

## Research Paper

# Modeling and Estimating Daily Reference Evapotranspiration Using Soft Computing Models (Case Study: Aligudarz Station)

Yaser Sabzevari <sup>1</sup>,Abbas Parsaie <sup>2\*</sup>,AmirHamzeh Haghiabi <sup>3</sup>

<sup>1</sup> PhD student, water department, Faculty of Agriculture, Isfahan University of Technology, Isfahan, Iran, y.sabzevari@ag.iut.ac.ir

<sup>2</sup> Assistant Professor, Department of Hydraulic Structures, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran, Parsaie@scu.ac.ir

<sup>3</sup> Professor, water department, Faculty of Agriculture, Lorestan university, Khoramabad, Iran, corresponding author email: Haghiabi.a@lu.ac.ir



10.22125/IWE.2023.173301

Received:  
**December 6, 2021**  
Accepted:  
**March 7, 2022**  
Available online:  
**June 25, 2023**

Keywords:  
arm Water Management,  
Gene Expression  
Programming, Group  
Method of Data Handling,  
Multivariate Adaptive  
Regression Splines, Step-  
By-Step Regression

## Abstract

Calculation and estimation of evapotranspiration is one of the most important parameters of water management in agricultural engineering projects. The aim of this study was to evaluate the models of gene expression programming (GEP), group method of data handling (GMDH), and Multivariate adaptive regression spline (MARS) to estimating daily reference evapotranspiration. For this purpose, daily data recorded during the last 25-year period (1993-2017) of Aligudarz region (located in the east of Lorestan province) were used. 80% of the data were used for training and the remaining 20% for testing the models. The modeling results showed that only with the maximum temperature and average wind speed can evapotranspiration be estimated with very good accuracy. The error indices of GEP model in testing stage are  $R^2=0.93$  and  $RMSE=0.84$ , the error indices of MARS and GMDH models are  $R^2=0.94$  and  $RMSE=0.76$ . Comparing the performance of the models showed that the March model performed better than the other models.

## 1. Introduction

Optimal management of water resources requires accurate determination of water balance components in each region. Evapotranspiration, which includes evaporation from the surface of water or wet soil and transpiration from the surface of vegetation is one of the most important and effective components of water balance in any region. Therefore, estimating the rate of evapotranspiration is an essential need for each region. To the reference evapotranspiration should be estimate by a reliable method. Worldwide, the FAO-Penman-Monteith model is considered as the basic model for calculating the reference evapotranspiration (Allen et al, 1998). This method requires a number of climatic variables including: minimum and maximum air temperature, solar radiation, relative humidity and sunny hours. Providing models and relationships to calculate and predict reference evapotranspiration in areas such as Iran, which is facing water shortage, is necessary to determine the pattern of cultivation and optimal management of soil and water resources. There

\* **Corresponding Author:** AmirHamzeh Haghiabi

**Address:** Water Department, Lorestan University, Iran,

**Email:** Haghiabi.a@lu.ac.ir

**Tel:** 09161610580

are various models for this purpose, from physical models to data driven models. So that the second category is widely expanded due to the need for minimal information and a lot of scalabilities. Artificial intelligence models are among these models. In various studies, these models have been used to model and predict evapotranspiration. Ledlani et al. (2012) modeled daily evapotranspiration in the Algerian Mediterranean region using fuzzy system methods and linear multivariate regression. The obtained results showed the acceptable performance of both models in predicting evapotranspiration.

## 2. Materials and Methods

The aim of the present study is to apply and compare three models of artificial intelligence: GEP, GMDH and MARS in estimating the reference evapotranspiration of Aligudarz station. Aligudarz is located between 49° and 42° east longitude and 33° and 24° north latitude with an area of 130 hectares in the east of Lorestan province. This region has an altitude of 2022 m above sea level and an average rainfall of 387.7 mm and an annual evaporation of 2048.2 mm. According to the De Martonne and Embereger climatic classification indices, it has a semi-arid and arid climate, with mild summers and very cold winters. In this study, to calculate the daily evapotranspiration of the reference plant in Aligudarz climatic conditions with FAO-Penman-Monteith model, the data of maximum temperature ( $T_{max}$ ), minimum temperature ( $T_{min}$ ), maximum relative humidity (RH $_{max}$ ), minimum humidity Relative (RH $_{min}$ ), sunshine (Sun) and wind speed (WS) of the Aligudarz synoptic station for a period of 25 years (1993-1997) were used.

## 3. Results

To investigate the effect of model input parameters including:  $T_{max}$ ,  $T_{min}$ , RH $_{max}$ , RH $_{min}$ , Sun and WS in estimating reference evapotranspiration and determining the input pattern of the models, stepwise regression analysis between these parameters was performed. Correlation of input parameters showed that maximum temperature has the greatest effect on reference evapotranspiration and is a stronger predictor than other variables. After that, WS Sun,  $T_{min}$ , RH $_{min}$  and RH $_{max}$  have the considerable effect in predicting and estimating evapotranspiration, respectively. The modeling results showed that only with the maximum temperature and average wind speed can evapotranspiration in Aligudarz can be estimated with appropriate accuracy. With these two variables and using the GEP model with the accuracy of and (mm / day) RMSE = 0.84, the MARS with the accuracy of and RMSE = 0.63(mm / day) and the GMDH with accuracy and RMSE = 0.63 mm / day can estimate the daily reference evapotranspiration in Aligudarz area.

## 4. Discussion and Conclusion

Comparison of the performance of the models showed that the MARS model with a Nash coefficient of 0.963 in the training phase and 0.962 in the test phase had a better performance than other models.

## 5. Six important references

- 1) Berti, A., Tardivo, G., Chiaudani, A., Rech, F. and Borin, M. 2014. Assessing reference Evapotranspiration by the Hargreaves method in north-eastern Italy. *Agric. Water Manage.* 140: 20-25.
- 2) Djaman, K.B., Balde, A., Sow, A., Muller, B., Irmak, S.K., N'Diaye, M., Manneh, B.D., Moukoumbi, Y., Futakuchi, K. and Saito, K. 2015. Evaluation of sixteen reference Evapotranspiration methods under sahelian conditions in the Senegal River Valley. *J. Hydrol. Region. Stud.* 3: 139-159.
- 3) Ferreira, C. 2001. Gene Expression Programming: A New Adaptive Algorithm for Solving Problems, *Complex Systems*, 13 (2), 87-129.
- 4) Kisi, O. 2009. Modeling monthly evaporation using two different neural computing techniques. *Irrigation Science*, 27:417-430.
- 5) 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: 5959-5969.
- 6) 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), D05109.



## مدل سازی و تخمین تبخیر و تعرق روزانه گیاه مرجع با مدل های محاسبات نرم (مطالعه موردی: ایستگاه الیگودرز)

یاسر سبزواری<sup>۱</sup>، عباس پارسائی<sup>۲</sup>، امیر حمزه حقی آبی<sup>۳\*</sup>

تاریخ ارسال: ۱۴۰۰/۰۵/۱۹

تاریخ پذیرش: ۱۴۰۰/۱۲/۱۶

مقاله پژوهشی

### چکیده

محاسبه و تخمین تبخیر و تعرق یکی از مهمترین پارامترهای مدیریت آب در پروژه های مهندسی کشاورزی است. هدف از این مطالعه ارزیابی مدل های برنامه ریزی بیان ژن، دسته بندی گروهی داده ها و اسپلین تطبیقی چندگانه در برآورد تبخیر-تعرق مرجع ایستگاه الیگودرز است. بدین منظور از اطلاعات روزانه مربوط به دوره زمانی ۲۵ ساله (۱۹۹۳-۲۰۱۷) منطقه الیگودرز واقع در شرق استان لرستان استفاده شد. ۸۰ درصد داده ها برای آموزش و ۲۰ درصد مابقی آنها برای آزمون مدل ها در نظر گرفته شد. براساس رگرسیون گام به گام درجه حرارت حداکثر دارای بیشترین تأثیر بر تبخیر-تعرق مرجع است. نتایج مدل سازی نشان داد تنها با داشتن درجه حرارت ماکزیمم و سرعت متوسط باد می توان تبخیر-تعرق را در الیگودرز برآورد کرد. با این دو متغیر و استفاده از مدل برنامه ریزی بیان ژن با دقت  $R^2 = 0/93$  و  $RMSE = 0/84$  (mm/day) با مدل اسپلین تطبیقی چندگانه هم با دقت  $0/96 = R^2$  و  $RMSE = 0/63$  (mm/day) و با مدل دسته بندی گروهی داده ها با دقت  $0/96 = R^2$  و  $RMSE = 0/63$  (mm/day) تبخیر و تعرق مرجع را برآورد نمود. مقایسه عملکرد مدل ها نشان داد که مدل اسپلین تطبیقی چندگانه با مقدار ضریب نش به میزان  $0/963$  در مرحله آموزش و  $0/962$  در مرحله آزمون، عملکرد بهتری نسبت به سایر مدل ها داشته است.

واژه های کلیدی: اسپلین تطبیقی چندگانه، برنامه ریزی بیان ژن، دسته بندی گروهی داده ها، رگرسیون گام به گام، مدیریت آبیاری.

<sup>۱</sup> دانشجوی دکتری، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران، ۰۹۳۵۷۳۴۱۵۱۱، y.sabzevari@ag.iut.ac.ir

<sup>۲</sup> استادیار، گروه سازه های آبی، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز ایران ۰۹۱۶۳۶۸۵۸۶۷، Parsaie@scu.ac.ir

<sup>۳</sup> استاد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه لرستان، ایران، ۰۹۱۶۱۶۱۰۵۸۰، Haghiaibi.a@lu.ac.ir (نویسنده مسئول)

## مقدمه

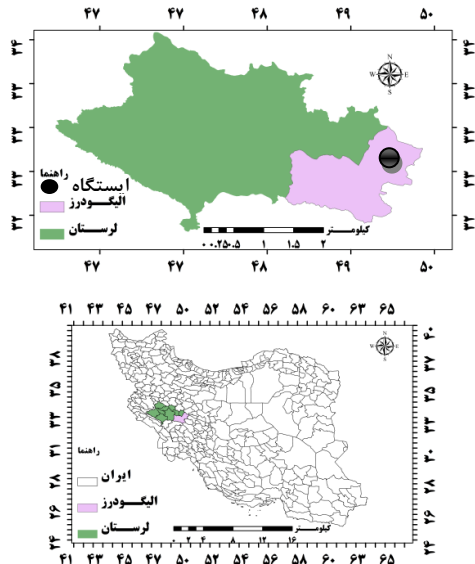
لازمه مدیریت بهینه منابع آب تعیین دقیق مؤلفه‌های بیلان آب در هر منطقه است. تبخیر - تعرق که شامل تبخیر از سطح آب یا خاک مرطوب و تعرق از سطح پوشش گیاهی یکی از مهم‌ترین و تأثیرگذارترین اجزای بیلان آب در هر منطقه است. علاوه بر این برآورد میزان دقیق تبخیر - تعرق از سطح پوشش گیاهی بخشی از اطلاعات لازم برای تعیین نیاز آبی و هیدرومدول آبیاری به منظور طراحی شبکه‌های آبیاری و زهکشی است (آزادلاتپه و همکاران، ۱۳۹۴). از این رو برآورد میزان تبخیر - تعرق یک نیاز ضروری برای هر منطقه است. برای برآورد تبخیر - تعرق از سطح هر پوشش گیاهی در ابتدا باید تبخیر - تعرق گیاه مرجع با یک روش قابل اطمینان محاسبه شود. در سراسر جهان مدل فائو-پنمن-مانتیت به عنوان مدل مرجع برای محاسبه تبخیر - تعرق گیاه مرجع در نظر گرفته می‌شود (Niaghi et al., ۲۰۱۳; Liu et al., ۲۰۱۷; Tabari et al., ۲۰۱۳; Trajkovic, 2007). تحقیقات زیادی در سراسر دنیا صحت نتایج مربوط به مدل فائو-پنمن-مانتیت در شرایط مختلف آب و هوایی را نشان می‌دهد (Berti et al., 2014; Djaman et al., ۲۰۱۵). با این حال، این روش به تعدادی از متغیرهای اقلیمی شامل: حداقل و حداکثر دمای هوا، تابش خورشیدی، رطوبت نسبی و ساعات آفتابی نیاز دارد. ارائه مدل‌ها و روابطی به منظور محاسبه و پیش‌بینی تبخیر - تعرق مرجع در مناطقی مانند ایران که با مشکل کمبود منابع آبی مواجه است، برای تعیین الگوی کشت و مدیریت بهینه منابع آب و خاک امری ضروری به نظر می‌رسد. مدل‌های مختلفی اعم از مدل‌های فیزیکی تا مدل‌های تحلیل داده‌ای برای این منظور وجود دارند. به طوری که دسته دوم به دلیل نیاز به حداقل اطلاعات و قابلیت توسعه‌ی فراوان به طور وسیعی در حال گسترش هستند. مدل‌های هوش مصنوعی از جمله این مدل‌ها هستند. در مطالعات مختلفی از این مدل‌ها برای مدل‌سازی و پیش‌بینی تبخیر-تعرق استفاده شده است. کولائیان و همکاران (۱۳۹۲) به ارزیابی روش‌های نوین محاسباتی در برآورد تبخیر و تعرق روزانه در دشت ناز ساری پرداختند. نتایج نشان داد با ۴ پارامتر ورودی، مدل‌های

شبکه عصبی MLP و رگرسیون چند متغیر خطی از دقت بالایی برخوردارند. سیفی و همکاران (۱۳۸۹) به توسعه مدل ترکیبی رگرسیون چند متغیره-تحلیل مؤلفه‌ها و عامل‌های اصلی در پیش‌بینی تبخیر-تعرق مرجع در ایستگاه کرمان پرداختند. نتایج نشان داد اختلاف معنی‌داری بین دو روش رگرسیون چند متغیره-تحلیل مؤلفه‌ها و رگرسیون چند متغیره وجود ندارد. حسینی و همکاران (۱۳۹۴) به ارزیابی و آنالیز 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 چند متغیره دارد. فیض‌الله‌پور و همکاران (۱۳۹۶) به ارزیابی و تحلیل عدم قطعیت برآورد تبخیر-تعرق گیاه مرجع با استفاده از برنامه‌ریزی ژنتیک پرداختند. نتایج نشان داد که مدل‌های برنامه‌ریزی ژنتیک دارای دقت بالاتری نسبت به روش‌های تجربی هستند. لدلانی و همکاران (۲۰۱۲) تبخیر-تعرق روزانه در منطقه مدیترانه‌ای کشور الجزایر را توسط روش‌های سیستم فازی<sup>۱</sup> و رگرسیون چند متغیره خطی مدل‌سازی کردند. نتایج به دست آمده نشان از عملکرد قابل قبول هر دو مدل در پیش‌بینی تبخیر-تعرق داشت. ملک و همکاران (۲۰۱۳) و کیشی (۲۰۰۹) برای برآورد تبخیر از تشت از روش‌های شبکه عصبی مصنوعی و رگرسیون چند متغیره خطی استفاده کردند. نتایج این دو مطالعه بیانگر کارایی خوب مدل رگرسیون چند متغیره خطی بوده است. لیما و همکاران (۲۰۱۵) به ارزیابی الگوریتم یادگیری ماشین در محیط زیست به عنوان یک مدل رگرسیون غیرخطی پرداخت و این الگوریتم نتایج بهتری نسبت به مدل رگرسیون خطی چندگانه نشان داد. یاسین و همکاران (۲۰۱۶) به ارزیابی شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن برای شبیه‌سازی تبخیر تعرق مرجع در اقلیم خشک پرداختند. نتایج نشان داد که ۸ مدل برآورد تبخیر تعرق مرجع با استفاده از شبکه عصبی مصنوعی نسبت به همان مدل‌ها با برنامه‌ریزی بیان ژن دارای دقت بالاتری هستند. کرباسی (۱۳۹۵) به پیش‌بینی تبخیر-تعرق مرجع ایستگاه سینوپتیک اهواز با استفاده از مدل ترکیبی

<sup>۱</sup> - Fuzzy



یک ایستگاه سینوپتیک با عرض شمالی ۳۳ درجه و ۲۴ دقیقه و طول شرقی ۴۹ درجه و ۴۲ دقیقه 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

در این پژوهش جهت محاسبه روزانه تبخیر و تعرق گیاه مرجع در شرایط آب و هوایی الیگودرز با مدل فائو-پنمن-مانتیت، 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 ( $WS$ ) مربوط به ایستگاه سینوپتیک الیگودرز مربوط به دوره زمانی ۲۵ ساله (۱۹۹۳-۲۰۱۷) با کمک نرم افزار RefET-۳/۲ استفاده شد. رابطه فائو-پنمن-مانتیت توسط (Allen et al., 1998) به صورت رابطه زیر ارائه شد.

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

در این رابطه:  $ET_0$ : تبخیر-تعرق گیاه مرجع، برحسب میلی‌متر بر روز  $R_n$ : تشعشع خالص برحسب مگا ژول بر مترمربع بر روز  $G$ : شار گرمایی خاک، برحسب مگا ژول بر مترمربع بر روز  $\Delta$ : شیب نمودار فشار بخار اشباع نسبت به

موجک-شبکه عصبی GMDH پرداخت. نتایج نشان داد این مدل هیبریدی توانایی بالایی در پیش‌بینی تبخیر-تعرق مرجع روزانه دارد. کیشی و علیضمیر (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  $ET_0$  بود. سبزواری و همکاران (۱۳۹۹) به کاربرد رگرسیون چندمتغیره و برنامه‌ریزی بیان ژن در برآورد تبخیر و تعرق مرجع در خرم‌آباد پرداختند که نتایج نشان دهنده عملکرد بهتر برنامه‌ریزی بیان ژن در برآورد تبخیر و تعرق مرجع بود. همانطور که ملاحظه می‌شود، مطالعات زیادی در ایران و جهان با استفاده از روش‌های هوشمند برای برآورد تبخیر-تعرق مرجع با مدل‌هایی که به اطلاعات ورودی کمتر نیاز دارند، انجام گرفته که معمولاً از یک یا دو روش برای این منظور استفاده شده است. مرور مطالعات گذشته نشان دهنده کارایی و قابلیت روش‌های هوش مصنوعی در برآورد تبخیر-تعرق مرجع است. بر این اساس، هدف از پژوهش حاضر کاربرد و مقایسه سه مدل هوش مصنوعی برنامه‌ریزی بیان ژن، دسته‌بندی گروهی داده‌ها و اسپلاین تطبیقی چندگانه<sup>۱</sup> در برآورد تبخیر-تعرق مرجع ایستگاه الیگودرز می‌باشد.

## مواد و روش‌ها

### منطقه مورد مطالعه

شهرستان الیگودرز بین ۴۹ درجه و ۴۲ دقیقه طول شرقی و ۳۳ 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 متوسط بارش ۳۸۷/۷ میلی‌متر و تبخیر سالانه ۲۰۴۸/۲ میلی‌متر می‌باشد. طبق روش‌های طبقه‌بندی اقلیمی دمارتن و آمبرژه دارای اقلیم نیمه‌خشک و خشک سرد، با تابستان معتدل و زمستان بسیار سرد است. شکل ۱ نشان دهنده موقعیت جغرافیایی الیگودرز در لرستان و کشور می‌باشد. این شهرستان دارای

<sup>۱</sup> - Multivariate Adaptive Regression Splines (MARS)

مطالعه جهت انجام تحلیل رگرسیون گام به گام از نرم افزار SPSS-24 استفاده شد.

### برنامه ریزی بیان ژن

برنامه ریزی بیان ژن (GEP)، توسعه یافته روش های برنامه ریزی ژنتیک (GP) و الگوریتم ژنتیک (GA) است که در سال ۱۹۹۹ توسط Ferreira بر اساس تئوری Darwin ابداع شد. در روش برنامه ریزی بیان ژن، کروموزوم های خطی و ساده با طول ثابت، مشابه با الگوریتم ژنتیک و ساختارهای شاخه ای با اندازه و شکل های متفاوت، مشابه با درختان تجزیه در برنامه ریزی ژنتیک ترکیب می شوند. در این روش فنوتیپ و ژنوتیپ از هم جدا شده و سیستم قادر خواهد بود از تمام مزایای تکاملی بهره گیرد.

در این پژوهش برای اجرای مدل GEP از نرم افزار Genexprotocols 5 استفاده شد. علاوه بر چهار عملگر اصلی (F1)، {+,-,\*,/}، حالت هایی بر اساس عملگرهای ریاضی پیش فرض برنامه (F2)، {Sin, Cos, Atan, √, Exp, Ln, 3√, +, -, \*, /} نیز در نظر گرفته شد. در ساخت درخت تجزیه و تولید جمعیت اولیه، از راه حل هایی که از ترکیب تصادفی مجموعه توابع (عملگرهای ریاضی مورد استفاده در روابط) و ترمینال ها (متغیرهای مسئله و اعداد ثابت) ایجاد می شوند، استفاده گردید.

در ابتدای اجرای مدل GEP، ابتدا باید تعداد ژن ها و کروموزوم ها مشخص شود. با توجه به مطالعات انجام شده قبلی، بهترین جمعیت از ۳۰ تا ۵۰ کروموزوم و پاسخ مناسب از انتخاب ۲ یا ۳ ژن حاصل می شود (Ferreira, 2001). 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 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 وارون سازی همان مقادیر پیش فرض مدل و مقدار مجموع نرخ سه عملگر ترکیب تک نقطه ای، ترکیب دو نقطه ای و ترکیب ژن، طبق نظر (Ferreira, 2001) برابر ۰/۷ تعیین شد.

دما، برحسب کیلو پاسکال بر درجه سلسیوس ۷؛ ثابت سایکرومتری، برحسب کیلو پاسکال بر درجه سلسیوس e<sub>a</sub>-e<sub>d</sub>؛ کمبود فشار بخار در ارتفاع ۲ متری، برحسب کیلو پاسکال U<sub>2</sub>؛ میانگین روزانه سرعت باد در ارتفاع ۲ متری از سطح زمین، برحسب متر بر ثانیه T؛ میانگین روزانه دمای هوا در ارتفاع ۲ متری از سطح زمین، برحسب درجه سلسیوس.

در این پژوهش، جهت بررسی اثرگذاری متغیرهای مختلف و تعیین الگوهای ورودی مدل ها و مدل سازی تبخیر و تعرق مرجع، از روش رگرسیون گام به گام نیز استفاده شده است.

### مدل رگرسیون گام به گام

با این روش می توان به طور هم زمان به تحلیل و بررسی چند متغیر مختلف پرداخت. رگرسیون گام به گام در واقع، رابطه بین چند متغیر پیش بین با متغیر پاسخ مورد نظر را بیان می کند. این گونه مدل ها دارای فرضیاتی هستند که آن ها را از رگرسیون ساده متمایز می نماید:

- تعداد متغیرهای پیش بین (مستقل) در رگرسیون باید کمتر از تعداد مشاهدات باشد.
- 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:

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

که  $\beta$  ماتریس ضرایب رگرسیونی که ضرایب نامشخصی هستند که در حقیقت، مسؤل برآورد پارامتر پاسخ هستند.  $e$  ماتریس 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 نمی توان معکوس کرد و باعث افزایش خطا در اثر گرد کردن داده ها و محاسبات می شود. در این



$Y = w_0 + w_1x_1 + w_2x_2 + w_3x_1^2 + w_4x_2^2 + w_5x_1x_2$  (۵)

ضرایب ۶ گانه هر نرون در شبکه از طریق رویکرد حداقل مربعات محاسبه می‌شوند که مراحل انجام این کار به صورت زیر است. فرض کنید  $N$  بردار ورودی  $X_n = (x_{n1}, x_{n2}, \dots, x_{np})$  که  $n = 1, \dots, N$  در مجموعه آموزشی وجود داشته که هر کدام از  $P$  مقدار صحیح تشکیل شده باشند. مقدار مطلوب خروجی  $n$  ام با  $\phi_n$  نشان داده می‌شود. می‌بایست مجموعه‌ای از ضرایب ۶ گانه برای هر نرون پیدا شود طوری که میانگین مربعات خطا بین خروجی‌های این نرون  $y_n$  و مقدار واقعی  $\phi_n$  حداقل باشد. با استفاده از معادلات نرمال گوسین، ضرایب به صورت زیر به دست می‌آیند:

$$\begin{aligned}\phi_1 &= w_0 + w_1x_{1i} + w_2x_{1j} + w_3x_{1i}^2 + w_4x_{1j}^2 + w_5x_{1i}x_{1j} \\ \phi_2 &= w_0 + w_1x_{2i} + w_2x_{2j} + w_3x_{2i}^2 + w_4x_{2j}^2 + w_5x_{2i}x_{2j}\end{aligned}$$
 (۶)

$$\phi_N = w_0 + w_1x_{Ni} + w_2x_{Nj} + w_3x_{Ni}^2 + w_4x_{Nj}^2 + w_5x_{Ni}x_{Nj}$$

که به صورت ماتریس کلی زیر نوشته می‌شود:

$$\phi = XW$$
 (۷)

ماتریس‌های  $X$ ،  $\phi$  و  $W$  دارای ابعاد  $N \times 1$ ،  $N \times 6$  و  $6 \times 1$  هستند. معادلات نرمال با استفاده از ضرب طرفین رابطه بالا در ترانپوز ماتریس  $X$  به دست می‌آیند:

$$X^T \phi = (X^T X)W \Rightarrow W = (X^T X)^{-1} X^T \phi$$
 (۸)

$X^T X$  یک ماتریس  $6 \times 6$  بوده و می‌توان ضرایب را با استفاده از روش معکوس به دست آورد. ماتریس  $W$  شامل مجموعه‌ای از ۶ ضریب می‌باشد که قادر به تقریب خروجی‌های صحیح با حداقل میانگین مربعات خطا می‌باشد.

### روش الگوریتم دسته‌بندی گروهی داده‌ها (GMDH<sup>1</sup>)

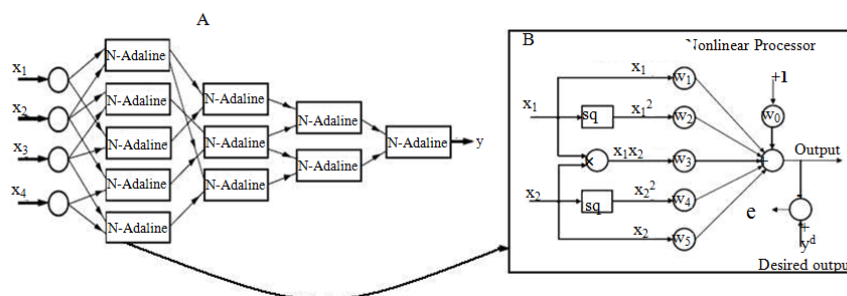
رزنبلات<sup>۲</sup> در سال ۱۹۵۸ مفهوم پرسپترون را بر اساس عملکرد نرون‌ها ارائه کرد. GMDH ترکیبی از نرون‌ها می‌باشد و نسخه‌های اصلاح‌شده‌ای از این روش برای کاربردهای مختلف مدل‌سازی مورد استفاده قرار گرفته است. در شکل (۲) دیاگرام شماتیک شبکه GMDH همراه با نمایی از ساختار نرون با تابع محرک چندجمله‌ای درجه‌دو نشان داده شده است. در این شکل  $X$  و  $sq$  به ترتیب بیانگر مجذور و حاصل ضرب،  $x_1$  و  $x_2$  ورودی‌ها و  $y$  خروجی می‌باشد. معیار خارجی برای تعیین ساختار شبکه به صورت زیر تعریف می‌شود:

$$R^2 = \frac{\sum_{t=1}^N (\hat{y}(t) - \bar{\hat{y}})^2}{\sum_{t=1}^N (y(t) - \bar{y})^2}$$
 (۴)

که  $\hat{y}$  خروجی مطلوب،  $\bar{\hat{y}}$  خروجی محاسباتی، میانگین خروجی‌های محاسباتی و  $\bar{y}$  میانگین خروجی‌های مطلوب است. ساختار شبکه از طریق الگوریتم ارائه شده تعیین می‌شود: انتخاب متغیرهای ورودی به مدل، جمع‌آوری مجموعه‌ای از داده‌های مشاهده‌ای مرتبط با موضوع مورد بررسی، تقسیم‌بندی داده‌ها به دو گروه داده‌های واسنجی و اعتبارسنجی، ساخت  $N$ -نرون برای تمام ترکیبات ۲ تایی از متغیرهای ورودی، انتخاب نوع تابع محرک برای هر نرون، تخمین ضرایب وزنی برای تمام  $N$ -نرون‌ها با روش تخمین حداقل مربعات و محاسبه معیار فیلتراسیون، مقایسه معیار محاسبه شده برای تمام  $N$ -نرون‌ها با مقدار آستانه. هر نرون در ساختار GMDH یک تابع غیرخطی از ورودی‌ها را اجرا می‌کند. این تابع غیرخطی به صورت زیر می‌باشد:

<sup>2</sup> - Rosenblatt

<sup>1</sup> Group Method of Data Handling



شکل (۱): (A) شماتیکی از ساختار N-adaline در GMDH و (B) ساختار هر adaline

افزوده می‌شوند. این عمل مدلی را تولید می‌کند که پیچیده بوده و موجب بیش برآزش می‌گردد. در مرحله دوم یا مرحله حذف پسرو، توابع مبنائی که از اهمیت و تأثیر گذاری کمتری در تخمین برخوردار هستند، حذف می‌شوند. سرانجام بهترین مدل بر مبنای کمینه بودن معیاری به نام «اعتبارسنجی تقاطعی تعمیم یافته (GCV) انتخاب می‌گردد. فرض کنید GCVk مقدار GCV برای k امین مدل در مرحله حذفی باشد. این کمیت توسط عبارت زیر تعریف می‌شود:

$$GCV_k = \frac{1}{n} \sum_{i=1}^n \left( y_i - \hat{f}_k(x_i) \right)^2 / \left( 1 - C(k)/n \right) \quad (11)$$

که  $\hat{f}_k$  مدل برآورد شده در گام k ام مرحله حذفی پسرو و (تعداد جملات مدل در مرحله k ام +  $\lambda \cdot m = C_k$ ) که m معرف تعداد گره‌های توابع اسپلاین موجود در مدل و در عمل معمولاً بین ۲ تا ۴ انتخاب می‌شود.

#### معیارهای ارزیابی:

به منظور ارزیابی دقت و کارایی مدل‌ها، از معیارهای ضریب تبیین ( $R^2$ )، ریشه میانگین مربعات خطا (RMSE) و ضریب نش (NS) به صورت زیر استفاده گردید. بهترین و بدترین مقادیر برای این سه معیار به ترتیب: یک، صفر و یک است.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2} \quad (12)$$

#### مدل اسپلاین تطبیقی چندگانه (MARS)

نخستین بار توسط فردمن در سال ۱۹۹۱ به منظور پیش‌بینی خروجی‌های عددی پیوسته ارائه شده است. مدل اسپلاین تطبیقی چندگانه یکی از مدل‌های ناپارامتریک محلی است. این مدل از تمام داده‌های مربوطه بصورت یکجا استفاده نمی‌کند بلکه داده‌ها را به زیر مجموعه‌هایی تقسیم کرده و سپس عمل مدل‌سازی را برای هر کدام از این زیر مجموعه‌ها انجام می‌دهد که اصطلاحاً به آن مدل محلی اطلاق می‌شود. اساس این روش مبتنی بر توابعی با نام توابع مبنا است که برای هر متغیر توضیحی به صورت زیر تعریف می‌شود:

$$\max \{0, x - t\} \quad \text{and} \quad \max \{0, t - x\} \quad (9)$$

که در آن t گره نامیده شده و در عمل یکی از مشاهدات آن متغیر توضیحی است. این توابع موسوم به توابع اسپلاین هستند که در گره t زوج منعکس یافته‌ای هستند. شکل کلی مدل مارس بصورت زیر تعریف می‌شود:

$$\hat{Y} = C_0 + \sum_{k=1}^M C_k B_k(X) \quad (10)$$

که در این معادله  $\hat{Y}$  مقدار تخمین متغیر مورد پاسخ، X بردار متغیرهای توضیحی،  $B_k$  تابع مبنا و  $C_k$  ضرایبی هستند که با کمینه‌سازی مجموع مربعات باقیمانده‌ها تعیین می‌شوند. هر تابع مبنا ممکن است به صورت یک تابع خطی اسپلاین یا حاصلضرب دو یا چند تا از آنها باشد که مبین اثرات متقابل است (فردمن، ۱۹۹۱). برازش مدل MARS در دو مرحله انجام می‌گیرد. در مرحله پیشرو، تعداد زیادی از توابع مبنا با گره‌های متفاوت بطور پی در پی به مدل





نشان دهنده این موضوع است که پارامترهای ورودی همگی بر متغیر تبخیر-تعرق اثرگذار بوده و در پیش‌بینی مورد استفاده قرار می‌گیرند. مقایسه ضریب استاندارد شده بدست آمده برای پارامترهای ورودی نشان می‌دهد که بیشینه درجه حرارت با ضریب ۰/۵۸۰ دارای بیشترین تأثیر بر تبخیر-تعرق مرجع است و پیش‌بینی کننده قوی‌تری نسبت به بقیه متغیرهاست. بعد از آن به ترتیب سرعت باد، ساعات آفتابی، درجه حرارت حداقل، رطوبت نسبی حداقل و رطوبت نسبی حداکثر دارای بیشترین تأثیر در پیش‌بینی و برآورد تبخیر-تعرق می‌باشند. حسینی و همکاران (۱۳۹۴) نتیجه گرفتند که بیشینه درجه حرارت و سرعت باد بیشترین تأثیر را بر تبخیر-تعرق مرجع داشته است که با نتایج پژوهش حاضر مطابقت دارد. مقدار  $t$  (چه مثبت و چه منفی) اگر بزرگ‌تر از ۱/۹۶ باشد و مقدار سطح معنی‌داری اگر کوچک‌تر از ۰/۰۱ 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.

جدول (۱): مقدار ضرایب دقت مدل و پارامترهای رگرسیون گام به گام

ضرایب استاندارد نشده	ضرایب استاندارد شده	t	سطح معنی‌داری	
B	خطای معیار	Beta		
-۰/۶۰۶	۰/۰۸۷		-۶/۹۴۶	• (Constant)
۰/۰۵۱	۰/۰۰۴	۰/۱۳۸	۱۴/۱۷۹	• Tmin
۰/۱۷۱	۰/۰۰۴	۰/۵۸	۴۸/۰۵۹	• Tmax
۰/۰۰۹	۰/۰۰۱	۰/۴۶	۸/۴۶۲	• Rhmin
-۰/۰۲۲	۰/۰۰۱	-۰/۱۶	-۳۰/۳۷۹	• Rhmax
۰/۴۲۷	۰/۰۰۵	۰/۲۳	۸۰/۳۲۶	• WS
۰/۱۶	۰/۰۰۴	۰/۱۶۹	۴۰/۷۲۸	• Sun

$$R^2 = \frac{\left( \sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y}) \right)^2}{\left( \sum_{i=1}^N (X_i - \bar{X})^2 \sum_{i=1}^N (Y_i - \bar{Y})^2 \right)} \quad \text{و} \quad 0 \leq R^2 \leq 1 \quad (13)$$

$$NS = 1 - \frac{\sum_{i=1}^N (X_i - Y_i)^2}{\sum_{i=1}^N (X_i - \bar{Y}_i)^2} \quad \text{و} \quad -\infty \leq NS \leq 1 \quad (14)$$

## نتایج و بحث

### رگرسیون گام به گام

جهت 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 مدل‌ها، تحلیل رگرسیونی گام به گام بین این پارامترها صورت گرفت که نتایج در جدول ۱ نشان داده شده است: یکی از بخش‌های مهم جدول (۱) مقادیر بتا و سطح معنی‌داری است. با مشاهده‌ی سطح معنی‌داری حاصل شده می‌توان گفت همه پارامترها در سطح یک درصد معنی‌دار بوده و

الگوی ورودی مدل‌ها برای پیش‌بینی تبخیر و تعرق مرجع در جدول ۲ نشان داده شده است. بر اساس این جدول، مدل‌سازی تبخیر-تعرق مرجع با شش الگو به صورت جدول زیر انجام شد.

جدول (۲): الگوهای ورودی مدل‌های MARS و GMDH، GEP

الگوهای ورودی	کد الگو	شماره
$ET_0=f\{T_{max}\}$	M1	۱
$ET_0=f\{T_{max}, WS\}$	M2	۲
$ET_0=f\{T_{max}, WS, SUN\}$	M3	۳
$ET_0=f\{T_{max}, WS, SUN, T_{min}\}$	M4	۴
$ET_0=f\{T_{max}, WS, SUN, T_{min}, Rh_{min}\}$	M5	۵
$ET_0=f\{T_{max}, WS, SUN, T_{min}, Rh_{min}, Rh_{max}\}$	M6	۶

پیش‌بینی شده بر روی خط نیمساز ( $X=Y$ ) یا در نزدیکی آن واقع شده‌اند که نشان دهنده نزدیک بودن مقادیر مشاهداتی و محاسباتی است. علاوه بر آن مشاهده می‌گردد که مدل در برآورد مقادیر حدی حداقل عملکرد مناسبی از خود نشان داده به طوری 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 می‌گردد که در صورت وجود یک رابطه غیرخطی بین ورودی و خروجی آن، دقت مدل کاهش پیدا نکند که در عملگر F2 حتی باعث بهبود عملکرد مدل هم شده است. سبزواری و همکاران (۱۳۹۹) نیز در پژوهش خود به این نتیجه دست یافتند که مدل GEP در برآورد مقادیر حدی حداقل دقت خوبی داشته ولی در برآورد مقادیر حدی حداکثر از دقت آن کاسته شده است که با نتایج پژوهش حاضر مطابقت دارد.

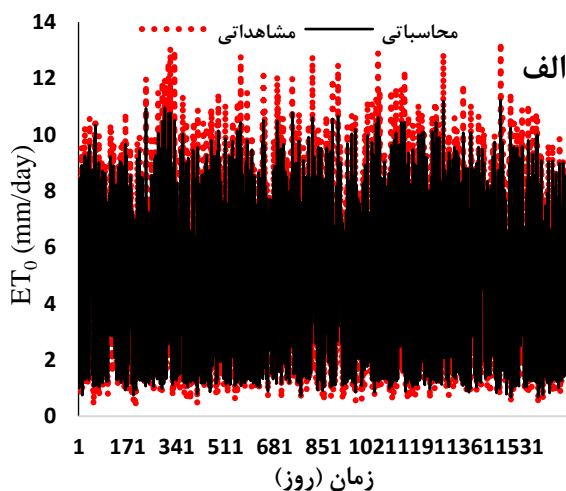
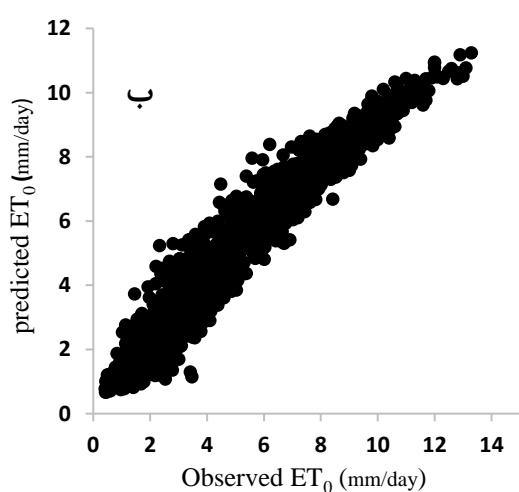
### مدل‌سازی تبخیر-تعرق مرجع به وسیله GEP

نتایج حاصل از اجرای مدل برنامه‌ریزی بیان ژن برای داده‌های آموزش و آزمون در شش الگوی ورودی به همراه دو نوع عملگر ریاضی، در جدول ۳ ارائه شده است. همانطور که مشاهده می‌شود، برنامه‌ریزی بیان ژن توانسته است در تمام سناریوها عملکرد و دقت مناسبی از خود ارائه دهد. در بین سناریوهای انجام شده، ساختار ترکیبی شماره ۲ با عملگر F2 در مرحله آزمون و ساختار ترکیبی شماره ۳ با عملگر F2 در مرحله آموزش در ایستگاه الیگودرز، دارای کم‌ترین خطا و بیشترین دقت بوده و توانسته‌اند میزان تبخیر-تعرق مرجع را با دقت مناسبی برآورد کنند. الگوی شماره ۲ متشکل از درجه حرارت حداکثر و سرعت باد، به دلیل کمتر بودن اطلاعات ورودی انتخاب شد. Yassin et al. (2016) در پژوهش خود الگوی شامل  $T_{min}$ ،  $T_{max}$ ،  $Rh_{min}$ ،  $Rh_{max}$  و Sun و WS با  $RMSE=1/14$  و  $R^2=0/95$  در مرحله آزمون و  $RMSE=1/12$  و  $R^2=0/95$  در مرحله آموزش را به عنوان مناسب‌ترین الگو انتخاب کردند.

شکل (۳- الف) مقادیر مشاهداتی و محاسباتی تبخیر و تعرق مرحله آزمون مدل را نشان می‌دهد. همانطور که مشاهده می‌گردد مدل توانسته است روند تغییرات  $ET_0$  را به طور 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 به طوری که در اکثر موارد، مقادیر محاسباتی و

جدول (۳): نتایج مدل GEP با استفاده از دو مجموعه عملگر ریاضی در مدل‌سازی تبخیر-تعرق مرجع

آزمون		آموزش		عملگر	الگو	
R <sup>2</sup>	RMSE(mm/day)	R <sup>2</sup>	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 مقادیر مشاهداتی و پیش‌بینی شده نسبت به زمان، (ب): نمودار پراکنش میان مقادیر مشاهداتی و پیش‌بینی شده در مرحله آزمون

$$ET_0 = ((-6.3 + (0.013)) / 2.0) + \text{atan}(\frac{\text{gepMax2}(\frac{\text{gepMax2}(\text{Tmax} + \text{WS})}{2.0}, \text{WS}) + ((\text{WS} + 6.7) / 2.0) / 2.0 - \text{WS}}{\text{gepMax2}(\frac{\text{gepMax2}(\text{WS}, \text{WS}) + ((\text{Tmax} + 0.2) / 2.0) / 2.0, \frac{(((-2.3 + 5.1) / 2.0) + ((8.4 + \text{WS}) / 2.0)) / 2.0}}{2.0}, \frac{(((-2.3 + 5.1) / 2.0) + ((8.4 + \text{WS}) / 2.0)) / 2.0}}{2.0}))} (15)$$

فرم ریاضی مدل استخراج شده با توجه به تابع پیوند جمع بین ژن‌ها، برای پیش‌بینی تبخیر-تعرق مرجع در ایستگاه مطالعاتی، با قرار دادن مقادیر ثابت و پارامترهایی که در ساختار بیان درختی مدل GEP به کار رفته، تعیین شد. رابطه ۱۵، نشان دهنده رابطه استخراج شده از GEP برای پیش-بینی تبخیر-تعرق مرجع می‌باشد:

نشده است. کرباسی (۱۳۹۵) در پژوهش خود به این نتیجه دست یافت که افزایش پارامترهای ورودی تأثیر ناچیزی بر روی دقت مدل GMDH داشته که با پژوهش حاضر مطابقت دارد، همچنین این مدل دارای  $RMSE=1/22$  در اهواز بوده است که دقت این مدل نسبت به پژوهش حاضر کمتر بوده است.

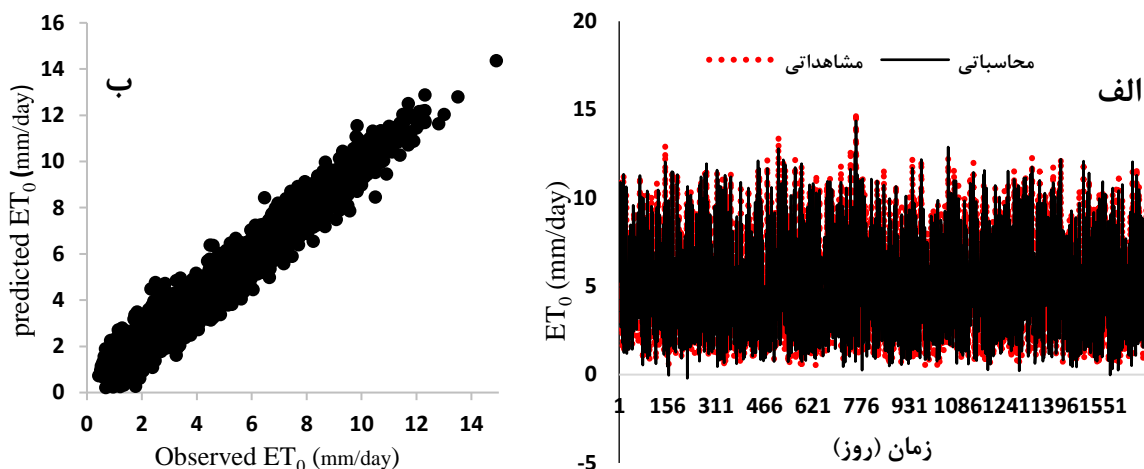
شکل ۴-الف نشان دهنده نمودار تغییرات زمانی تبخیر-تعرق مرجع محاسباتی و مشاهداتی در مرحله آزمون می-باشد. با توجه به این نمودار می‌توان گفت که مدل دسته-بندی گروهی داده‌ها روند تغییرات تبخیر-تعرق را به خوبی پیش‌بینی کرده است. همچنین مدل در برآورد مقادیر حدی حداقل و حداکثر عملکرد خوبی از خود نشان داده است. شکل ۴-ب نشان دهنده نمودار پراکنش مقادیر محاسباتی و مشاهداتی حول نیمساز ربع اول می‌باشد. با توجه به این نمودار نیز مدل، عملکرد مناسبی در برآورد تبخیر-تعرق مرجع روزانه در منطقه مطالعاتی داشته به‌طوری‌که در بیشتر موارد مقادیر محاسباتی و مشاهداتی منطبق بر خط نیمساز ربع اول یا در نزدیکی آن واقع شده‌اند.

### مدل‌سازی تبخیر-تعرق مرجع به وسیله GMDH

نتایج حاصل از مدل‌سازی تبخیر-تعرق مرجع با مدل دسته‌بندی گروهی داده‌ها برای مراحل آموزش و آزمون مدل، در جدول ۴ آورده شده است. بر اساس این جدول و میزان افزایش دقت با افزایش میزان اطلاعات ورودی و حجم مدل توسعه داده شده، افزایش پارامترهای ورودی تأثیر ناچیزی بر روی دقت مدل GMDH داشته است. از بین الگوهای ورودی مختلف، الگوی شماره دو شامل: درجه حرارت حداکثر و سرعت باد با میزان  $R^2=0/96$  و  $RMSE=0/63$  (mm/day) در مرحله آموزش و  $R^2=0/94$  و  $RMSE=0/76$  (mm/day) در مرحله آزمون بهترین عملکرد را نسبت به سایر ساختارها داشته و توانسته است میزان تبخیر-تعرق مرجع منطقه مورد مطالعه را با دقت مناسبی شبیه‌سازی نماید. لازم به ذکر است، چون حداقل داده ورودی جهت مدل‌سازی با GMDH، ۲ پارامتر است به همین دلیل الگوی شامل یک پارامتر در مدل‌سازی دخیل

جدول (۴): نتایج مدل GMDH در مدل‌سازی تبخیر-تعرق مرجع

ردیف	الگو	آموزش		آزمون	
		$R^2$	RMSE	$R^2$	RMSE
۱	M1	-	-	-	-
۲	M2	۰/۹۶	۰/۶۳	۰/۹۶	۰/۷۶
۳	M3	۰/۹۶	۰/۶۳	۰/۹۶	۰/۵۹
۴	M4	۰/۹۶	۰/۶۲	۰/۹۷	۰/۵۶
۵	M5	۰/۹۶	۰/۵۸	۰/۹۷	۰/۵۴
۶	M6	۰/۹۶	۰/۹۶	۰/۹۷	۰/۵۳



شکل (۴): (الف): نمودار مقادیر مشاهداتی و پیش‌بینی شده نسبت به زمان، (ب): نمودار پراکنش میان مقادیر مشاهداتی و پیش‌بینی شده GMDH در مرحله آزمون

برای هر دو مرحله آزمون و آموزش می‌باشد. با توجه به جدول ۵ و بر اساس میزان افزایش دقت با افزایش میزان اطلاعات ورودی و حجم مدل توسعه داده شده، الگوی شماره دو شامل: درجه حرارت حداکثر و سرعت باد با میزان  $R^2=0/96$  و  $RMSE=0/59$  (mm/day) در مرحله آموزش و  $R^2=0/96$  و  $RMSE=0/61$  (mm/day) در مرحله آزمون، دارای عملکرد بهتری نسبت به سایر ترکیب‌های ورودی بوده است.

رابطه ریاضی مستخرج از مدل دسته‌بندی گروهی داده-ها متشکل از درجه حرارت حداکثر و سرعت باد بر مبنای الگوی ورودی شماره ۲ به عنوان الگوی برتر، به صورت رابطه ۱۶ ارائه شده است.

$$ET_0 = 0.7561 + 0.0208T_{max} + 0.1137WS_{avg} + 0.0035T_{max}^2 - 0.0210WS_{avg}^2 + 0.0275WS_{avg}T_{max} \quad (16)$$

#### مدل‌سازی تبخیر-تعرق مرجع به وسیله MARS

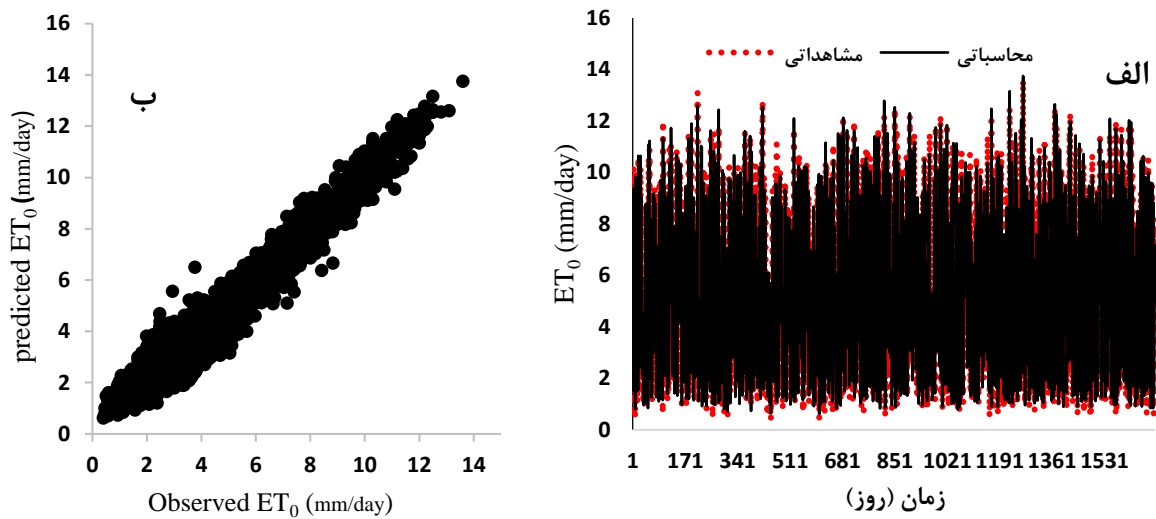
جدول ۵ نشان دهنده نتایج حاصل از مدل‌سازی تبخیر-تعرق گیاه مرجع به وسیله مدل اسپلاین تطبیقی چندگانه

جدول (۵): نتایج مدل MARS در مدل‌سازی تبخیر-تعرق مرجع

آزمون		آموزش		الگو	ردیف
RMSE	$R^2$	RMSE	$R^2$		
۱/۱۶	۰/۸۶	۱/۰۱	۰/۸۹	M1	۱
۰/۶۱	۰/۹۶	۰/۵۹	۰/۹۶	M2	۲
۰/۴۴	۰/۹۸	۰/۴۸	۰/۹۷	M3	۳
۰/۴۳	۰/۹۸	۰/۴۵	۰/۹۸	M4	۴
۰/۴۲	۰/۹۷	۰/۴۶	۰/۹۷	M5	۵
۰/۴۳	۰/۹۷	۰/۴۶	۰/۹۸	M6	۶

برآورد مقادیر حدی عملکرد خوبی داشته است. شکل ۵-ب نمودار پراکنش مقادیر محاسباتی و مشاهداتی تبخیر-تعرق را نسبت به نیمساز ربع اول نشان می‌دهد. با توجه به این نمودار هم مدل دارای دقت مناسبی در برآورد تبخیر-تعرق مرجع در منطقه مطالعاتی می‌باشد.

نمودار شکل ۵-الف نشان دهنده تغییرات مقادیر محاسباتی و مشاهداتی تبخیر-تعرق مرجع در دوره زمانی مورد مطالعه می‌باشد. بنابراین نمودار، مدل عملکرد خوبی در برآورد تبخیر-تعرق مرجع داشته و توانسته روند تغییرات آن را با دقت قابل قبولی پیش‌بینی نماید. همچنین مدل در



شکل (۵): (الف) نمودار مقادیر مشاهداتی و پیش‌بینی شده نسبت به زمان، (ب): نمودار پراکنش میان مقادیر مشاهداتی و پیش‌بینی شده در مرحله آزمون MARS

مرحله آموزش و  $R^2 = 0.96$  و  $RMSE = 0.61$  (mm/day) در مرحله آزمون، دارای بهترین عملکرد نسبت به دو مدل دیگر در برآورد تبخیر-تعرق مرجع روزانه‌ی ایستگاه الیگودرز بوده است. کرباسی (۱۳۹۵) در مقایسه بین مدل موجک-GMDH و مدل GMDH به این نتیجه رسید که مدل موجک-GMDH دقت بهتری نسبت به مدل GMDH داشته است. یاسین و همکاران (۲۰۱۶) در مقایسه بین مدل‌های شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن، کارایی بهتر شبکه عصبی مصنوعی را نشان دادند.

ارزیابی عملکرد مدل‌های برنامه‌ریزی بیان ژن، دسته‌بندی گروهی داده‌ها و اسپلاین تطبیقی چندگانه

در این بخش، به ارزیابی عملکرد و کارایی مدل‌های برنامه‌ریزی بیان ژن، دسته‌بندی گروهی داده‌ها و اسپلاین تطبیقی چندگانه پرداخته شده است. جدول ۶ دربردارنده مقادیر آماره‌های خطای هر سه مدل در مراحل آزمون و آموزش مدل‌ها می‌باشد. بر اساس مقادیر آماره‌ها، مدل MARS با  $R^2 = 0.96$  و  $RMSE = 0.59$  (mm/day) در

جدول (۶): آماره‌های خطاسنجی مدل‌های GEP، GMDH و MARS در برآورد تبخیر-تعرق مرجع

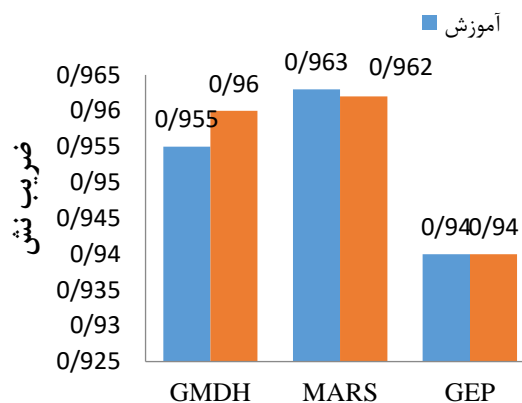
آزمون		آموزش		نام مدل
$R^2$	RMSE	$R^2$	RMSE	
۰/۹۴	۰/۷۶	۰/۹۳	۰/۸۴	برنامه‌ریزی بیان ژن
۰/۹۶	۰/۷۶	۰/۹۶	۰/۶۳	دسته‌بندی گروهی داده‌ها
۰/۹۶	۰/۶۱	۰/۹۶	۰/۵۹	اسپلاین تطبیقی چندگانه

در مرحله آزمون دقت بیشتری نسبت به دو مدل دیگر داشته است.  $NS = 0.962$

نمودار شکل ۶ نشان دهنده مقدار ضریب نش الگوهای منتخب هر سه مدل مورد استفاده در دو مرحله آموزش و آزمون می‌باشد. بر اساس این ضریب نیز مدل اسپلاین تطبیقی چندگانه با  $NS = 0.963$  در مرحله آموزش و



عملگرهای پیش‌فرض مدل با  $R^2=0/93$  و  $(\text{mm/day})$  و  $\text{RMSE}=0/84$  در مرحله آموزش و  $R^2=0/94$  و  $(\text{mm/day})$  و  $\text{RMSE}=0/76$  در مرحله آزمون، در مدل اسپلین تطبیقی چندگانه با دقت  $R^2=0/96$  و  $(\text{mm/day})$  و  $\text{RMSE}=0/63$  در مرحله آموزش و  $R^2=0/94$  و  $(\text{mm/day})$  و  $\text{RMSE}=0/76$  در مرحله آزمون و در مدل دسته‌بندی گروهی داده‌ها با دقت  $R^2=0/96$  و  $(\text{mm/day})$  و  $\text{RMSE}=0/63$  در مرحله آموزش و  $R^2=0/96$  و  $(\text{mm/day})$  و  $\text{RMSE}=0/76$  در مرحله آزمون توانایی خوبی در مدلسازی و تخمین تبخیر-تعرق روزانه منطقه الیگودرز دارد. بررسی عملکرد از بین مدل‌های مورد استفاده نشان داد، مدل اسپلین تطبیقی چندگانه با مقدار ضریب نش به میزان  $0/963$  در مرحله آموزش و  $0/962$  در مرحله آزمون، دارای دقت بیشتری نسبت به دو مدل دیگر می‌باشد. با توجه به نقش مهم پارامتر تبخیر-تعرق مرجع در مطالعات کشاورزی، طرح‌های مدیریت منابع آب، تعیین نیازآبی و تعیین هیدرومدول برای طراحی شبکه‌های آبیاری و زهکشی، نتایج حاصل از این تحقیق می‌تواند راهکار مناسبی برای محاسبه و برآورد این پارامتر مهم هیدرولوژیکی به مدیران و طراحان ارائه دهد مخصوصاً که هر سه مدل مورد استفاده با اطلاعات ورودی کمتر نسبت به روش فائو-پنمن-مانتیت عملکرد بهتری داشته‌اند. در پایان پیشنهاد می‌شود عملکرد مدل‌های مورد استفاده در این تحقیق نسبت به داده‌های لایسیمیتری ارزیابی شود و همچنین عملکرد این مدل‌ها در اقلیم‌های مختلف مورد مقایسه قرار گیرد



شکل (۶): نمودار مقدار ضریب نش مربوط به مدل‌های GEP، MARS و GMDH

## نتیجه‌گیری

در این پژوهش با استفاده از مدل‌های هوش مصنوعی برنامه‌ریزی بیان ژن، دسته‌بندی گروهی داده‌ها و اسپلین تطبیقی چندگانه، به مدل‌سازی تبخیر-تعرق مرجع ایستگاه الیگودرز و مقایسه نتایج با روش فائو-پنمن-مانتیت پرداخته شد. جهت بررسی اثرگذاری پارامترهای مختلف هواشناسی بر تبخیر-تعرق مرجع از رگرسیون گام به گام استفاده شد که بر اساس آن، ۶ ترکیب به عنوان الگوهای ورودی مدل‌سازی تعیین شد. نتایج مدل‌سازی نشان داد که در هر سه مدل مورد استفاده، الگوی شماره ۲ متشکل از درجه حرارت حداکثر و سرعت باد بهترین عملکرد را نسبت به سایر الگوها دارد. این الگو در مدل برنامه‌ریزی بیان ژن و

## منابع

- آزادطلاتیپه، ن.، بهمنش، ج.، منتصری، م. و رضوردی‌نژاد، و. ۱۳۹۴. مقایسه روش‌های سری زمانی و شبکه عصبی مصنوعی در پیش‌بینی تبخیر-تعرق مرجع (مطالعه موردی: ارومیه). علوم و مهندسی آبیاری (مجله علمی-پژوهشی)، جلد ۳۸، شماره ۴ زمستان ۹۴.
- حسینی، س.م.ر.، گنجی‌خرمدل، ن. و خلت‌آبادی‌فراهانی، ا.ج. ۱۳۹۴. ارزیابی و آنالیز حساسیت روش‌های مختلف تخمین تبخیر و تعرق مرجع روزانه در یک اقلیم سرد. نشریه پژوهش‌های کاربردی علوم آب، سال اول، شماره دوم، زمستان ۱۳۹۴، صفحات ۲۹-۴۰.
- سبزواری، ی.، نصرالهی، ع.ج.، شریفی‌پور، م. و شاهی‌نژاد، ب. ۱۳۹۹. کاربرد رگرسیون چندمتغیره و برنامه‌ریزی بیان ژن در برآورد تبخیر و تعرق. نشریه علوم و مهندسی آبیاری، مرداد ۱۳۹۹. شناسه دیجیتال (DOI): 10.22055/JISE.2020.31583.1890

سیفی، ا.، میرلطیفی، س.م. و ریاحی، ح. ۱۳۸۹. توسعه مدل ترکیبی رگرسیون چندگانه-تحلیل مؤلفه‌ها و عامل‌های اصلی (MLR-PCA) در پیش‌بینی تبخیر-تعرق مرجع (مطالعه موردی: ایستگاه کرمان). نشریه آب‌و خاک، جلد ۲۴، شماره ۶، صفحات ۱۱۸۶-۱۱۹۶.

فیض‌اله‌پور، ف. دلور، م. و حسامی افشار، م. ۱۳۹۶. ارزیابی و تحلیل عدم قطعیت برآورد تبخیر-تعرق گیاه مرجع با استفاده از برنامه‌ریزی ژنتیک. نشریه دانش آب‌و خاک، جلد ۲۷، شماره ۴، ۱۳۵-۱۴۷.

کرباسی، م. ۱۳۹۵. پیش‌بینی تبخیر-تعرق مرجع ایستگاه سینوپتیک اهواز با استفاده از مدل ترکیبی موجک-شبکه عصبی GMDH. نشریه پژوهش‌های حفاظت آب و خاک، جلد بیست و سوم، شماره چهارم، ۱۳۹۵.

کولائی، ع.، غلامی سفیدکوهی، م.ع. و ضیاتبارا احمدی، م.خ. ۱۳۹۲. ارزیابی روش‌های نوین محاسباتی در برآورد تبخیر و تعرق روزانه (مطالعه موردی دشت ناز ساری). دوازدهمین همایش ملی آبیاری و کاهش تبخیر. دانشگاه شهید باهنر کرمان، کرمان، ۵ و ۶ شهریور.

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), D05109.

Berti, A., Tardivo, G., Chiaudani, A., Rech, F. and Borin, M. 2014. Assessing reference Evapotranspiration by the Hargreaves method in north-eastern Italy. *Agric. Water Manage.* 140: 20-25.

Djaman, K.B., Balde, A., Sow, A., Muller, B., Irmak, S.K., N'Diaye, M., Manneh, B.D., Moukoumbi, Y., Futakuchi, K. and Saito, K. 2015. Evaluation of sixteen reference Evapotranspiration methods under sahelian conditions in the Senegal River Valley. *J. Hydrol. Region. Stud.* 3: 139-159.

Ferreira, C. 2001. Gene Expression Programming: *A New Adaptive Algorithm for Solving Problems, Complex Systems*, 13 (2), 87-129.

Kisi, O. 2009. Modeling monthly evaporation using two different neural computing techniques. *Irrigation Science*, 27:417-430.

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 (2018) 41-48.

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: 5959-5969.

Lima, A.R., Cannon, A.J. and Hsieh, W.W. 2015. Nonlinear regression in environmental sciences using extreme learning machines: a comparative evaluation. *Environ Model Software*. 73, 175-188.

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. *J. Agric. Water Manage.* 184: 145-155.

Mallik, A.P., Jyothy, S.A. and Sekhar Reddy, K.C. 2013. Daily reference evapotranspiration estimation using linear regression and Ann models. *The Institution of Engineers (India)*, 93(4):215-221.

Niaghi, A.R., Majnooni-Heris, A., Haghi, D.Z. and Mahtabi, G., 2013. Evaluate several potential Evapotranspiration methods for regional Use in Tabriz, Iran. *J. Appl. Environ. Biol.* 3: 31-41.

Tabari, H., Grismer, M.E. and Trajkovic, S. 2013. Comparative analysis of 31 reference Evapotranspiration methods under humid conditions. *Irrig. Sci.* 31: 107-117.

Trajkovic, S. 2007. Hargreaves versus Penman-Monteith under humid conditions. *J. Irrig. Drain. Eng.* 133: 38-47.





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, 110-124.



## Modeling and Estimating Daily Reference Evapotranspiration Using Soft Computing Models (Case Study: Aligudarz Station)

Yaser Sabzevari<sup>1</sup>, Abbas parsaei<sup>2</sup>, AmirHamzeh Haghiabi<sup>3\*</sup>

### Abstract

Calculation and estimation of evapotranspiration is one of the most important parameters of water management in agricultural engineering projects. The aim of this study was to evaluate the models of gene expression programming (GEP), group method of data handling (GMDH), and Multivariate adaptive regression spline (MARS) to estimating daily reference evapotranspiration. For this purpose, daily data recorded during the last 25-year period (1993-2017) of Aligudarz region (located in the east of Lorestan province) were used. 80% of the data were used for training and the remaining 20% for testing the models. The modeling results showed that only with the maximum temperature and average wind speed can evapotranspiration be estimated with very good accuracy. The error indices of GEP model in testing stage are  $R^2 = 0.93$  and  $RMSE = 0.84$ , the error indices of MARS and GMDH models are  $R^2 = 0.94$  and  $RMSE = 0.76$ . Comparing the performance of the models showed that the March model performed better than the other models.

**Keywords:** Farm Water Management, Gene Expression Programming, Group Method F Data Handling, Multivariate Adaptive Regression Splines, Step-By-Step Regressio

---

<sup>1</sup> PhD student, water department, Faculty of Agriculture, Isfahan University of Technology, Isfahan, Iran, y.sabzevari@ag.iut.ac.ir

<sup>2</sup> Assistant Professor, Department of Hydraulic Structures, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran, Parsaei@scu.ac.ir

<sup>3</sup> Professor, water department, Faculty of Agriculture, Lorestan university, Khoramabad, Iran, corresponding author email: Haghiabi.a@lu.ac.ir