

## برآورد روزانه تبخیر و تعرق مرجع در دشت سیستان با استفاده از الگوریتم‌های فراابتکاری

هادی سیاسر<sup>۱\*</sup>، امیر سالاری<sup>۲</sup>، ام‌البنین محمدرضاپور<sup>۳</sup>، حلیمه پیری<sup>۴</sup>

تاریخ دریافت: ۱۳۹۹/۱۰/۱

تاریخ پذیرش: ۱۴۰۰/۲/۲۱

### چکیده

تبخیر و تعرق یکی از مهم‌ترین پارامترهای مؤثر در اعمال صحیح مدیریت منابع آب بوده و روش‌های متعدد مستقیم و غیرمستقیمی برای اندازه‌گیری آن وجود دارد. این روش‌ها اصولاً وقت‌گیر، پرهزینه و نیازمند داده‌های هواشناسی زیادی هستند. هدف از اجرای این تحقیق، محاسبه تبخیر و تعرق گیاه مرجع در دشت سیستان با استفاده از الگوریتم‌های فراابتکاری است. الگوریتم‌های فراابتکاری از جمله روش‌های برآورد با دقت و سرعت بالا بدون نیاز به حجم زیادی داده است و تاکنون مطالعات فراوانی در خصوص ارائه روش‌های تخمین تبخیر-تعرق گیاه مرجع (ETO) با استفاده از سیستم‌های هوشمند صورت گرفته است. به این منظور، در این پژوهش به بررسی امکان پیش‌بینی این مؤلفه مهم در شمال استان سیستان و بلوچستان با استفاده از مدل‌های برنامه‌ریزی بیان ژن و یادگیری عمیق پرداخته شد. ابتدا بر اساس رابطه فائو-پنمن-مانتیت، میزان تبخیر و تعرق پتانسیل روزانه در ایستگاه سینوپتیک زابل با استفاده از داده‌های هواشناسی روزانه شامل دمای بیشینه، دمای کمینه، دمای میانگین، رطوبت نسبی حداکثر، رطوبت نسبی حداقل، رطوبت نسبی میانگین، ساعات آفتابی، سرعت باد، بارش و تبخیر از تشت در طول دوره آماری ۱۳۸۸ تا ۱۳۹۶ محاسبه شد. با ارائه الگوهای مختلف شامل ترکیبی از پارامترهای هواشناسی به‌عنوان ورودی‌های مدل در مقیاس زمانی روزانه، مقدار تبخیر و تعرق توسط مدل‌های پیشنهادی به‌عنوان خروجی مدل برآورد شد. همچنین قابلیت پیش‌بینی این مدل‌ها، در مقایسه نتایج آن‌ها با نتایج روش فائو-پنمن-مانتیت به‌عنوان روش مبنا ارزیابی شد و این مقادیر به‌عنوان مرجع برای مقایسه نتایج مدل‌های مورد مطالعه در تحقیق استفاده گردید. مقایسه نتایج در مدل‌های مختلف بر اساس آماره‌های ضریب تبیین و جذر میانگین مربعات خطا صورت گرفت. نتایج نشان داد مدل‌های برنامه‌ریزی بیان ژن و یادگیری عمیق دارای دقت بالایی در برآورد میزان تبخیر و تعرق بوده و در تمامی سناریوها، مدل یادگیری عمیق از دقت بالاتری نسبت به مدل بیان ژن برخوردار است. در مدل یادگیری عمیق در بین تمامی سناریوها، سناریوی M5 با متغیرهای دمای حداکثر، دمای حداقل، دمای میانگین، رطوبت حداکثر، رطوبت حداقل، رطوبت میانگین، سرعت باد و تبخیر از تشت با کمترین خطا ( $RMSE=0.517$ ) و بیشترین ضریب تبیین ( $R^2=0.996$ ) و در مدل برنامه‌ریزی بیان ژن سناریوی M1 با متغیرهای میانگین دما، دمای کمینه، دمای بیشینه و حداکثر رطوبت با بیشترین ضریب تبیین  $R^2=0.985$  و کمترین خطا  $RMSE=0.985$  حائز بیشترین دقت شدند. نتیجه کلی این تحقیق، توصیه کاربرد مدل یادگیری عمیق برای برآورد تبخیر و تعرق منطقه سیستان است.

**کلیدواژه‌ها:** تبخیر-تعرق، مدل یادگیری عمیق، مدل GEP.

۱. استادیار، گروه کشاورزی، دانشگاه پیام نور، ایران، نویسنده مسئول، hadisiasar@pnu.ac.ir

۲. استادیار، عضو هیئت‌علمی، گروه مهندسی آب، دانشکده کشاورزی، مجتمع آموزش عالی میناب، دانشگاه هرمزگان

۳. دانشیار، عضو هیئت‌علمی، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه زابل

۴. استادیار، عضو هیئت‌علمی، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه زابل

\* این مقاله برگرفته از پژوهش مستقل است.

## مقدمه

محدود بودن حجم منابع آبی، رشد جمعیت و افزایش تقاضای مواد غذایی لزوم بهینه‌سازی مصرف آب را ایجاب کرده است. تبخیر و تعرق به‌عنوان یکی از عوامل مهم چرخه هیدرولوژیکی نقش مهمی در بهینه‌سازی مصرف آب دارد. تاکنون روش‌های متعددی برای محاسبه تبخیر و تعرق بر مبنای داده‌های اقلیمی پیشنهاد و پژوهش‌ها و تحقیقات زیادی بر ارزیابی این روش‌ها متمرکز شده است (فنگ<sup>۱</sup> و همکاران، ۲۰۱۷؛ جاو<sup>۲</sup> و همکاران، ۲۰۱۷؛ گوسیک<sup>۳</sup> و همکاران، ۲۰۱۶). در این میان، ترکیب معادله‌های توازن انرژی و آترودینامیک به این لحاظ که اساس آن‌ها بر روابط منطقی و مبنای فیزیکی استوار است، دقیق‌ترین نتایج را به دست می‌دهد (جنسن<sup>۴</sup> و همکاران، ۱۹۹۰). سازمان خواربار جهانی روش فائو-پنمن-مانتیت<sup>۵</sup> را به‌عنوان روش استاندارد محاسبه تبخیر و تعرق گیاهان معرفی کرده است (آلن<sup>۶</sup> و همکاران، ۱۹۹۸). نتایج تحقیقات مختلف در ایران نیز نشان می‌دهد که روش فائو-پنمن-مانتیت بهترین روش برآوردی تبخیر و تعرق است (موسوی‌بایگی و همکاران، ۲۰۰۹). مهم‌ترین محدودیت کاربرد مدل مذکور نیاز به پارامترهای اقلیمی زیادی است که ممکن است برای یک منطقه خاص در دسترس نباشد. از این‌رو روش‌های تجربی و نیمه‌تجربی زیادی نیز برای تخمین  $ET_0$  با استفاده از داده‌های محدودتر، ابداع شده است. از نقاط ضعف روش‌های تجربی این است که این روش‌ها عموماً منطقه‌پایه بوده و فقط تحت شرایط اقلیمی و اگرونومیکی خاصی دارای اعتبار هستند (وانگ<sup>۷</sup> و همکاران، ۲۰۱۳). در سالیان اخیر ابزارهای یادگیری و روش‌های نوین هوشمند برای حل برخی از مشکلات ارائه شده و به‌طور گسترده‌ای مورد استفاده قرار گرفته است، پیش‌بینی‌های هواشناسی با استفاده از روش‌های عددی، راهکاری مناسب برای پیش‌بینی تبخیر و تعرق مورد استفاده قرار گیرد (تیان و مارتینز<sup>۸</sup>، ۲۰۱۲). یکی از روش‌های فراابتکاری

فراابتکاری در مدل‌سازی، مدل برنامه‌ریزی بیان ژن است که در مطالعات کشاورزی و هیدرولوژیکی به‌صورت موفقیت‌آمیزی مورد استفاده قرار گرفته است. نتایج مطالعات نشان می‌دهد که مدل GEP در مقایسه با مدل شبکه عصبی مصنوعی و روش‌های آماری از دقت بهتری در تخمین تبخیر و تعرق ساعتی برخوردارند (ایزدی‌فر و الشورباگی، ۲۰۱۰). نتایج مقایسه مدل‌های GEP، سیستم استنتاج تطبیقی عصبی-فازی و روش‌های تجربی هارگریوز-سامانی و پرستلی-تیلور برای تخمین تبخیر و تعرق مرجع در شمال اسپانیا حاکی از آن است که GEP به‌ترتیب دقیق‌تر از سیستم استنتاج عصبی-فازی، پرستلی-تیلور و هارگریوز-سامانی می‌باشد (شیری و همکاران، ۲۰۱۲). همچنین بررسی‌های محققان در مقایسه نتایج روش برنامه‌ریزی بیان ژن و روش‌های هارگریوز-سامانی و بلانی‌کریدل نشان داد که مدل برنامه‌ریزی بیان ژن می‌تواند به‌خوبی جایگزین روش‌های تجربی شود (تراوره و گوون<sup>۹</sup>، ۲۰۱۲).

در تحقیقی در جنوب غربی ترکیه با مقایسه برنامه‌ریزی بیان ژن (GEP) و سیستم تطبیقی مبتنی بر استنتاج عصبی فازی (ANFIS) به‌منظور برآورد تبخیر روزانه مشخص شد که مدل GEP نسبت به مدل ANFIS از دقت بالاتری برخوردار است (ترزی<sup>۱۰</sup>، ۲۰۱۲).

از دیگر روش‌های نوین هوشمند مدل‌سازی استفاده از الگوریتم یادگیری عمیق<sup>۱۱</sup> است. DL در واقع نگرش جدیدی به ایده شبکه‌های عصبی دارد. DL به مجموعه الگوریتم‌هایی گفته می‌شود که در تلاش‌اند تا مفاهیم انتزاعی و پیچیده را در سطوح و لایه‌های مختلف مدل کنند (گودفلو<sup>۱۲</sup> و همکاران، ۲۰۱۶؛ آرل<sup>۱۳</sup> و همکاران، ۲۰۱۰). مطالعات مختلفی در این زمینه انجام شده است که این روش‌های یادگیری جدید با بهره‌گیری از رگرسیون‌های گروهی در حال توسعه‌اند (بروتن<sup>۱۴</sup> و همکاران، ۲۰۰۰؛ کشکین<sup>۱۵</sup> و همکاران، ۲۰۰۹؛ پاراسورامن<sup>۱۶</sup> و همکاران، ۲۰۰۷).

9. Traore and Guven  
10. Terzi  
11. Deep learning (DL)  
12. Goodfellow  
13. Arel  
14. Bruton  
15. Keskin  
16. Parasuraman

1. Feng  
2. Gao  
3. Gocic  
4. Jensen  
5. PMF-56  
6. Allen  
7. Wang  
8. Tian and Martinez

بررسی سناریوهای مختلف از ترکیب پارامترهای هواشناسی به‌عنوان ورودی مدل برنامه‌ریزی بیان ژن و یادگیری عمیق و استخراج بهترین الگوها توسط مدل‌های برنامه‌ریزی بیان ژن و یادگیری عمیق بر اساس ارزیابی است.

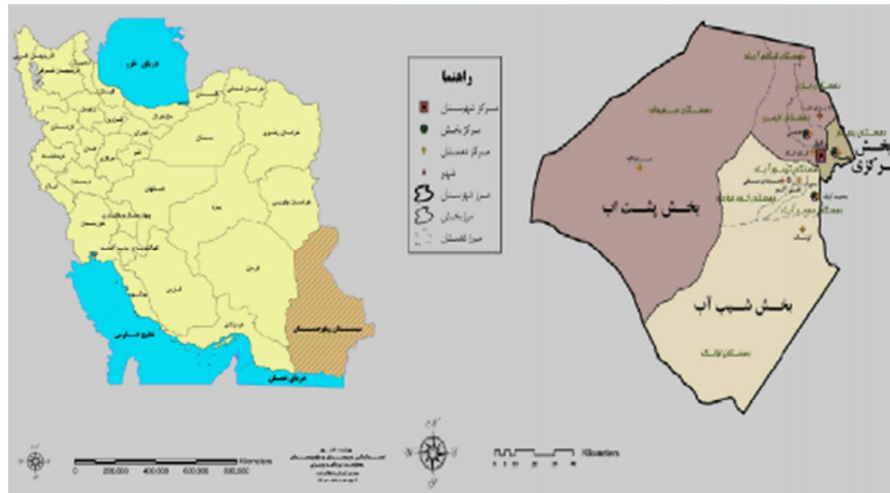
### مواد و روش‌ها

دشت سیستان در شمال استان سیستان و بلوچستان در عرض شمالی ۱۸'، ۳۰' تا ۲۰'، ۳۱° و طول شرقی ۱۰'، ۶۱° تا ۵۰'، ۶۱° واقع است. دشت سیستان با میانگین بارش سالیانه ۵۰ میلی‌متر و میزان تبخیر سالیانه ۴۰۰۰ تا ۵۰۰۰ میلی‌متر از شرایط محیطی کشت مناسبی برخوردار نبوده و بر اساس شاخص خشکی دومارتن جزء نواحی فراخشک محسوب می‌شود. موقعیت دشت مورد مطالعه در شکل (۱) آمده است.

ژو<sup>۱</sup> و همکاران (۲۰۱۷) از یک روش مبتنی بر مدل یادگیری عمیق برای شناسایی سطح زیر کشت بادام زمینی با ماهواره لندست ۸ در کشور چین استفاده کردند. نتایج تجربی به‌دست‌آمده در این تحقیق با دقت ۹۶/۴۲٪ نشان‌دهنده عملکرد خوب این مدل بوده است.

نتایج یک تکنیک یادگیری عمیق به نام SAE برای پیش‌بینی چهار پارامتر آب‌وهوایی دما، سرعت باد، فشار سطح متوسط دریاها و دمای نقطه شبنم با استفاده از کلان‌داده‌های هواشناسی حاکی از موفقیت‌آمیز بودن روش‌های مبتنی بر یادگیری عمیق در پیش‌بینی پارامترهای آب و هواشناسی است (لیو<sup>۲</sup> و همکاران، ۲۰۱۴).

هدف از انجام این تحقیق، شناسایی پارامترهای مؤثر در محاسبه تبخیر و تعرق مرجع روزانه در دشت سیستان و



شکل (۱): موقعیت جغرافیایی شهرستان زابل

Figure (1): geographic location of Zabol city

$$E_{PM56} = \frac{0.408\Delta(R_n - G)}{\Delta + \gamma(1 + 0.34u_2)} + \frac{900}{[1 + 0.34u_2]} \frac{u_2(e_s - e_a)}{(T + 273)} \quad (1)$$

در رابطه فوق  $ET_{PM56}$  تبخیر و تعرق گیاه مرجع ( $mm \cdot day^{-1}$ )، شیب منحنی فشار بخار اشباع ( $kpaC^{-1}$ )، تابش خالص در سطح پوشش گیاهی ( $MJ m^{-2} d^{-1}$ )، چگالی شار گرمای خاک ( $MJ m^{-2} d^{-1}$ )، ضریب سایکرومتری رطوبتی ( $Kpa C^{-1}$ )، میانگین دمای روزانه ( $^{\circ}C$ )،  $u_2$  سرعت باد در ارتفاع دومتری ( $m s^{-1}$ )،  $e_s$  فشار بخار اشباع ( $Kpa$ ) و  $e_a$  فشار بخار واقعی ( $Kpa$ ) است. در تحقیق حاضر با استفاده از داده‌های هواشناسی روزانه، تبخیر و تعرق مرجع روزانه تعیین شد.

### معادله پنمن-مانتیت

محققان بسیاری در سراسر دنیا روش فائو-پنمن-مانتیت را که به‌صورت تئوری بر اساس توازن انرژی روی یک سطح مرطوب پوشیده از گیاه ارائه شده و دارای مبنای فیزیکی است و پارامترهای فیزیولوژیکی و آئرودینامیکی را هم زمان در نظر می‌گیرد، به‌عنوان بهترین روش تعیین تبخیر و تعرق مرجع توصیه کرده‌اند (آلن و همکاران، ۱۹۹۸). بنابراین در این مطالعه، از این روش به‌عنوان روش مبنا (رابطه ۱) استفاده شد.

1. Zhu
2. Liu

برنامه‌ریزی بیان ژن<sup>۱</sup>

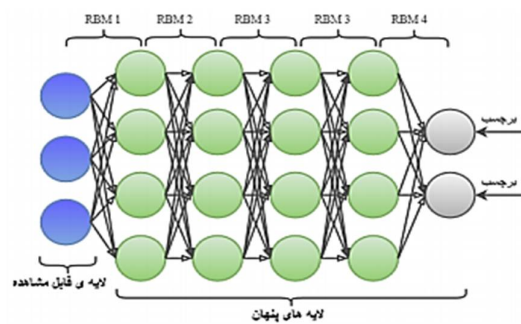
## یادگیری عمیق

در سال‌های اخیر، پیشرفت‌های زیادی در حوزه یادگیری عمیق صورت گرفته که سبب ایجاد روش‌های گوناگونی در این حوزه شده است. ایده اصلی شبکه‌های مبتنی بر یادگیری عمیق، نزدیک شدن هرچه بیشتر یادگیری ماشین به هدف اصلی‌اش، یعنی هوش مصنوعی، می‌باشد. شبکه باور عمیق یکی از قدرتمندترین الگوریتم‌های یادگیری عمیق (DL)<sup>۳</sup> است (هینتون<sup>۴</sup> و همکاران، ۲۰۰۶).

در نگاه اول، دو شبکه باور عمیق و پرسپترون چندلایه<sup>۵</sup> از لحاظ ساختار شبکه یکسان به نظر می‌آیند. این در حالی است که شبکه‌های باور عمیق از چند لایه ماشین بولتزمن محدود شده<sup>۶</sup> ساخته شده‌اند، در واقع ماشین بولتزمن نوع خاصی از میدان‌های تصادفی مارکوف هستند که از شبکه‌های متقارن با واحدهای تصادفی باینری تشکیل شده‌اند (کیوان‌راد و همایون‌پور، ۲۰۱۵).

با وجود  $p$  نرون در لایه ورودی و خروجی، برقراری تابع  $f: \mathbb{R} \rightarrow \mathbb{R}$  در لایه پنهان، وجود نرون  $q$  به عنوان تابع انتقال و  $x \in I$  به عنوان مقدار ویژگی به دست آمده، ورودی لایه  $\mathbb{R}^p$  آمین نرون در لایه پنهان برای تمام مقادیر از رابطه (۲) محاسبه می‌شود (کیوان‌راد و همایون‌پور، ۲۰۱۵) (شکل ۲).

$$h_i(h) = f(W_i^T x + b_i) \quad (2)$$



شکل (۲): ساختار شبکه باور عمیق

Figure (2): Deep belief network structure

روش برنامه‌ریزی بیان ژن را که شکل توسعه یافته برنامه‌ریزی ژنتیک است، فریرا<sup>۲</sup> ارائه کرد (فریرا، ۲۰۰۱). در این روش ابتدا برای مسئله یک تابع هدف مشخص شده، سپس با استفاده از یک فرایند گام به گام، با تغییر ساختار تابع، بهترین تابع هدف با کمترین خطا به دست می‌آید (گون و همکاران، ۲۰۰۸؛ شیر و همکاران، ۲۰۱۲). بر اساس این تحقیقات، کاربرد تابع برازش ریشه میانگین مربعات خطا دقیق‌ترین نتایج را در مدل‌سازی ET<sub>0</sub> حاصل کرده است (شیری و همکاران، ۲۰۱۲).

در پژوهش حاضر، برای اجرای مدل برنامه‌ریزی بیان ژن از برنامه GeneXproTools (4.0) استفاده شد، که در آن انتخاب جمعیت‌های اولیه که همان الگوهای ورودی می‌باشد، از اهمیت بالایی برخوردار است. پارامترها و مقادیر مورد استفاده برای تخمین مقدار تبخیر و تعرق پتانسیل با استفاده از روش برنامه‌ریزی بیان ژن در جدول (۱) ارائه شده است. انتخاب این عملگرها با توجه به مطالعات شیری و همکاران (۲۰۱۲) صورت گرفته است.

جدول (۱): توابع و مقادیر مورد استفاده در مدل GEP

Table (1): Functions and values used in the GEP model

مجموعه توابع	+	-	×	÷	$\sqrt{\quad}$	$\sqrt[3]{\quad}$	$\log$	$10^x$
تعداد کروموزوم‌ها	۳۰							
طول رأس	۸							
تعداد ژن‌ها در هر کروموزوم	۳							
تابع پیوند	جمع (+)							
تابع خطای برازش (RMSE)	ریشه میانگین مربعات خطا							
نرخ جهش	۰/۰۴۴							
نرخ وارون‌سازی	۰/۱							
نرخ ترانهش درجه متوالی	۰/۱							
نرخ ترانهش ریشه درجه متوالی	۰/۱							
نرخ ترکیب تک‌نقطه‌ای	۰/۳							
نرخ ترکیب دونقطه‌ای	۰/۳							
نرخ ترکیب ژنی	۰/۱							
نرخ ترانهش ژن	۰/۱							

3. Deep Learning
4. Hinton
5. Multi Layer Perceptron
6. Restricted Boltzmann Machine

1. GEP
2. Ferreira

ضریب تبیین که هرچه به یک نزدیکتر باشد، نشان‌دهنده همبستگی بالای مدل است. آماره RMSE (ریشه میانگین مربعات خطا)، آماره‌ای است که با مقایسه اختلاف بین مقادیر پیش‌بینی‌شده و واقعی، عملکرد مدل را نشان داده و مقادیر کمتر آن، نشان‌دهنده عملکرد بهتر مدل است، این آماره کم یا بیش برآوردی را نشان نداده و هم‌بعد با پارامترهایی است که از آن به دست آمده است. آماره MAE میانگین قدر مطلق خطا که هرچه این شاخص به صفر نزدیکتر باشد، نشان‌دهنده انحراف کمتر و دقت بالاتر مدل است. مقدار مثبت آماره MAE نشان‌دهنده تخمین بیشتر و مقدار منفی این آماره نشان‌دهنده تخمین کمتر از میزان حقیقی است. آماره MBE (میانگین خطای گرایش) نیز شاخصی است که به مقایسه بین مقادیر پیش‌بینی‌شده توسط مدل و مقادیر واقعی می‌پردازد، مقادیر نزدیک به صفر این شاخص نشان‌دهنده تطابق خوب مقادیر برآوردی و واقعی می‌باشد، مقادیر مثبت آن نشان‌دهنده بیش‌برآوردی و مقادیر کمتر آن نشان‌دهنده کم‌برآوردی است. برای انتخاب مناسب‌ترین مدل، بر اساس معیارهای خطای ذکر شده، از روش رتبه‌بندی برای داده‌های تست استفاده می‌شود. بدین معنا که به هر یک از مدل‌ها بر اساس معیارهای ذکر شده رتبه‌ای داده می‌شود و در پایان پس از مقایسه مجموع رتبه‌های به دست آمده، مناسب‌ترین مدل برای تخمین بهترین معادله انتخاب می‌شود. شیوه رتبه‌بندی این مدل‌ها بدین صورت است که در هر روش، مدلی که کمترین میزان خطای RMSE را دارد، پایین‌ترین رتبه، یعنی رتبه ۱ و به مدل‌ها یا ترکیبات ورودی دیگر در هر یک از دسته‌های ورودی برحسب میزان معیار خطاهای مذکور، رتبه‌های ۲ تا آخر تعلق می‌گیرد. اما در مورد معیار ضریب تبیین ( $R^2$ )، به مدل یا ترکیبی که بیشترین مقدار این معیارها را دارد، رتبه ۱ اختصاص می‌یابد. همچنین به مدلی که میزان معیار خطای آن‌ها یکسان باشند، رتبه یکسان تعلق می‌گیرد. در نهایت پس از انتخاب مدل‌ها یا ترکیبات مناسب هر روش و مقایسه بین آن‌ها، روش برتر در برآورد تبخیر و تعرق مرجع در منطقه مورد مطالعه انتخاب می‌شود (هیتون و همکاران، ۲۰۰۶). در این پژوهش، برای ارزیابی توانایی مدل‌های برنامه‌ریزی بیان ژن و یادگیری

در مدل‌هایی نظیر پرسپترون چندلایه، آموزش با کمک شیوه انتشار رو به عقب صورت می‌گیرد؛ در صورتی که استفاده از همین روش آموزشی در شبکه‌های باور عمیق، سبب ایجاد خطای گرادبان محوشونده<sup>۲</sup> می‌شود (گولوروت و بنجیو<sup>۳</sup>، ۲۰۱۰). شیوه آموزش در شبکه‌های باور عمیق به این صورت است که ابتدا اولین ماشین بولتزمن محدودشده آموزش می‌بیند، سپس لایه پنهان<sup>۴</sup> آن، نقش لایه قابل مشاهده دومین ماشین بولتزمن محدودشده را ایفا می‌کند و آموزش دومین ماشین بولتزمن محدودشده با استفاده از خروجی ماشین بولتزمن اول صورت می‌گیرد. این روند تا آموزش تک‌تک لایه‌های مدل تداوم می‌یابد (هیتون و همکاران، ۲۰۰۶).

### معیارهای آماری برای مقایسه مدل‌ها

در این تحقیق به منظور مقایسه مدل‌های مختلف از شاخص‌های آماری ضریب تبیین ( $R^2$ ) برای تعیین همبستگی بین مقادیر پیش‌بینی‌شده و اندازه‌گیری‌شده و همچنین میانگین قدر مطلق خطا (MAE) برای نشان دادن میزان همخوانی میان مجموعه‌های داده‌های مقادیر مشاهداتی و مدل‌سازی‌شده، جذر میانگین مربعات خطا (RMSE) که میزان خطا را بیان می‌کند و خطای میانگین گرایش (MBE) که میزان کم یا بیش‌برآوردی را نشان می‌دهد، به عنوان معیارهای ارزیابی استفاده شد (روابط ۳، ۴، ۵ و ۶).

$$R^2 = \left[ \frac{\sum_{i=1}^N (O_i - \bar{O}_i) \sum_{i=1}^N (P_i - \bar{P}_i)}{\sqrt{\sum_{i=1}^N (O_i - \bar{O}_i)^2 \sum_{i=1}^N (P_i - \bar{P}_i)^2}} \right]^2 \quad (3)$$

$$MAE = \frac{\sum_{i=1}^N (|O_i - P_i|)}{N} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_i - P_i)^2}{n}} \quad (5)$$

$$MBE = \frac{1}{N} \sum_{i=1}^N (P_i - O_i) \quad (6)$$

در این روابط،  $O_i$  تبخیر و تعرق به روش فائو-پنمن-مانتیش؛  $P_i$  تبخیر و تعرق محاسبه‌شده به هر روش،  $R^2$

1. Backward Propagation Machine
2. Vanishing Gradient
3. Glorot and Bengio
4. Visible Layer

عمیق، ۲۵ سناریو با ترکیب پارامترهای مختلف اقلیمی به عنوان برنامه‌ریزی بیان ژن و مدل یادگیری عمیق با سناریوهای ورودی مدل ایجاد شد (جدول ۳). همچنین با محاسبه مقادیر تعریف‌شده مختلف ایجاد شد و عملکرد آن‌ها مورد بررسی تبخیر و تعرق مرجع از روش فائو-پنمن-مانیتث و در نظر قرار گرفت. گرفتن این مقادیر به عنوان مقادیر هدف در مدل‌های

جدول (۳): سناریوهای در نظر گرفته شده برای ساختار شبکه

Table (3): Scenarios for network structure

سناریو	پارامترهای ورودی شبکه	تعداد متغیر
۱	دمای حداکثر، دمای حداقل، دمای میانگین، رطوبت حداکثر، رطوبت حداقل، رطوبت میانگین، بارش، ساعات آفتابی، سرعت باد، تبخیر از تشت	۱۰
۲	دمای حداقل، دمای میانگین، رطوبت حداکثر، رطوبت حداقل، رطوبت میانگین، بارش، ساعات آفتابی، سرعت باد، تبخیر از تشت	۹
۳	دمای حداکثر، دمای حداقل، دمای میانگین، رطوبت حداکثر، رطوبت حداقل، رطوبت میانگین، ساعات آفتابی، سرعت باد، تبخیر از تشت	۹
۴	دمای حداکثر، دمای حداقل، دمای میانگین، رطوبت حداکثر، رطوبت حداقل، رطوبت میانگین، ساعات آفتابی، تبخیر از تشت	۸
۵	دمای حداکثر، دمای حداقل، دمای میانگین، رطوبت حداکثر، رطوبت حداقل، رطوبت میانگین، سرعت باد، تبخیر از تشت	۸
۶	دمای حداکثر، دمای حداقل، دمای میانگین، رطوبت حداکثر، رطوبت حداقل، رطوبت میانگین، سرعت باد، تبخیر از تشت	۷
۷	دمای حداکثر، دمای حداقل، دمای میانگین، رطوبت میانگین، سرعت باد، تبخیر از تشت	۶
۸	دمای حداکثر، دمای میانگین، رطوبت میانگین، سرعت باد، تبخیر از تشت	۵
۹	دمای میانگین، رطوبت میانگین، سرعت باد، تبخیر از تشت	۴
۱۰	دمای حداکثر، رطوبت میانگین، سرعت باد، تبخیر از تشت	۴
۱۱	دمای حداکثر، رطوبت میانگین، تبخیر از تشت	۳
۱۲	دمای حداکثر، سرعت باد، تبخیر از تشت	۳
۱۳	دمای حداکثر، سرعت باد	۲
۱۴	دمای حداکثر	۱
۱۵	سرعت باد	۱
۱۶	تبخیر از تشت	۱
۱۷	دمای حداکثر، دمای حداقل	۲
۱۸	رطوبت میانگین	۱
۱۹	دمای حداکثر، تبخیر از تشت	۲
۲۰	دمای حداکثر، رطوبت میانگین	۲
۲۱	دمای حداکثر، رطوبت حداکثر	۲
۲۲	دمای میانگین، سرعت باد	۲
۲۳	دمای حداکثر، سرعت باد	۲
۲۴	ساعات آفتابی	۱
۲۵	دمای حداقل	۱

### برنامه‌ریزی بیان ژن

توانایی ارائه رابطه صریح برای تخمین تبخیر و تعرق مرجع

مدنظر قرار گرفت. برای استخراج روابط بین تبخیر و تعرق

مرجع و متغیرهای مورد نظر ابتدا مجموعه داده‌های تبخیر و

استفاده از برنامه‌ریزی بیان ژن به دلیل توان انتخاب متغیرهای

مؤثر در مدل و حذف متغیرهایی با تأثیر کمتر و همچنین

مقادیر MAE به ترتیب برابر با ۴/۶۲۱، ۴/۴۳۸، ۳/۱۹۸ و ۲/۳۵۵ و بیشترین دقت نیز مربوط به مدل‌های M1، M3، M13 و M7 با مقادیر MAE به ترتیب برابر با ۰/۶۸۳، ۰/۷۳۳، ۰/۷۸۰ و ۰/۹۹۱ می‌باشد، در مدل بیان ژن نیز تمامی سناریوها، دارای بیش‌برآورد به دست آمدند. شاخص‌های آماری RMSE و  $R^2$  نیز این نتایج را تأیید می‌کند. در مجموع، نتایج نشان داد مدل‌های برنامه‌ریزی بیان ژن و یادگیری عمیق از دقت بسیار بالا و میانگین خطای گرایش (MBE) اندکی در برآورد تبخیر و تعرق پتانسیل داشته و در تمامی سناریوها، مدل‌های یادگیری عمیق از دقت بیشتری نسبت به بیان ژن برخوردارند. یکی از قابلیت‌های مدل GEP این است که علاوه بر استخراج فرم صریح رابطه مورد نظر و ضرایب مرتبط به آن، متغیرهای مؤثر را نیز شناسایی کرده و متغیرهایی را که تأثیر کمتری بر خروجی دارند، حذف می‌کند. لذا با توجه به خروجی‌های مدل GEP به ترتیب متغیرهای میانگین دما، دمای کمینه، دمای بیشینه، حداکثر رطوبت مهم‌ترین پارامترها هستند که تأثیر بیشتری در پیش‌بینی مقادیر  $ET_0$  دارند (جدول ۴). پارامترهای مورد استفاده و نرخ آن‌ها در استخراج مدل‌هایی برای تخمین تبخیر و تعرق پتانسیل با استفاده از روش برنامه‌ریزی بیان ژن با استفاده از سایر تحقیقات به دست آمد (شیری و همکاران، ۲۰۱۲).

کمترین ضرایب تبیین ( $R^2$ ) در هر دو مدل بیان ژن و یادگیری عمیق در سناریو M24 به ترتیب برابر با ۰/۴۳۵ و ۰/۴۵۳ به دست آمد، این سناریو فقط از پارامتر ساعات آفتابی استفاده کرد و این نشان‌دهنده اهمیت به مراتب کمتر این پارامتر نسبت به سایر پارامترهای هواشناسی دارد. کمترین همبستگی بعد از این سناریو، در هر دو مدل بیان ژن و یادگیری عمیق مربوط به سناریو M15 با ضریب تبیین به ترتیب برابر با ۰/۴۶۹ و ۰/۵۰۶ به دست آمد. در این سناریو تنها پارامتر مورد استفاده سرعت باد است. در مجموع، ساعات آفتابی و سرعت باد دو پارامتری بودند که کمترین تأثیر در برآورد تبخیر و تعرق توسط مدل‌های بیان ژن و یادگیری عمیق را به خود اختصاص دادند (جدول ۳ و ۴).

تعرق مرجع و پارامترهای مورد نظر توسط نرم‌افزار فراخوانی شده و سپس با معرفی توابع و عملگرهای ریاضی مختلف و تعیین مشخصات مدل از قبیل اندازه جمعیت، تعداد نسل، ماکزیمم عمق درخت، حداکثر تعداد زیربرنامه‌ها در برنامه نهایی تولید شده و پس از اجرای برنامه برای داده‌های مورد نظر، کسب همگرایی و تولید بهترین برنامه با ساده‌سازی آن، رابطه مورد نظر به دست می‌آید.

## نتایج و بحث

نتایج آنالیز آماری سناریوهای مختلف مدل GEP مدل‌های انتخابی در ایستگاه مورد مطالعه در جدول (۴) ارائه شده است. همان‌طور که ملاحظه می‌شود، روابط صریح استخراج‌شده توسط مدل برنامه‌ریزی بیان ژن از دقت مطلوبی برخوردارند. بیشترین مقدار آماره  $R^2$  و کمترین مقدار RMSE (بیشترین دقت) در بین سناریوهای مختلف مورد مطالعه مدل یادگیری عمیق در سناریوی M5 به ترتیب برابر با ۰/۹۹۶ و ۰/۵۱۷ به دست آمد. در این سناریو نسبت به سناریوی یک که شامل مجموعه کامل پارامترهای هواشناسی می‌شود، بارش و ساعات آفتابی حذف شده است. با کاهش تعداد پارامترهای مورد استفاده دقت مدل افزایش می‌یابد اما این سطح از تعداد (۸ پارامتر) و چیدمان پارامترها (سناریوی M5) بیانگر اهمیت حداقل تعداد و نوع پارامترهای مورد نیاز است؛ به عبارتی با حذف بارش و میزان ساعات آفتابی در این منطقه (با تعداد کمتری پارامتر) با استفاده از مدل یادگیری عمیق می‌توان بهترین دقت برآوردی تبخیر و تعرق را به دست آورد. با توجه به نتایج شاخص میانگین قدر مطلق خطا (MAE)، در مدل یادگیری عمیق، کمترین دقت به ترتیب مربوط به مدل‌های M15، M18، M14 و M16 با مقادیر MAE به ترتیب برابر با ۴/۲۱۳، ۳/۱۳۱، ۲/۶۵۶ و ۲/۲۹۸ و بیشترین دقت نیز مربوط به مدل‌های M1، M6، M5 و M3 با مقادیر MAE به ترتیب برابر با ۰/۳۹۹، ۰/۴۰۲، ۰/۴۲۲ و ۰/۴۲۲ است. مقادیر مثبت آماره MAE در این مدل در بین تمامی سناریوها نشان‌دهنده بیش‌برآورد بودن است. در مورد مدل بیان ژن، کمترین دقت به ترتیب مربوط به مدل‌های M14، M15، M24 و M16 با

جدول (۴): نتایج آنالیز مدل‌های انتخابی در ایستگاه مورد مطالعه

Table (4): Results of analysis of selected models in the studied station

سناریو	یادگیری عمیق (DL)			بیان ژن (GEN)		
	$R^2$	MAE	RMSE	$R^2$	MAE	RMSE
M1	۰/۹۹۶	۰/۴۲۲	۰/۵۵۷	۰/۹۸۵	۰/۶۸۳	۰/۹۵۳
M2	۰/۹۹۴	۰/۴۶۲	۰/۶۳۴	۰/۹۸۴	۱/۳۹۴	۱/۶۵۱
M3	۰/۹۹۶	۰/۴۲۲	۰/۵۵۷	۰/۹۸۳	۰/۸۳۳	۱
M4	۰/۹۳۲	۱/۳۲۶	۲/۰۰۳	۰/۹۲۰	۱/۴۶۵	۲/۱۶۰
M5	۰/۹۹۶	۰/۳۹۹	۰/۵۱۷	۰/۹۴۳	۱/۲۷۳	۱/۸۲۲
M6	۰/۹۹۴	۰/۴۰۲	۰/۵۸۰	۰/۹۳۴	۱/۳۵۵	۱/۹۶۶
M7	۰/۹۹۴	۰/۴۶۳	۰/۶۲۸	۰/۹۶۵	۰/۹۹۱	۱/۴۱۷
M8	۰/۹۹۴	۰/۴۴۴	۰/۵۹۹	۰/۹۶۵	۱/۰۲۶	۱/۴۱۸
M9	۰/۹۸۷	۰/۶۰۷	۰/۸۹۱	۰/۹۴۵	۱/۳۷۹	۱/۸۰۳
M10	۰/۹۹۴	۰/۴۶۷	۰/۶۰۵	۰/۹۲۲	۱/۴۱۴	۲/۱۶۰
M11	۰/۹۱۸	۱/۴۷۷	۲/۲۱۵	۰/۸۸۰	۱/۹۰۳	۲/۸۶۳
M12	۰/۹۷۳	۱/۰۰۹	۱/۲۸۲	۰/۹۳۹	۱/۴۸۸	۱/۹۱۱
M13	۰/۹۶۹	۱/۰۴	۱/۳۴۷	۰/۹۵۶	۱/۲۵۰	۰/۸۸۰
M14	۰/۷۹۴	۲/۶۵۶	۳/۴۸۶	۰/۸۴۲	۳/۱۹۸	۴/۱۰۹
M15	۰/۵۰۶	۴/۲۱۳	۵/۳۸۲	۰/۴۶۹	۴/۴۳۸	۵/۵۷۰
M16	۰/۸۳۷	۲/۲۹۸	۳/۰۹۸	۰/۸۲۸	۲/۳۵۵	۳/۱۷۴
M17	۰/۸۸۴	۱/۹۱۷	۲/۶۲۹	۰/۸۷۹	۱/۹۵۸	۲/۶۶۱
M18	۰/۷۱۶	۳/۱۳۱	۴/۱۰۸	۰/۷۲۶	۳/۰۵۹	۳/۹۹۹
M19	۰/۸۹۷	۱/۷۴۲	۲/۴۶۱	۰/۸۹۴	۱/۷۸۶	۲/۴۹۵
M20	۰/۸۷۵	۲/۰۳۳	۲/۷۴	۰/۸۷۶	۱/۹۶۴	۲/۶۹۰
M21	۰/۸۸۹	۱/۸۲۶	۲/۵۵۳	۰/۸۸۹	۱/۸۲۹	۲/۵۵۱
M22	۰/۹۵۶	۱/۲۳۶	۱/۶۰۴	۰/۹۴۱	۱/۴۴۱	۱/۸۶۸
M23	۰/۹۶۹	۱/۰۴	۱/۳۴۷	۰/۹۴۷	۱/۳۷۶	۱/۷۷۴
M24	۰/۴۵۳	۴/۵۴	۵/۶۸۱	۰/۴۳۵	۴/۶۲۱	۵/۷۶۱
M25	۰/۸۶۲	۲/۰۶۵	۲/۸۵۷	۰/۸۶۵	۲/۰۳۲	۲/۸۲۰

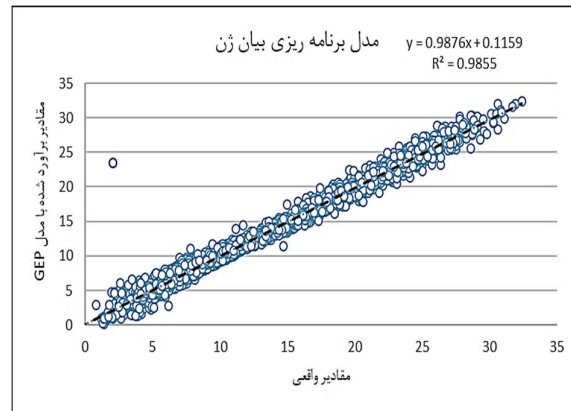
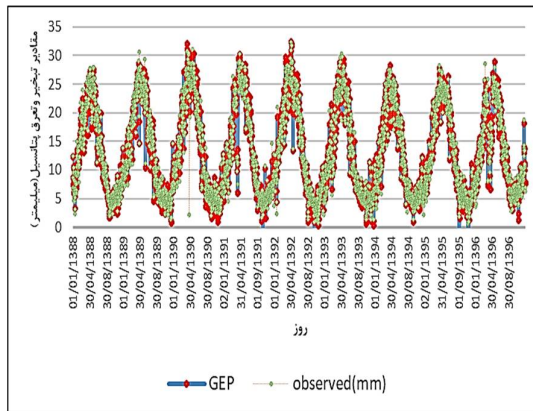
در معادله ۵،  $Tmean$  و  $E$  به ترتیب میانگین دما و تبخیر از تشت،  $RHmax$  حداکثر رطوبت،  $U_2$  سرعت باد در دو متری و  $SUN$  تعداد ساعات آفتابی هستند. در مدل برنامه‌ریزی بیان ژن، سناریوی M1 با بیشترین ضریب تبیین  $R^2=0.985$  و کمترین خطای  $RMSE=0.985$  و  $MAE=0.683$  به عنوان بهترین سناریوی برگزیده شد. تخمین نزدیک مقادیر محاسباتی با مقادیر واقعی، نشان‌دهنده دقت مناسب مدل در تخمین مقادیر تبخیر و تعرق پتانسیل است، محققان دیگری نیز نتایج مشابهی را گزارش کرده‌اند (شیری و همکاران، ۲۰۱۲؛ فریرا، ۲۰۰۶).

نتایج شکل (۳) نیز بیانگر قرابت زیاد مقادیر تخمینی و مشاهداتی و دقت تخمین بالای مدل‌های به‌کاررفته است. رابطه غیرخطی حاصل از مدل برنامه‌ریزی بیان ژن در معادله (۷) نشان داده شده است. این رابطه بر مبنای ترکیب کامل متغیرهای اقلیمی (سناریوی ۱) مورد مطالعه در این پژوهش استخراج شده است. توابع موجود در این رابطه بر اساس انتخاب کاربر قابل تغییرند:

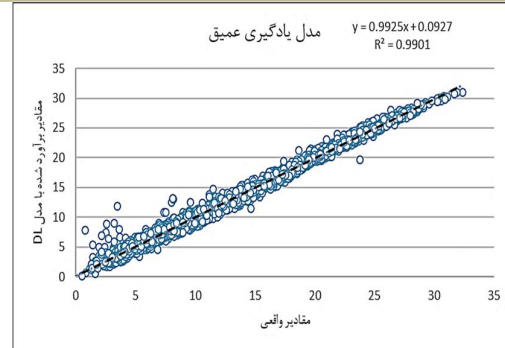
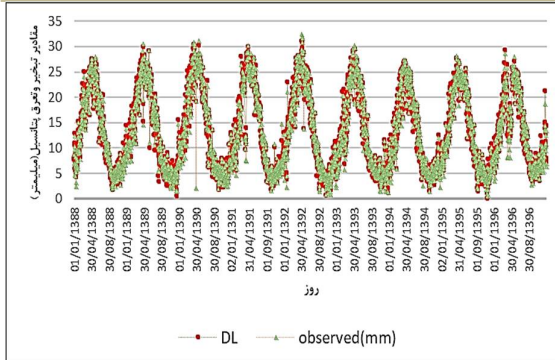
$$\begin{aligned}
 ET0 &= \{[(Tmax - E) + SUN \times 0.502] \times (0.502^2)\} - 1.016 \\
 &+ \left[ \sin(RHmax)^{\frac{1}{3}} \right]^3 - \cos \sqrt{(E + Tmax)} \\
 &+ \sqrt{\left\{ \left( \frac{E}{(RHmax + E)} \right) \times [(U_2 - Tmean) - Tmean] \right\}}
 \end{aligned} \quad (7)$$



قرار گرفتند. در مدل یادگیری عمیق نیز سناریوی M5 از بیشترین مقدار تأثیرگذاری برخوردار بوده و سناریوهای M1 و M3 در رتبه‌های بعدی قرار گرفتند. دیاگرام پراکنش منحنی پراکندگی یا مقادیر برآوردشده در برابر مقادیر مشاهداتی در مرحله آموزش و تست بر اساس نوع بر اساس نوع مدل در شکل (۴) آورده شده است.



شکل (۳): سمت راست، پراکنش مدل GEP با مدل فائو-پنمن-مانتیت؛ ب. سمت چپ: مقادیر تبخیر و تعرق مرجع برای ماه‌های آزمون  
Figure (3): Right, distribution of the GEP model with the FAO-Penman-Montith model b) Left: Reference evapotranspiration values for test months



شکل (۴): سمت راست، الف. پراکنش مدل DL با مدل فائو-پنمن-مانتیت؛ ب. سمت چپ: مقادیر تبخیر و تعرق مرجع برای ماه‌های آزمون  
Figure (4): Right, distribution of the DL model with the FAO-Penman-Montith model b) Left: Reference evapotranspiration values for test months

برای مثال، در مدل‌های برنامه‌ریزی بیان ژن و یادگیری عمیق، زمان محاسبات به ترتیب با ۲۳ ساعت و در مدل یادگیری عمیق این زمان سه دقیقه و ۲۶ ثانیه شدند که این موضوع به لحاظ صرفه‌جویی در زمان محاسبات برای کاربر بسیار مهم است؛ با این حال به دلیل اینکه در برآورد تبخیر و تعرق روزانه تغییرات کوتاه‌مدت پارامترهای هواشناسی مؤثر بر تبخیر و تعرق در نظر گرفته می‌شود، انتخاب پارامترهای مؤثر و در

نتایج مدل برنامه‌ریزی بیان ژن و یادگیری عمیق و انتخاب بهترین الگو

در مدل برنامه‌ریزی بیان ژن، بین تمامی سناریوها، سناریوی M1 با بالاترین ضریب همبستگی ( $R^2=0.985$ ) و کمترین خطا ( $RMSE=0.953$ ) به عنوان بهترین الگو انتخاب شد. پس از این الگو، الگوهای M3 و M7 در رتبه‌های دوم و سوم

همان طور که نتایج نشان دادند، همبستگی بالایی بین تبخیر و تعرق برآوردشده از این مدل‌ها و روش فائو-پنمن-مانتیت وجود دارد (شکل ۳ و ۴) و می‌توان از این مدل‌های محاسباتی، برای برآورد تبخیر و تعرق روزانه در زمانی که داده‌های محدودتری در دسترس باشند، استفاده کرد. نکته دیگر در بحث مدل‌سازی با مدل‌های مذکور، بحث زمان مورد نیاز برای محاسبات برای شبیه‌سازی در مدل‌های یادشده است.

عمیق الگوی M5 با  $RMSE=0.517$ ،  $MAE=0.399$  و بیشترین ضریب تبیین  $R^2 = 0.996$  به عنوان بهترین الگوهای هر مدل در منطقه مورد مطالعه به دست آمدند.

### نتیجه گیری

نتایج نشان داد که مدل یادگیری عمیق از دقت بیشتری در بین مدل‌های انتخابی برخوردار بوده و در این مدل سناریوی M5 با ۸ پارامتر ورودی دمای حداکثر، دمای حداقل، دمای میانگین، رطوبت حداکثر، رطوبت حداقل، رطوبت میانگین، سرعت باد و تبخیر از تشت با بیشترین ضریب همبستگی ( $R^2 = 0.996$ ) و کمترین خطا ( $RMSE=0.517$ ) از دقت بیشتری برای مدل‌سازی مقدار تبخیر و تعرق مرجع برخوردار شد. هرچند برآورد مدل‌ها با اندازه خطا همراه می‌باشد ولی کم‌برآوردی و بیش‌برآوردی در راستای خط همبستگی قرار داشته و پراکندگی زیادی در داده‌ها مشاهده نمی‌شود؛ این امر بیانگر همبستگی مناسب مدل‌ها و داده‌های مشاهده‌ای در منطقه مورد مطالعه است. در مدل یادگیری عمیق، پارامترهای دمای میانگین، دمای بیشینه، سرعت باد و رطوبت کمینه، بیشترین تأثیر را در تخمین تبخیر و تعرق روزانه داشتند. نتیجه کلی تحقیق کاربرد مدل یادگیری عمیق در برآورد تبخیر و تعرق منطقه سیستان با ضریب اطمینان بالا در برنامه‌ریزی‌های منابع آبی و شبکه‌های آبیاری و زهکشی است.

صورت امکان حذف برخی از آن‌ها به منظور کاهش متغیرهای ورودی، می‌تواند کمک زیادی در برآورد سریع و کم‌هزینه تبخیر و تعرق روزانه به خصوص مدل‌سازی نماید. یافته‌های محققان دیگر نیز این نتایج را تأیید می‌کند (کسکین و همکاران، ۲۰۰۹؛ گولوروت و بنجیو، ۲۰۱۰).

مدل یادگیری عمیق با شبیه‌سازی مستقیم و بلافاصل خروجی، نیازی به شکل محاسباتی نداشته و این یک مزیت ویژه نسبت به مدل‌های مشابه محسوب می‌شود، با این وجود باید توجه کرد که هرچند مدل‌های داده‌محور، نیازی به معادلات با پایه‌های فیزیکی، که بیشتر به صورت غیرخطی و پیچیده‌اند، ندارند ولی درک فرایندها و تأثیر عوامل مورد نظر در داده‌های ورودی بر تولید خروجی آن‌ها بسیار مهم است و در صورت عدم این درک، نمی‌توان انتظار نتایج مطلوبی را داشت.

### تحلیل حساسیت

در مدل برنامه‌ریزی بیان ژن پارامترهای میانگین دما، دمای کمینه، دمای بیشینه، تبخیر از تشت و میانگین رطوبت نسبی و در مدل یادگیری عمیق پارامترهای دمای میانگین، دمای بیشینه، سرعت باد و رطوبت حداقل به ترتیب بیشترین ضریب حساسیت را به خود اختصاص دادند. در مجموع در مدل برنامه‌ریزی بیان ژن، سناریوی M1 با  $RMSE=0.985$ ،  $MAE=0.683$  و ضریب تبیین  $R^2=0.985$  و در مدل یادگیری

### منابع

- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration guidelines for computing crop water requirements-fao irrigation and drainage paper 56. FAO, Rome 300 (9): D05109.
- Arel, I., Rose, D. C., Karnowski, T. P., 2010. Deep machine learning-a new frontier in artificial intelligence research. IEEE computational intelligence magazine 5(4), 13-18.
- Bruton, J.M., McClendon, R.W., Hoogenboom, G., 2000. Estimating daily pan evaporation with artificial neural.
- Feng, Y., Peng, Y., Cui, N., Gong, D., Zhang, K., 2017. Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. Comput. Electron. Agric 136, 71-78.
- Ferreira, C., 2001. Gene expression programming: a new adaptive algorithm for solving problems. arXiv preprint cs/0102027.
- Ferreira, C., 2006. Gene-expression programming; mathematical modeling by an artificial intelligence. Springer, Berlin, Heidelberg, New York.
- Gao, Z., He, J., Dong, K., Li, X., 2017. Trends in reference evapotranspiration and their Comput. Electron. Agric 113, 164-173.
- Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics. 249-256.
- Gocic, M., Petković, D., Shamsirband, S., Kamsin, A., 2016. Comparative analysis of reference evapotranspiration equations modelling by extreme learning machine. Comput. Electron. Agric 127, 56-63.

10. Goodfellow, I., Bengio, Y. Courville, A., 2016. Deep Learning.
11. Hinton, G. E., Osindero, S., Teh, Y. W., 2006. A fast learning algorithm for deep belief nets. *Neural computation* 18(7), 1527-1554.
12. Izadifar, Z., Elshorbagy, A., 2010. Prediction of hourly actual evapotranspiration using neural networks, genetic programming, and statistical models. *Hydrological Processes* (23)24, 3413-3425.
13. Jensen, M.E., Burman, R.D., Allen, R.G., 1990. *Evapotranspiration and Irrigation Water Requirement*. ASCE Manual, No.70, U.S.A.
14. Keskin, M. E., Terzi, O., Taylan, E. D., Ücukyaman, D. K., 2009. Meteorological drought analysis using data-driven models for the lakes district, Turkey. *Hydrological Sciences Journal* 54(6), 1114-1124.
15. Keyvanrad, M. A., Homayounpour, M. M., 2015. Deep belief network training improvement using elite samples minimizing free energy. *International Journal of Pattern Recognition and Artificial Intelligence* 29(05), 1551006.
16. Liu, J. N., Hu, Y., You, J. J., Chan, P. W., 2014. Deep neural network based feature representation for weather forecasting. In *Proceedings on the International Conference on Artificial Intelligence (ICAI)* (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
17. Musavi-Baygi, M., Erfanian, M., Sarmad, M., Khazaei, M., 2009. Estimation of reference crop Evapotranspiration using the least meteorological data (Case study: Khorasan Razavi province). *J. Water Soil*. 23: 1. 91-99. (In Persian)
18. Parasuraman, K., Elshorbagy, A., Carey, S. K., 2007. Modelling the dynamics of the evapotranspiration process using genetic programming. *Hydrological Sciences Journal* 52(3), 563-578.
19. Shiri, J., Kişi, Ö., Landaras, G., López, J. J., Nazemi, A. H., Stuyt, L. C., 2012. Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). *Journal of Hydrology* 414, 302-316.
20. Terzi, O. 2012. Daily pan evaporation estimation using gene expression programming and adaptive neural-based fuzzy inference system. *Journal Computing & Applications*, 23(3-4): 1035- 1044.
21. Terzi, Ö., Erol-Keskn, M., 2005. Modeling of daily pan evaporation. *J. Appl. Sci* 5, 368-372.
22. Tian, D., Martinez, C.G., 2012. Forecasting reference evapotranspiration using retro-spective forecast analogs in the South-eastern United States. *Journal of Hydrometeorol*, 1(3), 1874-1892.
23. Traore, S., Guven, A., 2012. Regional-specific numerical models of evapotranspiration using gene-expression programming interface in Sahel. *Water Resource Management* 26, 4367-4380.
24. Wang, Z., Wu, P., Zhao, X., Cao, X., 2013. GANN models for reference evapotranspiration estimation developed with weather data from different climate regions. *Theor & appl climatol*. DOI: 10.1007/s00704-013-0967.
25. Zhu, X. X., Tuia, D., Mou, L., Xia, G-S., Zhang, L., Xu, F., Fraundorfer, F. 2017. Deep learning in remote sensing: A comprehensive review and list of resources', *IEEE Geoscience and Remote Sensing Magazine*, 5(4): 8-36.

## Estimating Daily Reference Evapotranspiration in Sistan Plain Using Ultra-Innovative Algorithms

Hadi Siasar<sup>1\*</sup>, Amir Salari<sup>2</sup>, Omolbanin Mohamadrezapour<sup>3</sup>, Halimeh Piri<sup>4</sup>

Received: 21/12/2020

Accepted: 11/05/2021

### Extended Abstract

**Introduction:** Measuring the evapotranspiration rate plays an important role in the proper management of water resources, irrigation planning, and optimizing the allocation and distribution of water resources. There are different methods for measuring evapotranspiration, which is generally time-consuming and costly and requires a large bulk of meteorological data. However, new widely used methods have been introduced in recent years to solve this

1. Assist Professor Department of Agriculture, Payame Noor University(PNU), Iran, Corresponding author; hadisiasar@pnu.ac.ir

2. Assist Professor, Department of Water Science and Engineering, Minab Higher Education Center, University of Hormozgan

3. Associate Professor, Department of Water Engineering, Faculty of Agriculture, University of Zabol

4. Assist Professor, Department of Water Engineering, Faculty of Agriculture, University of Zabol

DOI: 10.22052/deej.2021.10.32.49

problem, among which are ultra-Innovative algorithms with high accuracy and speed that do not require extensive data. Therefore, this study sought to identify the most important parameters involved in measuring the daily evapotranspiration rate of the reference plant in Sistan plain using Gene Expression Programming and Deep Learning models.

**Material and methods:** located in the north of Sistan and Baluchestan province at the northern latitude  $30^{\circ}.18'$  to  $31^{\circ}.20'$  and the eastern longitude  $61^{\circ}.10'$  to  $61^{\circ}.50'$ , Sistan plain has an average annual precipitation rate of 50 mm and an annual evaporation rate of 4000- 5000 mm, being considered as one of the super-arid areas based on the Dumarten drought index, whose environmental conditions are not suitable for cultivation. The climatic data used in this study were collected from the Zabol synoptic station, including maximum temperature, minimum temperature, average temperature, maximum relative humidity, minimum relative humidity, average relative humidity, sunny hours, wind speed, precipitation, and pan evaporation during the statistical period of 2009-2017.

Moreover, the accuracy of Gene Expression Programming and Deep Learning was compared to the FAO-Penman-Montith method. Accordingly, the GeneXproTools software (4.0) was used to run the Gene Expression Programming model and MATLAB software was used to run the Deep Learning model. Also, the data were divided into two categories, 80% of which were used for training and 20% of which were used to validate the model. Considering the fact that selecting appropriate and effective initial inputs improves performance Since in smart models, different combinations of meteorological data were considered as model inputs. Then, the best scenario was selected to predict evapotranspiration by evaluating the results of different scenarios and combinations. Furthermore, the Coefficient of determination ( $R^2$ ) was used to calculate the correlation, mean absolute error value (MAE) was used to show the degree of consistency between the set of observed and predicted values, and the root mean square error (RMSE) (expressing the error intensity) was applied as evaluation criteria.

**Results:** The study's results indicated that Gene Expression Programming and Deep Learning Programming models were highly accurate in estimating evapotranspiration in all scenarios, with the Deep Learning model showing a higher accuracy in this regard than the Gene Expression one. Moreover, it was found that from among all the scenarios upon which the Deep Learning programming model was applied, the M5 scenario comprising of variables such as maximum temperature, minimum temperature, average temperature, maximum humidity, minimum humidity, average humidity, wind speed, and pan evaporation was the most accurate scenario with the lowest root mean square error (RMSE = 0.517) and the highest coefficient of determination ( $R^2 = 0.996$ ). On the other hand, out of all scenarios to which the Gene Expression model was applied, the M1 scenario containing variables such as mean temperature, minimum temperature, maximum temperature, and maximum humidity was the most accurate one, with the highest coefficient of determination ( $R^2 = 0.985$ ) and the lowest root mean square error (RMSE = 0.985).

However, in the deep learning model, the lowest accuracy belonged to M15, M18, M1, and M16 scenarios with MAE values of 4.213, 3.131, 2.656, and 2.298, respectively, and the highest accuracy belonged to the M5, M6, M1, and M3 scenarios with MAE values of 0.399, 0.402, 0.422 and 0.422, respectively. In this model, all scenarios were overestimated. On the other hand, In the Gene Expression model, the lowest accuracy is related to M24, M15, M14, and M16 scenarios with MAE values equal to 4.621, 4.438, 3.198, and 2.355, respectively, and the highest accuracy is also related to M1, M3, M13 and M7 scenarios with MAE values equal to 0.683, 0.733, 0.780 and 0.991, respectively. In this model, all scenarios are overestimated. On the other hand, in the Gene Expression model, the lowest accuracy belonged to M24, M15, M14, and M16 scenarios whose MAE values were 4.621, 4.438, 3.198, and 2.355, respectively, and the highest accuracy belonged to M1, M3, M13, and M7 scenarios whose MAE values were 0.683, 0.733, 0.780, and 0.991, respectively. In this model, all scenarios were overestimated.

According to the outputs of the GEP model, mean temperature, minimum temperature, maximum temperature, and maximum humidity were the most important parameters involved in the prediction of reference evapotranspiration values. In the Gene Expression Programming model, the M1 was selected as the best scenario with the highest coefficient of explanation  $R^2 = 0.985$ , the lowest error RMSE = 0.985, and MAE = 0.683, followed by the M3 and M7 scenarios. Moreover, in the Deep Learning model, the M5 scenario ranked first in predicting the reference evapotranspiration, followed by the M1 and M3 scenarios. Also, the high correlation between the estimated evapotranspiration of these models and the Fao-penman-montith method indicated that computational models can be used to estimate daily evapotranspiration when more limited data are available.

**Discussion and Conclusion:** The results showed that the evapotranspiration of the reference plant in the Sistan region can be determined in the shortest possible time (3 minutes and 26 seconds in the deep learning model) with acceptable accuracy using a few parameters (compared to the FAO method). Therefore, it is recommended that the Deep Learning model be applied in the Sistan region.

**Keywords:** Evapotranspiration, Deep learning model, GEP model.