

## بررسی عملکرد روش برنامه‌ریزی بیان ژن در پیش‌بینی تابش خورشیدی روزانه در گستره ایران

فهمیه خادم‌پور<sup>۱</sup>، عباس خاشعی سیوکی<sup>۲\*</sup>، مهدی امیرآبادی‌زاده<sup>۳</sup>

۱. دانشجوی دکتری گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران

۲. دانشیار، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران

۳. استادیار، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران

تاریخ وصول: ۱۳۹۶/۱۱/۷ تاریخ پذیرش: ۱۳۹۷/۶/۲۴

### چکیده

انرژی تابشی خورشید سرمنشأ انرژی‌های روی کره‌ی زمین است که یکی از منابع مهم انرژی پاک به ویژه در کشور ایران محسوب می‌شود. برآورد درست مقدار تابش خورشیدی به عنوان یکی از پارامترهای مهم در شبیه‌سازی رشد گیاهان و تخمین تبخیرتغرق اهمیت زیادی دارد. هدف از انجام این مطالعه، توانایی مدل شبیه‌سازی GeneXproTools 5.0 در پیش‌بینی تشعشعات خورشیدی بر اساس داده‌های هواشناسی (بارش، ساعات آفتابی، رطوبت نسبی، حداکثر دما، حداقل دما، میانگین دما و تابش خورشیدی، هم‌چنین با در نظر گرفتن ارتفاع، طول جغرافیایی و عرض جغرافیایی) در مقیاس روزانه، در ۳۱ ایستگاه در گستره‌ی ایران در بازه‌ی زمانی ۲۰۱۶-۲۰۰۷ می‌باشد. بهترین مدل بر اساس معیارهای ارزیابی  $RMSE$ ،  $MAE$ ،  $NS$  و  $R^2$  انتخاب شد. از نتایج تجزیه و تحلیل، مشخص شد که مقادیر پیش‌بینی شده سازگاری خوبی با مقادیر اندازه‌گیری شده در مدل برنامه‌ریزی بیان ژن (GEP) دارد. در مدل GEP، سناریوی e به دلیل در نظر نگرفتن ارتفاع، طول جغرافیایی، عرض جغرافیایی، بارش، رطوبت نسبی، حداکثر دما و حداقل دما، دارای ضریب همبستگی کمتر و خطای بیشتری می‌باشد و کارایی کمتری دارد. در این مدل، در سناریوی c با اضافه شدن پارامترهای میانگین دما و تابش خورشیدی، برآورد مطلوب‌تری از تابش خورشیدی حاصل می‌شود. نتایج سناریوهای c و e بسیار به هم نزدیک می‌باشد، اما سناریوی e در پیش‌بینی تابش خورشیدی ضعیف‌تر عمل می‌کند و بهترین مدل در این مطالعه، سناریوی c می‌باشد. زمانی که عرض جغرافیایی، بارش، رطوبت نسبی، حداکثر دما، حداقل دما، میانگین دما و تابش خورشیدی برای پیش‌بینی تابش خورشیدی در نظر گرفته شود، نتایج دقیق‌تری حاصل می‌شود. هم‌چنین، معیارهای ارزیابی تحت سناریوی c با  $R^2 = 0.72$ ،  $RMSE = 2.59$ ،  $MAE = 2.82$  و  $NSE = 0.72$  در بخش آموزش از سایر سناریوها بیشتر می‌باشد. در مجموع روش GEP دقیق‌ترین نتایج را در تخمین تابش خورشیدی روزانه در گستره‌ی ایران دارد.

**کلید واژگان:** تابش خورشیدی، سناریو، نرم‌افزار GeneXproTools

\*E- mail: abbaskhashei@birjand.ac.ir

## مقدمه

تشتک در خراسان جنوبی پرداختند. مقایسه نتایج آزمون دو مدل نشان داد که مدل برنامه‌ریزی بیان ژن کارایی بهتری نسبت به مدل عصبی- فازی در برآورد روزانه تبخیر از تشتک دارد. لاندرز و همکاران<sup>۲</sup> (۲۰۱۲)، به مقایسه برنامه-نویسی بیان ژن با شبکه‌ی عصبی مصنوعی و عصبی- فازی در برآورد تابش خورشیدی روزانه در شمال اسپانیا پرداختند. نتایج نشان داد که، شبکه عصبی مصنوعی بهترین عملکرد را در میان مدل‌های مورد مطالعه داشته است. کیسی و همکاران<sup>۳</sup> (۲۰۱۳)، برای مدل‌سازی بارش رواناب مدل‌های ANN و GEP و همچنین ANFIS را با هم مقایسه کردند و نشان داد که GEP برای مدل‌سازی بارش رواناب مناسب است. موزامیل و همکاران<sup>۴</sup> (۲۰۱۵)، به کاربرد برنامه‌ریزی بیان ژن در آنالیز فرکانس جریان و مقایسه‌ی آن با شبکه‌ی عصبی مصنوعی و توزیع گامبل پرداختند و دریافتند که برنامه‌ریزی بیان ژن نسبت به شبکه‌ی عصبی مصنوعی و توزیع گامبل مدلی برتر می‌باشد. ترائور و همکاران<sup>۵</sup> (۲۰۱۷)، به پیش‌بینی تبخیرتغرق مرجع روزانه با استفاده از اطلاعات آب و هوا و برنامه‌نویسی بیان ژن در چین پرداختند. نتایج نشان داد که، بیان ژن توانایی خوبی برای پیش‌بینی تبخیرتغرق روزانه با کمترین خطا را داشته است. با توجه به اینکه پیش‌بینی شدت تابش خورشیدی در گستره‌ی ایران در مقیاس روزانه با نرم افزار GeneXproTools تاکنون انجام نشده است و با توجه به گستردگی اقلیم‌های متفاوت در مناطق مختلف کشور بنابراین، این پژوهش در گستره‌ی ایران صورت گرفته است. لذا، توانایی مدل شبیه-سازی نرم‌افزار GeneXproTools در پیش‌بینی تشعشعات خورشیدی بر اساس ارتفاع، طول جغرافیایی، عرض جغرافیایی، بارش، رطوبت نسبی، حداکثر دما، حداقل دما، میانگین دما و تابش خورشیدی در مقیاس روزانه در ۳۱ ایستگاه در گستره ایران در بازه‌ی زمانی ۲۰۱۶-۲۰۰۷ انجام پذیرفته است.

انرژی تابشی خورشید سرمنشأ انرژی‌های روی کره‌ی زمین است که یکی از منابع مهم انرژی پاک به ویژه در کشور ایران محسوب می‌شود. برآورد درست مقدار تابش خورشیدی به عنوان یکی از پارامترهای مهم در شبیه‌سازی رشد گیاهان و تخمین تبخیرتغرق اهمیت زیادی دارد. اندازه‌گیری شدت تابش خورشیدی، اگر چه در ایران دارای سابقه‌ی نسبتاً طولانی است ولیکن به دلیل هزینه‌ی بالا، اکثر ایستگاه‌های موجود در کشور به دستگاه تابش‌سنج<sup>۱</sup> مجهز نیستند و این نقص حتی در کشورهای در حال توسعه نیز به چشم می‌خورد. در عوض، در قالب این ایستگاه‌ها پارامتر ساعت آفتابی (n) به طور روزانه اندازه‌گیری می‌شود. از آنجا که اطلاع از شدت تابش خورشیدی رسیده به سطح نسبت به زمان ساعات آفتابی اهمیت بیشتری دارد، لذا شناسایی روش‌های تخمین شدت تابش خورشیدی رسیده به سطح، از آمار موجود (ساعات آفتابی) از اهمیت ویژه‌ای برخوردار است (یزدان‌پناه و همکاران، ۱۳۸۸). مطالعات زیادی برای پیش‌بینی تابش خورشیدی انجام شده است که به برخی از آن‌ها اشاره می‌شود. قبادیان و همکاران (۱۳۹۱)، بررسی عملکرد روش برنامه‌ریزی بیان ژن در روندیابی سیلاب رودخانه‌ی زنگمار در مقایسه با روش موج دینامیکی را تجزیه و تحلیل کردند. نتایج بررسی آن‌ها نشان داد که مدل برنامه‌ریزی بیان ژن قادر است با دقت بیشتری حجم هیدروگراف خروجی را پیش‌بینی کند. مدل موج دینامیکی به طور متوسط با خطایی برابر ۷/۲۵ درصد و مدل برنامه-ریزی ژن با خطایی برابر ۱۷ درصد دبی پیک هیدروگراف خروجی را پیش‌بینی می‌کنند. احمدی و همکاران (۱۳۹۴)، به تجزیه و تحلیل کاربرد سیستم‌های استنتاجی عصبی- فازی تطبیقی و برنامه‌ریزی ژنتیک برای برآورد تبخیرتغرق ماهانه در شمال غرب ایران پرداختند. نتایج مطالعه‌ی آن‌ها نشان داد که هر دو مدل دقت بالا و خطای کمی در پیش-بینی تبخیرتغرق ماهانه گیاه مرجع داشتند. حقیقت‌جو و همکاران (۱۳۹۶)، به مقایسه روش‌های سیستم استنتاج عصبی- فازی و برنامه‌ریزی بیان ژن در برآورد تبخیر از

2 Landeras et al.

3 Kisi et al.

4 Muzzammil et al.

5 Traore et al.

1 Actinograph

## مواد و روش

زمانی مشترک ۱۰ ساله (۲۰۰۷-۲۰۱۶) می‌باشد و از داده‌های ارتفاع، طول جغرافیایی، عرض جغرافیایی، میانگین بارش، میانگین رطوبت نسبی، حداکثر دما، حداقل دما، میانگین دما و میانگین تابش خورشیدی در مقیاس روزانه استفاده شد. داده‌های روزانه از سازمان هواشناسی جمهوری اسلامی ایران IRIMO جمع‌آوری شده‌اند. مشخصات جغرافیایی ایستگاه‌های مطالعاتی و اقلیم آن‌ها در سیستم دوماترن در جدول ۱ و هم‌چنین مقادیر میانگین عوامل اقلیمی مذکور برای دوره‌ی مطالعاتی در جدول ۲ ارائه شده است. شکل ۱، موقعیت ایستگاه‌های مورد مطالعه را در سطح کشور نشان می‌دهد.

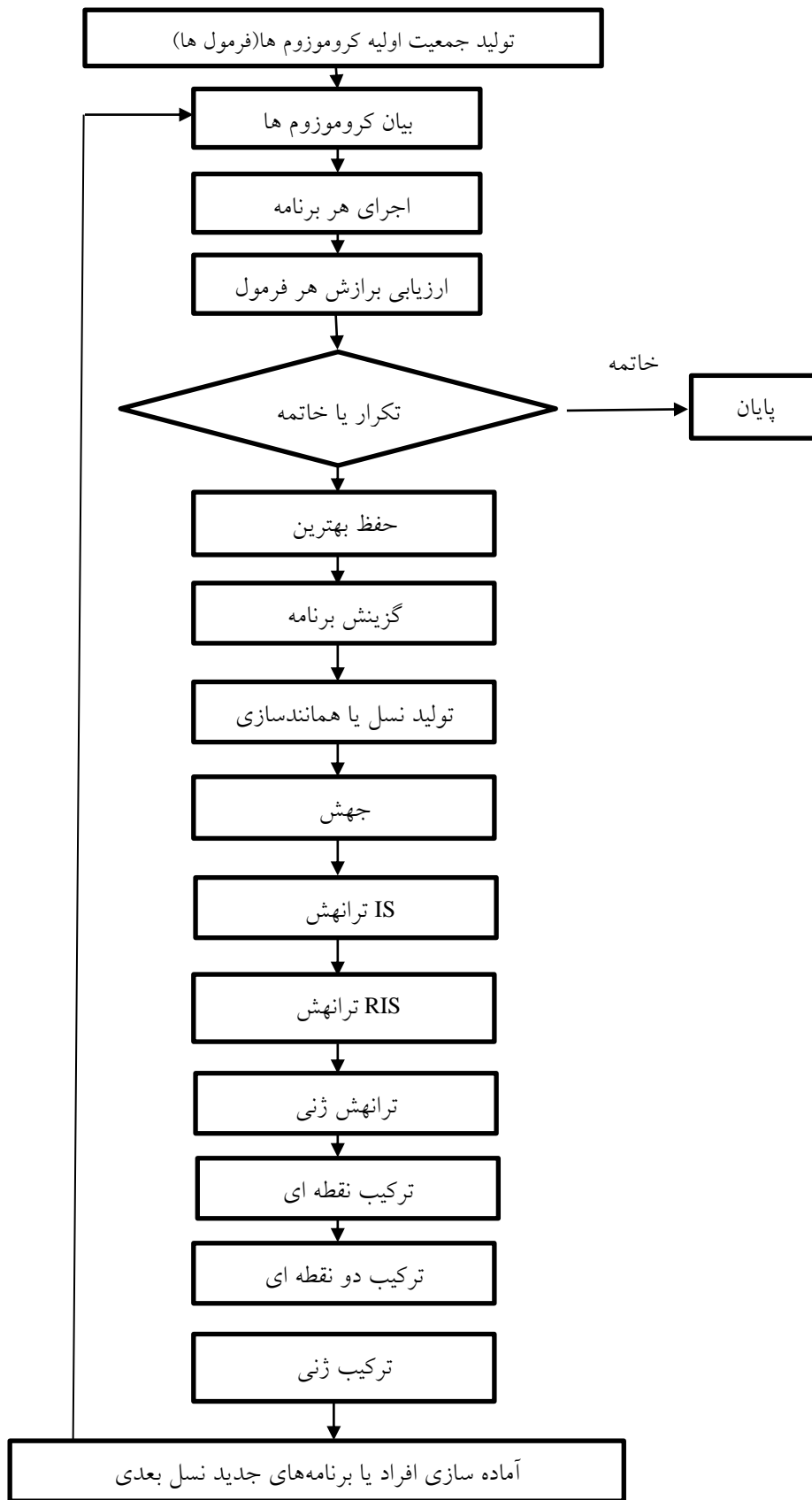
ایران با متوسط بارش سالانه ۲۴۱ میلی‌متر در عرض جغرافیایی خشک و نیمه‌خشک کره زمین و بین دو نصف النهار ۴۴° و ۶۴° شرقی و دو مدار ۲۵° و ۴۰° شمالی واقع شده است. حدود ۹۴/۸ درصد از سطح آن، در زمره‌ی مناطق خشک و نیمه‌خشک با ریزش‌های جوی کم و تبخیرتغرق زیاد قرار دارد. ایستگاه‌های هواشناسی در این تحقیق براساس تنوع اقلیمی انتخاب شدند. بدین ترتیب ایستگاه‌های مطالعاتی در سیستم طبقه‌بندی دوماترن به شش اقلیم فراهشک، خشک، نیمه‌خشک، مدیترانه‌ای، مرطوب و بسیار مرطوب (الف) تقسیم شد. ایستگاه‌های هواشناسی منتخب شامل ۳۱ ایستگاه سینوپتیک در گستره‌ی ایران با دوره‌ی



شکل ۱- موقعیت جغرافیایی ایستگاه‌های هواشناسی

جدول ۱ - مشخصات جغرافیایی ایستگاه‌های مطالعاتی و اقلیم آن‌ها در سیستم دوما رتن

نام ایستگاه	دوره‌ی آماری	طول جغرافیایی شرقی	عرض جغرافیایی شمالی	ارتفاع از سطح دریا متر	نوع اقلیم
اهواز	۲۰۰۷-۲۰۱۶	۴۸°۴۰'	۳۱°۲۰'	۲۲/۵	خشک
اراک	۲۰۰۷-۲۰۱۶	۴۹°۴۶'	۳۴°۰۶'	۱۳۱۲/۵	نیمه‌خشک
اردبیل	۲۰۰۷-۲۰۱۶	۴۸°۱۷'	۳۸°۱۵'	۱۳۳۲/۰	نیمه‌خشک
بندرعباس	۲۰۰۷-۲۰۱۶	۵۶°۲۲'	۲۷°۱۳'	۹/۸	فراخشک
بیرجند	۲۰۰۷-۲۰۱۶	۵۹°۱۲'	۳۲°۵۲'	۱۴۹۱/۰	خشک
بجنورد	۲۰۰۷-۲۰۱۶	۵۷°۱۹'	۳۷°۲۸'	۱۱۱۲/۰	نیمه‌خشک
بوشهر	۲۰۰۷-۲۰۱۶	۵۰°۵۰'	۲۸°۵۹'	۹/۰	خشک
اصفهان	۲۰۰۷-۲۰۱۶	۵۱°۴۰'	۳۲°۳۷'	۱۵۵۰/۴	خشک
قزوین	۲۰۰۷-۲۰۱۶	۵۰°۰۳'	۳۶°۱۵'	۱۲۷۹/۲	نیمه‌خشک
قم	۲۰۰۷-۲۰۱۶	۵۰°۵۱'	۳۴°۴۲'	۸۷۷/۴	خشک
گرگان	۲۰۰۷-۲۰۱۶	۵۴°۲۴'	۳۶°۵۴'	۰/۰	مدیترانه‌ای
همدان	۲۰۰۷-۲۰۱۶	۴۸°۳۲'	۳۴°۵۲'	۱۷۴۱/۵	نیمه‌خشک
ایلام	۲۰۰۷-۲۰۱۶	۴۶°۲۶'	۳۳°۳۸'	۱۳۳۷/۰	مدیترانه‌ای
کرج	۲۰۰۷-۲۰۱۶	۵۰°۵۴'	۳۵°۵۵'	۱۷۰۸/۰	خشک
کرمان	۲۰۰۷-۲۰۱۶	۵۶°۵۸'	۳۰°۱۵'	۱۷۵۳/۸	خشک
کرمانشاه	۲۰۰۷-۲۰۱۶	۴۷°۰۹'	۳۴°۲۱'	۱۳۱۸/۶	نیمه‌خشک
خرم‌آباد	۲۰۰۷-۲۰۱۶	۴۸°۱۷'	۳۳°۲۶'	۱۷۴۸/۸	نیمه‌خشک
مشهد	۲۰۰۷-۲۰۱۶	۵۹°۳۸'	۳۶°۱۶'	۹۹۹/۲	خشک
ارومیه	۲۰۰۷-۲۰۱۶	۴۵°۰۵'	۳۷°۳۲'	۱۳۱۵/۹	نیمه‌خشک
رشت	۲۰۰۷-۲۰۱۶	۴۹°۳۷'	۳۷°۱۹'	-۸/۶	بسیار مرطوب (الف)
سندج	۲۰۰۷-۲۰۱۶	۴۷°۰۰'	۳۵°۲۰'	۱۳۷۳/۴	نیمه‌خشک
ساری	۲۰۰۷-۲۰۱۶	۵۳°۰۰'	۳۶°۳۳'	۲۳/۰	مرطوب
سمنان	۲۰۰۷-۲۰۱۶	۵۳°۳۳'	۳۵°۳۵'	۱۱۲۷/۰	خشک
شهرکرد	۲۰۰۷-۲۰۱۶	۵۰°۵۱'	۳۲°۱۷'	۲۰۴۸/۹	نیمه‌خشک
شیراز	۲۰۰۷-۲۰۱۶	۵۲°۳۶'	۲۹°۳۲'	۱۴۸۴/۰	نیمه‌خشک
تبریز	۲۰۰۷-۲۰۱۶	۴۶°۱۷'	۳۸°۰۵'	۱۳۶۱/۰	نیمه‌خشک
تهران	۲۰۰۷-۲۰۱۶	۵۱°۱۹'	۳۵°۴۱'	۱۱۹۰/۸	خشک
یاسوج	۲۰۰۷-۲۰۱۶	۵۱°۳۳'	۳۰°۴۱'	۱۸۱۶/۳	مرطوب
یزد	۲۰۰۷-۲۰۱۶	۵۴°۱۷'	۳۱°۵۴'	۱۲۳۷/۲	فراخشک
زاهدان	۲۰۰۷-۲۰۱۶	۶۰°۵۳'	۲۹°۲۸'	۱۳۷۰/۰	فراخشک
زنجان	۲۰۰۷-۲۰۱۶	۴۸°۲۹'	۳۶°۴۱'	۱۶۶۳/۰	نیمه‌خشک



شکل ۲- فلوجارت الگوریتم بیان ژن

جدول ۲- خصوصیات آماری پارامترهای مورد استفاده

پارامتر	دوره‌ی آماری	حداقل	حداکثر	میانگین	انحراف از معیار	ضریب تغییرات
ارتفاع (متر)	۲۰۱۶-۲۰۰۷	-۸/۶	۱۵۵۰/۴۰	۱۰۱۴/۶۸	۷۰۰/۸۶	۰/۶۹
طول جغرافیایی (درجه)	۲۰۱۶-۲۰۰۷	۴۹/۶۲	۵۹/۲۰	۵۵/۸۴	۴/۵۱	۰/۰۸
عرض جغرافیایی (درجه)	۲۰۱۶-۲۰۰۷	۳۲/۶۲	۳۷/۳۲	۳۴/۲۸	۲/۰۸	۰/۰۶
بارش (میلی‌متر)	۲۰۱۶-۲۰۰۷	۰/۰۰	۶۸/۸۰	۰/۶۹	۳/۳۱	۴/۸۰
رطوبت نسبی (درصد)	۲۰۱۶-۲۰۰۷	۶/۵۰	۱۰۰/۰۰	۵۰/۰۴	۲۷/۵۹	۰/۵۵
حداکثر دما (°C)	۲۰۱۶-۲۰۰۷	-۳/۸۰	۴۲/۶۰	۲۳/۶۰	۹/۳۴	۰/۴۰
حداقل دما (°C)	۲۰۱۶-۲۰۰۷	-۱۶/۴۰	۲۷/۰۰	۱۰/۲۳	۸/۶۰	۰/۸۴
میانگین دما (°C)	۲۰۱۶-۲۰۰۷	-۱۴/۶۰	۳۴/۸۰	۱۶/۶۳	۹/۲۴	۰/۵۶
تابش خورشیدی (MJ.m <sup>-2</sup> .d <sup>-1</sup> )	۲۰۱۶-۲۰۰۷	۴/۷۵	۳۰/۱۶	۱۸/۹۷	۶/۷۷	۰/۳۶

به منظور مقایسه نتایج حاصل از الگوریتم بیان ژن در پیش-بینی تابش خورشیدی، ۸ سناریو با توجه به پارامترهای تأثیرگذار بر تابش خورشیدی مانند ارتفاع (متر)، طول جغرافیایی (درجه)، عرض جغرافیایی (درجه)، بارش (میلی-متر)، رطوبت نسبی (درصد)، حداکثر دما (°C)، حداقل دما (°C) و میانگین دما (°C) تعریف شد. در جدول ۳، پارامترهای ورودی در هر سناریو ارائه گردید. برای بهره-گیری از الگوریتم بیان ژن از نرم‌افزار GeneXproTools 5.0 استفاده شده است

جدول ۳- پارامترهای ورودی در سناریوهای مختلف

سناریو	پارامترهای ورودی
a	Altitude, Longitude, Latitude, p, RH, T <sub>max</sub> , T <sub>min</sub> , T <sub>mean</sub> , R <sub>s</sub>
b	Longitude, Latitude, p, RH, T <sub>max</sub> , T <sub>min</sub> , T <sub>mean</sub> , R <sub>s</sub>
c	Latitude, p, RH, T <sub>max</sub> , T <sub>min</sub> , T <sub>mean</sub> , R <sub>s</sub>
d	p, RH, T <sub>max</sub> , T <sub>min</sub> , T <sub>mean</sub> , R <sub>s</sub>
e	RH, T <sub>max</sub> , T <sub>min</sub> , T <sub>mean</sub> , R <sub>s</sub>
f	T <sub>max</sub> , T <sub>min</sub> , T <sub>mean</sub> , R <sub>s</sub>
g	T <sub>min</sub> , T <sub>mean</sub> , R <sub>s</sub>
h	T <sub>mean</sub> , R <sub>s</sub>

در جدول بالا، Altitude: ارتفاع (متر)، Longitude: طول جغرافیایی (درجه)، Latitude: عرض جغرافیایی (درجه)، p: بارش (میلی‌متر)، RH: رطوبت نسبی (درصد)، T<sub>max</sub>: حداکثر دما (°C)، T<sub>min</sub>: حداقل دما (°C)، T<sub>mean</sub>: میانگین دما (°C) و R<sub>s</sub>: تابش خورشیدی (MJ.m<sup>-2</sup>.d<sup>-1</sup>) می‌باشد.

انتخاب تابع برازش می‌تواند بر اساس روابط متعددی صورت گیرد و اگر این تابع به درستی انتخاب شود مدل از برازش بهتری برخوردار است که در این پژوهش، بعد از سعی و خطا کردن و انتخاب توابع مختلف، تابع خطای

مطلق نسبی یا RAE (Relative Absolute Error) به عنوان بهترین تابع برازش انتخاب گردید. خطای مطلق نسبی، بسیار شبیه به خطای مربع نسبی است به این معنا که فقط میانگین مقادیر واقعی را پیش‌بینی می‌کند. معادله RAE به صورت معادله (۱) می‌باشد.

$$E_i = \frac{\sum_{j=1}^n |P_{ij} - T_j|}{\sum_{j=1}^n |T_j - T_j^-|} \quad (1)$$

**ترکیب ژنی (Gene recombination)**

در این ترکیب، ابتدا دو کروموزوم والد انتخاب می‌شود و یکی از ژن‌های دو کروموزوم (در هر کروموزوم یک ژن) به صورت تصادفی انتخاب می‌شود، در نهایت ژن‌های انتخاب شده در دو کروموزوم با یکدیگر جابجا می‌شوند (فریرا، ۲۰۰۱).

ترکیب تک نقطه‌ای، دو کروموزوم به صورت تصادفی انتخاب می‌شوند و در دو کروموزوم به صورت تصادفی یک محل شکاف (محل شکاف در دو کروموزوم یکی باشد) انتخاب می‌شود و از محل شکاف تا پایین دست (انتهای کروموزوم) به عنوان توالی در نظر گرفته می‌شود و در نهایت توالی انتخاب شده در کروموزوم‌ها با یکدیگر عوض می‌شود. در ترکیب دو نقطه‌ای، از آن جایی که ترکیب دو نقطه‌ای قادر است نواحی کدگذاری نشده داخل کروموزوم-ها را به مراتب بیشتری روشن و خاموش کند، مطلوب‌تر است. در این عملگر ابتدا دو کروموزوم والد انتخاب می‌شود و در هر یک از آن‌ها یک توالی با طول تصادفی و برابر بین ابتدا و انتهای کروموزوم انتخاب می‌شود (جایگاه اول و آخر کروموزوم را شامل نمی‌شود) و در نهایت توالی انتخاب شده در کروموزوم‌ها با یکدیگر عوض می‌شود (فریرا، ۲۰۰۱).

که در آن،  $P$  مقدار پیش‌بینی شده توسط مدل،  $T_j$  مقدار تابع هدف و  $T^-$  نیز میانگین توابع هدف می‌باشد.

در جدول ۴، مقادیر مطلوب پارامترهای مورد استفاده در نرم‌افزار GEP که با آزمایش و سعی و خطا، مقدار مطلوب هر کدام حاصل گردید، نشان داده شده است. به برخی از آن‌ها اشاره می‌شود.

هدف عملگر جهش (Mutation Rate)، بهسازی تصادفی، داخل کروموزوم‌های معینی است. خصوصیت این عملگر آن است که برای جلوگیری از ایجاد افراد معیوب از نظر قواعد، برخی عملیات بدون نقص را اجرا می‌نماید. عملگر ترانهش (Transposition)، عناصر ترانهش از GEP قطعاتی از ژنوم هستند که می‌تواند فعال شده و در قسمت دیگری در کروموزوم کپی شوند. در GEP، سه نوع ترانهش که عبارتند از عملگر RIS، IS و عملگر ترانهش ژنی وجود دارد (فریرا، ۲۰۰۱).

**عملگر IS elements (insertion sequence elements)**  
هر توالی در ژنوم (با طول ۱۰ و ۳۰ به صورت تصادفی) ممکن است تبدیل به یک عنصر ترانهش شود، این توالی به صورت تصادفی در سراسر کروموزوم انتخاب می‌شود. یک کپی از توالی ترانهش ساخته شده است و در هر موقعیت در head یک ژن (به جز در موقعیت شروع) قرار می‌گیرد و برای اینکه طول کروموزوم ثابت بماند، یک توالی به طول همان توالی تصادفی از آخر بخش head در همان ژنی که توالی کپی می‌شود، حذف می‌شود (فریرا، ۲۰۰۱).

**عملگر RIS elements (root IS elements)**

هر توالی در ژنوم (با طول ۱۰ و ۳۰ به صورت تصادفی) به شرط اینکه توالی با تابع شروع شود، انتخاب می‌شود و در موقعیت شروع ژن (head) کپی می‌شود. برای اینکه طول کروموزوم ثابت بماند، یک توالی به طول همان توالی تصادفی از آخر بخش head در همان ژنی که توالی کپی می‌شود، حذف می‌شود (فریرا، ۲۰۰۱).

جدول ۴- پارامترهای مورد استفاده در GEP

مقدار	پارامتر
۱۰	اندازه سر (Head Size)
۹	اندازه دنباله (Tail Size)
۵۰	تعداد کروموزومها (Number of Chromosomes)
۴	تعداد ژن‌ها در هر کروموزوم (Number of Genes)
۰/۰۵	نرخ جهش (Mutation Rate)
۰/۰۵	نرخ وارونسازی (Inversion Rate)
۰/۳	نرخ ترکیب تک نقطه ای (One-Point Recombination Rate)
۰/۲	نرخ ترکیب دو نقطه ای (Two-Point Recombination Rate)
۰/۲	نرخ ترکیب ژن (Gene Recombination Rate)
۰/۸	نرخ ترانزیشن درجه متوالی (IS Transposition Rate)
۰/۸	نرخ ترانزیشن ریشه درجه متوالی (RIS Transposition Rate)
۰/۸	نرخ ترانزیشن ژن (Gene Transposition Rate)
RAE	معیار خطای تابع (Fitness Function Error Type)
تابع جمع "+"	تابع پیوند (Linking Function)
+,*, /,-, ln,sin,cos,arctan,x <sup>2</sup> ,√	مجموعه توابع (Function Set)
۲	تعداد اجرا (Number of Runs)
۴۰۰۰	تعداد جمعیت یا نسل (Number of Generation)
۰/۸	نرخ عملگر جهش در ثوابت تصادفی (R.C Mutation Rate)
۰/۰۵	نرخ عملگر ترانزیشن در دامنه DC (DC Specific IS Transposition Rate)

$$NSE = 1 - \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (۴)$$

$$R^2 = \frac{(\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}))^2}{(\sum_{i=1}^n (X_i - \bar{X})^2)(\sum_{i=1}^n (Y_i - \bar{Y})^2)} \quad (۵)$$

که در آن، N تعداد کل داده‌ها،  $X_i$  داده‌های شبیه‌سازی شده،  $Y_i$  داده‌های مشاهداتی،  $i$  گام‌های زمانی و  $\bar{X}$  و  $\bar{Y}$  میانگین داده‌های  $X$  و  $Y$  می‌باشد.

### نتایج و بحث

هدف اصلی این تحقیق، انتخاب بهترین مدل برای پیش‌بینی میانگین روزانه تابش خورشیدی در گستره ایران با استفاده از پارامترهای هواشناسی است. تعداد داده‌های آموزش و آزمایش برای مدل GEP در جدول ۵ ارائه شده است.

### ارزیابی عملکرد مدل

در پایان جهت ارزیابی و آزمون اعتبار مدل GEP، به ارزیابی عملکرد آن پرداخته شد. برای ارزیابی عملکرد این مدل از آماره‌های ریشه متوسط خطای مربعات (RMSE)، متوسط قدر مطلق خطا (MAE)، معیار نش- ساتکلیف (NSE) و ضریب تبیین ( $R^2$ ) مطابق معادلات (۲) تا (۵) استفاده گردید.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{N}} \quad (۲)$$

$$MAE = \frac{\sum_{i=1}^n |X_i - Y_i|}{N} \quad (۳)$$



جدول ۵- داده‌های آموزش و آزمایش ایستگاه‌های مطالعاتی برای مدل GEP

مدل	دوره‌ی آماری	آموزش		آزمایش	
		کل داده‌ها	داده‌های آموزش	کل داده‌ها	داده‌های آموزش
GEP	۲۰۱۶-۲۰۰۷	۸۰۰۰	۶۴۰۰	۸۰۰۰	۱۶۰۰

در میان داده‌های جمع‌آوری شده ۸۰٪ داده‌ها برای آموزش (کلاً ۸۰۰۰ داده، ۶۴۰۰ داده از هر پارامتر برای هر ۳۱ ایستگاه) برای مدل استفاده شد. آزمایش برای ۲۰٪ داده‌ها (کلاً ۱۶۰۰ داده از هر پارامتر برای هر ۳۱ ایستگاه) برای مدل انجام شد. عملکرد مدل بیان ژن از طریق ارزیابی آماره‌های بهترین برازش (Best fitness)، ریشه متوسط خطای مربعات (RMSE)، متوسط قدر مطلق خطا (MAE)، معیار نش-ساتکلیف (NSE) و ضریب تبیین ( $R^2$ ) که در بهترین حالت (حالت آرمانی مدل) مقادیر BestFitness، RMSE، MAE، NSE و  $R^2$  به ترتیب برابر ۱۰۰۰، صفر، صفر، ۱ و ۱ می‌باشند. جدول ۶، پارامترهای ارزیابی مدل بیان ژن را به ازای سناریوهای مختلف در دو مرحله‌ی آموزش و آزمایش نشان می‌دهد.

جدول ۶- نتایج برنامه GEP برای سناریوهای مختلف

دوره‌ی آماری	رتبه	سناریو	آموزش				آزمایش					
			NSE	MAE	$R^2$	Best Fitness	RMS E (mmd <sup>-1</sup> )	Best Fitness	$R^2$	MAE	RMSE (mmd <sup>-1</sup> )	NSE
۲۰۱۶-۲۰۰۷	۳	a	۰/۶۳	۳/۲۲	۰/۶۹	۶۴۷/۴۵	۴/۱۴	۶۴۵/۳۷	۰/۶۹	۳/۲۹	۴/۳۵	۰/۹۸
۲۰۱۶-۲۰۰۷	۴	b	۰/۶۰	۳/۹۰	۰/۶۸	۶۳۴/۷۰	۴/۲۹	۶۴۴/۳۵	۰/۶۹	۳/۳۲	۴/۲۶	۰/۹۸
۲۰۱۶-۲۰۰۷	۱	c	۰/۷۲	۲/۸۲	۰/۷۲	۶۷۶/۷۴	۳/۵۹	۶۷۹/۷۱	۰/۷۳	۲/۸۲	۳/۵۶	۰/۵۸
۲۰۱۶-۲۰۰۷	۵	d	۰/۵۷	۳/۴۶	۰/۶۷	۶۲۹/۲۷۶	۴/۴۲	۶۴۱/۸۶	۰/۶۹	۳/۳۷	۴/۳۹	۰/۹۸
۲۰۱۶-۲۰۰۷	۸	e	۰/۶۶	۳/۲۳	۰/۶۶	۶۴۷/۱۵	۳/۹۷	۶۴۳/۰۳	۰/۶۵	۳/۲۷	۴/۰۳	۰/۹۹
۲۰۱۶-۲۰۰۷	۷	f	۰/۶۷	۳/۱۸	۰/۶۷	۶۵۱/۰۹	۳/۹۳	۶۴۷/۲۷	۰/۶۶	۳/۱۸	۳/۹۴	۰/۹۹
۲۰۱۶-۲۰۰۷	۲	g	۰/۷۱	۲/۸۹	۰/۷۱	۶۷۱/۴۴	۳/۶۳	۶۶۹/۷۷	۰/۷۱	۲/۹۱	۳/۶۷	۰/۹۹
۲۰۱۶-۲۰۰۷	۶	h	۰/۶۷	۳/۱۹	۰/۶۷	۶۴۹/۹۷	۳/۹۶	۶۴۸/۳۳	۰/۶۵	۳/۱۷	۳/۹۹	۰/۹۹

### انتخاب بهترین مدل

برازش و کم بودن مقادیر خطا، مدل مطلوب می‌باشد. در مقایسه با نتایج تحقیقات دیگر، خاشعی سیوکی و همکاران (۱۳۹۲)، نتایج مدل‌های شبکه عصبی مصنوعی (ANN) و سیستم استنتاج فازی (ANFIS) را در برآورد سطح ایستابی آبخوان دشت نیشابور مقایسه کردند. نتایج روش‌های مذکور نشان داد که مدل ANN با  $R^2 = ۰/۹۴$  نسبت به مدل ANFIS از دقت بیشتری برخوردار بود. حقیقت‌جو و همکاران (۱۳۹۶)، به مقایسه روش‌های سیستم استنتاج عصبی- فازی و برنامه‌ریزی بیان ژن در برآورد تبخیر از تشک در خراسان جنوبی پرداختند. مقایسه نتایج آزمون دو مدل نشان داد که مدل برنامه‌ریزی بیان ژن کارایی بهتری نسبت به مدل عصبی- فازی در برآورد روزانه تبخیر از

از مقایسه‌ی نتایج آماره‌های ارزیابی مدل در سناریوهای مختلف در جدول ۷ مشاهده می‌شود که در تمام سناریوها، سناریوی c به دلیل در نظر گرفتن پارامترهای عرض جغرافیایی، بارش، میانگین رطوبت نسبی، حداکثر دما، حداقل دما، میانگین دمای هوا و میانگین تابش خورشیدی پیش‌بینی مطلوب‌تری را ارائه کرده است. همچنین سناریوی e دارای ضعیف‌ترین نتیجه می‌باشد (به دلیل در نظر نگرفتن پارامترهای ارتفاع، طول جغرافیایی، عرض جغرافیایی و بارش). از این رو می‌توان این چنین برداشت کرد که سناریوی c با توجه به بالا بودن ضریب همبستگی، بهترین

است. از این شکل می‌توان دید که مقادیر پیش‌بینی GEP بسیار نزدیک به مقادیر اندازه‌گیری تابش خورشیدی جهانی تقریباً برای تمام مجموعه داده‌هاست. بهترین مدل برای این مطالعه، دارای مقدار  $R^2$  برای آموزش و آزمایش داده‌ها به ترتیب  $0/72$  و  $0/73$  می‌باشد. به طور کلی، مقادیر  $R^2$  به دست آمده از شکل ۲ به وضوح نشان می‌دهد که سناریوی c بهترین مدل برای پیش‌بینی میانگین روزانه تابش خورشیدی جهانی در گستره‌ی ایران است. در شکل ۲- الف و ۲- ب، به ترتیب در بخش آموزش و آزمایش نمودار  $R^2$ ، نشان داده شده که در آن محور طول‌ها مربوط به مقادیر روزانه تابش خورشیدی است که از مدل به دست می‌آید و محور عرض‌ها مقادیر تابش خورشیدی واقعی را نشان می‌دهد. در سناریوی c که بهترین ورودی‌های مدل مربوط به این سناریو می‌باشد که شامل عرض جغرافیایی، بارش، رطوبت نسبی، حداکثر دما، حداقل دما، میانگین دما و خروجی آن تابش خورشیدی در نظر گرفته می‌شود. با توجه به شکل ۲، می‌توان گفت ورودی‌های سناریوی c بهترین ورودی برای این مدل می‌باشد. البته در سناریوهای دیگر هم نتیجه نزدیک به نتیجه سناریو c بوده به جز در سناریو e که نتایج ضعیفی حاصل شده، لذا می‌توان نتیجه گرفت در میان پارامترها، پارامتر حداکثر دما ( $R^2=0/67$ ) تأثیر بیشتری در بدست آوردن جواب بهینه دارد.

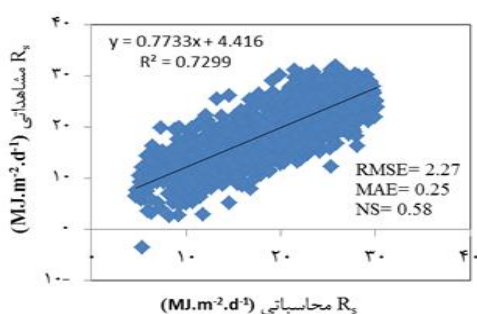
تشتک دارد. سیدو و آیتک<sup>۱</sup>، به بررسی توانایی عملکرد برنامه‌ریزی بیان ژن برای مدل‌سازی تبخیر پتانسیل با استفاده از داده‌های ده ساله کشور ساحلی، بورکینا فاسو پرداختند. نتایج نشان داد که، روش GEP توانایی خوبی برای ارائه مدل بر اساس داده‌های منطقه‌ای را دارد. نیا و همکاران<sup>۲</sup> (۲۰۱۳)، به اندازه‌گیری تابش خورشیدی در الجزایر پرداختند. نتایج نشان داد که بین مقادیر اندازه‌گیری شده و محاسبه شده یک سازگاری خوبی برقرار است و مدل‌ها برای پیش‌بینی میانگین ماهانه تابش خورشیدی در الجزایر و در هر مکان با اقلیم مشابه توصیه می‌شود. نیلامگام و امیرثام<sup>۳</sup> (۲۰۱۶)، به پیش‌بینی تابش خورشیدی برای سیستم‌های خورشیدی با استفاده از مدل‌های شبکه عصبی مصنوعی با الگوریتم‌های مختلف پس انتشار خطا پرداختند. نتایج حاکی از آن است که مدل ANN توسعه یافته برای پیش‌بینی میانگین ماهانه تابش خورشیدی برای طراحی یا ارزیابی تأسیسات انرژی خورشیدی در جایی که امکانات اندازه‌گیری اطلاعات هواشناسی در هند وجود ندارد، مناسب می‌باشد. ترائور و همکاران (۲۰۱۷)، به پیش‌بینی تبخیرتغرق مرجع روزانه با استفاده از اطلاعات آب و هوا و برنامه‌نویسی بیان ژن در چین پرداختند. نتایج نشان داد که، بیان ژن توانایی خوبی برای پیش‌بینی تبخیرتغرق روزانه با کمترین خطا ( $MAE = 0/902$ ,  $RMSE = 1/207$ ) و بیشترین همبستگی ( $R = 0/77$ ) را داشته است.

شکل ۲، عملکرد مقادیر اندازه‌گیری شده و پیش‌بینی شده را نشان می‌دهد که پیش‌بینی میانگین روزانه تابش خورشیدی برای مدل مطلوب GEP (سناریوی c) در مطالعه‌ی حاضر

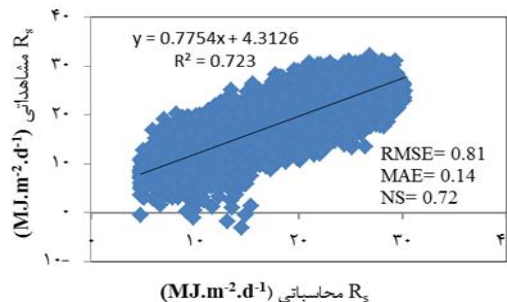
- 1 Seydou & Aytac
- 2 Nia et al.
- 3 Amirtham & Neelamegam

جدول ۷- نتایج GEP مربوط به مدل مطلوب و مدل نامرغوب

آزمایش					آموزش				سناریو	رتبه	دوره‌ی آماری	
NSE	RMSE (mm <sup>d<sup>-1</sup></sup> )	MAE	R <sup>2</sup>	Best Fitness	RMSE (mm <sup>d<sup>-1</sup></sup> )	NSE	MAE	R <sup>2</sup>				Best Fitness
۰/۵۸	۳/۵۶	۲/۸۲	۰/۷۳	۶۷۹/۷۱	۳/۵۹	۰/۷۲	۲/۸۲	۰/۷۲	۶۷۶/۷۴	c	۱	۲۰۱۶-۲۰۰۷
۰/۹۹	۴/۰۳	۳/۲۷	۰/۶۵	۶۴۳/۰۳	۳/۹۷	۰/۶۶	۳/۲۳	۰/۶۶	۶۴۷/۱۵	e	۸	۲۰۱۶-۲۰۰۷



شکل ۲- ب (بخش آزمایش)



شکل ۲- الف (بخش آموزش)

شکل ۲- نتایج برنامه برای مدل مطلوب در بخش آموزش و آزمایش

وجود می‌آیند. به این صورت که اگر هر یک از ثوابت در معادله‌ی نهایی مسئله بیاید علاوه بر اعدادی که خود مسئله ایجاد می‌کند این اعداد نیز در ساده‌تر شدن معادله کمک می‌کنند.

### ثوابت عددی (Numerical Constants)

در جدول ۸ ثابت‌های عددی که به صورت تصادفی در هر یک از ژن‌های کروموزوم برانزنده ایجاد شده، نشان داده شده است. این ثوابت بخاطر ایجاد عدد در معادله‌ی نهایی به

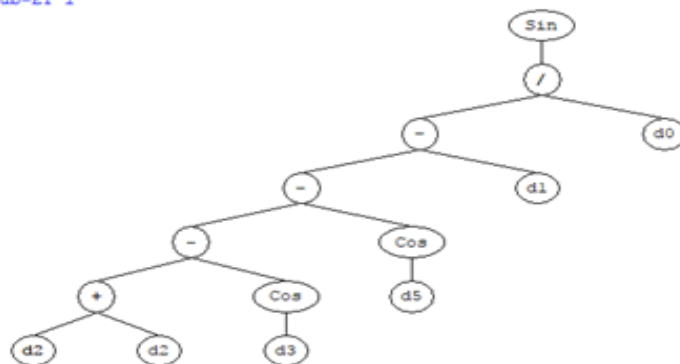
جدول ۸- ثوابت عددی مورد استفاده در ژن‌های کروموزوم (برنامه) مطلوب

ژن چهارم	ژن سوم	ژن دوم	ژن اول
$C_0 = ۳۶۳/۰۱۳۱۸۳۰۸۰۵۳۸$	$C_0 = ۵/۰۲۵۴۸۲۹۵۵۴۱۲۴۶$	$C_0 = -۴/۸۱۶۶۷۹۷۵۵۴۵۵$	$C_0 = ۹/۵۱۴۱۴۵۳۲۹۱۴۲۱۲E-۰۳$
$C_1 = ۶/۳۱۸۲۴۰۶۹۸۵۶۸۶۸ E-۰۲$	$C_1 = -۷/۶۹۰۹۶۹۵۷۳۰۴۶۰۵$	$C_1 = -۵/۲۸۹۱۶۲۸۷۷۲۸۵۰$	$C_1 = ۶/۸۲۶۸۵۳۸۴۶۸۵۸۱E-۰۳$
$C_2 = -۵/۷۶۳۴۲۰۵۱۴۵۴۲۰۷$	$C_2 = ۶/۰۱۴۲۸۲۶۶۲۴۳۴۷۷ E-۰۲$	$C_2 = -۱/۵۵۳۵۰۸۱۶۶۶۳۵۵۵ E-۰۲$	$C_2 = ۷۱/۰۸۶۷۶۴۱۲۲۴۴۰۳$
$C_3 = -۵/۷۱۲۷۵۹۷۸۸۱۱۹۱$	$C_3 = -۷/۹۲۴۷۴۱۳۵۵۶۳۲۱۹$	$C_3 = ۴۵/۷۳۱۹۸۶۴۴۹۷۸۱۸$	$C_3 = -۵/۴۳۵۰۴۱۳۵۲۵۸۰۳۴$
$C_4 = ۵/۸۶۳۵۲۱۲۲۵۶۲۳۳۴$	$C_4 = -۷/۶۷۰۶۹۴۵۸۵۳۰۲۳$	$C_4 = -۰/۷۲۰۸۱۲۷۳۸۴۲۵۸۵۵$	$C_4 = -۵/۷۶۲۲۳۰۲۹۲۶۷۲۵۱E-۰۳$
$C_5 = ۵/۱۴۶۲۳۵۵۴۱۸۵۶۱۴$	$C_5 = -۷/۷۵۰۷۸۵۸۵۱۶۱۹۰۱$	$C_5 = ۵/۴۱۹۷۸۲۰۹۷۸۴۲۳۴ E-۰۲$	$C_5 = ۴/۲۴۲۹۸۸۳۷۲۴۴۷۸۹$
$C_6 = -۸۴/۹۱۷۷۵۲۶۱۶۹۶۲۲$	$C_6 = -۸۴/۸۳۲۸۳۴۸۶۴۳۴۵۲$	$C_6 = ۲۹/۸۷۴۶۳۳۰۱۴۹۲۳۶$	$C_6 = ۱۴۸/۵۵۸۹۱۵۹۸۲۵۴۳$
$C_7 = -۸/۴۰۶۸۵۷۵۰۹۰۷۹۲۶$	$C_7 = -۷/۹۸۶۳۸۸۷۴۴۷۷۳۷۱ E-۰۲$	$C_7 = ۰/۳۱۶۲۵۸۱۱۲۲۱۹۶۱۱$	$C_7 = ۵/۵۴۶۷۳۹۰۹۷۲۶۲۴۹$
$C_8 = -۷/۳۰۰۶۷۷۲۲۹۴۸۳۰۲$	$C_8 = ۹/۴۸۷۲۸۹۰۴۰۸۰۳۲۵$	$C_8 = ۰/۱۸۳۴۱۶۲۴۱۹۵۰۷۴۳$	$C_8 = -۱۰۲/۶۹۵۰۸۹۵۷۱۸۲۵$
$C_9 = ۰/۹۴۳۳۲۷۱۲۷۹۰۳۰۷۳$	$C_9 = -۰/۳۰۱۹۸۰۶۵۱۲۶۴۹۹۲$	$C_9 = ۳/۲۳۹۵۳۹۷۸۰۸۷۷۱$	$C_9 = ۰/۷۲۹۳۹۲۲۶۷۴۷۶۳۳۳$
$C_{10} = -۴/۲۹۸۵۳۲۰۵۹۶۹۴۲$	$C_{10} = ۰/۴۷۹۹۰۳۵۶۱۵۱۰۰۵۶$	$C_{10} = ۸/۶۰۱۶۴۱۸۹۵۸۰۹۸۱$	$C_{10} = ۹/۲۲۶۶۶۰۹۶۹۸۷۸۲۳$

ساختار درختی مدل مطلوب

شکل ۳، ساختار مدل خروجی مطلوب را به صورت درختی نشان می‌دهد و همچنین معادله‌ی حاصل از این ساختار حاصل شده است. از آنجایی که چهار ژن در اینجا تشکیل شده، هریک از ژن‌ها یک زیردرخت و معادله‌ی مربوط به خود را دارد که در نهایت با توجه به تابع پیوند معادله‌ی نهایی حاصل می‌شود.

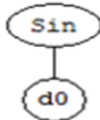
Sub-ET 1



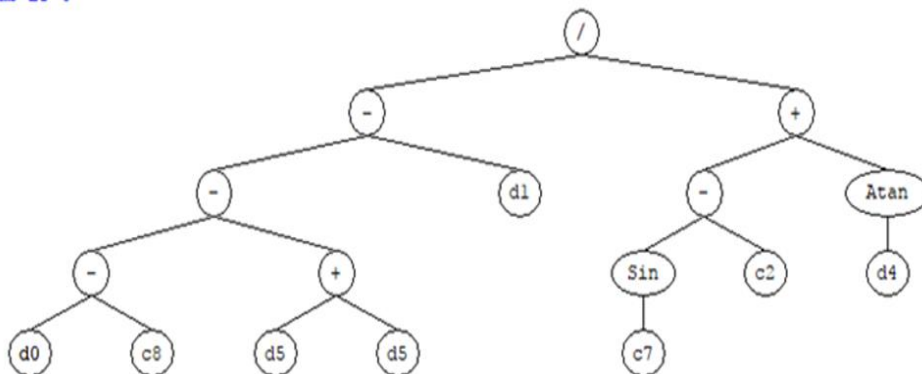
Sub-ET 2



Sub-ET 3



Sub-ET 4



شکل ۳- ساختار درختی مدل مطلوب

عرض جغرافیایی (درجه)، بارش (میلی‌متر)، رطوبت نسبی (درصد)، حداکثر دما (°C)، حداقل دما (°C) و میانگین دما (°C) هستند، ایجاد شده است که در زیر معادلات مربوط به هر زیر درخت نوشته شده است.

شکل ۳، زیردرخت مربوط به ژن اول (رابطه ۶)، دوم (رابطه ۷)، سوم (رابطه ۸) و چهارم (رابطه ۹) را نشان می‌دهد که در این زیردرخت ثوابت عددی تصادفی C7، C2 و C8 و ورودی‌های (d0)، اول (d1)، دوم (d2)، سوم (d3)، چهارم (d4) و پنجم (d5) که در سناریوی c به ترتیب

$$SUB ET_1 = \sin((((d[2]+d[2])-\cos(d[3]))-\cos(d[5]))-d[1])/d[0])) \quad (6)$$

$$SUB ET_2 = d[5] \quad (7)$$

$$SUB ET_3 = \sin(d[0]) \quad (8)$$

$$SUB ET_4 = (((d[0]-G4C8)-(d[5]+d[5]))-d[1])/((\sin(G4C7)-G4C2)+\text{atan}(d[4]))) \quad (9)$$

$$G4C8 = -7.30067722948302$$

$$G4C7 = -8.40685750907926$$

$$G4C2 = -5.76342051454207$$

$$R_s = SUB ET_1 + SUB ET_2 + SUB ET_3 + SUB ET_4 = \sin((((2RH)-\cos(T_{max}))-\cos(T_{mean}))-p/\text{altitude})) + T_{mean} + \sin(\text{altitude}) + (((\text{altitude} + 7.30)-(2 T_{mean}))-p)/((\sin(-8.41)+ 5.76)+\text{atan}(T_{min}))) \quad (10)$$

به دلیل در نظر نگرفتن ارتفاع، طول جغرافیایی، عرض جغرافیایی، بارش، رطوبت نسبی، حداکثر دما و حداقل دما، دارای ضریب همبستگی کمتر و خطای بیشتری می‌باشد و کارایی کمتری دارد. در این مدل، در سناریوی c با اضافه شدن پارامترهای میانگین دما و تابش خورشیدی، برآورد مطلوب‌تری از تابش خورشیدی حاصل می‌شود. نتایج سناریوهای c و e بسیار به هم نزدیک می‌باشد اما سناریوی e در پیش‌بینی تابش خورشیدی ضعیف‌تر عمل می‌کند و بهترین مدل در این مطالعه، سناریوی c می‌باشد. زمانی که عرض جغرافیایی، بارش، رطوبت نسبی، حداکثر دما، حداقل دما، میانگین دما و تابش خورشیدی برای پیش‌بینی تابش خورشیدی در نظر گرفته شود، نتایج دقیق‌تری حاصل می‌شود. هم‌چنین، معیارهای ارزیابی تحت سناریوی c با  $NSE = 0.72$  و  $MAE = 2.82$ ،  $RMSE = 3.59$ ،  $R^2 = 0.72$  بخش آموزش از سایر سناریوها بیشتر می‌باشد. در مجموع روش GEP دقیق‌ترین نتایج را در تخمین تابش خورشیدی روزانه در گستره‌ی ایران دارد.

#### منابع

1. Abarashi, M., H. Moftah Helaghi, A. Sani Khani A. Dehghani, 2013, Comparison of performance of three intelligent methods in

چون تابع پیوند، تابع جمع است باید ژن‌ها را با یکدیگر جمع کرده یا از نگاه دیگر برای به دست آوردن فرمول جواب، باید چهار زیردرخت حاصله از ژن‌ها را با هم جمع کرد. در نهایت رابطه (۱۰) حاصل می‌شود.

#### کیفیت مدل حاصل شده

در این پژوهش سعی بر آن شد عملکرد مدل برنامه‌ریزی بیان ژن، جهت شبیه‌سازی تابش خورشیدی روزانه در گستره‌ی ایران با استفاده از داده‌های موجود مورد ارزیابی قرار گیرد. در مجموع معیارهای ارزیابی مدل نشان داد که مدل پیشنهادی می‌تواند کاربرد موثری در پیش‌بینی تابش خورشیدی روزانه در گستره‌ی ایران داشته باشد.

#### نتیجه‌گیری

در این مطالعه، توانایی مدل برنامه‌ریزی بیان ژن برای پیش‌بینی میانگین روزانه‌ی تابش خورشیدی در ۳۱ ایستگاه در گستره‌ی ایران مورد ارزیابی قرار گرفت. بهترین مدل بر اساس معیارهای ارزیابی  $RMSE$ ،  $MAE$ ،  $NSE$  و  $R^2$  انتخاب شد. از نتایج تجزیه و تحلیل، مشخص شد که مقادیر پیش‌بینی شده سازگاری خوبی با مقادیر اندازه‌گیری شده در مدل برنامه‌ریزی بیان ژن (GEP) دارد. در مدل GEP، سناریوی e

- country (Northern Spain). *Energy Conversion and Management*. No. 62, pp. 1-13.
10. Muzzammil, M., J. Alam, M. Danish, 2015, Application of gene expression programming in flood frequency analysis. *Journal of Indian Water Resources Society*. No. 2, pp. 1-6.75.
  11. Neelamegam, P., V. Arasu Amirtham, 2016, Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. *Journal of Applied Research Technology*. pp. 206- 214.
  12. Nia, M., M. Chegaar, M.F. Benatallah, M. Aillerie, 2013, Contribution to the quantification of solar radiation in Algeria. *Journal of Energy Procedia*, pp. 730- 737.
  13. Roshangar, K. Sh. Mirahidaryan, 2014, Application of evolutionary method of gene expression planning in estimating the scour depth of bridge bases in non-adhesive substrates based on laboratory and field data. *The 8th National Civil Engineering Congress of Babylon*. (In Farsi)
  14. Roshangar, K., H. Bonakdari, S. Akhgar, F. Harsami, 2015, Investigating the Functioning of Gene Expression Programming (GEP) in predicting energy dissipation on stepper overflow in a skimming regime. *10th International Congress on Civil Engineering*.
  15. Seydou, T., G. Aytac, 2012, Regional-Specific numerical models of evapotranspiration using Gene-Expression Programming interface in Sahel. *Journal of Water Resources Management*. pp. 4367-4380.
  16. Traore, S., Y. Luo, G. Fipps, 2017, Gene-Expression Programming for Short-Term Forecasting of Daily Reference Evapotranspiration Using Public Weather Forecast Information. *Journal of Water Resources Management*.
  17. Yazdanpanah, H., R. Mirmaghbian, H. Beraki, H, 2009, Estimation of total solar radiation in the horizontal plane of the Earth in Isfahan. *Journal of Geography and Environmental Planning*, No. 1. order to predict static level fluctuations (Case study: Zarrin Golz plain). *Journal of Water and Wastewater Protection Research*. No. 1, pp. 180-163. (In Farsi)
  2. Ahmadi, F., S. Eisham, K. Khalili, C. Bahmanesh, 2014, Application of Adaptive Neuro-Fuzzy Inference Systems and Genetic Programming for estimating evapotranspiration in the northwest of Iran. *Journal of Water Research Agriculture*. No. 2. (In Farsi)
  3. Ferreira, C. 2001, "Gene expression programming: A new adaptive algorithm for solving problems," *Complex Systems*, No.2, pp. 87-129.
  4. Ghobadian, R., M.A.Ghorbani, M. Khalaj, 2013, Investigating the Functioning of Gene Expression Planning Method in Zangmar River Floodwatching Comparison with Dynamic Waveform. *Journal of Water and Soil*. No. 3, pp. 602-592. (In Farsi)
  5. Ghorbani, M.A. M.R. Salehi, 2011, Application of Gene Expression Program in Investigating Changes in Groundwater Quality Data Using Water Drops in Borkhar, Isfahan. *Sixth National Congress on Civil Engineering*. (In Farsi)
  6. Haghightajo, P., Z. Mohammadzadeh Shahroodi, A. Mohammad Reza Zapor, 2017, Comparison of neuro-fuzzy inference system and gene expression planning in estimating evaporation from pan (Case Study: Southern Khorasan Province). *Journal of Water and Soil Conservation*. No. 4. (In Farsi)
  7. KhasheiSiuki, A., B. Ghahraman, M. Kuchakzadeh, 2013, Comparison of Artificial Neural Network Models, ANFIS and Regression in Estimating Staging Level of the Aquifer in Neishabour Plain. *Journal of Irrigation and Drainage*. No.1, pp. 22-10. (In Farsi)
  8. Kisi, O., J. Shiri, M. Tombul, 2013, Modeling rainfall-runoff process using soft computing techniques. *Computers & Geosciences*. No. 51, pp. 108-117.
  9. Landaras, G., J. Javier Lopez, O. Kisi, J. Shiri, 2012. Comparison of Gene Expression Programming with neuro- fuzzy and neural network computing techniques in estimating daily incoming solar radiation in the Basque