



**EXTENDED ABSTRACT**

**Application of Multivariate Regression and Gene Expression Programming in Modeling Reference Evapotranspiration (Case Study: Khorramabad Station)**

Y. Sabzevari<sup>1</sup>, A.H. Nasrollahi<sup>2\*</sup>, M. Sharifipour<sup>3</sup> and B. Shahinejad<sup>4</sup>

1- MSc student, Department of Water Engineering, Faculty of Agriculture and Natural Resources, Lorestan University.

2\* - Corresponding Author, Assistant Professor, Department of Water Engineering, Faculty of Agriculture and Natural Resources, Lorestan University. ([nasrollahi.a@lu.ac.ir](mailto:nasrollahi.a@lu.ac.ir)).

3- Assistant Professor, Department of Water Engineering, Faculty of Agriculture and Natural Resources, Lorestan University.

4- Assistant Professor, Department of Water Engineering, Faculty of Agriculture and Natural Resources, Lorestan University.

**ARTICLE INFO**

*Article history:*

Received: 11 November 2019

Revised: 31 July 2020

Accepted: 2 August 2020

*Keywords:*

Correlation, FAO Penman Monteith, GEP, Regression.

**TO CITE THIS ARTICLE :**

Sabzevari, Y., Nasrollahi, A., Sharifipour, M., Shahinejad, B. (2022). 'Application of Multivariate Regression and Gene Expression Programming in Modeling Reference Evapotranspiration (Case Study: Khorramabad Station)', *Irrigation Sciences and Engineering*, 45(1), pp. 35-48. doi: 10.22055/jise.2020.31583.1890.

**Introduction**

Accurate estimation of water requirements of plants is a key factor in controlling several hydrological processes including: planning and management of water resources, especially in arid and semi-arid regions (Laaboudi et al., 2012; Wen et al., 2015) water pricing and water requirement for Irrigation (Yassin et al., 2016). In this study, multivariate regression methods and gene expression planning were evaluated to estimate reference evapotranspiration. For model input data, the Khorramabad Synoptic Station information including: maximum and minimum temperatures, maximum and minimum relative humidity, sunny hours and monthly wind speeds in the range of 1983-2017(420 months) were used. Based on the relationship between input and output parameters, six input patterns were determined for modeling. 70% of the data were used for training and 30% were used for model validation. The results of multivariate regression showed that the proposed model had acceptable accuracy with  $R^2 = 0.952$ . The analysis of model coefficients showed the greatest effect of maximum temperature with a coefficient of 0.604 on reference evapotranspiration. Gene expression planning results showed that the fifth pattern with four main operators was  $R^2 = 0.958$  and  $RMSE = 0.704$  in the training phase and  $R^2 = 0.977$  and  $RMSE = 0.615$  in the test phase had better performance.

**Materials and Methods**

The study area is the Khorramabad Plain, which is 2500 km<sup>2</sup> in elevation 1147.8 m above sea level, located in the middle of Zagros valley and is located in the center of Lorestan province. Climatically, it is a semi-arid region based on the Dumerton coefficient, and is semi-humid cold on the basis of the Ambergris climogram. The average annual rainfall in this region is approximately 508 mm. The reference evapotranspiration calculation was performed by FAO-Penman-Monteith method which requires data on radiation, temperature, humidity and wind speed and was able to estimate the reference evapotranspiration of the plant with high confidence level over a wide range of regions and climates. In this study this method was used as the standard

method of RefET 3.2 software. SPSS-24 software was used for multivariate regression analysis. In this study gene expression programming algorithm was used. The first step in GEP is to form an initial population of solutions. Then chromosomes are expressed by tree expression (ETs). In the next step, the compatibility or efficiency of each member of the population of chromosomes must be assessed by the fitting function. Although analysis of problems with GEP does not require a detailed understanding of the mathematical structure of the problems, but we need to determine the optimality of the potential answers of problem (chromosomes) and their compatibility with a method and by removing weaker ones and keeping the answers more consistent, getting closer to the optimal answer. In evolution and in the system of natural selection it is the role of nature to prefer more consistent responses to other responses. To evaluate the accuracy and efficiency of the models,  $R^2$ , RMSE and NS were used. The best values for these criteria are 1, 0 and 1, respectively.

### Results and discussion

To investigate the influence of model input parameters including: maximum temperature, minimum temperature, maximum relative humidity, minimum humidity, sunshine hours and wind speed on estimation of reference evapotranspiration and determination of input pattern of GEP model, multivariate regression analysis was performed between these parameters. Observing the significant level can be said that all parameters are significant at the 1% level, indicating that the input parameters all affect the evapotranspiration variable and will be used in prediction. Comparison of the standardized coefficient for the input parameters shows that maximum temperature with a coefficient of 0.604 has the most effect on the reference evapotranspiration and is a stronger predictor than the other variables. After that, respectively, wind speed, sunshine, minimum temperature, maximum and minimum relative humidity have the most influence on prediction and estimation of evapotranspiration. Hosseini *et al.* (2015) were concluded that maximum temperature and wind speed had the most effect on the reference evapotranspiration, which is in agreement with the results of the present study. According to studies in this field, the best population is from 30 to 50 chromosomes and the appropriate response comes from the selection of 2 or 3 genes (Ferreira, 2001); Therefore, 30 chromosomes and 3 genes in chromosome composition were used for the initial population in this study. Also the link function (+) was chosen because of its superiority over other functions. The maximum iteration of the fitting function was also set to 1000. The values of the genetic operators rate of consecutive insertion and consecutive insertion root equals 0.1, the mutation rate and the inversion rate of the same model default values, and the sum of the three rate operators of single-point combinations, two-point combinations, and gene combinations, respectively (Ferreira, 2001) was set at 0.7. In this study, gene expression planning has been able to perform well in all scenarios. Combined structure 5 with four main operators at Khorramabad station has the least error and most accuracy and was able to estimate the reference evapotranspiration rate with good accuracy.

### Conclusion

In this study, multivariate regression and gene expression programming methods were used to estimate reference evapotranspiration in Khorramabad. To determine the input pattern based on the correlation coefficients between input and output parameters, a decision was made. Six input patterns were determined for modeling and implemented with two types of mathematical operators, one as four main operators and one as default operators by gene expression programming. Due to the complex relationship between parameters affecting reference evapotranspiration, the efficiency of gene expression planning was evaluated. The results showed that gene expression programming method has acceptable ability to estimate reference-evapotranspiration under Khorramabad climatic conditions and can be used with high confidence in estimating reference-evapotranspiration and as an acceptable model. Finally, it is suggested to use newer methods such as game theory in determining effective variables and modeling input patterns and to use new intelligent methods for modeling reference evapotranspiration and Compare with older models.

**Gratitude**

In the end, we thank and appreciate the cooperation of the meteorological organization.

**Acknowledgment**

In the end, the authors thank the cooperation of the Meteorological Organization.

**References**

- 1- Ferreira C., 2001. Gene expression programming: a new adaptive algorithm for solving problems, *Complex Systems*, 13 (2), pp. 87-129.
- 2- Hosseini, S., Ganji Khoramdel, N. and Khalat Abadi Farahani, 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*, 1 (2), pp. 29-40. (In Persian).
- 3- 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.
- 4- 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), pp. 3195-3209.
- 5- Yassin, M.A., Alazba, A.A. and Mattar, M.A., 2016. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agricultural Water Management*, 163, pp. 110-124.



© 2022 Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0 license) (<http://creativecommons.org/licenses/by/4.0/>).

## کاربرد رگرسیون چندمتغیره و برنامه‌ریزی بیان ژن در مدل‌سازی تبخیر تعرق مرجع (مطالعه موردی: ایستگاه خرم‌آباد)

یاسر سبزواری<sup>۱</sup>، علی حیدر نصرالهی<sup>۲\*</sup>، مجید شریفی پور<sup>۳</sup> و بابک شاهی نژاد<sup>۴</sup>

۱- دانشجوی کارشناسی ارشد مهندسی آبیاری و زهکشی، دانشکده کشاورزی، دانشگاه لرستان.

۲- نویسنده مسئول، استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه لرستان، nasrolahi.a@lu.ac.ir

۳- استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه لرستان.

۴- استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه لرستان.

پذیرش: ۱۳۹۹/۵/۱۲

بازنگری: ۱۳۹۹/۵/۱۰

دریافت: ۱۳۹۸/۸/۲۰

### چکیده

در این پژوهش به منظور تخمین تبخیر تعرق مرجع در ایستگاه خرم‌آباد، روش‌های رگرسیون چندمتغیره و برنامه‌ریزی بیان ژن مورد بررسی و استفاده قرار گرفتند. برای اطلاعات ورودی مدل از اطلاعات ایستگاه سینوپتیک خرم‌آباد شامل: درجه حرارت حداکثر و حداقل، رطوبت نسبی حداکثر و حداقل، ساعات آفتابی و سرعت باد ماهانه در بازه زمانی ۱۳۹۵-۱۳۶۱ (به تعداد ۴۲۰ ماه) استفاده شد. بر اساس رابطه بین پارامترهای ورودی و خروجی، شش الگوی ورودی برای مدل‌سازی تعیین گردید. ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد داده‌ها برای صحت‌سنجی مدل‌ها به کار گرفته شد، همچنین در روش برنامه‌ریزی بیان ژن دو نوع عملگر ریاضی شامل چهار عملگر اصلی و عملگرهای پیش فرض مدل مورد بررسی قرار گرفتند. نتایج حاصل از رگرسیون چندمتغیره نشان داد که مدل پیشنهادی با مقدار  $R^2 = 0/952$  از دقت قابل‌قبولی برخوردار است. تحلیل ضرایب مدل حاکی از بیشترین تأثیر حداکثر درجه حرارت با ضریب  $0/604$  بر تبخیر تعرق مرجع بود. نتایج برنامه‌ریزی بیان ژن نشان داد الگوی پنجم با عملگرهای چهار اصلی، در مرحله آموزش با  $R^2 = 0/958$ ،  $RMSE = 0/704$  و  $NS = 0/97$  و در مرحله آزمون با  $R^2 = 0/977$ ،  $RMSE = 0/615$  و  $NS = 0/977$  عملکرد بهتری را داشته و نتایج به‌دست‌آمده نشان داد که برنامه‌ریزی بیان ژن دارای توانایی قابل‌قبولی در تخمین تبخیر تعرق مرجع تحت شرایط آب‌وهوایی خرم‌آباد بوده و به‌عنوان مدل قابل‌استفاده در این زمینه معرفی کرد.

کلید واژه‌ها: رگرسیون، فائو پنمن مانیت، همبستگی، GEP.

### مقدمه

برآورد دقیق نیاز آبی گیاهان، عاملی اصلی در کنترل چندین فرایند هیدرولوژیکی از جمله: برنامه‌ریزی و مدیریت منابع آب خصوصاً در مناطق خشک و نیمه‌خشک (Laaboudi ; (2012) Tabari et al)؛ (2012) et al.؛ (2012) Wen et al) تعیین قیمت آب (Pour Yassin et al. (2013) Ali Baba et al) و نیاز آب آبیاری (2016) است. تبخیر تعرق مرجع ( $ET_0$ ) یکی از مهم‌ترین اجزای چرخه هیدرولوژی می‌باشد که تابع متغیرهای اقلیمی از جمله: سرعت باد در نزدیکی سطح زمین، دمای هوا، تابش خورشیدی و رطوبت نسبی است. این متغیرها به عرض و موقعیت جغرافیایی هر منطقه وابسته هستند. در نتیجه مقادیر  $ET_0$  را می‌توان با استفاده از مدل‌های آب و هواشناسی که بر پایه معادله‌های فیزیکی یا رابطه تجربی متغیرهای هواشناسی هستند، برآورد و شبیه‌سازی نمود. یکی از روش‌هایی که مبنای فیزیکی دارد، روش فائو-پنمن-مانیت (PMF-56) است که  $ET_0$  را به‌صورت روزانه و ماهانه برآورد می‌کند.

سازمان خواروبار جهانی (FAO) این روش را به‌عنوان معادله استاندارد توصیه کرده است (Allen et al., 1998). این روش به تعدادی از متغیرهای اقلیمی نظیر حداقل و حداکثر دمای هوا، تابش خورشیدی، رطوبت نسبی و ساعات آفتابی نیاز دارد. با این وجود، این متغیرهای هواشناسی خصوصاً در کشورهای درحال توسعه، یا در دسترس نبوده یا معمولاً غیرقابل‌اعتماد هستند (Almorox et al (2015) بنابراین ارزیابی روشی که بتواند با داده‌های ورودی کمتر، دقت مناسبی را به همراه داشته باشد، خصوصاً در این مناطق الزامی است. در سال‌های اخیر، روش‌های هوش مصنوعی (Artificial Intelligence) با داده‌های هواشناسی محدود، به‌طور موفقیت‌آمیزی در شبیه‌سازی  $ET_0$  مورد استفاده قرار گرفته‌اند. از جمله روش‌های هوش مصنوعی به کار گرفته شده برای شبیه‌سازی  $ET_0$  می‌توان به برنامه‌ریزی بیان ژن اشاره کرد. (Shiri et al., 2014)؛ (Marti et al., 2015) Seifi et al. (2010) به توسعه مدل ترکیبی رگرسیون چندمتغیره تحلیل مؤلفه‌ها و عامل‌های اصلی در پیش‌بینی تبخیر تعرق مرجع در

ماهانه در مصر پرداخت. نتایج نشان داد که برنامه‌ریزی بیان ژن برای مدل‌سازی تبخیرتغرق مرجع دقیق‌تر از مدل‌های تجربی است. هم‌چنان که ملاحظه می‌شود تاکنون مطالعه‌های مختلفی در زمینه کاربرد و ارزیابی مدل‌های هوش مصنوعی برای برآورد تبخیرتغرق مرجع در نقاط مختلف جهان انجام شده است. نتایج تحقیقات مختلف نشان می‌دهد که مدل‌های هوش مصنوعی در بسیاری از موارد توانسته‌اند با داده‌های ورودی کمتر ولی با دقت بالاتر در برآورد تبخیرتغرق مورد استفاده قرار گیرند. از آنجایی که تا به حال در استان لرستان تحقیقی در مورد روش‌های برآورد تبخیرتغرق مرجع صورت نگرفته لذا هدف از پژوهش حاضر بررسی و تعیین پارامترهای مؤثر در برآورد ET<sub>0</sub> با استفاده از روش رگرسیون چندمتغیره و ارایه یک رابطه مناسب و کاربردی توسط برنامه‌ریزی بیان ژن (Gene Expression Programming) در ایستگاه خرم‌آباد است.

## مواد و روش‌ها

### منطقه مورد مطالعه

دشت خرم‌آباد با وسعتی حدود ۲۵۰۰ کیلومترمربع در ارتفاع ۱۱۴۷/۸ متری از سطح دریا و در میان دره‌ی زاگرس قرار دارد که در مرکز استان لرستان واقع گردیده است. به‌طور کلی از نظر اقلیمی بر اساس ضریب دوما رتن جز مناطق نیمه‌خشک و بر اساس کلیموگرام آمبرژه نیمه‌مرطوب سرد محسوب می‌شود. میانگین مجموع بارش سالیانه این منطقه حدود ۵۰۸ میلی‌متر است. آمار و اطلاعات هواشناسی مورد نیاز برای محاسبه‌ی تبخیرتغرق گیاه مرجع در این پژوهش، از ایستگاه سینوپتیک خرم‌آباد با مختصات جغرافیایی طول ۴۷ درجه و ۱۷ دقیقه، عرض ۳۳ درجه و ۲۶ دقیقه و ارتفاع از سطح دریا ۱۱۴۷/۸، اخذ گردید. شکل (۱) نشان‌دهنده‌ی موقعیت دشت خرم‌آباد و ایستگاه سینوپتیک خرم‌آباد است.

در این پژوهش برای محاسبه تبخیرتغرق مرجع با استفاده از روش فائو-پنمن-مانتیت از اطلاعات: درجه حرارت حداکثر، درجه حرارت حداقل، رطوبت نسبی حداکثر، رطوبت حداقل، ساعات آفتابی و سرعت باد در ارتفاع دو متری در یک بازه زمانی ۳۵ ساله (۱۳۹۵-۱۳۶۱) و به‌صورت ماهانه استفاده گردید.

### محاسبه تبخیرتغرق مرجع به روش فائو-پنمن-مانتیت

در سال ۱۹۹۰ از سوی کمیسیون بین‌المللی آبیاری و زهکشی (ICID) و سازمان خواروبار جهانی (FAO) روش فائوپنمن‌مانتیت به‌عنوان تنها روش استاندارد برای محاسبه تبخیرتغرق گیاه مرجع با توجه به داده‌های اقلیمی و هم‌چنین برای ارزیابی سایر روش‌های تجربی پی‌شهاد گردید (Hargreaves, 1994). این روش نیازمند داده‌های تابش، دما، رطوبت و سرعت باد بوده و توانسته است با درجه اعتماد بالایی و در دامنه وسیعی از مناطق و اقلیم‌ها برآورد صحیحی از تبخیرتغرق گیاه مرجع ارایه دهد (Allen et al., 1998). در این پژوهش برای محاسبه تبخیرتغرق مرجع از روش

ایستگاه کرمان پرداختند. نتایج نشان داد اختلاف معنی‌داری بین دو روش رگرسیون چندمتغیره-تحلیل مؤلفه‌ها و رگرسیون چندمتغیره وجود ندارد. Hosseini et al. (2015) به ارزیابی و تحلیل حساسیت روش‌های مختلف تخمین تبخیرتغرق 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. Ahmadi et al. (2015) در پژوهشی کاربرد سیستم‌های استنتاج عصبی-فازی تطبیقی و برنامه‌ریزی ژنتیک را برای برآورد تبخیرتغرق ماهانه در شمال غرب ایران بررسی کردند. نتایج این پژوهش حاکی از دقت بالا و خطای کم هر دو مدل در پیش‌بینی تبخیرتغرق ماهانه گیاه مرجع بوده و کارایی بهتر مدل ANFIS نسبت به روش GP بود. Mohammadrezapour et al. (2015) روند تبخیرتغرق پتانسیل ماهانه سیستان و بلوچستان در یک دوره ۴۰ ساله را در قالب ۱۶ الگوی زمانی تأخیری با استفاده از برنامه‌ریزی ژنتیک مورد بررسی قرار دادند. نتایج نشان داد تمامی مدل‌های مناسب مورد قبول، در الگوی ۱۰ به بالا به‌دست آمده است. Sattari و Esmailzadeh (2016) به مقایسه نتایج مدل درختی M5 و برنامه‌ریزی ژنتیک با روش فائوپنمن‌مانتیت برای تخمین تبخیرتغرق مرجع پرداختند. نتایج نشان داد که هر دو روش مذکور نتایج دقیقی را جهت پیش‌بینی تبخیرتغرق مرجع ماهانه در منطقه تبریز ارایه می‌دهند، ولی مدل درختی M5 روابط خطی ساده، قابل‌فهم و کاربردی‌تری برای پیش‌بینی تبخیرتغرق مرجع ارایه می‌کند. Feyzollahpour et al. (2017) به ارزیابی و تحلیل عدم قطعیت برآورد تبخیرتغرق گیاه مرجع با استفاده از برنامه‌ریزی ژنتیک پرداختند. نتایج نشان داد که مدل‌های برنامه‌ریزی ژنتیک دارای دقت بالاتری نسبت به روش‌های تجربی هستند. Ladlani et al. (2012) تبخیرتغرق روزانه در منطقه مدیرانه‌ای کشور الجزایر را توسط روش‌های سیستم فازی (Fuzzy) و رگرسیون چندمتغیره خطی مدل‌سازی کردند. نتایج به‌دست‌آمده نشان از عملکرد قابل‌قبول هر دو مدل در پیش‌بینی تبخیرتغرق داشت. Yassin et al. (2016) به ارزیابی شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن برای شبیه‌سازی تبخیرتغرق مرجع در اقلیم خشک پرداختند. نتایج نشان داد که هشت مدل برآورد تبخیرتغرق مرجع با استفاده از شبکه عصبی مصنوعی نسبت به همان مدل‌ها با برنامه‌ریزی بیان ژن دارای دقت بالاتری هستند. Shiri (2017) به ارزیابی مدل‌های فائوپنمن‌مانتیت ۵۶، تجربی، نیمه‌تجربی و برنامه‌ریزی بیان ژن در شبیه‌سازی تبخیرتغرق مرجع روزانه برای مناطق فراخشک ایران پرداخت. نتایج حاکی از برتری مدل برنامه‌ریزی بیان ژن نسبت به روش‌های تجربی و نیمه‌تجربی است. Kisi و Alizamir (2018) به مدل‌سازی تبخیرتغرق مرجع با استفاده از یک روش اکتشافی پیوستگی مویک پرداختند و نتایج نشان داد که مدل‌های شامل چهار متغیر ورودی، دقت بهتری را نسبت به مدل‌های با ورودی کمتر ارایه می‌دهند. تشعشعات خورشیدی مؤثرترین متغیر در ET<sub>0</sub> بود. Mattar (2018) با استفاده از برنامه‌ریزی بیان ژن به مدل‌سازی تبخیرتغرق مرجع

۱) تعداد متغیرهای پیش‌بین (مستقل) در رگرسیون باید کمتر از تعداد مشاهدات باشد.

۲) همبستگی خطی کامل بین متغیرهای پیش‌بین و پاسخ وجود داشته باشد.

در صورت نقض این دو فرض معادله رگرسیونی را نمی‌توان تخمین زد. مدل رگرسیونی به شکل ماتریسی را می‌توان به صورت معادله زیر نشان داد:

$$Y = X\beta + e \quad (2)$$

که  $\beta$  ماتریس ضرایب رگرسیونی که ضرایب نامشخصی هستند که در حقیقت، مسئول برآورد پارامتر پاسخ هستند (Sarmadian et al. (2010) و Gholami et al. (2012)،  $e$  ماریس خطای برازش و  $Y$  نیز ماتریس پاسخ است. با حل معادله (۲) برحسب  $\beta$  خواهیم داشت:

$$\beta = (X'X)^{-1}(X'Y) \quad (3)$$

که در رابطه بالا،  $X'$  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 نمی‌توان معکوس کرد و باعث افزایش خطا در اثر گرد کردن داده‌ها و محاسبات می‌شود. در این مطالعه جهت انجام تحلیل رگرسیون چندمتغیره از نرم‌افزار SPSS-24 استفاده شد.

فائو-پنمن-مانتیت به‌عنوان روش استاندارد از نرم‌افزار RefET-۳/۲ استفاده شد. این نرم‌افزار با استفاده از متوسط روزانه درجه حرارت حداقل و حداکثر، متوسط روزانه رطوبت نسبی حداقل و حداکثر، تعداد ساعات آفتابی و سرعت باد در ارتفاع دو متری، تبخیر تعرق مرجع را محاسبه می‌کند. رابطه مورد استفاده برای محاسبه تبخیر تعرق مرجع به صورت زیر است:

$$ET_0 = \frac{[0.408(R_n - G) + [(900\gamma)/(T + 273)] \times U_2 \times (e_a - e_d)]}{\Delta + \gamma \times (1 + 0.34U_2)} \quad (1)$$

در این رابطه:  $ET_0$ : تبخیر تعرق گیاه مرجع، برحسب  $R_n$  (mm/day): تشعشع خالص برحسب  $G$  ( $MJ/m^2 \cdot day$ ); شار گرمایی خاک، برحسب  $\Delta$  ( $MJ/m^2 \cdot day$ ); شیب نمودار فشار بخار اشباع نسبت به دما، برحسب  $\gamma$  ( $kp/^\circ C$ ); ثابت سایکرومتری، برحسب  $e_a - e_d$  ( $kp/^\circ C$ ); کمبود فشار بخار در ارتفاع ۲ متری، برحسب  $U_2$ : میانگین روزانه سرعت باد در ارتفاع ۲ متری از سطح زمین، برحسب  $T$  (m/s): میانگین روزانه دمای هوا در ارتفاع ۲ متری از سطح زمین، برحسب ( $^\circ C$ ).

### مدل رگرسیونی خطی-چند متغیره

با این روش می‌توان به‌طور هم‌زمان به تحلیل و بررسی چند متغیر مختلف پرداخت. رگرسیون چندمتغیره در واقع، رابطه بین چند متغیر پیش‌بین با متغیر پاسخ موردنظر را بیان می‌کند. این‌گونه مدل‌ها دارای فرضیاتی هستند. فرض‌هایی که رگرسیون چندمتغیره را از رگرسیون ساده متمایز می‌نماید این است که:

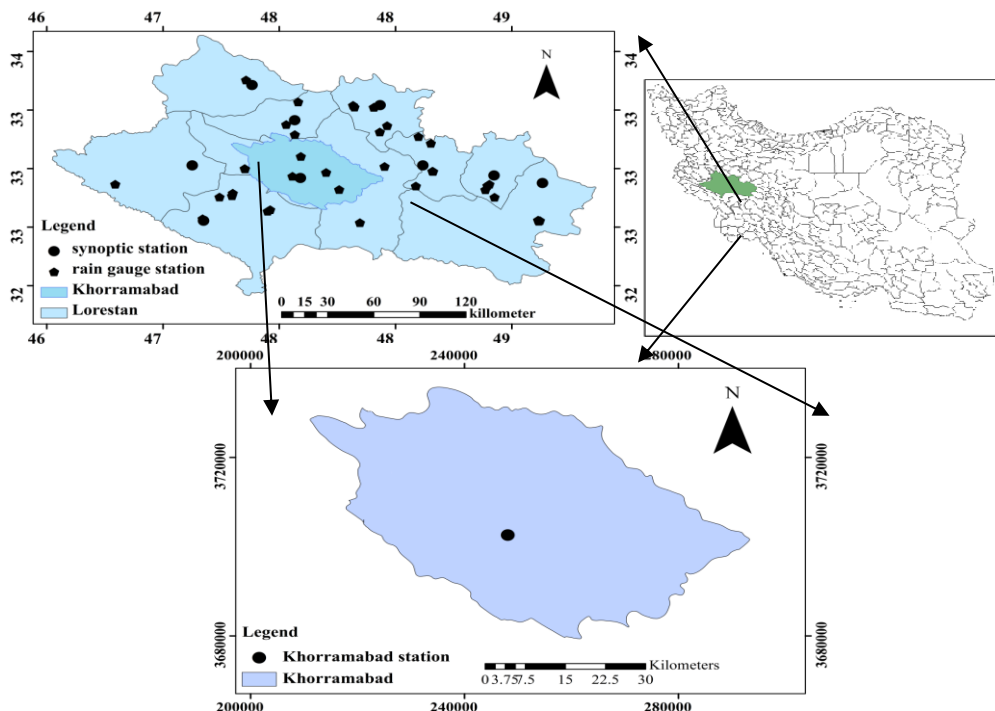


Fig. 1- Geographical position of Khorramabad plain in Lorestan and Iran

شکل ۱- موقعیت جغرافیایی دشت خرم‌آباد در لرستان و کشور



$$t=h(n-1)+1 \quad (۴)$$

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  $\{x^2, \exp, \log, \sin, \cos, \dots\}$  و یا توابع تعریف‌شده توسط کاربر است. توابع و ترمینال‌ها در بخش سر ژن وجود دارند و دنباله تنها شامل ترمینال‌هاست (Lopes and Weinert, 2004) در این پژوهش، برای مدل‌سازی با GEP، از نرم‌افزار GeneXproTools-5 استفاده گردید.

### معیارهای ارزیابی

به‌منظور ارزیابی دقت و کارایی مدل‌ها، از معیارهای ضریب تعیین ( $R^2$ )، ریشه میانگین مربعات خطا (RMSE) و ضریب نش‌سانکلیف (NS) به‌صورت زیر استفاده گردید. بهترین مقادیر برای این معیارها به ترتیب: یک، صفر و یک است.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2} \quad (۵)$$

$$R^2 = \frac{\left( \sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y}) \right)^2}{\left( \sum_{i=1}^N (X_i - \bar{X})^2 \sum_{i=1}^N (Y_i - \bar{Y})^2 \right)} \quad \text{و} \quad 0 \leq R^2 \leq 1 \quad (۶)$$

$$NS = 1 - \frac{\sum_{i=1}^N (X_i - Y_i)^2}{\sum_{i=1}^N (X_i - \bar{Y}_i)^2} \quad \text{و} \quad -\infty \leq NS \leq 1 \quad (۷)$$

در این روابط،  $Y_i$  و  $X_i$  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 است.

### نتایج و بحث

#### رگرسیون چندمتغیره

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 تعرق مرجع و تعیین الگوی ورودی مدل GEP، تحلیل رگرسیونی چندمتغیره بین این پارامترها صورت گرفت که نتایج در جدول (۱) نشان داده شده است. یکی از بخش‌های مهم این جدول مقادیر بتا و سطح معنی‌داری است. با

### برنامه‌ریزی بیان ژن

روش برنامه‌ریزی بیان ژن، ترکیب و توسعه‌یافته‌ی روش‌های GA و GP است که در سال ۱۹۹۹ توسط فریرا ابداع شد. 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. به‌طور خلاصه می‌توان بیان داشت که در این روش ژنوتیپ و فنوتیپ از هم جدا شده و سیستم قادر خواهد بود از تمام مزایای تکاملی بهره‌گیرد. با وجود اینکه فنوتیپ در GEP مشابه ساختار شاخه‌ای GP است، اما ساختار شاخه‌ای در GEP که بیان درختی نیز نامیده می‌شود، بیانگر تمام ژنوم‌های مستقل است. در مجموع به‌طور خلاصه می‌توان بیان داشت که در GEP به‌سازی در یک ساختار خطی اتفاق می‌افتد و سپس به‌صورت ساختار درختی بیان می‌شود، این امر سبب خواهد شد که تنها ژنوم اصلاح‌شده به نسل بعد انتقال یابد و نیازی به ساختارهای سنگین تکثیر و جهش نباشد (Ferreira, 2001).

### الگوریتم برنامه‌ریزی بیان ژن

نخستین گام در GEP، تشکیل جمعیت اولیه از راه‌حل‌هاست. سپس کروموزوم‌ها با بیان درختی (ETs) به نمایش در می‌آیند. در گام بعد، باید سازگاری یا کارایی هر عضو از جمعیت کروموزوم‌ها به‌وسیله تابع برازش ارزیابی شود. گرچه در تحلیل مسئله‌ها با GEP به دانستن دقیق ساختار ریاضی مسئله‌ها نیازی نیست، اما باید میزان مطلوب بودن جواب‌های بالقوه مسئله (کروموزوم‌ها) و میزان سازگاری آن‌ها را به روشی مشخص نمود و با حذف جواب‌های ضعیف‌تر و حفظ جواب‌های سازگارتر، به پاسخ بهینه مسئله نزدیک شد. 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، پاسخ بهینه به دست آید، فرایند تولید نسل متوقف خواهد شد.

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 می‌گردد و هر ژن شامل یک سر و یک دنباله است. اندازه سر ( $h$ )، به‌وسیله کاربر تعیین می‌شود اما اندازه دنباله ( $t$ )، به‌صورت تابعی از  $h$  و پارامتر  $n$  محاسبه می‌گردد که پارامتر  $n$  حداکثر تعداد آرگومان توابع موجود است. رابطه (۴) طول دنباله را با توجه به پارامترهای دیگر تعیین می‌کند (Lopes and Weinert, 2004):

انتهاپی جدول (۱) نیز نشان دهنده رابطه رگرسیونی بین پارامترهای ورودی و تبخیرتغرق مرجع بر اساس شش الگوی ورودی است. بر اساس شاخص‌های ارزیابی، الگوی شماره شش با  $RMSE=0.7$  و  $R^2=0.97$  به‌عنوان مناسب‌ترین الگو معرفی شد.

الگوی ورودی مدل‌ها برای پیش‌بینی تبخیرتغرق مرجع در جدول (۲) نشان داده شده است.

بر اساس این جدول، مدل‌سازی تبخیرتغرق مرجع با شش الگو به‌صورت جدول زیر انجام شد.

شکل (۲-ا) مقادیر مشاهداتی و محاسباتی تبخیرتغرق رگرسیون چند متغیره را در طول دوره آماری ۳۵ ساله نشان می‌دهد. همان‌طور که مشاهده می‌گردد مدل توانسته است روند تغییرات  $ET_0$  را به‌طور مناسبی پیش‌بینی نماید. در شکل (۲-ب) مقادیر مشاهداتی در مقابل مقادیر محاسباتی آموزش مدل نسبت به نیمساز ناحیه اول ترسیم شده‌اند. بر اساس نمودار پراکنش مشاهده می‌شود که عملکرد مدل مناسب بوده به‌طوری که در اکثر موارد، مقادیر مشاهداتی در نزدیکی مقادیر محاسباتی واقع شده است.

مشاهده‌ی سطح معنی‌داری حاصل‌شده می‌توان گفت همه پارامترها در سطح یک درصد معنی‌دار بوده و نشان دهنده این موضوع است که پارامترهای ورودی همگی بر متغیر تبخیرتغرق اثرگذار بوده و در پیش‌بینی مورد استفاده قرار می‌گیرند. مقایسه ضریب استانداردشده به‌دست‌آمده برای پارامترهای ورودی نشان می‌دهد که حداکثر درجه حرارت با ضریب  $0.64$  دارای بیشترین تأثیر بر تبخیرتغرق مرجع است و پیش‌بینی‌کننده قوی‌تری نسبت به بقیه متغیرهاست. بعد از آن به‌ترتیب سرعت باد، ساعات آفتابی، درجه حرارت حداقل، رطوبت نسبی حداکثر و حداقل دارای بیشترین تأثیر در پیش‌بینی و برآورد تبخیرتغرق می‌باشند. Hosseini et al. (2015) نتیجه گرفتند که حداکثر درجه حرارت و سرعت باد بیشترین تأثیر را بر تبخیرتغرق 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 معنی‌داری اگر کوچک‌تر از  $0.1$  باشد، نشان می‌دهد متغیر پیش‌بین بر متغیر پاسخ تأثیر معنی‌دار دارد. با توجه به مقادیر مربوط به هر شش متغیر، همه‌ی متغیرها دارای تأثیر معنی‌دار هستند. قسمت

#### جدول ۱- مقادیر ضرایب و پارامترهای مدل رگرسیون چندمتغیره

Table 1- Model accuracy coefficients and multivariate regression parameters

|                  | Unstandardized Coefficients |            | Standardized Coefficients   | t       | Sig.  |
|------------------|-----------------------------|------------|---|---------|-------|
|                  | B                           | Std. Error | Beta  |         |       |
| (Constant)       | -.382                       | .114       |   | -3.357  | 0.001 |
| $T_{max}$        | .250                        | .004       | .604  | 65.827  | 0     |
| $T_{min}$        | .033                        | .004       | .055  | 7.876   | 0     |
| $RH_{max}$       | -.034                       | .001       | -.186   | -42.032 | 0     |
| $RH_{min}$       | -.009                       | .001       | -.038   | -7.434  | 0     |
| WS               | .399                        | .005       | .212  | 84.196  | 0     |
| SUN              | .133                        | .005       | .108  | 29.189  | 0     |
| Evaluation index |                             |            |   |         |       |
| NS               | RMS E                       | $R^2$      | Equation  |         |       |
| 0.80             | 1.94                        | 0.938      | $ET_0 = -0.382 + 0.25T_{max}$   |         |       |
| 0.82             | 1.85                        | 0.957      | $ET_0 = -0.382 + 0.25T_{max} + 0.399WS$   |         |       |
| 0.72             | 2.54                        | 0.96       | $ET_0 = -0.382 + 0.25T_{max} + 0.399WS + 0.133Sun$  |         |       |
| 0.69             | 2.72                        | 0.961      | $ET_0 = -0.382 + 0.25T_{max} + 0.399WS + 0.133Sun + 0.033T_{min}$                                 |         |       |
| 0.73             | 2.48                        | 0.962      | $ET_0 = -0.382 + 0.25T_{max} + 0.399WS + 0.133Sun + 0.033T_{min} - 0.009RH_{min}$                 |         |       |
| 0.97             | 0.7                         | 0.962      | $ET_0 = -0.382 + 0.25T_{max} + 0.399WS + 0.133Sun + 0.033T_{min} - 0.009RH_{min} - 0.034RH_{max}$ |         |       |



جدول ۲- الگوهای ورودی مدل‌های رگرسیون چند متغیره و برنامه‌ریزی بیان ژن

Table 2- Input Patterns of Multivariate Regression Models and Gene Expression Programming

| Input Patterns of Models |   |
|--------------------------|---|
| 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                        | $ET_0 = f\{T_{max}, WS, Sun, T_{min}, RH_{min}, RH_{max}\}$ |

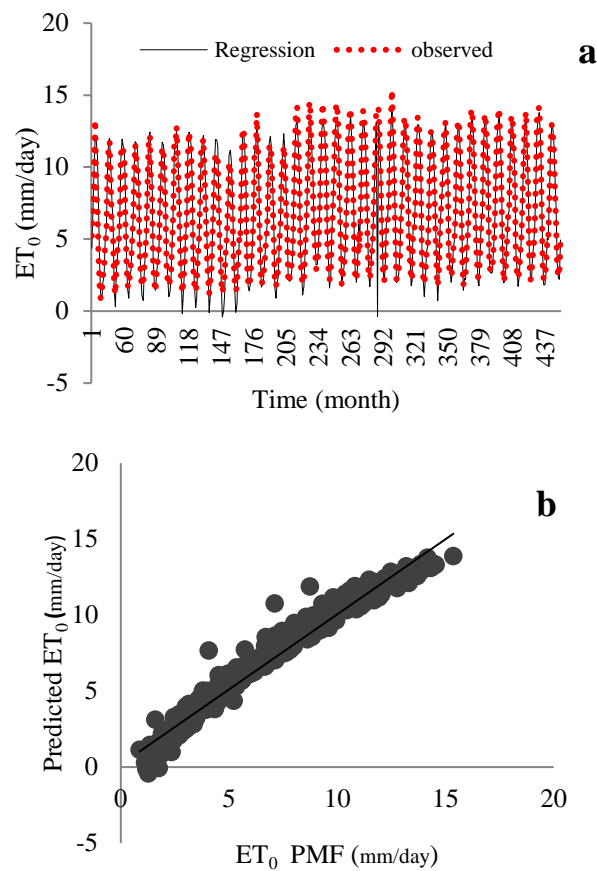


Fig. 2- (a) Diagram of calculated and predicted values over time; (B) Diagram of the distribution of calculated and predicted Regression values

شکل ۲- (a): نمودار مقادیر محاسباتی و پیش‌بینی شده نسبت به زمان، (b): نمودار پراکنش میان مقادیر محاسباتی و پیش‌بینی شده GEP در مرحله آموزش

جدول ۳- مقادیر پارامترهای مورد استفاده در مدل GEP

Table 3- The values of the parameters used in the GEP model

| Parameters                      | Value | Genetic operator              | Value   |
|---------------------------------|-------|-------------------------------|---------|
| Number of chromosome            | 30    | Mutation rate                 | 0.00138 |
| Number of gene in chromosome    | 3     | Inversion rate                | 0.00546 |
| Link function                   | (+)   | IS Transposition rate         | 0.1     |
| Error Criterion of Fit Function | RMSE  | RIS Transposition rate        | 0.1     |
| Maximum Repeat Fit Function     | 1000  | One-Point Recombination rate  | 0.2     |
|                                 |       | Two-Points Recombination rate | 0.2     |
|                                 |       | Gene Recombination rate       | 0.3     |

جدول ۴ - روش برنامه‌ریزی بیان ژن با استفاده از سه مجموعه عملگر ریاضی

Table 4 - Gene expression programming method using three sets of mathematical operators

| Pattern | operator | Testing      |                |       | Training     |                |       |       |
|---------|----------|--------------|----------------|-------|--------------|----------------|-------|-------|
|         |          | RMSE(mm/day) | R <sup>2</sup> | NS    | RMSE(mm/day) | R <sup>2</sup> | NS    |       |
| 1       | M1       | F1           | 1.01           | 0.94  | 0.939        | 0.822          | 0.959 | 0.945 |
|         |          | F2           | 1.007          | 0.942 | 0.94         | 0.818          | 0.961 | 0.946 |
| 2       | M2       | F1           | 0.842          | 0.96  | 0.957        | 0.857          | 0.948 | 0.949 |
|         |          | F2           | 0.786          | 0.966 | 0.965        | 0.954          | 0.954 | 0.953 |
| 3       | M3       | F1           | 1.05           | 0.958 | 0.936        | 1.06           | 0.93  | 0.927 |
|         |          | F2           | 0.674          | 0.982 | 0.973        | 0.775          | 0.959 | 0.966 |
| 4       | M4       | F1           | 0.815          | 0.971 | 0.961        | 0.802          | 0.955 | 0.956 |
|         |          | F2           | 0.954          | 0.949 | 0.947        | 0.955          | 0.934 | 0.941 |
| 5       | M5       | F1           | 0.625          | 0.977 | 0.977        | 0.741          | 0.958 | 0.97  |
|         |          | F2           | 0.654          | 0.976 | 0.975        | 0.773          | 0.953 | 0.968 |
| 6       | M6       | F1           | 0.851          | 0.959 | 0.958        | 0.993          | 0.933 | 0.942 |
|         |          | F2           | 0.643          | 0.983 | 0.976        | 0.867          | 0.953 | 0.964 |

می‌شود Ferreira (2001) بنابراین برای جمعیت اولیه در این پژوهش از ۳۰ کروموزوم و از سه ژن در ترکیب کروموزوم استفاده شد. هم‌چنین تابع پیوند (+) به علت برتری بر دیگر توابع انتخاب گردید. حداکثر تکرار تابع برازش هم ۱۰۰۰ تعیین شد. مقادیر نرخ عملگرهای ژنتیکی ترانهش درج متوالی و ترانهش ریشه درج متوالی برابر ۰/۱، نرخ جهش و نرخ وارون‌سازی همان مقادیر پیش‌فرض مدل و مقدار مجموع نرخ سه عملگر ترکیب تک نقطه‌ای، ترکیب دونقطه‌ای و ترکیب ژن، طبق نظر Ferreira (2001) برابر ۰/۷ تعیین شد. در جدول (۳)، پارامترهای استفاده‌شده در مدل به همراه مقادیر آن‌ها، ارایه 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 (۴) ارایه شده است. همان‌طور که مشاهده می‌شود، برنامه‌ریزی بیان ژن توانسته است در تمام سناریوها عملکرد و دقت مناسبی از خود ارایه دهد. در بین سناریوهای انجام‌شده، ساختار ترکیبی شماره پنج با چهار عملگر اصلی در ایستگاه خرم‌آباد، دارای کم‌ترین خطا و بیشترین دقت بوده و توانسته است میزان تبخیر تعرق مرجع را با دقت مناسبی برآورد کند.

#### مدل‌سازی تبخیر تعرق مرجع به وسیله GEP

بر اساس تحلیل همبستگی بین پارامترهای ورودی و خروجی، شش الگو جهت مدل‌سازی تبخیر تعرق مرجع به صورت جدول (۲) پیشنهاد گردید. همچنین برای اجرای مدل، از ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد داده‌ها برای صحت‌سنجی و آزمون انتخاب شد. در برآورد تبخیر تعرق مرجع، علاوه بر چهار عملگر اصلی (F1)، حالت‌هایی بر اساس عملگرهای ریاضی پیش‌فرض برنامه (F2) نیز در نظر گرفته شد (روابط ۸ و ۹). برای ساخت درخت تجزیه و تولید جمعیت اولیه، از راه‌حل‌هایی که از ترکیب تصادفی مجموعه توابع (عملگرهای ریاضی مورد استفاده در روابط) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) ایجاد می‌شوند، استفاده گردید.

$$F1 = \{+, /, -, *, \sqrt{\quad}\} \quad (8)$$

$$F2 = \{+, /, -, *, \sqrt{\quad}, \sqrt[3]{\quad}, \text{Exp}, \text{Ln}, \text{Sin}, \text{Cos}, \text{Atan}\} \quad (9)$$

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 دو یا سه ژن حاصل

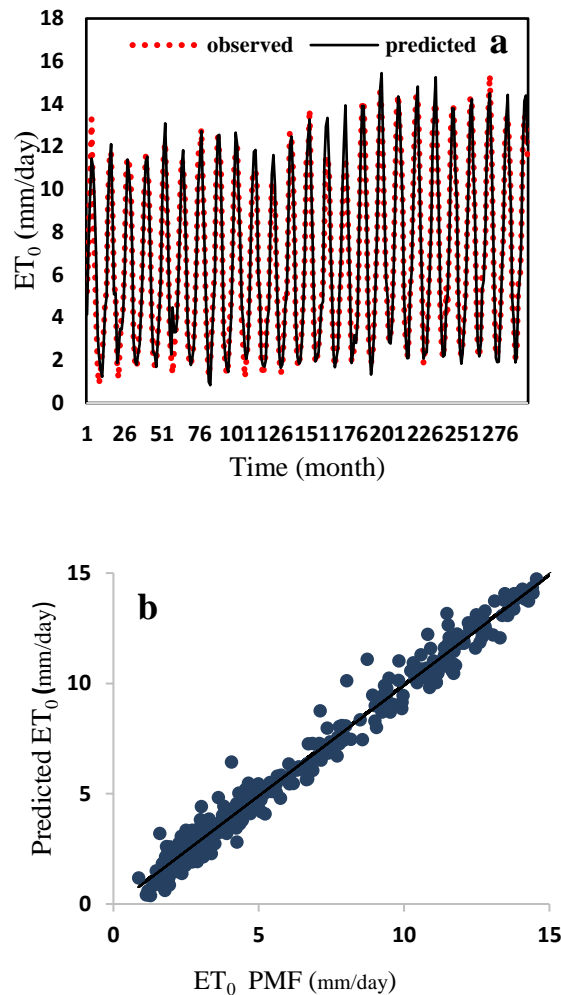
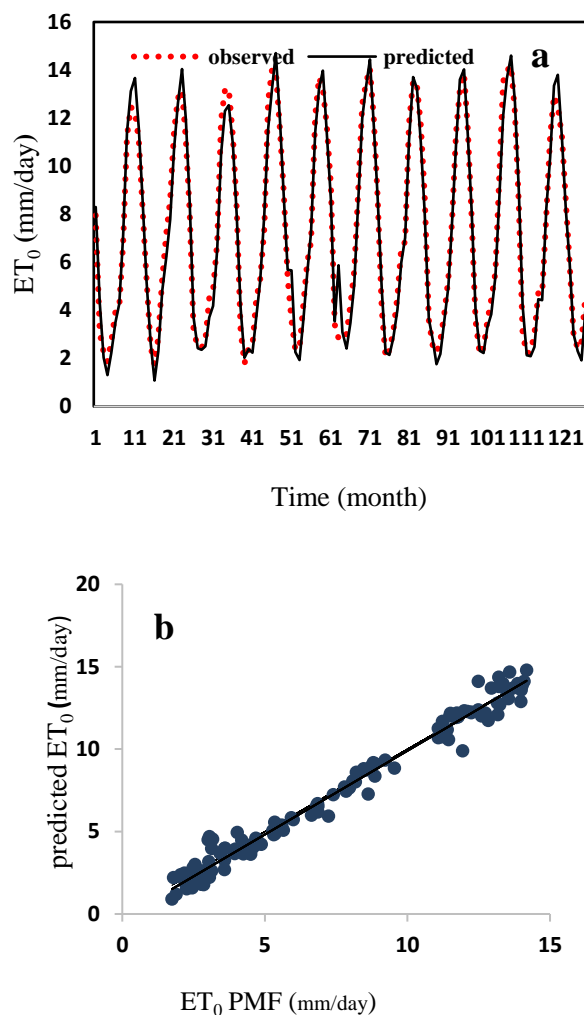


Fig. 3- (a) Diagram of calculated and predicted values over time; (B) Diagram of the distribution of calculated and predicted GEP values in the training phase

شکل ۳- (a): نمودار مقادیر محاسباتی و پیش‌بینی شده نسبت به زمان، (b): نمودار پراکنش میان مقادیر محاسباتی و پیش‌بینی شده در مرحله آموزش

مقدار آن زیاد نیست. در مجموع استفاده از عملگرهایی مانند:  $\sin$ ،  $\cos$ ،  $\ln$  و ... برای استخراج رابطه باعث می‌گردد که در صورت وجود یک رابطه غیرخطی بین ورودی و خروجی آن، دقت مدل کاهش پیدا نکند که در عملگر F2 حتی در مواردی باعث بهبود عملکرد مدل هم شده است. شکل (۴-ا) نیز نشان‌دهنده مقادیر مشاهداتی و محاسباتی تبخیرتقرق مربوط به داده‌های آزمون مدل است. بر اساس این شکل نیز می‌توان گفت مدل توانسته است روند تغییرات  $ET_0$  را به‌طور مناسبی پیش‌بینی کند. شکل (۴-ب) نشان‌دهنده مقادیر مشاهداتی در مقابل مقادیر محاسباتی آزمون مدل است که بر اساس آن مدل توانسته تخمین قابل قبولی از مقادیر محاسباتی نسبت به مقادیر مشاهداتی خصوصاً در مقادیر حداقل و حداکثر داشته باشد.

شکل (۳-ا) مقادیر مشاهداتی و محاسباتی تبخیرتقرق مرحله آموزش مدل را در طول دوره آماری نشان می‌دهد. همان‌طور که مشاهده می‌گردد مدل توانسته است روند تغییرات  $ET_0$  را به‌طور مناسبی پیش‌بینی نماید. 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 به‌طوری‌که در اکثر موارد، مقادیر محاسباتی و پیش‌بینی شده 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 می‌گردد که مدل در برآورد مقادیر حدی شامل حداقل و حداکثر تقریباً عملکرد مناسبی از خود نشان داده به‌طوری‌که مقادیر تخمینی نزدیک به مقادیر واقعی هستند. هرچند در برخی نقاط اختلافاتی دیده می‌شود که



**Fig. 4- (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**

شکل ۴- (a): نمودار مقادیر محاسباتی و پیش‌بینی شده نسبت به زمان، (b): نمودار پراکنش میان مقادیر محاسباتی و پیش‌بینی شده GEP در مرحله آزمون

همبستگی بین پارامترهای ورودی و خروجی تصمیم‌گیری به عمل آمد. بر اساس رگرسیون گام‌به‌گام ترتیب اثرگذاری مشخصه‌های اقلیمی بر تبخیرتعرق مرجع به صورت: حداکثر درجه حرارت، سرعت باد، ساعات آفتابی، درجه حرارت حداقل، رطوبت نسبی حداکثر و رطوبت نسبی حداقل مشخص گردید. در مرحله بعدی شش الگوی ورودی برای مدل‌سازی تعیین و به همراه دو نوع عملگر ریاضی، یکی به صورت چهار عملگر اصلی و یکی به صورت عملگرهای پیش‌فرض با روش برنامه‌ریزی بیان ژن اجرا شدند. برنامه‌ریزی بیان ژن به دلیل دارا بودن ساختار ریاضی غیرخطی این قابلیت را دارد که فرایندهای غیرخطی پیچیده‌ای را که بین ورودی و خروجی هر سیستمی برقرار باشد را توصیف کند. با توجه به ارتباط پیچیده پارامترهای مؤثر بر تبخیرتعرق مرجع، کارایی برنامه‌ریزی بیان ژن

فرم ریاضی مدل استخراج‌شده با توجه به تابع پیوند جمع بین ژن‌ها، برای پیش‌بینی تبخیرتعرق مرجع در ایستگاه مطالعاتی، با قرار دادن مقادیر ثابت و پارامترهایی که در ساختار بیان درختی مدل GEP به کاررفته، تعیین شد. رابطه (۱۰)، نشان‌دهنده رابطه استخراج‌شده از GEP برای پیش‌بینی تبخیرتعرق مرجع است:

$$ET_0 = (WS / G1C5) + ((SUN / ((RHmin + G2C5) + SUN)) * SUN) + (Tmin / G3C0) \quad (10)$$

### نتیجه‌گیری

در این مطالعه از روش‌های رگرسیون چندمتغیره و برنامه‌ریزی بیان ژن برای برآورد تبخیرتعرق گیاه مرجع در شرایط آب‌وهوایی خرم‌آباد استفاده شد، برای تعیین الگوی ورودی نیز بر اساس ضرایب

مورد استفاده مدیران و طراحان قرار گیرند. بنابراین با روابط ارایه شده در این تحقیق، با اطلاعات ورودی کمتر و دقتی نزدیک به روش فائوینمن مانتیث، می توان تبخیرتغرق مرجع را در شرایط آب و هوایی خرم آباد محاسبه نمود. همچنین روابط به دست آمده با ورودی های کمتر نتایج رضایت بخشی ارایه نموده اند که می توانند در مناطقی که با مشکل نقص آماری مواجه هستند، مورد استفاده قرار گیرند. در پایان پیشنهاد می شود، از روش های جدیدتر مانند تئوری بازی ها در تعیین متغیرهای مؤثر و الگوهای ورودی مدل سازی و از روش های نوین هوشمند برای مدل سازی تبخیرتغرق مرجع استفاده و با مدل های قدیمی تر مقایسه شود.

### سپاسگزاری

در پایان نویسندگان از همکاری سازمان هواشناسی تشکر و قدردانی می نمایند.

مورد بررسی و ارزیابی قرار گرفت. بر پایه نتایج به دست آمده می توان بیان نمود که روش برنامه ریزی بیان ژن دارای توانایی قابل قبولی در تخمین تبخیرتغرق مرجع تحت شرایط آب و هوایی خرم آباد بوده و می توان با اطمینان بالا از آن در برآورد تبخیرتغرق مرجع بهره برده و به عنوان مدل قابل قبول در این زمینه معرفی کرد. نتایج به دست آمده از اجرای مدل با ساختارهای ورودی مختلف نشان داد که ساختارهای ترکیبی نسبت به ساختارهای دیگر از عملکرد بهتری برخوردار است. همچنین ضرایب تعیین به دست آمده برای سناریوهای مختلف نشان دهنده انتخاب مناسب پارامترهای ورودی و تأثیرگذاری آن ها در مدل سازی است. همچنین مدل های مبتنی بر برنامه ریزی بیان ژن راه حل های صریحی را ارائه می کنند که می توان رابطه بین متغیرهای ورودی و خروجی را معین نمود. لذا این نتایج می توانند با توجه به اهمیتی که تبخیرتغرق مرجع در مطالعات و طرح هایی مانند مطالعات کشاورزی، طرح های مدیریت منابع آب، تعیین نیاز آبی و تعیین هیدرومدول برای طراحی شبکه های آبیاری و زهکشی ایفا می کند.

### References

- 1- Ahmadi, F., Aisham, S., Khalili, K. And Bahman, c., 2015. Application of adaptive neuro-fuzzy inference system and genetic programming to estimate monthly transpiration evaporation northwest of iran. *Journal of Water Research in Agriculture*, 29 (2), pp. 235-247. (In Persian).
- 2- 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), pp. 300.
- 3- Almorox, J., Quej, V.H. and Marti, P., 2015. Global performance ranking of temperature-based approaches for evapotranspiration estimation considering Koppen climate classes. *Journal of Hydrology*, 528, pp. 514-522.
- 4- Ferreira C., 2001. Gene expression programming: a new adaptive algorithm for solving problems, *Complex Systems*, 13 (2), pp. 87-129.
- 5- Feyzollahpour, F., Delaware, M. and Hesami Afshar, M., 2017. Evaluation and analysis of uncertainty estimation of reference evapotranspiration using reference genetics. *Journal of Soil and Water Science*, 27 (4), pp. 135-147. (In Persian).
- 6- 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*, 26 (3), pp. 79-100. (In Persian).
- 7- Hargreaves, G. H., 1994. Defining and using reference evapotranspiration. *Irrigation and Drainage Engineering*, 120(6), pp. 1132-1139.
- 8- Hosseini, S., Ganji Khoramdel, N. and Khalat Abadi Farahani, 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*, 1 (2), pp. 29-40. (In Persian).
- 9- 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.
- 10- 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.

- 11- Ladlani, I., Hauichi, L., Dhemili, L., Heddem, S. and Blouze, K.H., 2012. 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*, 39, pp. 5959-5969.
- 12- Lopes, H.S. and Weinert, W.R., 2004. EGYPSIS: An enhanced gene expression programming approach for symbolic regression problems. *Journal of Applied Mathematics and Computer Science*, 14 (3), pp. 375-384.
- 13- Marti, P., Gonzalez-Altozano, P., Lopez-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.
- 14- Mattar, M. A., 2018. Using gene expression programming in monthly reference evapotranspiration modeling: a case study in Egypt. *Agricultural Water Management*, 198, pp. 28-38.
- 15- MohammadRezapour, A., Amini, A. and Karandish, F., 2015. Modeling monthly potential evapotranspiration using genetic programming in sistán baluchestan province. *Journal of Soil and Water Conservation Research*. 22 (5), pp. 307-313. (In Persian).
- 16- 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.
- 17- 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*, 9, pp. 11-20.
- 18- 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*, 41 (1), pp. 211-220. (In Persian).
- 19- Seifi, A., Mir Latifi, S. M. and Riahi, H., 2010. Development of multiple regression-principal component and factor analysis (mlr-pca) hybrid model in predicting reference evapotranspiration (Case Study: Kerman Station). *Watershed Magazine*, 24 (6), pp. 1186-1196. (In Persian).
- 20- 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.
- 21- 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.
- 22- 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, pp. 78-89.
- 23- 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), pp. 3195-3209.
- 24- Yassin, M.A., Alazba, A.A. and Mattar, M.A., 2016. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agricultural Water Management*, 163, pp. 110-124.