



Ability of Gene Expression Programming Model to Estimate Reference Evapotranspiration with Minimal Meteorological Data

Mohammad Karmi^{*1}, Farzin Sayad Beiranvand²

1. Master's student, Department of Engineering and Water Resources Management, Faculty of Civil Engineering, Isfahan University of Technology, Isfahan, Iran
2. Former Master's student, Department of Water Engineering, Faculty of Agriculture, Isfahan University of Technology, Isfahan, Iran

Received: 17 October 2024/ **Revised:** 21 October 2024/ **Accepted:** 22 October 2024

<https://doi.org/10.22034/arwe.2024.2043340.1029>

Abstract

amount of reference evapotranspiration using meteorological data, the period of 30-year of Ahvaz meteorological station was used. The reference method for calculating evapotranspiration was the Penman-Monteith method. In this research, four scenarios with different combinations of model input parameters were examined. Finally, the results showed that the best combination of input parameters for the gene expression programming model included parameters of minimum temperature, maximum temperature, relative humidity and sunshine hours. For these input parameters, R2, MAE, and RMSE statistics in the training phase are 0.964, 0.421 mm/d, and 0.507 mm/d, respectively, and in the test phase, they are 0.965, 0.419 mm/d, and 0.506 mm/d, respectively. was obtained Rainfall was the only parameter that showed the least effect on the rate of evaporation and transpiration, so that the increase of this parameter to the input parameters of the model caused a slight change (less than one percent) in the values of the evaluation indices. It seems that the reason for this is the low amount of rainfall in this area, so that in most months, the amount of rainfall is very small. Finally, the results of this research showed that the gene expression programming model can be used as a suitable tool with good accuracy for estimating reference evaporation and transpiration.

Keywords: Artificial intelligence, Evaporation and transpiration, GEP model, Penman Mantis.



توانایی مدل برنامه‌ریزی بیان ژن در برآورد میزان تبخیر و تعرق مرجع با حداقل داده‌های هواشناسی

محمد کرمی^{۱*}، فرزین صیاد بیرانوند^۲

۱. دانشجوی کارشناسی ارشد، گروه مهندسی و مدیریت منابع آب، دانشکده مهندسی عمران، دانشگاه صنعتی اصفهان، اصفهان، ایران
۲. دانشجوی سابق کارشناسی ارشد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران

دریافت: ۲۲ مهر ماه ۱۴۰۳ / اصلاحات: ۳۰ مهر ماه ۱۴۰۳ / پذیرش: ۰۱ آبان ماه ۱۴۰۳

<https://doi.org/10.22034/arwe.2024.2043340.1029>

چکیده

هوش مصنوعی به عنوان یکی از ابزار معتبر در مدل‌سازی فرایندهای پیچیده غیرخطی، شناخته شده است. در این تحقیق، به منظور بررسی توانایی مدل برنامه‌ریزی بیان ژن (Gene Expression Programming (GEP)) در برآورد میزان تبخیر و تعرق با استفاده از داده‌های هواشناسی، از دوره آماری ۳۰ سال هواشناسی ایستگاه هواشناسی اهواز، استفاده گردید. روش مرجع برای محاسبه تبخیر و تعرق، روش پنمن مانیتیت در نظر گرفته شد. در این تحقیق چهار سناریو با ترکیب‌های مختلف پارامترهای ورودی مدل مورد بررسی قرار گرفت. در نهایت نتایج نشان داد، بهترین ترکیب پارامترهای ورودی برای مدل برنامه‌ریزی بیان ژن، شامل پارامترهای دمای حداقل، دمای حداکثر، رطوبت نسبی و ساعات آفتابی بود. به ازای این پارامترهای ورودی، مقادیر MAE ، $R2$ و $RMSE$ در مرحله آموزش به ترتیب برابر $0/964$ ، $0/421$ میلی متر بر روز و $0/507$ میلی متر بر روز و در مرحله آزمون (تست) به ترتیب برابر $0/965$ ، $0/419$ و $0/506$ به دست آمد. بارش تنها پارامتری بود که کمترین اثر را بر میزان تبخیر و تعرق نشان داد، به گونه‌ای که افزایش این پارامتر به پارامترهای ورودی مدل، باعث تغییر ناچیز (کمتر از یک درصد) در مقادیر شاخص‌های ارزیابی گردید. به نظر می‌رسد علت این امر به میزان بارندگی کم این منطقه باشد، به گونه‌ای که در اکثر ماهها، میزان بارندگی بسیار ناچیز است. در نهایت نتایج این تحقیق نشان داد مدل برنامه‌ریزی بیان ژن می‌تواند به عنوان یک ابزار مناسب با دقت خوب برای برآورد میزان تبخیر و تعرق مرجع قرار بگیرد.

کلمات کلیدی: پنمن مانیتیت، تبخیر و تعرق، مدل برنامه‌ریزی بیان ژن، هوش مصنوعی.

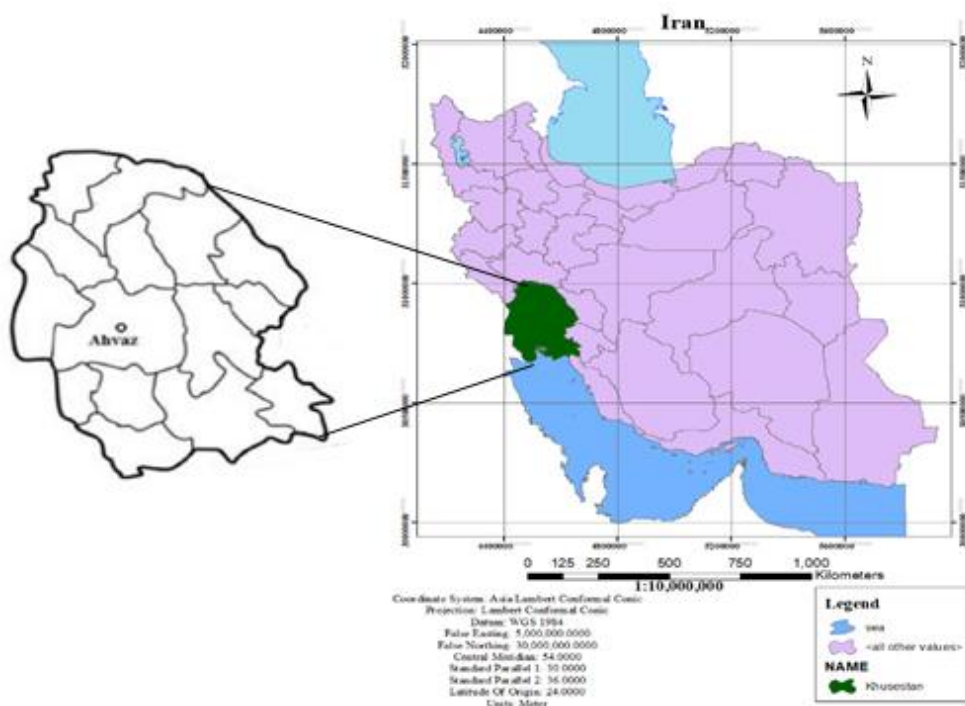
مقدمه

فرآیند تبخیر- تعرق توسط عوامل متعددی که در انتقال آب از خاک به گیاه و نهایتاً به اتمسفر و تبخیر از منطقه تحت کشت مؤثرند، کنترل می شود. مدل‌سازی چنین فرآیند پیچیده‌ای که بسیاری از پارامترهای مؤثر بر آن بر یکدیگر نیز اثر متقابل می گذارند، امری بسیار مشکل بوده که بدون فرضهای متعدد جهت ساده نمودن موضوع امکانپذیر نمی باشد (Monavar Sabegh et al., 2023). روش های متعددی برای محاسبه تبخیر-تعرق وجود دارد که در یک تقسیم بندی کلی، می توان آنها را به دو دسته روشهای مستقیم و غیر مستقیم تقسیم نمود. یکی از رایجترین روشهای غیر مستقیم، استفاده مقادیر برآورد شده تبخیر-تعرق گیاه مرجع با استفاده از معادله پنمن-مانتیت می باشد. در مناطق مختلف با شرایط آب و هوایی متفاوت، نتایج این مدل ها ممکن است با دقت های متفاوت باشد (Shiri et al., 2014). اخیراً کاربرد هوش مصنوعی و روش های مبتنی بر داده کاوی در حل مسائل مهندسی آب، کاربرد چشمگیری داشته است (Siasar and honar, 2019; Landeras et al., 2008). استفاده از مدل هایی مانند ماشین بردار پشتیبان، درخت تصمیم، جنگل تصادفی، برنامه ریزی بیان ژن و ... به عنوان روش های معتبر در مدل‌سازی فرایندهای پیچیده غیرخطی، ثابت شده است. با توجه به اینکه تبخیر و تعرق نیز یک فرآیند پیچیده و غیر خطی می باشد، مدل های مذکور می توانند به عنوان ابزار مناسبی جهت مدل سازی فرآیند تبخیر و تعرق مورد استفاده قرار بگیرد (Siasar and honar, 2019). اساساً روش های مبتنی بر هوش مصنوعی جزو روش های جعبه سیاه بوده و بدون در نظر گرفتن فرآیند فیزیکی حاکم بر سیستم و بر اساس داده های ورودی و خروجی قادر به ایجاد رابطه با دقت مشخص می گردند (Siasar and honar, 2019). محققین زیادی در نقاط مختلف دنیا از مدل های مختلف هوش مصنوعی برای شبیه‌سازی و پیش‌بینی فرایندهای مختلف پیچیده طبیعی استفاده کرده‌اند. شیری (Shiri, 2017) برای مناطق خشک و نیمه خشک به ارزیابی مدل‌های فائو-پنمن-مانتیت ۵۶، تجربی، نیمه تجربی و برنامه‌ریزی بیان ژن در شبیه‌سازی تبخیر تعرق مرجع روزانه پرداختند. نتایج نشان داد که مدل برنامه‌ریزی بیان ژن نسبت به روش‌های تجربی و نیمه تجربی عملکرد بهتری را نشان داده است. مهدی زاده و همکاران گزارش کردند که روش رگرسیون چند متغیره تطبیقی و ماشین‌بردار پشتیبان با هسته توابع پایه شعاعی به طور کلی بهتر از برنامه ریزی بیان ژن و ماشین بردار پشتیبان با توابع چندجمله‌ای مدل‌سازی تبخیر و تعرق را انجام می‌دهند. . مатар (Mattar, 2018) با استفاده از برنامه‌ریزی بیان ژن به مدل‌سازی تبخیر تعرق مرجع ماهانه در مصر پرداخت. نتایج نشان داد که برنامه‌ریزی بیان ژن با ورودی‌های بیشینه و کمینه درجه حرارت، رطوبت نسبی و سرعت باد دارای بیشترین دقت در برآورد ET_0 می باشد. ولیپور و همکاران (Valipour et al., 2019) در نشان دادند که مدل برنامه‌ریزی بیان ژن یک ابزاری قدرتمند در پیش‌بینی تبخیر-تعرق در اقلیم نیمه خشک و مدیترانه‌ای است اما در اقلیم‌های بسیار مرطوب و بعضی مناطق خشک باید با احتیاط مورد استفاده قرار گیرد. محققینی دیگری همچون (Shiri et al., 2021; Izadyar et al., 2023; Karimi et al., 2013; Landeras et al., 2008; Shiri et al., 2014) نیز به مدل سازی تبخیر و تعرق با استفاده از روش های مختلف هوش مصنوعی پرداخته اند. همان طور که نتایج محققین نشان داده است استفاده از روش های هوش مصنوعی برای محاسبه میزان تبخیر و تعرق کاربردی می باشد. لذا هدف از این تحقیق بررسی توانایی برنامه ریزی بیان ژن در برآورد میزان تبخیر و تعرق مرجع در منطقه اهواز می باشد.

مواد و روش‌ها

منطقه مورد مطالعه

استان خوزستان با مساحتی حدود ۶۴۲۳۶ کیلومتر مربع، وسیع ترین استان در نیمه غربی کشور می باشد. این استان بین ۴۷ درجه و ۴۱ دقیقه تا ۵۰ درجه و ۳۹ دقیقه طول شرقی از نصف النهار گرینویچ و ۲۹ درجه و ۵۸ دقیق تا ۳۳ درجه و ۴ دقیقه عرض شمالی از خط استوا قرار دارد. در این تحقیق داده های هواشناسی از ایستگاه هواشناسی اهواز با طول جغرافیایی ۴۸ درجه و ۴۰ دقیقه و عرض جغرافیایی ۳۱ درجه و ۲۰ دقیقه و ارتفاع از سطح دریا ۲۲/۵ متر اخذ گردید. میانگین وضعیت بارندگی و درجه حرارت در مقیاس بلند مدت در جدول (۱) ارائه شده است.



شکل ۱. منطقه مورد مطالعه

Fig.1. Case study

جدول ۱. میانگین بلند مدت دما و بارندگی در ماههای مختلف شهر اهواز

Table 1. Long-term average temperature and rainfall in different months of Ahvaz

عنوان	بارندگی (میلیمتر)	درجه حرارت (درجه سانتی گراد)
فروردین	۲۲/۷	۲۳
اردیبهشت	۶/۸	۲۹/۹
خرداد	۰/۱	۳۵/۱
تیر	۰/۰	۳۷/۴
مرداد	۰/۱	۳۷/۹
شهریور	۰/۲	۳۴/۹
مهر	۳/۳	۲۹/۷
آبان	۲۶	۲۲/۵
آذر	۵۹/۵	۱۵/۷
دی	۴۵/۴	۱۲/۸
بهمن	۲۶/۳	۱۳/۹
اسفند	۲۷/۷	۱۷/۷
میانگین	۱۸/۲	۲۵/۹

معرفی مدل و پارامترهای ورودی:

در این تحقیق، از داده‌های متداول هواشناسی شامل داده‌های دمای حداکثر، دمای حداقل، ساعات آفتابی، رطوبت نسبی و بارندگی در گام‌های زمانی روزانه استفاده شده است. در ابتدا میزان تبخیر و تعرق روزانه بر اساس روش پنمن مانیتث با استفاده از نرم‌افزار ET_0 Calculator محاسبه گردید. در این تحقیق دوره آماری از ۳۰ سال، از سال ۱۹۹۱ تا ۲۰۱۰ که در دسترس بود مورد استفاده قرار گرفت. روش پنمن مانیتث برای اکثر مناطق جهان در محدوده وسیعی برای محاسبه تبخیر و تعرق مرجع به کار برده شده است و از سوی

سازمان خوار و بار جهانی (FAO) به عنوان یک روش استاندارد برای محاسبه تبخیر و تعرق مرجع معرفی شده است. این روش همچنین به عنوان یک روش مرجع برای مقایسه سایر روش‌ها پیشنهاد شده است (Siasar and honar, 2019).

معرفی روش برنامه‌ریزی بیان ژن برای محاسبه تبخیر و تعرق

برنامه‌ریزی بیان ژن (Gene Expression Programming (GEP))، یک روش توسعه‌یافته بر اساس برنامه‌ریزی ژنتیک (GP) و الگوریتم ژنتیک (GA) است که در سال ۱۹۹۹ توسط Ferreira بر اساس تئوری Darwin ابداع شد. در GEP، از کروموزوم‌های خطی و ساده با طول ثابت استفاده می‌شود که مشابه الگوریتم ژنتیک هستند و ساختارهای شاخه‌ای با اندازه و شکل‌های متفاوت، آن را مشابه درختان تجزیه در برنامه‌ریزی ژنتیک می‌کند. در این روش، ژنوتیپ و فنوتیپ جداگانه مدیریت می‌شوند تا سیستم بتواند از تمامی مزایای الگوریتم‌های تکاملی بهره‌بردارد. برنامه‌ریزی بیان ژن برخلاف الگوریتم ژنتیک بر روی ساختار درختی رابطه‌ها عمل می‌کند که این ساختارهای درختی از مجموعه توابع (عملگرهای ریاضی مورد استفاده در رابطه‌ها) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) تشکیل شده‌اند. در این تحقیق برای اجرای مدل GEP از نرم‌افزار Genexprotools 5 استفاده شد. در این مدل علاوه بر چهار عملگر اصلی $\{+, -, \sqrt{\quad}, \ln(\quad)\}$ ، حالت‌هایی بر اساس عملگرهای ریاضی پیش‌فرض برنامه مانند $\sqrt{\quad}$ ، \ln و ... نیز در نظر گرفته شد (۱۲ و ۱۳). در 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 گردید. در ابتدای اجرای مدل GEP، باید تعداد ژن‌ها و کروموزوم‌ها مشخص شود. مقادیر مربوط به پارامترهای مدل که در این تحقیق مورد استفاده قرار گرفتند در جدول (۲) ارائه شده است. برای پارامترهایی مانند نرخ جهش، نرخ وارون‌سازی، نرخ جابجایی و نرخ جفت‌گیری از مقادیر پیش‌فرض نرم‌افزار Genexprotools 5 استفاده گردیده است.

جدول ۲. بهترین مقادیر مربوط به پارامترهای مدل

Table 2. Best values of model parameters

Values	پارامتر
30	تعداد کروموزوم (Number of chromosomes)
3	تعداد ژن (Number of genes)
Ta, Tw, es, ea, VPD	متغیرهای مستقل (Independent variables)
Tc	خروجی وابسته (Dependent output)
+, -, *, /, $\sqrt{\quad}$, $\ln(\quad)$, X2	عملگرهای ریاضی (Mathematical operators)
RMSE	تابع ارزیابی (Fitness function)
Addition	تابع پیوند (Linking functions)
0.00138	نرخ جهش (Mutation rate)
0.00546	نرخ وارون‌سازی (Inversion rate)
0.00277	نرخ جابجایی ژن (Gene transposition rate)
0.00277	نرخ جفت‌گیری ژن (Gene recombination rate)

مدلسازی با استفاده از برنامه‌ریزی بیان ژن

در این تحقیق اطلاعات مورد استفاده به دو بخش داده‌های آموزش و داده‌های آزمون تقسیم شدند. همان‌طور که از نامش مشخص است، داده‌های آموزش برای آموزش مدل مورد استفاده قرار می‌گیرند. داده‌های آزمون (تست) پس از اتمام فرآیند آموزش و برای ارزیابی عملکرد مدل‌ها در تعمیم نتایج، مورد استفاده قرار می‌گیرند. در این تحقیق ۷۰٪ داده‌های مورد استفاده به بخش آموزش و ۳۰٪ باقیمانده برای بخش آزمون تعلق گرفت. جدول ۲، نشان‌دهنده پارامترهای آماری داده‌های مورد استفاده می‌باشد. ترکیب‌های ورودی مدل‌سازی که در این تحقیق مورد بررسی قرار گرفته‌اند (سناریو) در جدول ۳ ارائه شده‌اند. به عبارت دیگر این ترکیب‌ها در هر سناریو به عنوان پارامترهای ورودی مدل برنامه‌ریزی بیان ژن، در نظر گرفته شد و در نهایت خروجی مدل که میزان تبخیر و تعرق مرجع می‌باشد، استخراج گردید.

جدول ۳. آمار توصیفی داده‌های مورد استفاده
Table 3. Descriptive statistics of the data used

متغیر	حدافل	حداکثر	میانگین	میان	انحراف معیار	ضرب همبستگی پارامترهای ورودی مدل با ET ₀
T _{max}	6.6	51.6	33.45	34.8	10.92	0.96
T _{min}	-1.2	38	19.32	20	8.24	0.90
RH	8	98	42.86	39.5	19.07	-0.88
Sun	0	13.4	8.63	9.8	3.51	0.67
Rain	0	107.4	0.57	0	3.46	-0.218
ET ₀	0.6	9.9	5.57	5.7	2.69	1.00
T _{max}	7.8	52.2	33.20	34	10.81	0.96
T _{min}	-0.6	35.4	19.11	19.6	8.24	0.90
RH	11.5	95	43.13	41	18.72	-0.88
Sun	0	13.6	8.61	9.7	3.47	0.65
Rain	0	53.7	0.52	0	3.03	-0.22
ET ₀	0.7	9.8	5.53	5.5	2.68	1

جدول ۴. سناریوهای مختلف ورودی مدل برنامه‌ریزی بیان ژن
Table 4. Different input scenarios of the gene expression programming model

ردیف	کد الگو	الگوی ورودی مدل
1	S1	ET ₀ =f{T _{max} + T _{min} }
2	S2	ET ₀ =f{ T _{max} + T _{min} +RH}
3	S3	ET ₀ =f{ { T _{max} + T _{min} +RH +Sun}
4	S4	ET ₀ =f{ { { T _{max} + T _{min} +RH +Sun =Rain}

معیارهای ارزیابی

به منظور ارزیابی توانایی مدل برنامه‌ریزی ژن، در این تحقیق از سه فاکتور ضریب تعیین (R²)، جذر میانگین مربعات خطا (RMSE) و میانگین قدر مطلق خطا (MAE) استفاده گردید. بهترین مقدار ضریب تعیین عدد یک و بهترین مقادیر RMSE و MAE برابر صفر می‌باشد.

$$R^2 = \frac{(\sum_{i=1}^n (ET_{0,i} - \overline{ET_0})(ET_{s,i} - \overline{ET_s}))^2}{\sum_{i=1}^n (ET_{0,i} - \overline{ET_0})^2 \cdot \sum_{i=1}^n (ET_{s,i} - \overline{ET_s})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (ET_{0,i} - ET_{s,i})^2}{n}} \quad (2)$$

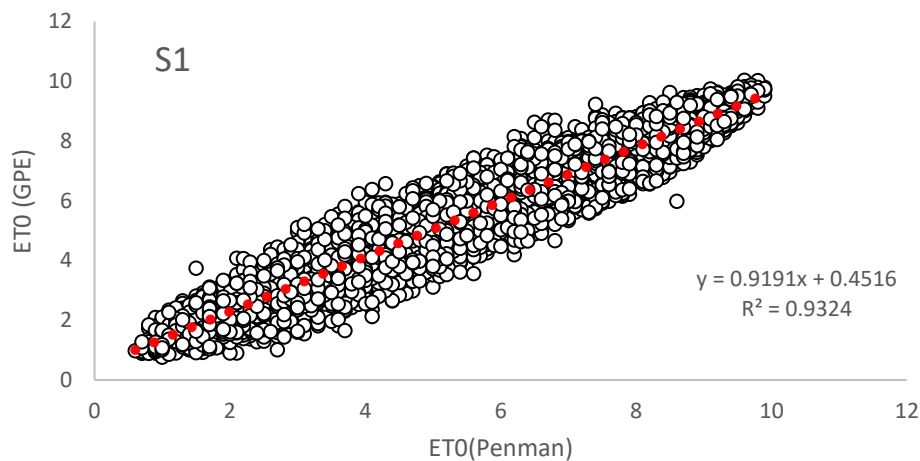
$$MAE = \frac{1}{N} \sum_{i=1}^n |ET_{0,i} - ET_{s,i}| \quad (3)$$

در روابط بالا، ET_{0,i} و ET_{s,i} به ترتیب مقادیر شبیه‌سازی و مشاهداتی در گام زمانی i ام، n تعداد گام‌های زمانی، $\overline{ET_0}$ و $\overline{ET_s}$ نیز به ترتیب میانگین مقادیر مشاهداتی و شبیه‌سازی می‌باشد. شاخص ضریب تبیین نشانگر دقت تخمین است که مقداری همواره بین ۰-۱ و +۱ دارد که علامت + و - بیانگر همبستگی مثبت یا منفی است و هرچه به ۱ نزدیک باشد نشان از همبستگی بیشتر بین مقادیر مشاهداتی و شبیه‌سازی دارد. آماره ریشه میانگین مربعات (RMSE) همواره مثبت بوده و بهترین حالت عملکرد زمانی است که مقدار به صفر نزدیک شود. مقدار میانگین قدرمطلق خطا بیان‌گر آن است که مدل مورد بررسی -تبخیر- تعرق مرجع را بیشتر با کمتر برآورد کرده است و هرچه مقدار آن به صفر نزدیک باشد حاکی از دقت بالاتر است.

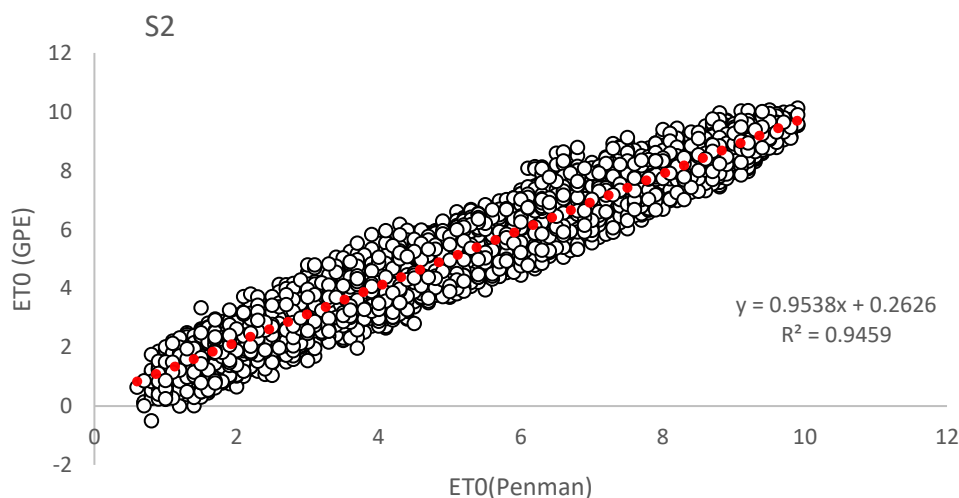
نتایج و بحث

نتایج حاصل از اجرای مدل GEP برای داده‌های آموزش به شرح زیر به دست آمد. نمودارهای یک به یک مربوط به مرحله آموزش در شکل‌های (۲ تا ۵) و نتایج ارزیابی در جدول (۳) ارائه شده‌اند. همان‌طور که نمودارها نشان می‌دهند، برنامه‌ریزی بیان ژن در همه سناریوها، به خوبی توانسته است برای شبیه‌سازی میزان تبخیر و تعرق آموزش ببینند. به عبارت دیگر با حداقل داده، مثل دمای حداکثر و دمای حداقل، آموزش با R²، MAE و RMSE برابر ۰/۹۳۲، ۰/۵۷۳ و ۰/۷ میلی متر بر روز، دقت نسبتاً خوبی انجام شده است. در

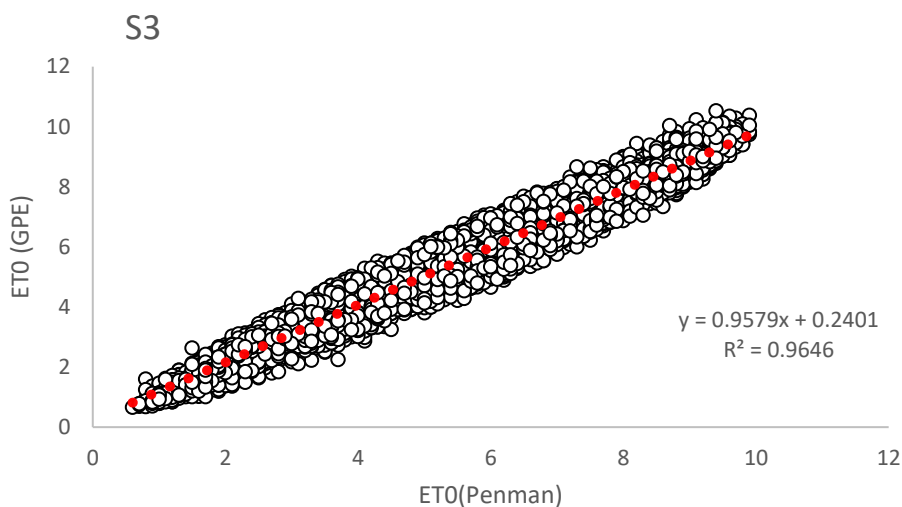
سناریوی دوم، پارامتر رطوبت نسبی، به پارامترهای ورودی اضافه گردید. با اضافه شدن رطوبت نسبی، آماره R^2 نسبت به سناریوی اول ۱/۵ درصد افزایش، آماره MAE، نسبت به سناریوی اول ۹/۳ درصد کاهش و آماره RMSE، نسبت به سناریوی اول ۱۰ درصد کاهش را نشان داد. در سناریوی سوم با افزایش ساعات آفتابی، آماره R^2 نسبت به سناریوهای اول و دوم ۳/۴ و ۱/۹ درصد افزایش، آماره MAE، نسبت به سناریوهای اول و دوم ۲۶/۵ و ۱۹/۰۳ درصد کاهش و آماره RMSE، نسبت به سناریوهای اول و دوم ۲۷/۵ و ۱۹/۵ درصد کاهش را نشان داد. در سناریوی آخر، پارامتر بارندگی به ترکیبات ورودی اضافه گردید. نتایج نشان داد در این سناریو، آماره R^2 نسبت به سناریوهای اول، دوم و سوم ۳/۳، ۴/۸ و ۶/۵ درصد کاهش، آماره MAE، نسبت به سناریوهای اول، دوم و سوم ۴۹/۴، ۶۴/۶ و ۱۰۳ درصد افزایش و آماره RMSE، نسبت به سناریوهای اول و دوم ۱۳/۱ و ۳/۵ درصد کاهش و نسبت به سناریوی سوم، حدود ۲۰ درصد افزایش را نشان داد. نتایج نشان می‌دهد بهترین آموزش زمانی حاصل گردید که از ترکیب دمای حداقل هوا، دمای حداکثر هوا، رطوبت نسبی و ساعات آفتابی به عنوان ورودی استفاده شود. به نظر می‌رسد بارش پارامتر موثری در مرحله آموزش نباشد. البته قابل ذکر است که میزان همبستگی این پارامتر با تبخیر و تعرق مرجع در این منطقه پایین می‌باشد. لذا در مرحله آموزش نیز اثر منفی خود را نشان داده است.



شکل ۲. نمودار یک به یک مربوط به سناریوی S1 در مرحله آموزش
 Fig. 2. Scatter plot related to the S1 scenario in the training phase

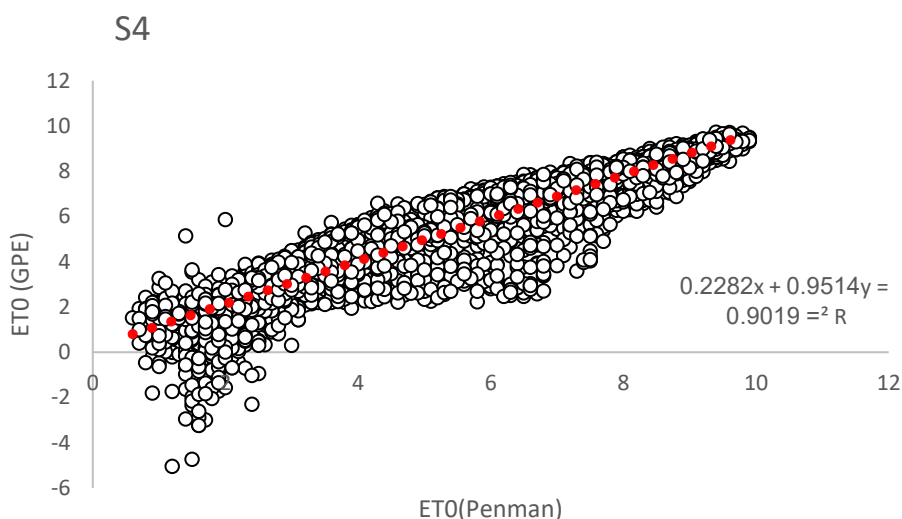


شکل ۳. نمودار یک به یک مربوط به سناریوی S2 در مرحله آموزش
 Fig. 3. Scatter plot related to the S2 scenario in the training phase



شکل ۴. نمودار یک به یک مربوط به سناریو S3 در مرحله آموزش

Fig 4. Scatter plot related to the S3 scenario in the training phase



شکل ۵. نمودار مربوط به سناریو S4 در مرحله آموزش

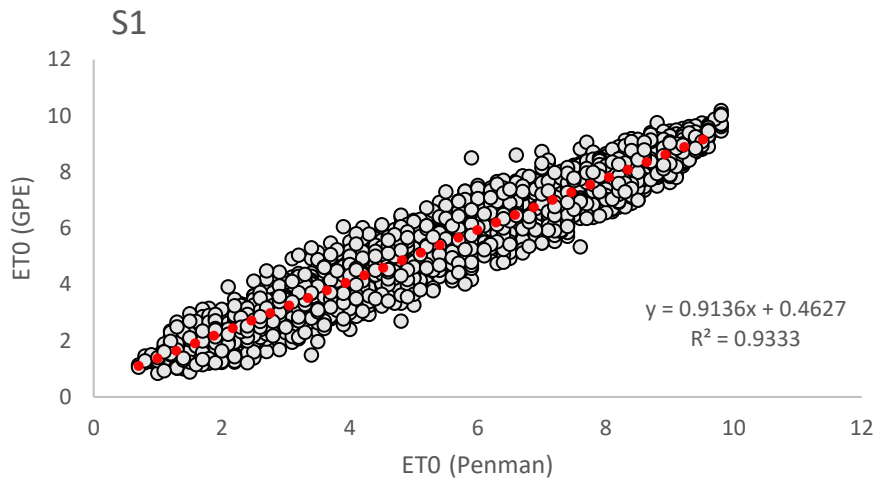
Fig 5. Scatter plot related to the S4 scenario in the training phase

جدول ۵. شاخص‌های ارزیابی برای مرحله آموزش

Table 5. Evaluation indices for the training stage

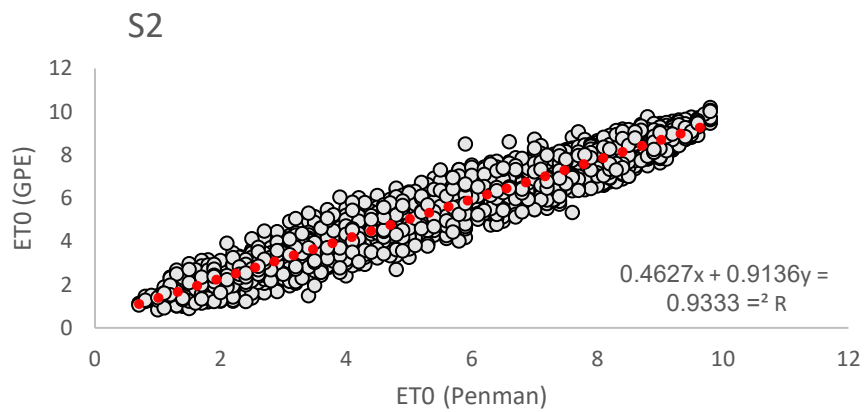
RMSE	MAE	R ²	پارامتر	سناریو
0.700	0.573	0.932	Tmax+Tmin	S1
0.63	0.52	0.946	Tmax+Tmin+RH	S2
0.507	0.421	0.964	Tmax+Tmin+RH+SUN	S3
0.608	0.856	0.901	Tmax+Tmin+RH+SUN+Rain	S4

نتایج مربوط به مرحله آزمون (Test) مدل به صورت نمودار یک به یک، در شکل‌های (۹ تا ۶) و نتایج ارزیابی با شاخص‌های R²، MAE و RMSE در جدول (۴) ارائه شده‌اند.



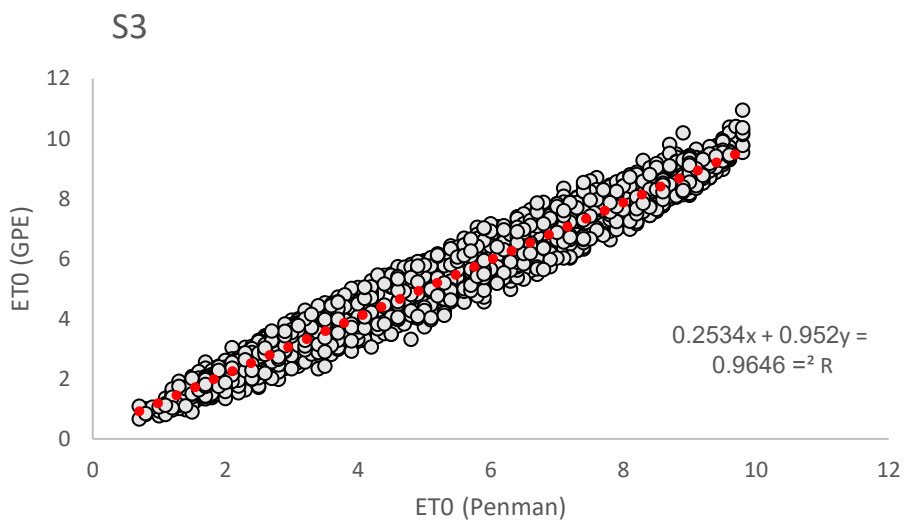
شکل ۶. نمودار یک به یک مربوط به سناریو اول در مرحله آزمایش (Test)

Fig. 6. Scatter plot related to the S1 scenario in the test phase



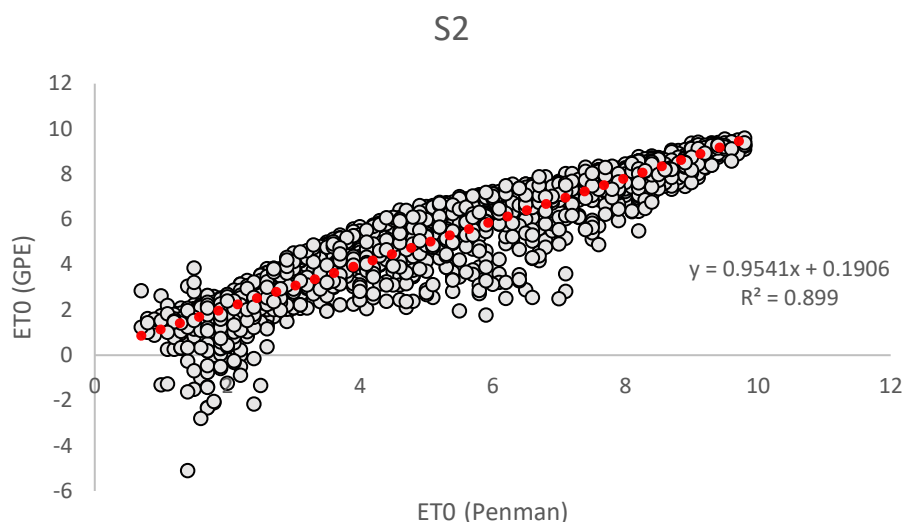
شکل ۷. نمودار یک به یک مربوط به سناریو S2 در مرحله آزمایش (Test)

Fig. 7. Scatter plot related to the S2 scenario in the test phase



شکل ۸. نمودار یک به یک مربوط به سناریو S3 در مرحله آزمایش (Test)

Fig. 8. Scatter plot related to the S3 scenario in the test phase



شکل ۹. نمودار یک به یک مربوط به سناریو S4 در مرحله آزمایش (Test)
 Fig. 9. Scatter plot related to the S4 scenario in the test phase

جدول ۶. شاخص‌های ارزیابی برای مرحله آزمایش

Table 6. Evaluation indices for the test stage

سناریو	پارامتر	R ²	MAE	RMSE
۱	Tmax+Tmin	0.933	0.569	0.695
۲	Tmax+Tmin+RH	0.946	0.514	0.621
۳	Tmax+Tmin+RH+SUN	0.965	0.419	0.506
۴	Tmax+Tmin+RH+SUN+Rain	0.899	0.868	0.609

همان‌طور که نمودارها نشان می‌دهند، برنامه ریزی بیان ژن در همه سناریوها، در شبیه‌سازی میزان تبخیر و تعرق به خوبی عمل کرده است. به عبارت دیگر با حداقل داده، مثل دمای حداکثر و دمای حداقل، آزمون با R^2 ، MAE و RMSE برابر ۰/۹۳۳، ۰/۵۶۹ و ۰/۷ میلی‌متر بر روز ۰، دقت نسبتاً خوبی انجام شده است. در سناریوی دوم، پارامتر رطوبت نسبی، به پارامترهای ورودی اضافه گردید. با اضافه شدن رطوبت نسبی، آماره R^2 نسبت به سناریوی اول ۱/۴ درصد افزایش، آماره MAE نسبت به سناریوی اول ۹/۷ درصد کاهش و آماره RMSE نسبت به سناریوی اول ۱۰/۶ درصد کاهش را نشان داد. در سناریوی سوم با افزایش ساعات آفتابی، آماره R^2 نسبت به سناریوهای اول و دوم ۳/۴ و ۲ درصد افزایش، آماره MAE نسبت به سناریوهای اول و دوم ۲۶/۴ و ۱۸/۵ درصد کاهش و آماره RMSE نسبت به سناریوهای اول و دوم ۲۷/۲ و ۱۸/۵ درصد کاهش را نشان داد. در سناریوی آخر، پارامتر بارندگی به ترکیبات ورودی اضافه گردید. نتایج نشان داد در این سناریو، آماره R^2 نسبت به سناریوهای اول، دوم و سوم ۳/۶، ۵ و ۶/۸ درصد کاهش، آماره MAE نسبت به سناریوهای اول، دوم و سوم ۵۲/۵، ۶۸/۹ و ۱۰۷/۲ درصد افزایش و آماره RMSE نسبت به سناریوهای اول و دوم ۱۲/۴ و ۱/۹ درصد کاهش و نسبت به سناریوی سوم، حدود ۲۰/۳ درصد افزایش را نشان داد. نتایج نشان می‌دهد بهترین نتایج تست نیز، زمانی حاصل گردید که از ترکیب دمای حداقل هوا، دمای حداکثر هوا، رطوبت نسبی و ساعات آفتابی به عنوان ورودی استفاده شود. به نظر می‌رسد بارش پارامتر موثری در شبیه‌سازی میزان تبخیر و تعرق مرجع در منطقه اهواز نباشد. علت این امر به بارندگی کم در اهواز می‌باشد. به گونه‌ای که میانگین بارش کمتر از یک میلی‌متر (جدول (۳)) می‌باشد. لذا بارندگی اثر مثبتی در تغییرات میزان تبخیر و تعرق نداشته است. فیض‌اله پور و همکاران (Feizolahpour et al., 2017)، سبزواری و سعیدی نیا (Sabzevari and Saeidinia., 2021) و سیاسر و همکاران (Siasar et al., 2022) نیز به بررسی توانایی مدل GEP در برآورد میزان تبخیر و تعرق در دشت سیستان، اصفهان، کرج، یزد و بجنورد پرداختند. نتایج نشان داد مدل برنامه ریزی بیان ژن قابلیت خوبی برای مدل‌سازی میزان تبخیر و تعرق را دارد.

نتیجه‌گیری

این تحقیق به منظور بررسی توانایی مدل برنامه‌ریزی بیان ژن در برآورد کردن میزان تبخیر و تعرق مرجع در اهواز انجام گرفت. در این تحقیق چهار سناریو با چهار ترکیب متفاوت از پارامترهای ورودی مورد بررسی قرار گرفت. در نهایت نتایج نشان داد مدل برنامه‌ریزی بیان ژن، توانایی بالایی در مدل کردن میزان تبخیر و تعرق بر اساس پارامترهای هواشناسی دارد. بر اساس شاخص‌های ارزیابی بهترین ورودی‌های مدل، پارامترهای دمای حداقل، دمای حداکثر، رطوبت نسبی و ساعات آفتابی تشخیص داده شد. به ازای این پارامترهای ورودی، مقادیر R^2 ، MAE و RMSE در مرحله آزمون به ترتیب برابر ۰/۹۶۵، ۰/۴۱۹ و ۰/۵۰۶ میلی متر بر روز و در مرحله آزمون (تست) به ترتیب برابر ۰/۹۶۵، ۰/۴۱۹ و ۰/۵۰۶ میلی متر بر روز به دست آمد.

منابع

- Houborg, R., & McCabe, M. F. (2018). A hybrid training approach for leaf area index estimation via Cubist and random forests machine-learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 135, 173-188.
- Feizolahpour, F., Delavar, M., & Hesami Afshar, M. (2017). Evaluation and Uncertainty Analysis of Reference Crop Evapotranspiration Estimation Using Genetic Programming. *Water and Soil Science*, 27(4), 135-147.
- Izadyar, M., hashemi, S., & Samadianfard, S. (2023). Estimating reference evapotranspiration in three arid, semi-arid and humid climates using gradient boosted tree, generalized linear model and random forest. *Water and Soil Science*, 33(3), 1-19. doi: 10.22034/ws.2021.49033.2450
- Karimi, S., Shiri, J., & Nazemi, A. H. (2013). Estimating daily reference crop evapotranspiration using artificial intelligences-based ANFIS and ANN techniques and empirical models. *Water and Soil Science*, 23(2), 139-158.
- Landeras, G., Ortiz-Barredo, A., & López, J. J. (2008). Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). *Agricultural water management*, 95(5), 553-565.
- Mattar, M.A. 2018. Using gene expression programming in monthly reference evapotranspiration modeling: a case study in Egypt. *Agricultural water management*, 198, 28-38.
- Monavar Sabegh, S., ZAREHAGHI, D., Samadianfard, S., Neishabouri, M. R., & Mikaeili, F. (2023). Estimation of Daily Reference Evapotranspiration Using Random Forest Optimized by Genetic Algorithm. *Water and Soil Science*, 33(4), 33-53. doi: 10.22034/ws.2021.48756.2449.
- Sabzevari, Y., & Saeidinia, M. (2021). Evaluation of experimental models and artificial intelligence in estimation of reference evapotranspiration (case study: Boroujerd station).
- Saggi, M. K., & Jain, S. (2020). Application of fuzzy-genetic and regularization random forest (FG-RRF): Estimation of crop evapotranspiration (ETc) for maize and wheat crops. *Agricultural Water Management*, 229, 105907.
- 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, 101-114.
- Shiri, J., Nazemi, A. H., Sadraddini, A. A., Landeras, G., Kisi, O., Fard, A. F., & Marti, P. (2014). Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Computers and Electronics in Agriculture*, 108, 230-241.
- Shiri, J., Sadraddini, A. A., Nazemi, A. H., Kisi, O., Landeras, G., Fard, A. F., & Marti, P. (2014). Generalizability of gene expression programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran. *Journal of hydrology*, 508, 1-11.
- Shiri, N., Kazemi, M. H., & Shiri, J. (2021). Application of empirical and random forest models in estimating reference evapotranspiration of humid regions. *Water and Soil Science*, 31(2), 1-11. doi: 10.22034/ws.2021.12356
- Siasar, H., & honar, T. (2019). Application of Support vector machine, CHAID and Random Forest models, in estimated daily Reference evapotranspiration in northern Sistan and Baluchestan province. *Iranian Journal of Irrigation & Drainage*, 13(2), 378-388.
- Siasar, H., Salari, A., Mohamadrezapour, O., & Piri, H. (2022). Estimating Daily Reference Evapotranspiration in Sistan Plain Using Ultra-Innovative Algorithms. *Desert Ecosystem Engineering*, 10(32), 85-96. doi: 10.22052/deej.2021.10.32.49
- Valipour, M., G. Sefidkouhi, M. Ali, M. Raeini-Sarjaz and S.M. Guzman. 2019. A Hybrid Data-Driven Machine Learning Technique for Evapotranspiration Modeling in Various Climates. *Atmosphere*, 10(6), 311.