ارزیابی مدل‌های فائوپنمن‌مانتیت ۵۶، تجربی، نیمه‌تجربی و برنامه‌ریزی بیان ژن در شبیه‌سازی تبخیر و تعرق مرجع روزانه برای مناطق فراخشک ایران پرداخت. نتایج حاکی از برتری مدل برنامه‌ریزی بیان ژن نسبت به روش‌های تجربی و نیمه‌تجربی است. کیشی و علی‌ضمیر^۳ (۲۰۱۸) به مدل‌سازی تبخیر و تعرق مرجع با استفاده از یک روش اکتشافی پیوستگی موجک پرداختند و نتایج نشان داد که مدل‌های شامل چهار متغیر ورودی، دقت بهتری را نسبت به مدل‌های با ورودی کمتر ارائه می‌دهند. تشعشعات خورشیدی مؤثرترین متغیر در ET₀ بود. با بررسی مطالعات انجام‌شده می‌توان به ضرورت ارزیابی روش‌های مختلف برآورد تبخیر و تعرق مرجع و ارائه روابطی با داده‌های ورودی محدودتر و دقت نزدیک به مدل فائوپنمن‌مانتیت در مناطق مختلف پی برد. لذا هدف از پژوهش حاضر ارزیابی هفت مدل تجربی برآورد تبخیر و تعرق مرجع و مدل‌سازی آن با مدل برنامه‌ریزی بیان ژن (GEP^۴) تحت شرایط آب‌وهوایی الیگودرز است.

مواد و روش‌ها

منطقه مورد مطالعه

شهرستان الیگودرز بین ۴۹ درجه و ۴۲ دقیقه طول شرقی و ۳۳

1. Ladlani
2. Fuzzy
3. Kisi and Alizamir
4. Gene Expression Programming

روش پژوهش

در این پژوهش برای محاسبه روزانه تبخیر و تعرق گیاه مرجع در شرایط آب و هوایی الیگودرز، از اطلاعات بیشینه درجه حرارت (T_{max})، کمینه درجه حرارت (T_{min})، بیشینه رطوبت نسبی (RH_{max})، کمینه رطوبت نسبی (RH_{min})، ساعات آفتابی (Sun) و سرعت باد (WS) ایستگاه سینوپتیک الیگودرز مربوط به دوره زمانی ۳۵ ساله (۱۹۸۳-۲۰۱۷) استفاده شد. به منظور برآورد ET_0 از روش های تعریف شده در نرم افزار RefET-۳/۲ استفاده شد. در این نرم افزار، روش های محاسباتی با توجه به نوع داده های ورودی (بیشینه و کمینه درجه حرارت، درجه حرارت متوسط روزانه، بیشینه و کمینه رطوبت نسبی، رطوبت نسبی متوسط روزانه، سرعت باد، میزان بارش، ارتفاع و مختصات جغرافیایی ایستگاه) شامل ۷ روش ترکیبی بر پایه پنمن، دو روش دمایی، یک روش تشعشعی و سه روش ترکیبی تشعشعی-دمایی است. در این پژوهش از دو روش ترکیبی فائوپنمن مانیتث و کیمبرلی پنمن (۱۹۸۲)، دو روش دمایی بلانی کریدل فائو ۲۴ و هارگریوز سامانی، روش تشعشعی فائو ۲۴ و سه روش تشعشعی دمایی شامل معادلات پرستلی تیلور، مکینک (۱۹۵۷) و تورک استفاده شد. روش محاسبه هر کدام از روش ها در جدول (۱) آورده شده است.

در این روابط، ET_0 تبخیر-تعرق گیاه مرجع برحسب میلی متر بر روز و R_n تشعشع خالص برحسب مگاژول بر مترمربع بر روز (مقدار R_n برای عرض های جغرافیایی مختلف برای هر ماه به دست می آید)، G شار گرمایی خاک برحسب مگاژول بر مترمربع بر روز، Δ شیب نمودار فشار بخار اشباع نسبت به دما برحسب کیلو پاسکال بر درجه سلسیوس، γ ثابت سایکرومتری برحسب کیلو پاسکال بر درجه سلسیوس، $e_a - e_d$ کمبود فشار بخار در ارتفاع ۲ متری برحسب کیلو پاسکال، U_2 میانگین روزانه سرعت باد در ارتفاع ۲ متری از سطح زمین برحسب متر بر ثانیه، T_{mean} میانگین روزانه دمای هوا در ارتفاع ۲ متری از سطح زمین برحسب درجه سلسیوس، T_D میانگین دامنه تغییرات دمای روزانه (اختلاف حداکثر و حداقل دما) در دوره مذکور، a_w و b_w ضرایب تجربی (که بنا بر پیشنهاد اولیه پنمن در صورتی که سرعت باد برحسب متر بر ثانیه باشد،

مقادیر عددی آنها به ترتیب واحد و 0.537 است) و k_w ضریب ثابت برابر $6/43$ برحسب میلی متر بر روز است. در این پژوهش، برای بررسی اثرگذاری متغیرهای مختلف و تعیین الگوهای ورودی مدل ها و مدل سازی تبخیر و تعرق مرجع، از روش های رگرسیون چندمتغیره (MLR^1) و برنامه ریزی بیان ژن استفاده شد.

مدل رگرسیونی خطی-چندمتغیره

با این روش می توان به طور هم زمان به تحلیل و بررسی چند متغیر مختلف پرداخت. رگرسیون چندمتغیره در واقع، رابطه بین چند متغیر پیش بین (مستقل) با متغیر پاسخ مورد نظر را بیان می کند. این گونه مدل ها دارای فرضیاتی هستند؛ فرضی که رگرسیون چندمتغیره را از رگرسیون ساده متمایز می کند؛ این است که: ۱. تعداد متغیرهای پیش بین در رگرسیون باید کمتر از تعداد مشاهدات باشد؛ ۲. همبستگی خطی کامل بین متغیرهای پیش بین و پاسخ وجود داشته باشد.

در صورت نقض این دو فرض، معادله رگرسیونی را نمی توان تخمین زد. مدل رگرسیونی به شکل ماتریسی را می توان به صورت معادله زیر نشان داد:

$$Y = X\beta + e \quad (1)$$

که β ماتریس ضرایب رگرسیونی که ضرایب نامشخصی هستند، مسئول برآورد پارامتر پاسخ اند (سرمدیان و همکاران، ۲۰۱۰؛ غلامی و همکاران، ۲۰۱۲)، e ماریس خطای برآزش و Y نیز ماتریس پاسخ است. با حل معادله (۲) برحسب β خواهیم داشت:

$$\beta = (X'X)^{-1}(X'Y) \quad (2)$$

که X' ترانهاده ماتریس X است. برای محاسبه معکوس $(X'X)$ لازم است متغیرهای مستقل همبستگی زیادی نداشته باشند، زیرا در این صورت ماتریس $(X'X)$ را نمی توان معکوس کرد و باعث افزایش خطا در اثر گرد کردن داده ها و محاسبات می شود. در این مطالعه برای انجام تحلیل رگرسیون چندمتغیره از نرم افزار SPSS-24 استفاده شد.

جدول (۱): مدل‌های تجربی برآورد تبخیر و تعرق مرجع

Table (1): Experimental Models of Reference ET_0

ردیف	نام روش محاسبه ET_0	رفرنس	شکل معادله
۱	هارگریوز سامانی	(Hargreaves, 1994)	$ET_0 = 0.0023 R_n (T_{mean} + 17.8) T_D^{0.5}$
۲	بلانی کریدل فائو ۲۴	(Blaney and Criddle, 1950)	$ET_0 = a + b [P(0.46 T_{mean} + 8.13)]$
۳	مکینک	(Makkink, 1957)	$ET_0 = 0.61 \frac{\Delta}{\Delta + \gamma} \cdot \frac{R_s}{2.45} - 0.12$
۴	تورک-۱۹۶۱	(Turc, 1961)	$ET_0 = a_T \times 0.013 \frac{T_{mean}}{T_{mean} + 15} \times \frac{23.8856 R_s + 50}{\lambda}$
۵	پریستلی تیلور	(Allen et al., 1998)	$ET_0 = 1.26 \frac{\Delta}{\Delta + \gamma} \cdot \frac{R_n - G}{\lambda}$
۶	معادله اصلاح شده پنمن توسط کیمبرلی ۱۹۸۲	(Allen et al., 1998)	$ET_0 = \frac{\Delta}{\Delta + \gamma} (R_n - G) + k_w \frac{\gamma}{\gamma + \Delta} (a_w + b_w u_2) (e_s - e_o) / \lambda$
۷	تابشی فائو	(Allen et al., 1998)	$ET_0 = a + b \left[\frac{\Delta}{\Delta + \gamma} R_s \right]$
۸	فائوپنمن مانتیث	(Allen et al., 1998)	$ET_0 = \frac{[0.408(R_n - G) + [(900\gamma)/(T + 273)] \times U_2 \times (e_a - e_d)]}{\Delta + \gamma \times (1 + 0.34U_2)}$

برنامه‌ریزی بیان ژن

معیار زمان: اگر پس از گذشت زمان معینی، در بهترین

جواب حاصل از تولید نسل بهبودی حاصل نشود، محاسبات متوقف خواهد شد.

ساختار ژن‌ها در برنامه‌ریزی بیان ژن

یک کروموزوم از یک یا چند ژن تشکیل می‌شود و هر ژن شامل یک سر و یک دنباله است. اندازه سر (h) به وسیله کاربر تعیین شده اما اندازه دنباله (t) به صورت تابعی از h و پارامتر n محاسبه می‌شود که پارامتر n حداکثر تعداد آرگومان توابع موجود است. رابطه (۳) طول دنباله را با توجه به پارامترهای دیگر تعیین می‌کند (لوپز و وینرت، ۲۰۰۴):

$$t = h(n-1) + 1 \quad (3)$$

مسائل با استفاده از مجموعه‌ای از توابع و ترمینال‌ها مدل‌سازی می‌شوند. مجموعه توابع معمولاً شامل توابع حسابی {+،*،/،-}، توابع مثلثاتی یا هر نوع توابع ریاضی دیگر مانند { x^2 , exp, log, sin, cos, ...} یا توابع تعریف شده توسط کاربر است. توابع و ترمینال‌ها در بخش سر ژن وجود دارند و دنباله تنها شامل ترمینال‌هاست (لوپز و وینرت، ۲۰۰۴).

روش برنامه‌ریزی بیان ژن، ترکیب و توسعه یافته روش‌های GA و GP است که فریرا در سال ۱۹۹۹ ابداع کرد. در این روش کروموزوم‌های خطی و ساده با طول ثابت، مشابه با الگوریتم ژنتیک و ساختارهای شاخه‌ای با اندازه و اشکال مختلف، مشابه با درختان تجزیه در برنامه‌ریزی ژنتیک ترکیب می‌شوند (فریرا، ۲۰۰۱).

الگوریتم برنامه‌ریزی بیان ژن

نخستین گام در GEP، تشکیل جمعیت اولیه از راه‌حل‌هاست؛ سپس کروموزوم‌ها با بیان درختی (ET_s) به نمایش درمی‌آیند. در گام بعد، باید سازگاری یا کارایی هر عضو از جمعیت کروموزوم‌ها به وسیله تابع برآزش ارزیابی شود. در پدیده تکامل و در نظام انتخاب طبیعی این نقش بر عهده طبیعت است که پاسخ‌های سازگارتر را بر دیگر جواب‌ها ترجیح دهد. در سیستم GEP معیارهای توقف به شرح زیر است:

معیار دستیابی به هدف: عدم ملاحظه تفاوت محسوس بین پاسخ‌های بهینه طی چند نسل می‌تواند معیاری برای دستیابی به پاسخ بهینه باشد.

معیار تکرار: چنانچه بعد از تعداد نسل معین، به پاسخ بهینه

برسیم، فرایند تولید نسل متوقف خواهد شد.

$$R^2 = \left(\frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2 \sum_{i=1}^N (Y_i - \bar{Y})^2}} \right)^2 \quad \text{و} \quad 0 \leq R^2 \leq 1 \quad (5)$$

$$MAE = \frac{\sum_{i=1}^n |x_i - y_i|}{n} \quad (6)$$

$$PE = \frac{|x_{avg} - y_{avg}|}{y_{avg}} * 100 \quad (7)$$

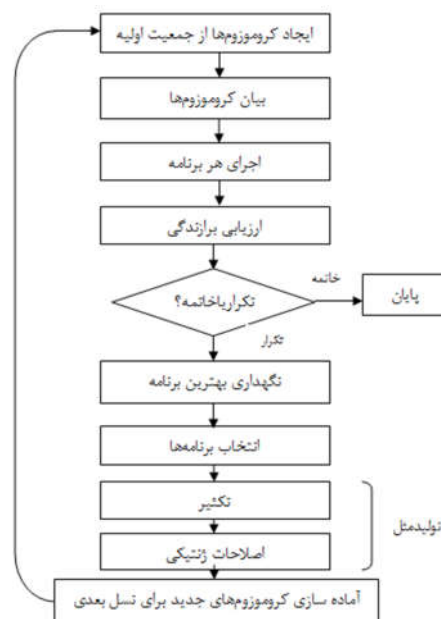
$$MR = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{y_i} \quad (8)$$

در روابط بالا، X_i و Y_i به ترتیب مقادیر محاسباتی و مشاهداتی در گام زمانی i ام، N تعداد گام‌های زمانی، \bar{X} و \bar{Y} نیز به ترتیب میانگین مقادیر مشاهداتی و محاسباتی هستند. مقادیر آماره‌های ریشه میانگین مربعات خطا و میانگین خطای مطلق همواره مثبت بوده و بهترین حالت عملکرد زمانی است که مقدار به صفر نزدیک شود. مقدار مثبت آماره درصد خطای تخمین بیانگر آن است که مدل مورد بررسی تبخیر و تعرق مرجع را بیشتر برآورد کرده و مقدار منفی آن بیانگر آن است که مدل مقدار تبخیر و تعرق را کمتر برآورد می‌کند. مقدار ضریب تبیین همواره بین -1 و $+1$ قرار دارد و هرچه به 1 نزدیک‌تر باشد، بیانگر همبستگی بیشتر و مثبت یا منفی بودن علامت آن نشان‌دهنده هم‌جهت یا خلاف جهت بودن همبستگی است. شاخص نسبت میانگین به منظور بیش‌برآوردی یا کم‌برآوردی مدل‌ها نسبت به مقدار اندازه‌گیری شده، مورد استفاده قرار گرفت. علاوه بر محک‌های آماری مورد اشاره، به منظور بررسی میزان همبستگی نتایج مدل‌ها از ضریب همبستگی اسپیرمن استفاده شد.

نتایج و بحث

جدول (۲) دربردارنده مقادیر شاخص‌های MAE ، $RMSE$ ، PE ، MR و ضریب همبستگی اسپیرمن مدل‌های پریستلی تیلور، هارگریوزسامانی، بلانی کریدل، مدل تشعشعی فائو، کیمبرلی پنمن، تورک و مکینک، نسبت به مدل مرجع فائوپنمن مانتیث است. بر اساس شاخص‌های MAE و $RMSE$ مناسب‌ترین مدل، به ترتیب کیمبرلی پنمن با مقادیر $1/0.9$ و $0/72$

به طور خلاصه، مراحل اصلی الگوریتم GEP در شکل زیر نشان داده شده است:



شکل (۲): الگوریتم برنامه‌ریزی بیان ژن (فریرا، ۲۰۰۱)

Figure (2): Gene Expression Programming Algorithm (Ferreira, 2001)

در این پژوهش، برای مدل‌سازی با GEP، از نرم‌افزار GeneXproTools-5 استفاده شد.

معیارهای ارزیابی

برای مقایسه روش‌های تجربی پریستلی تیلور، مکینک، هارگریوزسامانی، بلانی کریدل، مدل تشعشعی فائو ۲۴، اصلاح‌شده پنمن توسط کیمبرلی (۱۹۸۲) و مدل تورک نسبت به مدل فائوپنمن مانتیث، میانگین مربعات خطا ($RMSE$)، میانگین خطای مطلق (MAE)، درصد خطای تخمین (PE) و نسبت میانگین (MR) (رئوف و عزیزمی مبصر، ۲۰۱۷) و به منظور ارزیابی دقت و کارایی مدل‌ها، از معیارهای ضریب تبیین (R^2) و ریشه میانگین مربعات خطا ($RMSE$) به صورت زیر استفاده شد. بهترین و بدترین مقادیر برای این دو معیار به ترتیب، یک و صفر است.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2} \quad (4)$$

1. Root Mean Square Error
2. Mean Absolute Error
3. Percentage Error of Estimation
4. Mean Ratio

خطای تخمین در برآورد تبخیر و تعرق مرجع و در مقابل مدل کیمبرلی پنمن با $۰/۵/۳۷$ ، مدل تشعشعی فائو با $۰/۱۷/۱۳$ ، مدل تورک با $۰/۴۶/۱۱$ و مدل بلانی کریدل با $۰/۴۷/۳۶$ خطای تخمین، مناسب‌ترین مدل در برآورد تبخیر و تعرق مرجع نسبت به مدل مرجع هستند. با در نظر گرفتن مقدار آماره P در ضریب همبستگی اسپیرمن، این مقدار برای همه مدل‌ها کمتر از مقدار حدی $۰/۰۵$ است. بنابراین نتایج مدل‌های مورد استفاده در مقایسه با نتایج مدل فائوپنمن مانتیث از همبستگی قابل قبولی برخوردارند. بر اساس ضریب تبیین و ضریب همبستگی اسپیرمن، مدل‌های کیمبرلی پنمن، تشعشعی فائو و بلانی کریدل مناسب‌ترین مدل‌ها معرفی شدند و در مقابل مدل‌های تورک، پرستلی تیلور و مکینک به‌عنوان نامناسب‌ترین مدل‌ها شناخته شدند. در یک جمع‌بندی کلی و با در نظر گرفتن همه‌ی شاخص‌های ارزیابی در محل پژوهش نسبت به مدل فائوپنمن مانتیث، ترتیب مناسب بودن مدل‌ها این‌گونه است: مدل کیمبرلی پنمن، تشعشعی فائو، بلانی کریدل، تورک، پرستلی تیلور، هارگریوزسامانی و مکینک.

$(mm.day^{-1})$ ، مدل تشعشعی فائو با مقادیر $۱/۱$ و $۰/۹$ $(mm.day^{-1})$ ، مدل بلانی کریدل با مقادیر $۲/۳۱$ و $۱/۸۵$ $(mm.day^{-1})$ و مدل تورک با مقادیر $۲/۳۲$ و $۱/۸۶$ بودند. بر همین اساس و با در نظر گرفتن شاخص‌های RMSE و MAE، در محل پژوهش به‌ترتیب مدل مکینک با مقادیر $۳/۲۵$ و $۲/۵$ $(mm.day^{-1})$ ، هارگریوزسامانی با مقادیر $۳/۲۴$ و $۲/۴۲$ $(mm.day^{-1})$ و مدل پرستلی تیلور با مقادیر $۳/۱۸$ و $۲/۳۸$ $(mm.day^{-1})$ در مقایسه با مدل مرجع فائوپنمن مانتیث ضعیف‌ترین نتایج را داشتند. با توجه به شاخص MR، مدل کیمبرلی پنمن با $۰/۴$ نسبت به مدل فائوپنمن مانتیث، مقدار تبخیر و تعرق مرجع روزانه را کمتر برآورد کرده است. همچنین مدل مکینک با $۰/۶۶$ ، هارگریوزسامانی با $۰/۶۳$ ، مدل پرستلی تیلور با $۰/۵۴$ ، مدل تورک با $۰/۳۸$ ، مدل‌های بلانی کریدل با $۰/۳۶$ و تشعشعی فائو با $۰/۳۰$ نسبت به مدل فائوپنمن مانتیث، مقدار تبخیر و تعرق مرجع روزانه را بیش‌تر برآورد کردند. بر اساس شاخص PE، بدترین مدل‌ها به‌ترتیب، مدل مکینک با $۸۰/۵۳$ ، هارگریوزسامانی با $۷۸/۵۱$ و مدل پرستلی تیلور با $۷۱/۶$

جدول (۲): ارزیابی عملکرد شش مدل برآورد تبخیر و تعرق مرجع نسبت به مدل فائوپنمن مانتیث

ردیف	مدل	ضریب اسپیرمن		MR	R ²	PE (%)	MAE (mm.day ⁻¹)	RMSE (mm.day ⁻¹)
		R	P-value					
۱	پرستلی - تیلور	۰	۰/۸۸	۱/۵۴	۰/۷۹	۷۱/۶	۲/۳۸	۳/۱۸
۲	هارگریوز - سامانی	۰	۰/۹۰	۱/۶۳	۰/۸۱	۷۸/۵۱	۲/۴۲	۳/۲۴
۳	بلانی کریدل	۰	۰/۹۳	۱/۳۶	۰/۸۶	۴۷/۳۶	۱/۸۵	۲/۳۱
۴	تشعشعی فائو	۰	۰/۹۸	۱/۳	۰/۹۵	۱۷/۱۳	۰/۹	۱/۱
۵	کیمبرلی - پنمن	۰	۰/۹۸	۰/۹۶	۰/۹۵	۵/۳۷	۰/۷۲	۱/۰۹
۶	تورک	۰	۰/۸۸	۱/۳۸	۰/۷۸	۴۶/۱۱	۱/۸۶	۲/۳۲
۷	مکینک	۰	۰/۸۹	۱/۶۶	۰/۸۰	۸۰/۵۳	۲/۵	۳/۲۵

رگرسیون چندمتغیره

(۳) نشان داده شده است. در این جدول، مقدار ضریب همبستگی چندگانه (R)، ضریب تعیین (R²) و ضریب تعیین تصحیح‌شده (R² Adjusted)، تجزیه واریانس و ضرایب رگرسیون نشان داده شده است. در این مطالعه مقدار R² هم برابر $۰/۹۳۹$ است که نشان‌دهنده کارایی و دقت مناسب مدل در پیش‌بینی تبخیر و تعرق مرجع است.

برای بررسی تأثیر پارامترهای ورودی مدل شامل درجه حرارت حداکثر، درجه حرارت حداقل، رطوبت نسبی حداکثر، رطوبت حداقل، ساعات آفتابی و سرعت باد در برآورد تبخیر و تعرق مرجع و تعیین الگوی ورودی مدل GEP، تحلیل رگرسیونی چندمتغیره بین این پارامترها صورت گرفت که نتایج در جدول

جدول (۳): مقدار ضرایب دقت مدل و پارامترهای رگرسیون چندمتغیره

Table (3): Model accuracy coefficients and multivariate regression parameters

خطای معیار شبیه‌سازی	ضریب تعیین تصحیح شده	ضریب تعیین	ضریب همبستگی چندگانه
۰/۷۷۲۳	۰/۹۳۹	۰/۹۳۹	۰/۹۶۹
سطح معنی داری	F	میانگین مربعات	درجه آزادی
۰	۲۰۷۶۵/۷۱۱	۱۲۳۸۵/۴۳۸	۶
		مجموع مربعات	مجموع مربعات
		۷۴۳۱۲/۶۳	۷۴۳۱۲/۶۳
		۰/۵۹۶	۴۸۱۶/۲۲۹
		۸۰۸۱	۷۹۱۲۸/۸۵۹
سطح معنی داری	T	ضرایب استاندارد نشده	ضرایب استاندارد نشده
		Beta	B
		خطای معیار	خطای معیار
۰	-۶/۹۴۶	۰/۰۸۷	-۰/۶۰۶
۰	۱۴/۱۷۹	۰/۰۰۴	۰/۰۵۱
۰	۴۸/۰۵۹	۰/۰۰۴	۰/۱۷۱
۰	۸/۴۶۲	۰/۰۰۱	۰/۰۰۹
۰	-۳۰/۳۷۹	۰/۰۰۱	-۰/۰۲۲
۰	۸۰/۳۲۶	۰/۰۰۵	۰/۴۲۷
۰	۴۰/۷۲۸	۰/۱۶۹	۰/۱۶

بیشترین تأثیر را بر تبخیر و تعرق مرجع داشته است که با نتایج پژوهش حاضر مطابقت دارد. مقدار t (چه مثبت و چه منفی) اگر بزرگ‌تر از $1/96$ باشد و مقدار سطح معنی داری اگر کوچک‌تر از $0/01$ باشد، نشان می‌دهد متغیر پیش‌بین بر متغیر پاسخ تأثیر معنی دار دارد. با توجه به مقادیر مربوط به هر شش متغیر، همه متغیرها دارای تأثیر معنی دارند.

الگوی ورودی مدل‌ها برای پیش‌بینی تبخیر و تعرق مرجع در جدول (۴) نشان داده شده است. بر اساس این جدول، مدل‌سازی تبخیر و تعرق مرجع با شش الگو به صورت جدول زیر انجام شد. جدول (۵) نیز نشان‌دهنده رابطه رگرسیونی بین پارامترهای ورودی و تبخیر و تعرق مرجع بر اساس شش الگوی ورودی است. بر اساس شاخص‌های ارزیابی، الگوی شماره ۳ با $RMSE=1/3$ و $R^2=0/90$ به عنوان مناسب‌ترین الگو معرفی شد.

یکی از بخش‌های مهم جدول (۳) مقادیر بتا و سطح معنی داری است. با مشاهده سطح معنی داری حاصل شده می‌توان گفت همه پارامترها در سطح $0/1$ معنی دار بوده و نشان‌دهنده این موضوع است که پارامترهای ورودی همگی بر متغیر تبخیر و تعرق اثرگذار بوده و در پیش‌بینی مورد استفاده قرار می‌گیرند. مقایسه ضریب استاندارد شده به دست آمده برای پارامترهای ورودی نشان می‌دهد که بیشینه درجه حرارت با ضریب $0/580$ ، بیشترین تأثیر را بر تبخیر و تعرق مرجع دارد و پیش‌بینی‌کننده قوی‌تری نسبت به بقیه متغیرهاست. بعد از آن به ترتیب سرعت باد، ساعات آفتابی، درجه حرارت حداقل، رطوبت نسبی حداکثر و حداقل دارای بیشترین تأثیر در پیش‌بینی و برآورد تبخیر و تعرق هستند. حسینی و همکاران (۲۰۱۵) نتیجه گرفتند که بیشینه درجه حرارت و سرعت باد

جدول (۴): الگوهای ورودی مدل‌های رگرسیون چندمتغیره و برنامه‌ریزی بیان ژن

Table (4): Input Patterns of Multivariate Regression Models and Gene Expression Programming

الگوهای ورودی مدل	
1	$ET_0=f\{T_{max}\}$
2	$ET_0=f\{T_{max}, WS\}$
3	$ET_0=f\{T_{max}, WS, SUN\}$
4	$ET_0=f\{T_{max}, WS, SUN, T_{min}\}$
5	$ET_0=f\{T_{max}, WS, SUN, T_{min}, Rh_{min}\}$
6	$ET_0=f\{T_{max}, WS, SUN, T_{min}, Rh_{min}, Rh_{max}\}$

جدول (۵): روابط رگرسیونی بین تبخیر و تعرق مرجع و پارامترهای هواشناسی

Table (5): Regression relationships between reference evapotranspiration and meteorological parameters

شاخص ارزیابی	رابطه
$R^2 = 0.86$ و $RMSE = 2.94$	$ET_0 = -0.606 + 0.171T_{max}$
$R^2 = 0.88$ و $RMSE = 1.54$	$ET_0 = -0.606 + 0.171T_{max} + 0.427WS$
$R^2 = 0.90$ و $RMSE = 1.3$	$ET_0 = -0.606 + 0.171T_{max} + 0.427WS + 0.16Sun$
$R^2 = 0.91$ و $RMSE = 1.33$	$ET_0 = -0.606 + 0.171T_{max} + 0.427WS + 0.16Sun + 0.051T_{min}$
$R^2 = 0.91$ و $RMSE = 1.55$	$ET_0 = -0.606 + 0.171T_{max} + 0.427WS + 0.16Sun + 0.051T_{min} + 0.009 Rh_{min}$
$R^2 = 0.82$ و $RMSE = 1.226$	$ET_0 = -0.606 + 0.171T_{max} + 0.427WS + 0.16Sun + 0.051T_{min} + 0.009 Rh_{min} - 0.022 Rh_{max}$

مدل‌سازی تبخیر و تعرق مرجع به وسیله GEP

بر اساس تحلیل همبستگی بین پارامترهای ورودی و خروجی، شش الگو برای مدل‌سازی تبخیر و تعرق مرجع به صورت جدول (۴) پیشنهاد شد. همچنین برای اجرای مدل، از ۷۰٪ داده‌ها برای آموزش و ۳۰٪ داده‌ها برای صحت‌سنجی و آزمون انتخاب شد. در برآورد تبخیر و تعرق مرجع، علاوه بر چهار عملگر اصلی (F1)، حالت‌هایی بر اساس عملگرهای ریاضی پیش‌فرض برنامه (F2) نیز در نظر گرفته شد (روابط ۹ و ۱۰). برای ساخت درخت تجزیه و تولید جمعیت اولیه، از راه‌حلی که از ترکیب تصادفی مجموعه توابع (عملگرهای ریاضی مورد استفاده در روابط) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) ایجاد می‌شوند، استفاده شد.

$$F1 = \{+, /, -, * \} \quad (9)$$

$$F2 = \{+, /, -, *, \sqrt{\quad}, \sqrt[3]{\quad}, \text{Exp}, \text{Ln}, \text{Sin}, \text{Cos}, \text{Atan} \} \quad (10)$$

در اجرای مدل GEP، ابتدا باید تعداد ژن‌ها و کروموزوم‌ها مشخص شود. با توجه به مطالعات انجام‌شده قبلی، بهترین جمعیت از ۳۰ تا ۵۰ کروموزوم و پاسخ مناسب از انتخاب ۲ یا ۳ ژن حاصل می‌شود (فریرا، ۲۰۰۱)؛ بنابراین برای جمعیت اولیه در این پژوهش از ۳۰ کروموزوم و از ۳ ژن در ترکیب کروموزوم استفاده شد. همچنین تابع پیوند (+) به علت برتری بر دیگر توابع انتخاب شد. حداکثر تکرار تابع برآزش هم ۱۰۰۰ تعیین شد. مقادیر نرخ عملگرهای ژنتیکی ترانش درج متوالی و ترانش ریشه درج متوالی برابر ۰/۱، نرخ جهش و نرخ وارون‌سازی همان مقادیر پیش‌فرض مدل و مقدار مجموع نرخ سه عملگر ترکیب تک‌نقطه‌ای، ترکیب دونقطه‌ای و ترکیب ژن، طبق نظر فریرا (۲۰۰۱) برابر ۰/۷ تعیین شد. در جدول (۶) پارامترهای استفاده‌شده در مدل به همراه مقادیر آن‌ها ارائه شده است.

نتایج حاصل از اجرای مدل GEP برای داده‌های آموزش و آزمون در شش الگوی ورودی به همراه دو نوع عملگر ریاضی، در جدول (۷) ارائه شده است. همان‌طور که مشاهده می‌شود، برنامه‌ریزی بیان ژن توانسته است در تمام سناریوهای عملکرد و دقت مناسبی از خود ارائه دهد. در بین سناریوهای انجام‌شده، ساختار ترکیبی شماره ۲ با عملگر F2 در مرحله آزمون و ساختار ترکیبی شماره ۳ با عملگر F2 در مرحله آموزش در ایستگاه خرم‌آباد، دارای کمترین خطا و بیشترین دقت بوده و توانسته‌اند میزان تبخیر و تعرق مرجع را با دقت مناسبی برآورد کنند. الگوی شماره ۲ به دلیل کمتر بودن اطلاعات ورودی انتخاب شد.

شکل (۳-الف) مقادیر مشاهداتی و محاسباتی تبخیر و تعرق مرحله آزمون مدل را در طول دوره آماری نشان می‌دهد. همان‌طور که مشاهده می‌شود، مدل توانسته است روند تغییرات ET_0 را به‌طور مناسبی پیش‌بینی کند. در شکل (۳-ب) مقادیر مشاهداتی در مقابل مقادیر محاسباتی آموزش مدل نسبت به نیمساز ناحیه اول ترسیم شده‌اند. بر اساس نمودار پراکنش (قسمت ب) مشاهده می‌شود که عملکرد مدل مناسب بوده به‌طوری که در اکثر موارد، مقادیر محاسباتی و پیش‌بینی‌شده بر روی خط نیمساز ($X=Y$) یا در نزدیکی آن واقع شده‌اند که نشان‌دهنده نزدیک بودن مقادیر مشاهداتی و محاسباتی است. علاوه بر آن مشاهده می‌شود که مدل در برآورد مقادیر حدی حداقل عملکرد مناسبی از خود نشان داده به‌طوری که مقادیر تخمینی نزدیک به مقادیر واقعی هستند. ولی در برآورد مقادیر حدی حداکثر از دقت آن کاسته شده است. در مجموع استفاده از عملگرهایی مانند Sin، Cos، Ln و... برای استخراج رابطه باعث می‌شود که در صورت وجود

یک رابطه غیرخطی بین ورودی و خروجی آن، دقت مدل مدل هم شده است. کاهش پیدا نکند که در عملگر F2 حتی باعث بهبود عملکرد

جدول (۶): مقادیر پارامترهای مورد استفاده در مدل GEP

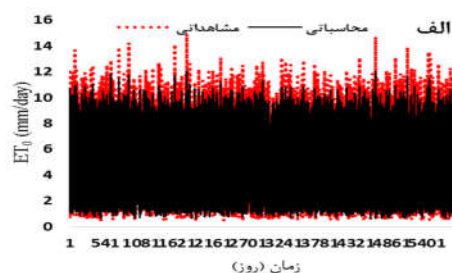
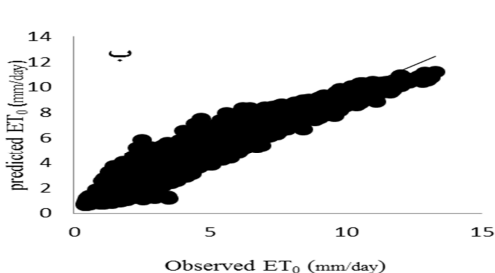
Table (6): The values of the parameters used in the GEP model

مقدار	عملگر ژنتیکی	مقدار	پارامتر
۰/۰۰۱۳۸	نرخ جهش ^۱	۳۰	تعداد کروموزومها
۰/۰۰۵۴۶	نرخ وارون سازی ^۲	۳	تعداد ژن در هر کروموزوم
۰/۱	نرخ ترانزیشن درج متوالی ^۳	(+)	تابع پیوند
۰/۱	نرخ ترانزیشن ریشه درج متوالی ^۴	RMSE	معیار خطای تابع برازش
۰/۲	نرخ ترکیب تک نقطه ای ^۵	۱۰۰۰	حداکثر تکرار تابع برازش
۰/۲	نرخ ترکیب دو نقطه ای ^۶		
۰/۳	نرخ ترکیب ژن ^۷		

جدول (۷): روش برنامه ریزی بیان ژن با استفاده از دو مجموعه عملگر ریاضی

Table (7): Gene expression programming method using two sets of mathematical operators

آزمون		آموزش		عملگر	الگو	
R ²	RMSE(mm/day)	R ²	RMSE(mm/day)			
۰/۸۶۴	۱/۱۶۳	۰/۸۵۸	۱/۱۹۱	F1	M1	۱
۰/۸۸۷	۱/۰۵۳	۰/۸۸	۱/۰۹۶	F2		
۰/۹۰۶	۰/۹۴۸	۰/۹۰	۱	F1	M2	۲
۰/۹۴۱	۰/۷۶	۰/۹۳۲	۰/۸۴۲	F2		
۰/۸۵۱	۱/۵۹۷	۰/۸۶۳	۱/۵۴۵	F1	M3	۳
۰/۹۲۶	۰/۸۶۱	۰/۹۴	۰/۷۶۶	F2		
۰/۹۱۵	۰/۹۰۳	۰/۹۱۳	۰/۹۳۳	F1	M4	۴
۰/۹۳	۰/۸۲۱	۰/۹۲۸	۰/۸۴۷	F2		
۰/۹۳۲	۰/۸۴۱	۰/۹۳۲	۰/۸۵۴	F1	M5	۵
۰/۹۳۳	۰/۷۹۷	۰/۹۲۸	۰/۸۵	F2		
۰/۹۳۴	۰/۸۰۳	۰/۹۳۱	۰/۸۴۴	F1	M6	۶
۰/۹۱۳	۰/۹۳۷	۰/۹۲۲	۰/۸۷۱	F2		

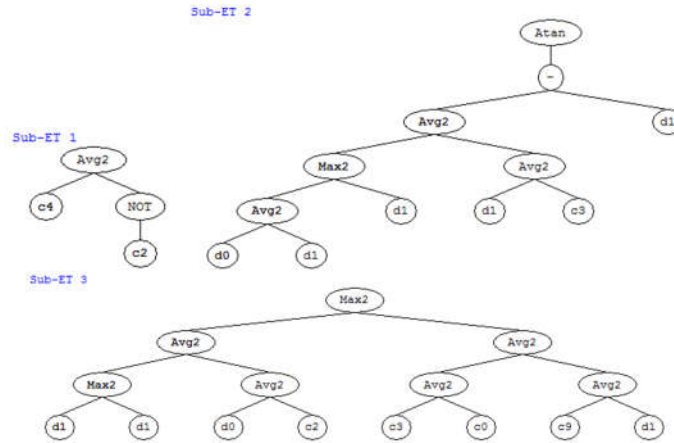


شکل (۳): الف: نمودار محاسباتی و پیش بینی شده نسبت به زمان؛ ب: نمودار پراکنش محاسباتی و پیش بینی شده GEP در مرحله آزمون

Figure (3): (a) Diagram of calculated and predicted values over time; (B) Diagram of the distribution between the calculated and predicted GEP values in the test phase

1. Mutation
2. Inversion
3. IS Transposition
4. RIS Transposition
5. One-Point Recombination
6. Two-Points Recombination
7. Gene Recombination

آن، $d0=Tmax$ و $d1=WS$ هستند. جدول (۸) نشان‌دهنده مقادیر ثابت در شکل (۵) است. در این جدول، $G1$ ، $G2$ و $G3$ به ترتیب بیانگر زیر درخت یا ژن‌های ۱، ۲ و ۳ است و C مقدار ثابت عددی است.



شکل (۵): بیان درختی ژن‌های ۱، ۲ و ۳ در پیش‌بینی تبخیر و تعرق مرجع ایستگاه الیگودرز

Figure (5): Tree Expression of 1,2 and 3 Genes in Predicting Reference Evapotranspiration of Aligoodarz Station

جدول (۸): مقادیر ثابت در بیان درختی مدل پیش‌بینی تبخیر و تعرق مرجع ایستگاه الیگودرز

Table (8): Constant values in the tree expression of the reference evapotranspiration model of the Aligoodarz station reference

Parameters	G3C9	G3C0	G3C3	G3C2	G2C3	G1C2	G1C4
Value	8.42	5.101	-2.339	0.188	6.727	1.013	-6.295

مدل‌های تجربی بیشتر است که با نتایج پژوهش حاضر مطابقت دارد. بعد از آن مدل‌های تجربی کیمبرلی-پنمن و تشعشعی فائو به ترتیب با $RMSE=1/09$ و $R^2=0/95$ و $RMSE=1/1$ و $R^2=0/95$ و مدل رگرسیون چندمتغیره با $RMSE=1/3$ و $R^2=0/90$ دارای بیشترین دقت در برآورد تبخیر و تعرق مرجع بوده‌اند. شایان ذکر است که مدل برنامه‌ریزی بیان ژن در همه الگوهای ورودی، دقت قابل قبولی داشته که در بسیاری از الگوها دقت مدل، بیشتر از سایر روش‌های تجربی و رگرسیون چندمتغیره بوده است.

نتیجه‌گیری

از بین روش‌های تجربی، مدل‌های کیمبرلی-پنمن و تشعشعی فائو نسبت به سایر روش‌های تجربی دارای دقت بیشتری بوده است. بر پایه نتایج به دست آمده می‌توان بیان کرد که روش برنامه‌ریزی بیان ژن دارای توانایی قابل قبولی در تخمین تبخیر و تعرق مرجع تحت شرایط آب‌وهوایی الیگودرز بوده و می‌توان با اطمینان بالا از آن در برآورد تبخیر

مدل‌های درختی مربوط به GEP در پیش‌بینی تبخیر و تعرق مرجع، شامل سه زیر درخت (ژن) است که در شکل (۴) نشان داده شده است. در این شکل، SUB-ET1، SUB-ET2 و SUB-ET3 به ترتیب ژن‌های اول تا سوم و پارامترهای

فرم ریاضی مدل استخراج شده با توجه به تابع پیوند جمع بین ژن‌ها، برای پیش‌بینی تبخیر و تعرق مرجع در ایستگاه مطالعاتی، با قرار دادن مقادیر ثابت و پارامترهایی که در ساختار بیان درختی مدل GEP به کار رفته است، تعیین شد. رابطه (۱۱) نشان‌دهنده رابطه استخراج شده از GEP برای پیش‌بینی تبخیر و تعرق مرجع است:

$$ET0 = \left(\frac{((G1C4 + (1.0 - G1C2)) / 2.0) + \text{atan}(\frac{((\text{gepMax2}(((Tmax + WS) / 2.0, WS) + ((WS + G2C3) / 2.0)) / 2.0) - WS))}{((\text{gepMax2}(((\text{gepMax2}(WS, WS) + ((Tmax + G3C2) / 2.0)) / 2.0), ((G3C3 + G3C0) / 2.0) + ((G3C9 + WS) / 2.0)) / 2.0))}}{2.0)} \right) \quad (11)$$

ارزیابی عملکرد مدل‌های تجربی، رگرسیون و برنامه‌ریزی

بیان ژن

از بین روش‌های تجربی، رگرسیون چندمتغیره و برنامه‌ریزی بیان ژن، مدل برنامه‌ریزی بیان ژن با مقدار $RMSE=0/843$ و $R^2=0/932$ در مرحله آموزش و $RMSE=0/76$ و $R^2=0/941$ در مرحله آزمون دارای دقت بیشتری نسبت به سایر مدل‌های برآورد تبخیر و تعرق مرجع بود. شیری (۲۰۱۷) در پژوهش خود به این نتیجه رسید که دقت برنامه‌ریزی بیان ژن نسبت به

خروجی را معین کرد. از بین دوازده الگوی برنامه‌ریزی بیان ژن، الگوهای شماره ۲ و شماره ۳ با عملگرهای پیش‌فرض مدل، دارای بهترین عملکرد در برآورد تبخیر و تعرق مرجع بودند. مقایسه عملکرد مدل‌های برآورد تبخیر و تعرق مرجع نشان‌دهنده برتری مدل برنامه‌ریزی بیان ژن نسبت به سایر مدل‌ها بود.

و تعرق مرجع بهره برد و به‌عنوان مدل قابل قبول در این زمینه معرفی کرد. همچنین ضرایب تبیین به‌دست‌آمده برای سناریوهای مختلف نشان‌دهنده انتخاب مناسب پارامترهای ورودی و تأثیرگذاری آن‌ها در مدل‌سازی است. همچنین مدل‌های مبتنی بر برنامه‌ریزی بیان ژن راه‌حل‌های صریحی را ارائه می‌کنند که می‌توان رابطه بین متغیرهای ورودی و

منابع

- Allen, R.G., Pereira, L.S., Raes, D. and Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome*, 300(9), p.D05109.
- Almorox, J., Quej, V.H. and Martí, P., 2015. Global performance ranking of temperature-based approaches for evapotranspiration estimation considering Köppen climate classes. *Journal of Hydrology*, 528, pp.514-522.
- Berti, A., Tardivo, G., Chiaudani, A., Rech, F. and Borin, M., 2014. Assessing reference evapotranspiration by the Hargreaves method in north-eastern Italy. *Agricultural Water Management*, 140, pp.20-25.
- 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; Soil Conservation Service, US Department of Agriculture: Washington, DC, USA.
- Djaman KB, Balde A, Sow A, Muller B, Irmak SK, N'Diaye M, Manneh BD, Moukoumbi Y, Futakuchi K, Saito K. Evaluation of sixteen reference Evapotranspiration methods under sahelian conditions in the Senegal River Valley. *Journal of Hydrology Regional Study*, 2015; 3: 139-159.
- Fallah, F. and Tahmasebipour, N. 2016. Meteorological drought analysis of Aligoudarz synoptic station using DIP software. 11th National Seminar on Watershed Management In date 2016-04-19 by Yasouj University, Watershed Management Society of Iran
- Ferreira, C., 2001. "Gene Expression Programming: A New Adaptive Algorithm for Solving Problems" *Complex Systems*, 13 (2), 87-129.
- Feyzollahpour, F., Delavar, M. and Hesamifshar, M. 2017. Evaluation and Analysis of Uncertainty Estimation of Reference Evapotranspiration Using Reference Genetics. *Journal of Soil and Water Science*, Volume 27, Number 4, 147-135. [Persian]
- Gholami, V., Derakhshan, Sh. and Darwari, Z. 2012. Investigation of Multivariate Regression and Artificial Neural Network in Simulation of Groundwater Salinity in Mazandaran Province. *Journal of Water Research in Agriculture*, Volume 26, Number 3, pages 79-100. [Persian]
- Grismer, M.E., Orang, M., Snyder, R. and Matyac, R., 2002. Pan evaporation to reference evapotranspiration conversion methods. *Journal of irrigation and drainage engineering*, 128(3), pp.180-184.
- Hargreaves, G. H., 1994. Defining and using reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, 120(6), 1132-1139..
- Hojbar, H., Moazed, H. and Shokrikoochak, S. 2017. Simulation of Reference Evapotranspiration using empirical methods and artificial neural network and compare with leisisimetre data in Urmia Kherizi station. *Journal of Irrigation and Water Engineering*, 4(15), 13-25. [Persian]
- Hosseini, S.M.R., Ganjikhoramdel, N. and Khalatabadifarahani, A.H. 2015. Evaluation and sensitivity analysis of different methods of estimating daily reference evapotranspiration in a cold climate. *Journal of Applied Research in Water Sciences*, Vol. 1, No. 2, Winter 2015, pp. 29-40 . [Persian]
- Khoshhal, J., Zareabyane, H., Joshani, A.R. and Khazae, M. 2016. Evaluation of different potential evapotranspiration methods with pan-FAO model in East and Northeast of Iran. *Journal of Natural Geography*, 8(28), 1-16. [Persian]
- Kisi, O. and Alizamir, M., 2018. Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: Wavelet extreme learning machine vs wavelet neural networks. *Agricultural and forest meteorology*, 263, pp.41-48.
- Laaboudi, A., Mouhouche, B. and Draoui, B., 2012. Neural network approach to reference evapotranspiration modeling from limited climatic data in arid regions. *International journal of biometeorology*, 56(5), pp.831-841.
- Ladlani, I., Hauichi, L., Dhemili, L., Heddem, S. and Blouze, KH., Estimation of daily reference evapotranspiration in the north of Algeria using adaptive neuro-Fuzzy inference system (ANFIS) and multiple linear regression Models: a comparative study. *Arabian Journal for Science and Engineering*, 2014; 39: 5959-5969. DOI 10.1007/s13369-014-1151-2
- Liu, X., Xu, C., Zhong, X., Li, Y., Yuan, X. and Cao, J., 2017. Comparison of 16 models for reference crop evapotranspiration against weighing

- lysimeter measurement. *Agricultural water management*, 184, pp.145-155.
19. Lopes, H.S. and Weinert, W.R., 2004. EGIPSYS: an enhanced gene expression programming approach for symbolic regression problems. *International Journal of Applied Mathematics and Computer Science*, 14(3), pp.375-384.
 20. Makkink, G.F., 1957. Testing the Penman formula by means of lysimeters. *J. Instit. Water Eng.*, 11, 277-288.
 21. Martí, P., González-Altozano, P., López-Urrea, R., Mancha, L.A. and Shiri, J., 2015. Modeling reference evapotranspiration with calculated targets. Assessment and implications. *Agricultural Water Management*, 149, pp.81-90.
 22. Niaghi, A.R., Majnooni-Heris, A., Haghi, D.Z. and Mahtabi, G., 2013. Evaluate several potential evapotranspiration methods for regional use in Tabriz, Iran. *Journal of Applied Environmental and Biological Sciences*, 3(6), pp.31-41.
 23. Pour-Ali Baba, A., Shiri, J., Kisi, O., Fard, A.F., Kim, S. and Amini, R., 2013. Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). *Hydrology Research*, 44(1), pp.131-146.
 24. Raouf, M. and Azizimobser, J. 2017. Evaluation of eighteen Reference Evapotranspiration model in Ardabil climatic conditions. *Water and Soil Conservation Journal*, Volume 24, Number6. [Persian]
 25. Sarmadian, F., Mehrjerdi, R., Asgari, H. and Akbarzadeh, A. 2010. Comparison of Neuro-Fuzzy Neural Network and Multivariate Regression in Predicting Some Soil Properties, *Journal of Watershed Research*, Volume 41, Number 1, pages 211 to 220. [Persian]
 26. Sattari, M.H. and Esmailzadeh, B. 2016. Comparison of the results of M5 tree model and genetic programming with FAO-Penman-Monteith method for reference evapotranspiration reference. *Journal of Water Resources Engineering*, Ninth Year, Winter 2016. [Persian]
 27. Shiri, J., Nazemi, A.H., Sadraddini, A.A., Landaras, G., Kisi, O., Fard, A.F. and Marti, P., 2014. Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Computers and Electronics in Agriculture*, 108, pp.230-241.
 28. Shiri, J., 2017. Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. *Agricultural water management*, 188, pp.101-114.
 29. Tabari, H., Grismer, M.E. and Trajkovic, S. 2013. Comparative analysis of 31 reference Evapotranspiration methods under humid conditions. *Irrigation Science*, 31: 107-117.
 30. Tabari, H., Kisi, O., Ezani, A. and Talaei, P.H. 2012. SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *Journal of Hydrology*, 444, 78-89.
 31. Turc, L., 1961. Estimation of irrigation water requirements, potential evapotranspiration: A simple climatic formula evolved up to date. *Ann. Agron.*, 12, 13-49.
 32. Wen, X., Si, J., He, Z., Wu, J., Shao, H. and Yu, H. 2015. Support-vector-machine-based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions. *Water resources management*, 29(9), 3195-3209.
 33. Yassin, M.A., Alazba, A.A. and Mattar, MA. 2016. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agricultural Water Management*, 163, 110-124.

Assessing Experimental and Intelligent Models in Estimating Reference Evapotranspiration

Yaser Sabzevari¹, Morad Ali Ghanbarpouri^{2*}

Received: 10/8/2021

Accepted: 6/10/2022

Extended Abstract

Introduction: As the most important element in the hydrologic cycle which depends on climate variables such as near-ground wind speed, air temperature, solar radiation, and relative humidity, reference evapotranspiration (ET₀) is normally computed through a variety of methods, each of which requires different and in some cases extensive data that are unavailable in many circumstances, especially in arid and semi-arid regions like Iran, providing different results based on the type of meteorological assumptions and data considered.

1. Doctoral student of irrigation and drainage, Department of Water Engineering, Isfahan University of Technology

2. Researcher of Soil and Water Research Department, Lorestan Agriculture and Natural Resources Research and Education Center; moradghanbari@gmail.com

DOI: 10.22052/deej.2021.11.36.11

ET₀ values can be estimated and simulated via meteorological models using physical equations or the empirical relationship existing between meteorological variables. In this regard, the FAO model is commonly used as a reference method worldwide for estimating evapotranspiration, requiring large sets of input data, including maximum/minimum air temperature, maximum/minimum relative humidity, solar radiation, and wind speed. However, as all of such information may sometimes be impossible to access, it is necessary to select models with fewer input data and reasonable accuracy.

Therefore, this study set out to investigate the applicability of seven reference evapotranspiration estimation models with fewer input data, including Kimberley Penman, FAO radiation model, Hargreaves Samani, Makkink, Belany Kridle FAO 24, Turc, and Priestly Teylor. The optimal model for estimating reference evapotranspiration was then determined using Gene Expression Programming Model (GEP), whose results were compared with the experimental methods. On the other hand, multivariate regression was used to identify the model's input patterns and to examine the influence of climatic parameters on ET₀.

Materials and methods: Expressing the relationship between several predictor variables and the response variable in question, multivariate regression models are based on two basic assumptions that distinguish them from the simple regression model:

1) The number of predictor (independent) variables in the regression should be less than the number of observations.

2) There is a complete linear correlation between predictor and response variables. Accordingly, in cases where the two assumptions are violated, the regression equation cannot be estimated.

Gene expression planning is a combination of the GA and GP methods developed by Ferreira in 1999, where, similar to decomposition trees in genetic programming, linear and simple chromosomes of constant length are combined, taking the genetic algorithm and branch structures of different sizes and shapes into account. Therefore, the genotype and phenotype are separated in this method, and the system is unable to enjoy all the evolutionary benefits.

While the GEP's phenotype is similar to that of the branched structure of GP, the tree expression in GEP (i.e., its branched structure) represents all independent genomes. Thus, in GEP, refinement occurs in a linear structure and is expressed as a tree structure, merely resulting in a modified genome that is transferred to the next generation.

Considering the fact that GEP requires no heavy structures to reproduce and mutate, this study used the meteorological data collected from Aligodarz Synoptic Station for a 35-year period (1983–2017), including the maximum and minimum daily temperature, maximum and minimum relative humidity, wind speed, and sunlight hours, 70% of which were used for training and the remaining 30% for testing the model. Moreover, two types of mathematical operators were used in the GEP method, including four-element operations and default model operators.

Results and discussion: The study's results suggested that the Kimberley Penman and FAO Radiation models were more accurate than the other experimental models. Furthermore, multivariate regression results indicated acceptable modeling accuracy, with the R² being 0.95. Also, the analysis of the model's coefficients revealed that the maximum temperature had the highest influence on the prediction and estimation of reference evapotranspiration, with its correlation coefficient being 0.58, followed by wind speed, sunshine hours, minimum temperature, and, maximum and minimum relative humidity, respectively.

Taking the above-mentioned results into account, this study found six models for estimating reference evapotranspiration that required fewer input data. Accordingly, in Gene Expression Programming, model 2 comprising default operators with RMSE of 0.843 and R² of 0.932, and the one with RMSE of 0.76 and R² of 0.941 performed better at training and test phases, respectively.

Moreover, the comparison of the reference evapotranspiration estimation models indicated that the Gene Expression Programming model outperformed the other ones. The GEP model was also found to have acceptable applicability for estimating reference evapotranspiration in Aligodarz city under climatic conditions, being a readily applicable model in this field.

Keywords: Radiation Model, Temperature Model, Radiation-temperature Model, Regression Model, GEP.