

مقایسه روش‌های شبکه‌های عصبی مصنوعی، شبکه بیزین و برنامه‌ریزی بیان ژن در پیش‌بینی خشکسالی (مطالعه موردی: ایستگاه سینوپتیک مراغه)

عباس عباسی^۱، کیوان خلیلی^۲، جواد بهمنش^۳ و اکبر شیرزاد^۴

۱- دکتری مهندسی منابع آب، دانشگاه ارومیه
۲- استادیار گروه مهندسی آب، دانشگاه ارومیه (نویسنده مسوول: Khalili2006@gmail.com)
۳- استاد گروه مهندسی آب، دانشگاه ارومیه
۴- استادیار گروه مهندسی عمران، دانشگاه صنعتی ارومیه
تاریخ ارسال: ۹۷/۷/۹ تاریخ پذیرش: ۹۸/۳/۵
صفحه: ۵۹ تا ۷۱

چکیده

خشکسالی جزء جدایی‌ناپذیر هر اقلیمی محسوب می‌گردد که تأثیرات مهمی بر بخش‌های مختلف جامعه دارد و سبب افزایش فشار بر منابع آبی می‌گردد. لذا پیش‌بینی وضعیت آبی آن می‌تواند به برنامه‌ریزان و تصمیم‌گیران در بخش‌های مختلف کمک شایانی نماید. در این تحقیق جهت پیش‌بینی خشکسالی در مقیاس‌های زمانی مختلف شاخص خشکسالی بارش و تبخیر تعرق استاندارد (SPEI) از ۵ ورودی مختلف شامل مقادیر شاخص SPEI با تأخیرهای یک تا ۵ ماهه استفاده و سپس از سه روش هوشمند شامل برنامه‌ریزی بیان ژن (GEP)، شبکه بیزین (BN) و شبکه عصبی مصنوعی (ANN) جهت پیش‌بینی مقادیر آبی استفاده گردید. نتایج نشان داد که هر سه روش در مقیاس زمانی کوتاه‌مدت شاخص SPEI از دقت مناسب برخوردار نمی‌باشند به طوری که بهترین عملکرد در مقیاس زمانی یک ماهه مربوط به مدل شبکه بیزین با ضریب همبستگی ۰/۱۴۲ و در مقیاس سه ماهه مربوط به روش ANN با ضریب همبستگی ۰/۷۰۴ می‌باشد. نتایج همچنین نشان داد که دقت پیش‌بینی مدل با افزایش مقیاس محاسبه SPEI رابطه مستقیم دارد و با افزایش مقیاس زمانی SPEI، دقت پیش‌بینی افزایش پیدا می‌کند. همچنین هر سه روش در مقیاس‌های زمانی بلندمدت از عملکرد مناسبی برخوردار می‌باشند.

واژه‌های کلیدی: پیش‌بینی، خشکسالی، مراغه، مدل هوشمند، شاخص SPEI

مقدمه

گرفتن عامل بارش در تعیین خشکسالی یک منطقه، از جمله انتقاداتی است که بر شاخص‌های خشکسالی مبتنی بر بارش خصوصاً شاخص SPI وارد است و بسیاری از محققان بر اهمیت درجه حرارت بر شرایط خشکسالی تأکید نموده‌اند که منجر به افزایش تقاضای آب در اثر افزایش تبخیر تعرق می‌گردد (۱۱). بر همین اساس می‌توان شاخص خشکسالی استاندارد شده بارش و تبخیر تعرق (SPEI)^۱ را که ویسنته سرانو در سال ۲۰۱۰ عرضه کرد در حکم شاخصی مناسب در تعیین خشکسالی مورد استفاده قرار داد (۲۱). در این شاخص سه متغیر بارش، دما و تبخیر تعرق پتانسیل (PET) لحاظ شده است و بر اساس میزان بارندگی و درجه حرارت و محاسبه تبخیر تعرق و بیلان آب می‌تواند اثرات تغییرات درجه حرارت را در ارزیابی خشکسالی لحاظ نماید. ذبیحی و همکاران (۲۳) دوره‌های خشک و مرطوب را در هفت ایستگاه سینوپتیک در غرب ایران با استفاده از شاخص‌های SPI و SPEI مورد بررسی قرار دادند. نتایج این تحقیق نشان داد که هر دو شاخص در دوره‌های مرطوب از همبستگی بالایی برخوردار می‌باشند. نتیجه این برنامه‌ریزی بیان ژن شاخه‌ای از الگوریتم‌های تکاملی است که توانایی مدل‌سازی فرآیندهای غیرخطی و پویا را دارد. قربانی و همکاران (۶) جهت پیش‌بینی دمای بیشینه، متوسط و کمینه ایستگاه سینوپتیک تبریز از روش‌های ANFIS، ANN و GEP استفاده کردند که نتایج این تحقیق نشانگر ارجحیت روش GEP بر دو روش دیگر در تخمین دمای هوا بود. زارع امینی

با توجه به فرارگیری ایران در یک منطقه خشک و نیمه‌خشک تاکنون خشکسالی‌های پی در پی، اثرات زیان‌باری بر معیشت و زندگی مردم به وجود آورده که نمونه بارز آن وضعیت حاکم بر دریاچه ارومیه در سال‌های اخیر می‌باشد که به عنوان یکی از عوامل اصلی کاهش سطح و حجم دریاچه از آن یاد می‌شود. لذا پایش و پیش‌بینی خشکسالی برای مدیران و تصمیم‌گیران بخش آب و سایر بخش‌های اقتصادی این امکان را فراهم می‌کند که برنامه‌ریزی بهتری برای مقابله با این بلای طبیعی داشته باشند. وقوع خشکسالی‌های شدید در قاره آفریقا، هند، آمریکای شمالی، چین، استرالیا، خاورمیانه و ایران در دهه‌های پیشین نشان از آسیب‌پذیری جوامع توسعه یافته و در حال توسعه در برابر خشکسالی است (۲۳). به دلیل افزایش جمعیت و تقاضا برای منابع آب و نقش آن در توسعه، آسیب‌پذیری در جوامع نسبت به خشکسالی در حال افزایش است. اصولاً برای تحلیل کمی خشکسالی، وجود یک شاخص مناسب جهت تعیین دقیق دوره‌های تر و خشک بسیار ضروری است (۳). اما کمبود آب در یک حوضه علی‌رغم بارش کافی می‌تواند به واسطه تبخیر بالا، تغییر در کاهش ذخایر آب سطحی و حتی مصرف بیش از حد رخ دهد. کاهش میزان بارندگی و افزایش دما هر یک به تنهایی و یا به کمک هم می‌توانند موجب خشکسالی شوند (۱۲). در محاسبه شاخص‌های رایج تعیین خشکسالی، مانند SPI صرفاً از میزان بارندگی روزانه و یا ماهانه استفاده می‌شود. اما صرفاً در نظر

1- Standardized Precipitation Evapotranspiration Index

2- Potential Evapotranspiration

آلودگی آبهای زیرزمینی منطقه دارد. خشکسالی‌های مستمر در سال‌های اخیر نقش مهمی در کاهش منابع آب در ایران و به تبع آن در حوزه آبخیز دریاچه ارومیه داشته است که وضعیت پیش آمده برای این دریاچه گواه این ادعا می‌باشد. لذا پیش‌بینی خشکسالی و وضعیت آن در آینده می‌تواند نقش مهمی در برنامه‌های تسکین خشکسالی بر عهده داشته باشد. در این تحقیق، پیش‌بینی خشکسالی در ایستگاه سینوپتیک مراغه با استفاده از شاخص SPEI و روش‌های هوشمند شبکه بیزین، برنامه‌ریزی بیان ژن و شبکه عصبی مصنوعی در مقیاس‌های زمانی مختلف شاخص مذکور مورد بررسی و نتایج این مدل‌ها با یکدیگر مورد مقایسه قرار گرفت.

مواد و روش‌ها منطقه مورد مطالعه

مراغه یکی از شهرهای استان آذربایجان شرقی است که در عرض جغرافیایی ۳۷ درجه و ۲۱ دقیقه و طول جغرافیایی ۴۶ درجه و ۱۲ دقیقه و در بخش شرقی دریاچه ارومیه واقع شده است. ارتفاع این ایستگاه ۱۳۴۴ متر از سطح آزاد دریا و سال تأسیس آن ۱۳۶۲ می‌باشد. در این تحقیق از داده‌های بلندمدت بارش و دمای ماهانه این ایستگاه در یک دوره آماری ۲۶ ساله مربوط به سال‌های ۲۰۰۹-۱۹۸۳ استفاده شده است.

شاخص SPEI

این شاخص مبتنی بر محاسبه بیلان آب است و برای محاسبه آن به داده‌های ماهانه بارش و دمای هوا نیاز است. معادله بیلان آب به صورت رابطه (۱) تعریف شده است.

$$D_i = P_i - PET_i \quad (1)$$

که در این رابطه، P و PET به ترتیب بارش و تبخیر تعرق پتانسیل، D اختلاف آنها و i شماره ماه مورد نظر است. معادلات مختلفی برای محاسبه PET وجود دارد که در محاسبه SPEI محدودیتی برای استفاده از هر یک از آنها وجود ندارد. در نسخه نخست SPEI از روش ساده تورنت-وایت (۲۰) استفاده شده، ولی در نسخه دوم آن (در صورت عدم محدودیت داده) توصیه شده است که از روش پنمن-مانتیت فائو (۲) استفاده شود که در این تحقیق نیز از این رابطه به منظور محاسبه PET استفاده شده است. پس از محاسبه مقادیر D ، از رهیافتی همانند شاخص SPI برای محاسبه SPEI استفاده می‌شود. برای این منظور ابتدا مجموع مقادیر متوالی سری D در پنجره‌های زمانی مختلف محاسبه می‌شود. اگر x سری تجمعی D در پنجره زمانی معین باشد، در مرحله بعد یک توزیع احتمال مناسب بر سری x برازش می‌یابد. طبق بررسی‌های ویسنه‌سرانو (۲۰) توزیع لوگ لجستیک سه پارامتری کارایی مناسبی در مدل‌سازی سری x دارد. فرم تابع چگالی احتمال توزیع مذکور به صورت رابطه (۲) بیان شده است.

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x-\gamma}{\alpha}\right)^{\beta-1} \left[1 + \left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right]^{-2} \quad (2)$$

و همکاران (۲۴) جهت تخمین دمای خاک از روش GEP استفاده نمودند و به این نتیجه رسیدند که این روش به دلیل ارائه راه‌حل‌های صریح بر روش ANN برتری دارد. مهدی‌زاده و همکاران (۱۵) ارزیابی دقت دو روش GEP و شبکه بیزین را در پیش‌بینی دمای روزانه هوا مورد بررسی قرار دادند و به این نتیجه رسیدند که هر دو روش از دقت قابل قبولی در پیش‌بینی دمای هوا برخوردار هستند، ولی روش شبکه بیزین دقت بیش‌تری نسبت به روش GEP دارد. با توجه به ماهیت غیرخطی پدیده‌ها در علوم جوی و هیدرولوژی و همچنین در مدیریت منابع آب در دهه‌های اخیر، استفاده از روش‌های شبکه عصبی مصنوعی کاربرد گسترده‌ای داشته است. مدل شبکه عصبی مصنوعی، مدلی قوی با توانمندی بالاست که در پیش‌بینی مسائل هیدرولوژیکی-اقلیمی می‌توان از آن استفاده نمود. مرید و همکاران (۱۹) در پژوهشی از شبکه عصبی مصنوعی برای پیش‌بینی شاخص‌های خشکسالی استفاده کردند و به این نتیجه رسیدند که مدل شبکه عصبی در پیش‌بینی شاخص‌ها دارای دقت مناسبی می‌باشد. دستورانی و افخمی (۴) از مدل شبکه عصبی مصنوعی برای پیش‌بینی خشکسالی در استان یزد استفاده نمودند و به این نتیجه رسیدند که این روش مدلی مناسب برای پیش‌بینی خشکسالی بوده و متغیرهای درجه حرارت و بارش مهم‌ترین ورودی برای شبکه محسوب می‌شوند. ماکا و پیچ (۱۳) با استفاده از دو روش شبکه عصبی مصنوعی نسبت به پیش‌بینی شاخص‌های خشکسالی SPI و SPEI اقدام نمودند که نتایج این تحقیق نشان داد که روش شبکه عصبی مصنوعی تجمعی نسبت به روش شبکه عصبی چند لایه پیشخور نتایج بهتری را داشته است. شبکه‌های بیزین که اصولی از نظریه گراف، نظریه احتمالات، علوم کامپیوتر و آمار را با هم ترکیب می‌کند، از جمله مدل‌های مطرح در زمینه پیش‌بینی در دهه‌های اخیر می‌باشد. حبیبی‌پور و همکاران (۷) جهت بررسی تعدیل دامنه تغییرات داده‌ها در بحث پیش‌بینی خشکسالی در ایستگاه سینوپتیک یزد، از دو روش شبکه عصبی مصنوعی و درخت تصمیم رگرسیونی استفاده کردند که نتایج این تحقیق نشان داد اعمال میانگین لغزان بر روی داده‌های اصلی سبب بهبود مناسب عملکرد هر دو مدل پیش‌بینی در مقیاس زمانی ۱۲ ماهه می‌گردد. مددگر و مرادخوانی (۱۴) جهت بررسی تغییرات مکانی خشکسالی و توسعه مدلی بر اساس خشکسالی‌های گذشته، در رودخانه گانسین آمریکا از روش شبکه بیزین استفاده نمودند. در این تحقیق یک مدل پیش‌بینی آماری چندمتغیره با توجه به میزان رواناب در حوضه طرح‌ریزی شد که مهم‌ترین مزیت آن مدل، داشتن ویژگی‌های احتمالی در تجزیه و تحلیل خشکسالی در سال‌های آتی بوده است. در این مدل بالاترین پیش‌بینی احتمالی با یک ارزیابی از عدم قطعیت‌ها مورد تحلیل قرار گرفت و شبکه‌های بیزین، احتمال خشکسالی در سال‌های آینده را با شدت‌های مختلف نشان دادند. فرمانی و همکاران (۵) برای مدیریت بهینه آلودگی آبهای زیرزمینی در منطقه کهنه‌گاز از شبکه‌های بیزین استفاده کردند و نتایج این تحقیق نشان داد که روش شبکه بیزین کارایی مناسب در مدیریت

با پردازش داده‌ها، قوانین نهفته در انبوه داده‌ها را به ساختار شبکه عصبی انتقال می‌دهد. امروزه در علوم مختلف، شبکه عصبی مصنوعی کاربرد وسیعی در شبیه‌سازی، بهینه‌سازی و طبقه‌بندی ایفا می‌کند که در این بخش، قابلیت شبیه‌سازی در پیش‌بینی، مورد استفاده قرار خواهد گرفت. نقش اصلی یک نرون بیولوژیکی عمل جمع ورودی‌ها تا جایی است که مجموع ورودی‌ها از حدی که به آن آستانه می‌گویند تجاوز نکند تا آن‌گاه یک خروجی تولید شود. شبکه‌های عصبی مصنوعی از سیستم‌های عصبی بیولوژیکی استنتاج شده‌اند. این مدل‌ها با ساختار طبیعی و هوشمند خود و نیز با الگوبرداری مناسبی از نرون‌های موجود در مغز انسان سعی می‌کنند که رفتار نرون‌های مغز را از طریق توابع ریاضی تعریف شده شبیه سازی نمایند. عملکرد سیناپسی در نرون‌های طبیعی از طریق وزن‌های محاسبه شده در خطوط ارتباطی نرون‌های مصنوعی مدل‌سازی می‌شوند. ماهیت و ذات تجربی و انعطاف‌پذیر این مدل‌ها سبب می‌شوند در مسائلی مانند مقوله پیش‌بینی و مدل سازی که چنین نگرشی در ساختار آن‌ها مشاهده می‌شود و از رفتاری غیرخطی و پیچیده برخوردارند به خوبی قابل استفاده باشند. همچنین در بسیاری از مسائل پیچیده ریاضی که به حل معادلات بفرنج غیرخطی منجر می‌شوند یک شبکه عصبی مصنوعی می‌تواند مورد استفاده قرار گیرد. هر لایه شامل تعدادی اجزای پردازشی به هم متصل شده تحت عنوان نرون‌ها (n) هستند. وزن‌ها در این شبکه‌ها برای تغییر توان عملیاتی پارامترها و نیز تغییر ارتباطات به نرون‌ها استفاده می‌شوند. همچنین بایاس‌ها اجزای اضافی در نرون‌های لایه‌های پنهان و خروجی هستند. در لایه پنهان، مقدار انتقال نرون (h_j) به صورت ریاضی با استفاده از رابطه زیر توصیف می‌شود؛ (۹)

$$n_{ij} = \sum_{i=1}^N W_{ij} X_i + B_j \quad h_j = f(n_{ij})$$

$$h_j = f(n_{ij})$$

مقدار نرون لایه خروجی نیز توسط رابطه (۱۰) بیان می‌شود.

$$n_{jk} = \sum_{i=1}^N W_{jk} h_j + B_k$$

$$h_k = f(n_{jk})$$

در تحقیق حاضر، از شبکه‌های عصبی مصنوعی پیشخور^۱ با ساختار سه لایه برای مدل‌سازی متغیرهای در نظر گرفته شده استفاده شد. تعداد نرون‌های هر لایه نیز به این صورت است که تعداد نرون‌های لایه ورودی همان متغیرهای ورودی در نظر گرفته شده یا مستقل برای برآورد یا مدل‌سازی متغیر خروجی و تعداد نرون لایه خروجی برابر با یک یا همان متغیر وابسته در نظر گرفته شده است. تعداد نرون‌های لایه پنهان نیز از طریق فرآیند آزمون و خطا برای حصول کمترین خطا در مدل‌سازی تعیین می‌گردد. توابع انتقال تانژانت-سیگموئید و خطی به سبب کاربرد گسترده این توابع در مطالعات هیدرولوژیکی در لایه‌های پنهان و خروجی مورد استفاده قرار گرفتند. برای آموزش شبکه‌های عصبی مصنوعی توسعه داده شده در تحقیق حاضر نیز از روش الگوریتم آموزشی لونبرگ-مارکواریت^۲ استفاده شد. زانتی و همکاران (۲۵) بیان نمودند که الگوریتم مورد اشاره به سبب همگرایی سریع و در نتیجه نیاز

که در این رابطه، β ، α و γ به ترتیب پارامترهای مقیاس، شکل و مبدأ و x سری تجمعی مقادیر D در پنجره زمانی معین است. پارامترهای این تابع با استفاده از روش گشتاورهای خطی (۱۰) از معادلات (۳) الی (۵) به دست می‌آیند.

$$\beta = \frac{2w_1 - w_0}{6w_1 - w_0 - 6w_2} \quad (۳)$$

$$\alpha = \frac{(w_0 - 2w_1)\beta}{\Gamma(1 + \frac{1}{\beta})\Gamma(1 - \frac{1}{\beta})} \quad (۴)$$

$$\gamma = w_0 - \alpha\Gamma(1 + \frac{1}{\beta})\Gamma(1 - \frac{1}{\beta}) \quad (۵)$$

که در این رابطه، $\Gamma(\cdot)$ تابع گاما و w_0 ، w_1 و w_2 گشتاورهای وزنی احتمال هستند. بدین ترتیب پس از محاسبه تابع توزیع تجمعی لوگ - لوجستیک مطابق رابطه (۶) و تبدیل آن به توزیع نرمال استاندارد با استفاده از تقریب آبرامویچ و استگان(۱)، شاخص SPEI مطابق رابطه (۷) استخراج می‌گردد.

$$F(x) = \left[1 + \left(\frac{\alpha}{x - \gamma} \right)^\beta \right]^{-1} \quad (۶)$$

$$SPEI = W - \frac{C_0 + C_1 W + C_2 W^2}{1 + d_1 W + d_2 W^2 + d_3 W^3} \quad (۷)$$

که در این رابطه، C_0 ، C_1 ، C_2 ، d_1 ، d_2 و d_3 مقادیر ثابت معادله SPEI می‌باشد و W از رابطه (۸) به دست می‌آید.

$$W = \sqrt{-2 \ln(P)} \quad \text{for} \quad P \leq 0.5 \quad (۸)$$

و $P = 1 - F(x)$. اگر $P > 0.5$ آنگاه عبارت $I - P$ در معادله (۸) جایگزین P می‌گردد. این شاخص می‌تواند برای پایش دوره‌های خشک و تر مورد استفاده قرار گیرد. خشکسالی زمانی شروع می‌گردد که مقادیر شاخص به منفی یک برسد و با مثبت شدن آن خاتمه می‌یابد.

شبکه‌های عصبی مصنوعی

ماهیت روابط تشریح‌کننده بسیاری از فرایندها و سری‌های زمانی اغلب غیرخطی است. لذا پیش‌بینی رفتار چنین فرایندها و سری‌هایی نیازمند ابزارهای پیش‌بینی دقیق است. در واقع، در بسیاری از سیستم‌های پیچیده و خصوصاً غیرخطی که مدل‌سازی و به دنبال آن پیش‌بینی و کنترل آنها از طریق روش‌های کلاسیک رگرسیونی امری بسیار دشوار و حتی بعضاً غیرممکن می‌باشد، از روش‌های غیر کلاسیک که عمدتاً برگرفته از رفتار طبیعت می‌باشد، استفاده می‌گردد. از جمله این روش‌ها، شبکه عصبی مصنوعی است که یکی از این روش‌های در حال تحول محسوب می‌شود و در موضوعات متنوعی از قبیل شناخت الگو، خوشه‌بندی، پیش‌بینی و ... به کار رفته و نتایج مفیدی داشته است. شبکه عصبی مصنوعی که الهام گرفته از شبکه عصبی بیولوژیکی مغز انسان است، شامل شبکه‌ای از نرون‌ها بوده که قادر است رفتار پیچیده‌ی بین نرون‌ها را نشان دهد. توانایی یادگیری، مهم‌ترین ویژگی شبکه عصبی و هر سیستم هوشمند است که

راه‌حل مسئله یافت شود. در روش GEP، پدیده‌های مختلف با استفاده از مجموعه‌ای از توابع و مجموعه‌ای از ترمینال‌ها برآورد و مدل‌سازی می‌شوند. مجموعه توابع شامل توابع اصلی حسابی و یا توابع تعریف شده توسط کاربر هستند. مجموعه ترمینال‌ها نیز از متغیرهای مستقل مسئله و مقادیر ثابت تشکیل شده‌اند. در این تحقیق از نرم‌افزار GeneXproTools 4.0 به منظور اعمال روش برنامه‌ریزی بیان ژن بر روی داده‌ها استفاده شده است.

فرآیند گام به گام برای مدل‌سازی متغیرهای در نظر گرفته شده در تحقیق حاضر با استفاده از مدل GEP به صورت زیر است؛

گام نخست: انتخاب تابع برازش که می‌تواند بر اساس روابط متعددی صورت گیرد که در این تحقیق از تابع برازش ریشه میانگین مربعات خطا (RMSE) استفاده شده است.

گام دوم: مجموعه متغیرهای ورودی و مجموعه توابع به منظور تولید کروموزوم‌ها انتخاب و معرفی می‌شوند. در این تحقیق، مجموعه متغیرهای ورودی شامل مقادیر شاخص SPEI در گام‌های زمانی مختلف می‌باشد.

گام سوم: ساختار و معماری کروموزوم‌ها انتخاب شدند. به طوری که اندازه طول هر رأس و تعداد ژن‌ها به ترتیب برابر ۸ و ۳ مورد استفاده قرار گرفت.

گام چهارم: در این گام انتخاب تابع پیوندی صورت گرفت که تابع جمع برای ایجاد پیوند بین زیرشاخه‌ها استفاده شد.

گام پنجم: در این مرحله عملگرهای ژنتیک و نرخ هر یک از آن‌ها انتخاب گردید.

روش شبکه بیزین

شبکه بیزین یک مدل گرافیکی احتمالاتی است که مجموعه‌ای از متغیرها و احتمالات مرتبط به هر کدام را نشان می‌دهد. این شبکه، یک گراف مستقیم و غیر چرخه‌ای است که در آن گره‌ها در حکم متغیرهای مسئله هستند. شبکه‌های بیزین محدودیتی در نمایاندن متغیرهای تصادفی ندارند. ساختار یک شبکه بیزین در واقع یک نمایش گرافیکی از اثرات متقابل متغیرهایی است که باید مدل شوند و علاوه بر این که کیفیت رابطه بین متغیرهای مسئله را نشان می‌دهد، کمیت ارتباط بین این متغیرها را نیز به نمایش می‌گذارد که به صورت عددی از توزیع احتمال مشترک آنها استفاده می‌کند. شبکه‌های بیزین را با نام‌هایی چون شبکه‌های تصمیم، شبکه‌های تصادفی، نمودارهای تأثیر و شبکه‌های بیز نیز می‌خوانند. این روش بر مبنای محاسبات احتمالات وابسته (قانون بیز) می‌باشد. آموزش شبکه بیزین شامل دو مرحله آموزش ساختاری و پارامتری می‌باشد. در بخش آموزش ساختاری فضای ممکن شکل ارتباطات با توجه به داده‌های مورد استفاده مشخص می‌شود که در این تحقیق از روش شرط مسیر (PC) استفاده شده است و در بخش آموزش پارامتری تخمین توزیع‌های احتمالاتی شرطی (θ) مدنظر می‌باشد که بدین منظور نخست یک توزیع یکنواخت اولیه بر پارامترهای ناشناخته برازش داده شده و سپس مقادیر (θ) با روش حداکثر درست‌نمایی داده‌های مورد استفاده مشخص می‌شود که برای این بخش نیز در این تحقیق از روش

به سیکل‌های آموزش کمتر بر روش پس‌انتشار^۱ برتری دارد. برای توسعه مدل‌های شبکه عصبی مصنوعی در تحقیق حاضر، از نرم‌افزار متلب (MATLAB) استفاده شد. اطلاعات با جزئیات بیشتر در خصوص شبکه‌های عصبی مصنوعی در منبع هایکین (۸) ارائه گردیده است.

برنامه‌ریزی بیان ژن

برنامه‌ریزی بیان ژن تعمیم یافته روش الگوریتم ژنتیک می‌باشد که نخستین بار توسط کوزا (۱۲) ارائه شد و بر مبنای نظریه تکامل داروین بنا نهاده شده است. GEP یک تکنیک برنامه‌ریزی خودکار است که راه حل مسئله را با استفاده از برنامه کامپیوتری ارائه می‌دهد. در این روش در ابتدای فرآیند رابطه تابعی بین متغیرها در نظر گرفته نمی‌شود و این روش قادر به بهینه‌سازی ساختار مدل و مؤلفه‌های آن می‌باشد. ساختارهای درختی از مجموعه توابع (عملگرهای ریاضی مورد استفاده در فرمول‌ها) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) تشکیل می‌شوند. نحوه مدل‌سازی به روش بیان ژن به صورت زیر می‌باشد.

با استفاده از ترکیب مجموعه توابع و ترمینال‌ها یک جمعیت اولیه ایجاد می‌شود.

هر یک از افراد جمعیت مذکور با استفاده از توابع برازش مورد ارزیابی قرار می‌گیرند.

تولید یک جمعیت جدید از فرمول‌ها که در این بخش از سه عمل ژنتیکی اصلی، تلاقی، جهش و تولید مثل یکی انتخاب می‌شود و سپس تعداد مناسبی از افراد جمعیت حاضر انتخاب می‌شوند و از عمل ژنتیکی انتخاب شده برای تولید فرزند (فرمول جدید) استفاده می‌شود و این فرمول جدید در یک جمعیت جدید وارد می‌شود و مدل مورد نظر با استفاده از تابع برازش مورد ارزیابی واقع می‌شود و این امر تا زمانی ادامه پیدا می‌کند که حداکثر تعداد تولید، تکرار خواهد شد.

مزیت برنامه‌ریزی بیان ژن نسبت به مدل‌های دیگر از جمله شبکه عصبی مصنوعی این است که در برنامه‌ریزی بیان ژن، ابتدا ساختار بلوک‌ها (متغیرهای ورودی، هدف و مجموعه توابع) تعریف شده و سپس ساختار بهینه مدل و ضرایب طی فرآیند آموزش تعیین می‌شوند، در حالی که در شبکه‌های عصبی، ابتدا باید ساختار تعیین شده و فقط ضرایب مدل، طی فرآیند آموزش حاصل شوند. همچنین این الگوریتم به طور خودکار می‌تواند متغیرهای ورودی که در مدل بیشترین تأثیر را دارند، انتخاب کند. در روش GEP کروموزوم‌ها به صورت بیان درختی نشان داده می‌شوند. سپس برازش‌های صورت گرفته با استفاده از توابع برازش مورد ارزیابی قرار می‌گیرند. اگر یک کیفیت مطلوبی از راه‌حل مسئله یافت شود و یا اینکه نسل‌ها به تعداد معینی برسند، فرآیند تکامل (تکرار) متوقف می‌شود و بهترین راه‌حل یافت شده گزارش می‌شود. با این حال، اگر شرایط توقف برنامه GEP حاصل نگردد بهترین راه‌حل یافت شده نگه داشته می‌شود که به فرآیند نخبه‌گزینی معروف است. در حقیقت، افراد شایسته شانس بیشتری برای تولید فرزندان دارند. پس از آن تغییرات ژنتیکی با استفاده از عملگرهای ژنتیکی مختلف اعمال می‌شود. این روند برای نسل‌های موجود تکرار می‌شود تا کیفیت رضایت‌بخشی از

که در این روابط، Q_i مقدار داده مشاهداتی در گام زمانی t ام، \bar{Q}_i مقدار داده پیش‌بینی شده در همان زمان، n تعداد داده‌ها، \bar{Q}_i میانگین مقادیر مشاهداتی و \bar{Q}_i میانگین مقادیر پیش‌بینی شده می‌باشد. هرچه مقادیر R به یک نزدیکتر باشد و $RMSE$ و MAE به صفر نزدیکتر باشد، نتایج دقیق‌تر و قابل اعتمادتر می‌باشد.

نتایج و بحث

جهت پیش‌بینی شاخص SPEI در مقیاس‌های زمانی مختلف با استفاده از مدل‌های هوشمند ابتدا داده‌های شاخص مطابق روابط (۱۶) و (۱۷) نرمال گردیدند.

$$Y_i = \frac{X_{oi}}{X_{o \max}}, X_{oi} \geq 0 \quad (16)$$

$$Y_i = \frac{X_{oi}}{|X_{o \min}|}, X_{oi} < 0 \quad (17)$$

که در این روابط، Y_i مقادیر استاندارد شده شاخص، X_{oi} مقادیر شاخص مشاهداتی، $X_{o \max}$ حداکثر مقدار شاخص مشاهداتی و $X_{o \min}$ حداقل مقدار شاخص مشاهداتی می‌باشد. سپس ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد باقی‌مانده آن برای صحت‌سنجی انتخاب گردید. به منظور پیش‌بینی خشکسالی با روش‌های هوشمند در این تحقیق از ۵ مدل مختلف استفاده شد. سپس بر اساس معیارهای ارزیابی بهینه‌ترین مدل در پیش‌بینی شاخص خشکسالی در مقیاس‌های زمانی مختلف انتخاب گردید. در مدل ۱، برای پیش‌بینی شاخص SPEI در ماه بعد از سری زمانی SPEI در ماه جاری استفاده گردید. در مدل ۲ پیش‌بینی شاخص در ماه بعد بر اساس مقادیر شاخص تا یک ماه قبل صورت پذیرفت و به طور مشابه در مدل ۳، ۴ و ۵ پیش‌بینی بر اساس مقادیر شاخص تا دو، سه و چهار ماه قبل انجام شد.

$$SPEI_{(t+1)} = f(SPEI_{(t)}) \quad \text{مدل ۱} \quad (18)$$

$$SPEI_{(t+1)} = f(SPEI_{(t)}, SPEI_{(t-1)}) \quad \text{مدل ۲} \quad (19)$$

$$SPEI_{(t+1)} = f(SPEI_{(t)}, SPEI_{(t-1)}, SPEI_{(t-2)}) \quad \text{مدل ۳} \quad (20)$$

$$SPEI_{(t+1)} = f(SPEI_{(t)}, SPEI_{(t-1)}, SPEI_{(t-2)}, SPEI_{(t-3)}) \quad \text{مدل ۴} \quad (21)$$

$$SPEI_{(t+1)} = f(SPEI_{(t)}, SPEI_{(t-1)}, SPEI_{(t-2)}, SF) \quad \text{مدل ۵} \quad (22)$$

نتایج روش برنامه‌ریزی بیان ژن

در این مدل ابتدا باید بهترین الگوی ورودی به مدل مشخص گردد. از بین الگوهای مختلف، الگویی که بهترین کارایی را ایجاد می‌کند به عنوان مناسب‌ترین الگوی ورودی در نظر گرفته می‌شود. در ادامه با در دست بودن سری زمانی شاخص SPEI اقدام به پیش‌بینی خشکسالی با استفاده از روش برنامه‌ریزی ژنتیک شد. برتری مدل برنامه‌ریزی ژنتیک نسبت به مدل‌های دیگر را می‌توان در خروجی صریح آن به صورت فرمول جبری در قیاس با ماتریس ضرایب وزنی شبکه عصبی یا عصبی-فازی برشمرد. به دلیل اینکه روند تکاملی

تخمین حداکثر (EM) استفاده شده است. الگوریتم EM در مواردی که هدف تخمین مجموعه‌ای از پارامترهای (θ) که مبتنی بر یک توزیع احتمالاتی بوده و تنها بخشی از داده‌ها که توسط این توزیع احتمالاتی تولید شده‌اند در دسترس است، مورد استفاده قرار می‌گیرد. در صورتی که مجموعه داده‌های مشاهداتی به صورت $X = \{x_1, x_2, \dots, x_m\}$ و با فرض استقلال از هم بوده و مجموعه داده‌های نامرئی به صورت $Z = \{z_1, z_2, \dots, z_m\}$ و کل داده‌ها به صورت $Y = X \cup Z$ باشد الگوریتم EM با فرض بیشینه نزدیکی (ML) تبدیل به h' می‌گردد به نحوی که $P(Y|h')$ بیشینه گردد در این صورت این الگوریتم برای تخمین توزیع احتمالی روی Y به جای پارامترهای واقعی (θ) از فرض فعلی h استفاده می‌کند. تابع $Q(h|h')$ را بنحوی تعریف می‌کنیم که $E(\ln P(Y|h'))$ به صورت تابعی از h' تبدیل شود. لذا با فرض $h = \theta$ و اینکه داده‌های X بخشی از کل داده‌های مشاهداتی (Y) می‌باشد خواهیم داشت:

$$Q(h|h') = E[\ln P(Y|h')|h, X] \quad (11)$$

مرحل الگوریتم EM به صورت زیر می‌باشد:
گام ۱ مرحله تخمین (E) : با استفاده از مقدار فرض فعلی h و داده‌های مشاهداتی X مقدار $Q(h|h')$ را محاسبه می‌کند تا توزیع احتمالی روی Y تخمین زده شود.
گام ۲ مرحله بیشینه‌سازی (M) : فرض h را با فرض h' جایگزین می‌کند تا تابع Q بیشینه شود.

$$\max_{h \rightarrow h'} Q(h|h') \quad (12)$$

در تحقیق حاضر، سه گام زیر برای مدل‌سازی پیش‌بینی خشکسالی با استفاده از شبکه بیزین انجام گرفت.

- ایجاد مدل مفهومی بر اساس ورودی‌های در نظر گرفته شده برای مدل‌سازی متغیر مستقل شاخص SPEI
- ایجاد ارتباطات مابین متغیرهای مدل مفهومی ایجاد شده با استفاده از کمان‌های ارتباطی. در ایجاد این ارتباط مقادیر شاخص SPEI در گام زمانی آینده متأثر از مقادیر این شاخص با تأخیرهای زمانی مشخص می‌باشد.
- آموزش شبکه بیزین توسعه داده شده. پس از آموزش شبکه، مدل بیزین رابطه ریاضی ارائه می‌نماید که این معادله برای بدست آوردن مقادیر پیش‌بینی شده توسط شبکه‌های بیزین مورد استفاده قرار می‌گیرد.

معیارهای ارزیابی

در این تحقیق، به منظور ارزیابی مدل‌های موردنظر از معیارهای ضریب همبستگی، میانگین قدر مطلق خطا و جذر میانگین مربعات خطا استفاده گردید که روابط آنها به ترتیب به صورت روابط (۱۳) الی (۱۵) می‌باشد.

$$R = \frac{\sum_{i=1}^n (Q_i - \bar{Q}_i)(\hat{Q}_i - \bar{\hat{Q}}_i)}{\sqrt{\sum_{i=1}^n (Q_i - \bar{Q}_i)^2 \sum_{i=1}^n (\hat{Q}_i - \bar{\hat{Q}}_i)^2}} \quad (13)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Q_i - \hat{Q}_i| \quad (14)$$

$$RMSE = \left(\frac{\sum_{i=1}^n (Q_i - \hat{Q}_i)^2}{n} \right)^{0.5} \quad (15)$$

در این تحقیق علاوه بر عملگرهای ریاضی پیش‌فرض برنامه F_1 از ترکیب ۴ عملگر اصلی (F_2) نیز استفاده شده است که در روابط (۲۳) و (۲۴) ارائه شده‌اند.

$$F_1: \{+, -, \times, \div, \sqrt{\quad}, \text{Exp}, \ln, x^2, x^3, \sqrt[3]{\quad}, \sin, \cos, \text{Atan}\} \quad (23)$$

$$F_2: \{+, -, \times, \div\} \quad (24)$$

در این مدل دارای یک فرآیند از پیش تعیین شده و مشخص نمی‌باشد، با یک بار اجرای برنامه نمی‌توان به جواب مطلوب و بهینه دست یافت. به همین دلیل جهت تنظیم بهینه و دقیق پارامترهای مدل باید از روش سعی و خطا با تکرارهای مختلف استفاده کرد. در این تحقیق، از حداقل سی بار اجرای برنامه برای یافتن مقادیر پارامترها و تنظیمات مختلف استفاده گردید که نتایج بهترین ساختارها و تنظیمات آن در استخراج مدل پیش‌بینی خشکسالی در جدول (۱) نشان داده شده است.

جدول ۱- مشخصات مدل برنامه‌ریزی بیان ژن مورد استفاده برای پیش‌بینی مقادیر SPEI
Table 1. Specifications of gene expression programming model used for SPEI values prediction

عملگرهای ژنتیکی	تنظیمات کلی
۰/۰۴۴	تعداد کروموزومها
۰/۱	اندازه رأس
۰/۱	تعداد ژن‌ها در هر کروموزوم
۰/۱	تعداد جمعیت تولیدی
۰/۳	تابع پیوند
۰/۳	جمع (+)
۰/۱	تابع برازش معیار خطا
۰/۱	RMSE
۰/۱	نرخ جهش
۰/۱	نرخ وارون‌سازی
۰/۱	نرخ ترانهش درج متوالی
۰/۱	نرخ ترانهش ریشه درج متوالی
۰/۳	نرخ ترکیب تک نقطه‌ای
۰/۳	نرخ ترکیب دو نقطه‌ای
۰/۱	نرخ ترکیب ژن
۰/۱	نرخ ترانهش ژن

می‌یابد. در این خصوص نتایج مشابهی توسط میشرا و دسای (۱۸،۱۷) و حسینی موغاری و عراقی‌نژاد (۹) برای پیش‌بینی شاخص SPI ارائه شده است. همچنین در جدول (۲) مقادیر معیارهای ارزیابی مربوط به دوره آزمون در مقیاس‌های زمانی مختلف و برای دو مجموعه عملگر ریاضی مختلف نشان داده شده است. بر اساس این جدول مشخص می‌گردد که استفاده از ترکیب هر دو نوع تابع F_1 و F_2 در مقیاس‌های زمانی بالا، از دقت مناسب برخوردار است و نتایج آنها به هم نزدیک است و به دلیل اینکه روابط توابع ریاضی (F_1) منجر به ایجاد فرمول‌های پیچیده‌ای می‌گردد که درک روابط میان متغیرها و نیز استفاده از آن برای کاربران کمی دشوار خواهد بود و همچنین روابط حسابی (F_2) دارای ساختار ساده ریاضی و سهولت در محاسبه می‌باشد لذا ترکیب F_2 به عنوان عملگر مناسب در پیش‌بینی شاخص SPEI در این تحقیق انتخاب گردید.

مدل مناسب پیش‌بینی بر اساس معیارهای ارزیابی R، RMSE و MAE برای هر دو ترکیب عملگر ریاضی و در تمامی مقیاس‌های زمانی استخراج شد. بر اساس آماره‌های ارزیابی ارائه شده در جدول (۲) در مرحله آزمون بهترین الگوی ورودی برای پیش‌بینی شاخص SPEI در گام زمانی ۱ تا ۴۸ ماهه انتخاب گردید که آماره‌های ارزیابی برجسته (بولد) شده در این جدول مربوط به مدل برتر می‌باشد. همچنین نتایج حاکی از عملکرد مناسب روش GEP در مقیاس‌های بالای شاخص SPEI می‌باشد. در خصوص شاخص SPEI1، روش GEP توانایی قابل قبولی در پیش‌بینی خشکسالی را ندارد که علت این موضوع را می‌توان در پراکندگی و تغییرات شدید شاخص SPEI در مقیاس یک ماهه جستجو کرد. بهبود عملکرد روش GEP با افزایش مقیاس زمانی محاسبه SPEI رابطه مستقیمی دارد. بدین معنی که با افزایش مقیاس محاسبه SPEI، مقادیر معیارهای ارزیابی مورد استفاده بهبود

جدول ۲- نتایج روش GEP در دوره صحت‌سنجی

Table 2. Result of GEP method in the test period

معیارهای ارزیابی	مدل	تابع F ₁						تابع F ₂					
		SPEI1	SPEI3	SPEI6	SPEI12	SPEI24	SPEI48	SPEI1	SPEI3	SPEI6	SPEI12	SPEI24	SPEI48
RMSE	۱	۰/۴۱۶	۰/۴۹۳	۰/۲۱۳	۰/۱۲	۰/۰۶۹	۰/۰۵۹	۰/۳۹۱*	۰/۳۴۷	۰/۲۴۱	۰/۱۱۹	۰/۰۶	۰/۰۶۸
	۲	۰/۴۰۸	۰/۳۱۱	۰/۲۱۱	۰/۱۲۳	۰/۰۶۹	۰/۰۵۹	۰/۴۱۶	۰/۵۴۶	۰/۲۳۸	۰/۱۲۱	۰/۰۶	۰/۰۶۹
	۳	۰/۴۲۷	۰/۳۳	۰/۲۰۸	۰/۱۲	۰/۰۶۹	۰/۰۶	۰/۴۰۷	۰/۲۸۸	۰/۲۰۸	۰/۱۱۹	۰/۰۶	۰/۰۷۱
	۴	۰/۴۳	۰/۲۸۹	۰/۲۱۷	۰/۱۲	۰/۰۷۳	۰/۰۵۸	۰/۴۱۳	۰/۲۷۳*	۰/۲۱	۰/۱۲۱	۰/۰۶۱	۰/۰۷
	۵	۰/۴۳۲	۰/۲۸۷	۰/۲۱۵	۰/۱۲۲	۰/۰۶۹	۰/۰۵۹	۰/۳۹۷	۰/۲۸۲	۰/۲۱۱	۰/۱۲۱	۰/۰۶۲	۰/۰۶۹
MAE	۱	۰/۳۵۸	۰/۲۱۳	۰/۱۵۲	۰/۰۸۵	۰/۰۵۱	۰/۰۴۴	۰/۳۲۷*	۰/۲۳۱	۰/۱۶۲	۰/۰۸۵	۰/۰۴	۰/۰۵
	۲	۰/۳۳۴	۰/۲۲۱	۰/۱۴۱	۰/۰۸۶	۰/۰۵۱	۰/۰۴	۰/۳۴۲	۰/۲۸۶	۰/۱۶۳	۰/۰۸۴	۰/۰۴۱	۰/۰۵
	۳	۰/۳۶۳	۰/۲۳۸	۰/۱۴۲	۰/۰۸۵	۰/۰۵	۰/۰۴	۰/۳۳۵	۰/۱۹۸	۰/۱۶۶	۰/۰۸	۰/۰۴۳	۰/۰۵۲
	۴	۰/۳۵۱	۰/۲۰۸	۰/۱۵۱	۰/۰۸۵	۰/۰۵۴	۰/۰۴۱	۰/۳۳۶	۰/۲	۰/۱۴۳	۰/۰۸۷	۰/۰۴۳	۰/۰۵۳
	۵	۰/۳۴	۰/۱۹۹	۰/۱۶۹	۰/۰۸۴	۰/۰۵۱	۰/۰۴۳	۰/۳۱۵	۰/۳۱	۰/۱۴۳	۰/۰۸۵	۰/۰۴۳	۰/۰۵
R	۱	۰/۱۳۴	۰/۶۴۱	۰/۸۴۷	۰/۹۲۴	۰/۹۷	۰/۹۵۶	۰/۱۳۳*	۰/۵۲۹	۰/۸۰۲	۰/۹۲۵	۰/۹۵۸	۰/۹۷
	۲	۰/۰۵۴	۰/۶۵۵	۰/۸۴۸	۰/۹۲۱	۰/۹۷	۰/۹۶	۰/۰۲۶	۰/۲۹۸	۰/۸۰۵	۰/۹۲۴	۰/۹۵۹	۰/۹۷
	۳	۰/۰۳۴	۰/۶۲۲	۰/۸۵۲	۰/۹۲۴	۰/۹۷	۰/۹۶	۰/۰۵۶	۰/۶۳۹	۰/۸۵۴	۰/۹۲۶	۰/۹۵۷	۰/۹۷
	۴	۰/۰۳۶	۰/۶۵۵	۰/۸۴۹	۰/۹۲۴	۰/۹۶	۰/۹۶	۰/۰۳۷	۰/۶۶۳	۰/۸۵۱	۰/۹۲۲	۰/۹۵۷	۰/۹۷
	۵	۰/۲۱۸	۰/۶۴۲	۰/۸۴۴	۰/۹۲۵	۰/۹۷	۰/۹۶	۰/۰۷۶	۰/۲۰۹	۰/۸۵	۰/۹۲۵	۰/۹۵۴	۰/۹۷

*: مقادیر برجسته شده در این جدول نشان دهنده مدل برتر انتخابی می‌باشد.

یکی از قابلیت‌های مهم روش GEP، ارائه روابط ریاضی برای متغیرهای هیدرولوژیکی می‌باشد. روابط نهایی برای مجموعه عملگر F₂ و برای ترکیب منتخب شاخص SPEI در مقیاس‌های زمانی مختلف در جدول (۳) نشان داده شده است.

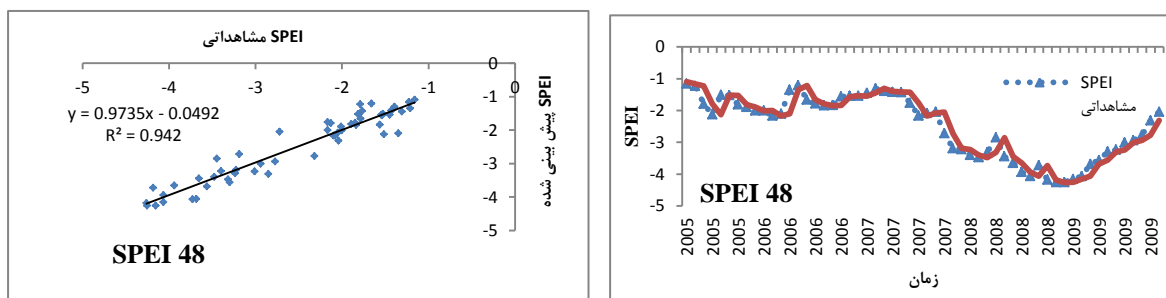
جدول ۳- روابط ریاضی برای بهترین مدل پیش‌بینی برنامه‌ریزی بیان ژن

Table 3. Mathematical relations for the best GEP prediction model

مقیاس زمانی	مدل برتر	رابطه ریاضی مدل برنامه‌ریزی بیان ژن
۱	۱	$SPEI(t+1) = \frac{SPEI(t)}{3SPEI(t) + 6.053} + \frac{SPEI(t)}{(3.02 - SPEI(t)) \times 4 SPEI(t)}$
۳	۴	$SPEI(t+1) = SPEI(t-1) \times SPEI(t-3) \times SPEI(t)^2 + SPEI(t) + \frac{6 SPEI(t-1) \times SPEI(t)}{-5.6 SPEI(t-3) \times SPEI(t-2)}$
۶	۳	$SPEI(t+1) = SPEI(t) + \frac{SPEI(t-3)^2}{-9.36 - 2 SPEI(t-2)} + \frac{2.38 (SPEI(t-3)^2 - (SPEI(t) \times SPEI(t-1)))}{9.234}$
۱۲	۳	$SPEI(t+1) = SPEI(t) + \frac{(SPEI(t) - SPEI(t-1)) \times SPEI(t-2)}{SPEI(t-2) \times 8.53} + \frac{SPEI(t-2)}{SPEI(t-1) - 3.56 + SPEI(t-2)}$
۲۴	۲	$SPEI(t+1) = SPEI(t) + \frac{-0.00018}{SPEI(t-1)} + SPEI(t) - 0.00006 SPEI(t-1)$
۴۸	۱	$SPEI(t+1) = 2SPEI(t) + \frac{SPEI(t)}{2SPEI(t) - 6.34}$

و پیش‌بینی شده بسیار به هم نزدیک بوده و مدل GEP از دقت بالایی برخوردار است.

پراکنش و سری زمانی مقادیر پیش‌بینی شده و مشاهداتی برای مدل برتر و برای مقیاس زمانی ۴۸ ماهه به عنوان نمونه در شکل (۱) ارائه شده است که بر اساس آن مشاهده می‌گردد در مقیاس‌های زمانی بالای شاخص SPEI مقادیر مشاهداتی



شکل ۱- سری زمانی و پراکنش مقادیر مشاهداتی و پیش‌بینی شده شاخص SPEI با استفاده از مدل برتر GEP در مقیاس زمانی ۴۸ ماهه
Figure 1. Time series and scatter plot of observational and predicted values of the SPEI index using the superior GEP model in the 48-month time-scale

می‌گردد که در مقیاس زمانی کوتاه مدت عملکرد شبکه بیزین مناسب نبوده و با افزایش مقیاس زمانی شاخص مذکور میزان عملکرد شبکه بیزین نیز افزایش پیدا کرده و عملکرد مدل پیش‌بینی با مقیاس زمانی شاخص SPEI رابطه مستقیم دارد به نحوی که مدل BN در مقیاس زمانی یک ماهه بدترین عملکرد و در مقیاس ۴۸ ماهه مناسب‌ترین عملکرد را داشته است. بهترین مدل برای مقیاس‌های زمانی مختلف در ایستگاه‌های مختلف متفاوت بوده و اغلب از ترتیب خاصی نسبت به افزایش تأخیر زمانی تبعیت نمی‌کند و ترکیبات مختلفی از تأخیرها به عنوان مدل برتر انتخاب می‌گردد. مقادیر آماره‌های ارزیابی برجسته (بولد) شده در جدول (۴) نشانگر مقادیر آماره‌های ارزیابی مدل برتر می‌باشد. بر اساس این جداول مشخص می‌گردد که روش شبکه بیزین در مقیاس زمانی یک ماهه در مدل یک با ضریب همبستگی ۰/۱۴۲، در مقیاس زمانی ۳ ماهه در مدل ۴ با ضریب همبستگی ۰/۶۶۷، در مقیاس زمانی ۶ ماهه در مدل ۱ با ضریب همبستگی ۰/۸۴۸، در مقیاس زمانی ۱۲ ماهه مدل ۲ با ضریب همبستگی ۰/۹۲۶، در مقیاس زمانی ۲۴ ماهه در مدل ۲ با ضریب همبستگی ۰/۹۵۹ و در مقیاس زمانی ۴۸ ماهه در مدل ۲ با ضریب همبستگی ۰/۹۷۱ بهترین عملکرد را داشته است.

نتایج روش شبکه بیزین

در این تحقیق با فرض وابستگی شاخص SPEI به مقادیر قبلی خود و انتخاب الگوریتم PC^۱ با سطح معنی‌داری ۵ درصد ساختار شبکه با توجه به تأثیر پارامترها بر روی یکدیگر و ۵ مدل ورودی در مقیاس‌های زمانی مختلف ایجاد گردید. جهت ایجاد و توسعه شبکه بیزین نخست چارچوب آن به معنی تعریف متغیرها و بررسی روابط موجود بین آنها بر اساس یک مدل مفهومی ایجاد گردید و سپس نوع ارتباط بین متغیرهای مختلف در شبکه که در این تحقیق مقادیر شاخص در یک ماه آینده متأثر از مقادیر شاخص با تأخیرهای زمانی مختلف می‌باشد تعیین گردید. در ادامه آموزش شبکه و به دست آوردن رابطه بین متغیرها مورد بررسی قرار گرفت که در این مرحله با استفاده از مقادیر شاخص SPEI با تأخیرهای زمانی مختلف به عنوان ورودی شبکه بیزین سعی در شناسایی رابطه ریاضی بین ورودی و خروجی مدل گردید که بر اساس آن می‌توان داده‌های مربوط به شاخص SPEI را به صورت مصنوعی تولید کرده و عملکرد مدل را ارزیابی کرد. در خاتمه با استفاده از معیارهای ارزیابی مختلف عملکرد مدل پیش‌بینی مورد مقایسه قرار می‌گیرد و مدل برتر بر اساس آماره‌های ارزیابی انتخاب می‌گردد. در جدول (۴) نتایج مربوط به الگوهای مختلف ورودی شبکه بیزین در مقیاس‌های زمانی مختلف شاخص SPEI در دوره آزمون نشان داده شده است. بر اساس آماره‌های ارزیابی موجود در این جدول مشخص

جدول ۴- نتایج روش شبکه بیزین در دوره آزمون

Table 4. Result of BN method in the test period

معیارهای ارزیابی	مدل	مقیاس زمانی					
		SPEI1	SPEI3	SPEI6	SPEI12	SPEI24	SPEI48
RMSE	۱	$+0.378^*$	-0.283	$+0.209^*$	-0.119	-0.069	-0.052
	۲	-0.384	-0.287	-0.211	$+0.118^*$	$+0.068^*$	$+0.052^*$
	۳	-0.382	-0.288	-0.212	-0.119	-0.07	-0.061
	۴	-0.382	$+0.277^*$	-0.212	-0.119	-0.071	-0.061
	۵	-0.395	-0.278	-0.212	-0.169	-0.132	-0.061
MAE	۱	$+0.294^*$	-0.199	$+0.141^*$	-0.084	-0.051	-0.044
	۲	-0.292	-0.199	-0.141	$+0.083^*$	$+0.049^*$	$+0.04^*$
	۳	-0.291	-0.201	-0.143	-0.083	-0.051	-0.043
	۴	-0.291	$+0.193^*$	-0.143	-0.083	-0.051	-0.042
	۵	-0.293	-0.193	-0.144	-0.132	-0.104	-0.043
R	۱	$+0.142^*$	-0.639	$+0.848^*$	-0.925	-0.958	-0.97
	۲	-0.047	-0.64	-0.847	$+0.926^*$	$+0.959^*$	$+0.971^*$
	۳	-0.056	-0.642	-0.847	-0.926	-0.956	-0.969
	۴	-0.045	$+0.667^*$	-0.847	-0.926	-0.957	-0.968
	۵	-0.096	-0.664	-0.847	-0.922	-0.956	-0.963

*: مقادیر برجسته شده در این جدول نشان دهنده مدل برتر انتخابی می‌باشد.

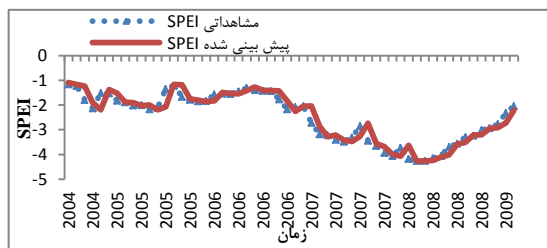
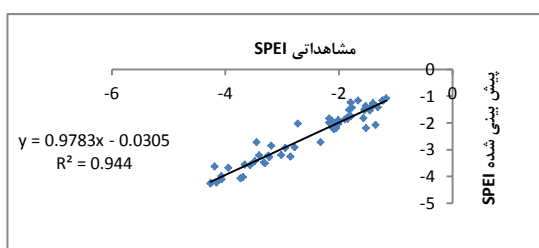
در شکل (۲) سری زمانی و پراکنش مقادیر مشاهداتی و پیش‌بینی شده شاخص SPEI برای مقیاس زمانی ۴۸ ماهه در مرحله آزمون ارائه شده است که بر اساس آن مشخص می‌گردد که مدل BN دارای دقت مناسب در مقیاس زمانی بلندمدت شاخص SPEI می‌باشد.

رابطه ریاضی حاصل از مدل برتر برای مقیاس‌های زمانی مختلف برای ایستگاه مراغه در جدول ۵ ارائه شده است که بر اساس رابطه خروجی شبکه بیزین در گام‌های زمانی مختلف مشاهده می‌گردد که متغیر شاخص SPEI با یک گام زمانی تأخیر بیش‌ترین ضریب را در معادله ریاضی به خود اختصاص داده است و بیشترین تأثیر را در مدل پیش‌بینی بر عهده دارد.

جدول ۵- روابط ریاضی برای بهترین مدل پیش‌بینی شبکه بیزین

Table 5. Mathematical relations for the best BN prediction model

مقیاس زمانی	مدل برتر	رابطه ریاضی
۱	۱	$SPEI_{t+1} = 0.1714 SPEI_t$
۳	۴	$SPEI_{t+1} = 0.31 SPEI_{t-3} - 0.333 SPEI_{t-2} + 0.49 SPEI_{t-1} + 0.799 SPEI_t$
۶	۱	$SPEI_{t+1} = 0.924 SPEI_t$
۱۲	۲	$SPEI_{t+1} = -0.156 SPEI_{t-1} + 1.13 SPEI_t$
۲۴	۲	$SPEI_{t+1} = -0.11 SPEI_{t-1} + 1.1 SPEI_t$
۴۸	۲	$SPEI_{t+1} = -0.23 SPEI_{t-1} + 1.225 SPEI_t$



شکل ۲- سری زمانی و پراکنش مقادیر مشاهداتی و پیش‌بینی شده شاخص SPEI با استفاده از مدل برتر شبکه بیزین در مقیاس زمانی ۴۸ ماهه

Figure 2. Time series and scatter plot of observational and predicted values of the SPEI index using the superior BN model in the 48-month time-scale

مختلفی از این شبکه که دارای ورودی‌های متفاوتی است به صورت کلی (P,n,1) مورد آزمایش قرار گرفت که در این ساختار P تعداد ورودی‌های شبکه (نرون‌های ورودی) و n

نتایج روش شبکه عصبی
در گام نخست به منظور بررسی کارایی روش شبکه عصبی در مدل‌سازی داده‌های سری زمانی خشکسالی، آرایش‌های

MATLAB استفاده شد. در این جعبه ابزار داده‌ها به صورت تصادفی و خودکار به سه قسمت تقسیم می‌گردد که حدود ۷۰ درصد داده‌ها برای آموزش و حدود ۱۵ درصد از داده‌ها هر کدام برای بخش آزمون و اعتبارسنجی مورد استفاده قرار می‌گیرد.

تعداد نرون‌های لایه میانی است که به صورت متغیر می‌تواند انتخاب گردد. به این منظور ۵ ساختار برای حالت‌هایی که به تعداد ۱، ۲، ۳، ۴ و ۵ وابستگی در شبکه وجود دارد با تعداد ۱ تا ۲۰ نرون برای لایه میانی مورد آزمایش قرار گرفت که نتایج ساختارهای بهینه در جدول (۶) آورده شده است. در این تحقیق از جعبه ابزار شبکه عصبی مصنوعی نرم‌افزار

جدول ۶- نتایج ساختار بهینه شبکه عصبی مصنوعی

Table 6. Results of optimal structure of ANN

مقیاس زمانی	SPEI 1	SPEI 3	SPEI 6	SPEI 12	SPEI 24	SPEI 48
مدل برتر	۱	۴	۴	۲	۴	۵
ساختار بهینه	۱-۱۲-۱	۴-۸-۱	۴-۱۲-۱	۲-۸-۱	۴-۸-۱	۵-۶-۱

ANN با ضریب همبستگی ۰/۳۸ و مدل ۱ بوده که عملکرد مناسبی نداشته است و با افزایش مقیاس زمانی، عملکرد مدل پیش‌بینی نیز بهبود پیدا کرده است به نحوی که در مقیاس زمانی ۴۸ ماهه ضریب همبستگی در مدل برتر (مدل ۵) به ۰/۹۷ رسیده است.

نتایج مربوط به آماره‌های ارزیابی برای مقیاس‌های زمانی مختلف و ورودی‌های مختلف (از یک تا ۵ تأخیر) در جدول (۷) ارائه شده است. در این جدول مقادیر برجسته شده نشانگر مدل برتر در هر مقیاس زمانی می‌باشد. بر اساس آماره‌های ارزیابی در مقیاس زمانی یک ماهه بهترین عملکرد مدل

جدول ۷- نتایج روش شبکه عصبی مصنوعی در دوره آزمون

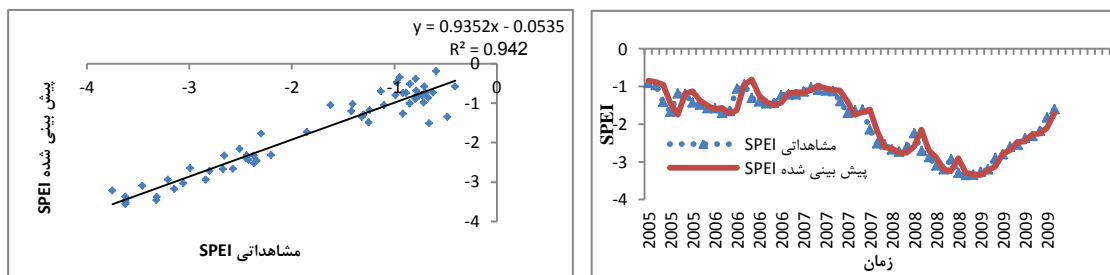
Table 7. Result of ANN method in the test period

معیارهای ارزیابی	مدل	مقیاس زمانی					
		SPEI1	SPEI3	SPEI6	SPEI12	SPEI24	SPEI48
RMSE	۱	۰/۴۱۸*	۰/۳۶۸	۰/۲۲۰	۰/۱۶۳	۰/۰۸۴	۰/۰۶۱
	۲	۰/۴۱۹	۰/۳۳۶	۰/۲۳۲	۰/۱۳۰*	۰/۰۷۶	۰/۰۶۴
	۳	۰/۴۲۳	۰/۳۰۹	۰/۲۸۱	۰/۱۳۴	۰/۰۷۹	۰/۰۶۴
	۴	۰/۴۲۲	۰/۲۷۱*	۰/۲۱۶*	۰/۱۵۵	۰/۰۶۷	۰/۰۶۰
	۵	۰/۴۵۳	۰/۲۹۲	۰/۲۴۶	۰/۱۶۰	۰/۰۶۸	۰/۰۶۰
MAE	۱	۰/۳۲۶*	۰/۲۷۵	۰/۱۵۶	۰/۱۲۹	۰/۰۶۶	۰/۰۴۳
	۲	۰/۳۲۷	۰/۲۳۲	۰/۱۷۴	۰/۰۹۴*	۰/۰۵۶	۰/۰۴۶
	۳	۰/۳۳۵	۰/۲۳۴	۰/۲۱۰	۰/۱۰۳	۰/۰۵۶	۰/۰۴۵
	۴	۰/۳۴۲	۰/۱۹۲*	۰/۱۵۶*	۰/۱۱۵	۰/۰۵۱*	۰/۰۴۳
	۵	۰/۳۸۱	۰/۲۱۷	۰/۱۶۱	۰/۱۲۰	۰/۰۵۱	۰/۰۴۳*
R	۱	۰/۰۳۸*	۰/۴۳۷	۰/۸۳۸	۰/۸۵۸	۰/۹۵۷	۰/۹۶۵
	۲	۰/۰۰۱	۰/۴۷۷	۰/۸۲۹	۰/۹۲۰*	۰/۹۵۰	۰/۹۶۴
	۳	۰/۰۲۴	۰/۶۵۲	۰/۷۲۴	۰/۹۱۳	۰/۹۵۰	۰/۹۶۴
	۴	۰/۰۰۱	۰/۷۰۴*	۰/۸۴۲*	۰/۸۹۴	۰/۹۵۹*	۰/۹۶۴
	۵	۰/۰۱۵	۰/۶۶۰	۰/۸۰۰	۰/۸۹۳	۰/۹۵۸	۰/۹۶۷*

*: مقادیر برجسته شده در این جدول نشان دهنده مدل برتر انتخابی می‌باشد.

قابل قبول روش شبکه عصبی در پیش‌بینی مقادیر شاخص خشکسالی در مقیاس زمانی بلندمدت می‌باشد.

در شکل (۳) سری زمانی و پراکنش مقادیر مشاهداتی و پیش‌بینی شده شاخص SPEI با مدل برتر به عنوان نمونه برای مقیاس زمانی ۴۸ ماهه ارائه شده است که نشان از دقت



شکل ۳- سری زمانی و پراکنش مقادیر مشاهداتی و پیش‌بینی شده شاخص SPEI با استفاده از مدل برتر شبکه عصبی مصنوعی در مقیاس زمانی ۴۸ ماهه

Figure 3. Time series and scatter plot of observational and predicted values of the SPEI index using the superior ANN model in the 48-month time- scale

RMSE و MAE به ترتیب ۰/۱۴۲، ۰/۳۶۸ و ۰/۲۹۴ و در مقیاس زمانی ۳ ماهه بهترین عملکرد مربوط به مدل شبکه عصبی مصنوعی با ضریب همبستگی، RMSE و MAE به ترتیب ۰/۷۰۴، ۰/۲۷۱ و ۰/۱۹۲ می‌باشد. همچنین در مقیاس زمانی بلندمدت عملکرد هر سه روش تقریباً یکسان می‌باشد و نتایج نشان داد که هر سه روش در مقیاس‌های زمانی بلندمدت شاخص SPEI از دقت مناسب برخوردار هستند و عملکرد مدل‌ها با افزایش مقیاس زمانی شاخص SPEI رابطه مستقیم دارد. همچنین یکی از مزیت‌های روش شبکه بیزین و برنامه‌ریزی بیان ژن، ارائه رابطه ریاضی بین متغیرهای مستقل و وابسته می‌باشد که می‌توان از این رابطه ریاضی جهت پیش‌بینی‌های آتی استفاده کرد.

تحقیق حاضر تلاشی بود جهت پیش‌بینی خشکسالی در ایستگاه سینوپتیک مراغه بر اساس شاخص SPEI و استفاده از مدل‌های هوشمند شبکه عصبی- مصنوعی، برنامه‌ریزی بیان ژن و شبکه بیزین که بدین منظور از ۵ مدل ورودی شامل مقادیر شاخص SPEI با تأخیرهای زمانی مختلف استفاده گردید. در روش GEP از ترکیب دو تابع مختلف شامل ۴ عملگر اصلی و پیش‌فرض برنامه جهت پیش‌بینی نمودن مقادیر شاخص خشکسالی استفاده گردید که نتایج نشان داد تابع ۴ عملگر اصلی دارای دقت مناسب می‌باشد و همچنین سبب سهولت در محاسبات می‌گردد. نتایج همچنین نشان داد هر سه روش در مقیاس زمانی کوتاه مدت از دقت مناسب برخوردار نیستند و بهترین عملکرد در مقیاس زمانی ۱ ماهه مربوط به مدل شبکه بیزین با ضریب همبستگی،

منابع

1. Abramowitz, M., I.A. Stegun, and D. Miller. 1965. Handbook of mathematical functions with formulas, graphs and mathematical tables (National Bureau of Standards Applied Mathematics Series No. 55). 239-239.
2. Allen, R.G., L.S. Pereira, D. Raes and M. Smith. 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56 FAO, Rome 300:D05109.
3. Da Silva, V.d.P.R. 2004. On climate variability in Northeast of Brazil. Journal of Arid Environments, 58(4): 575-596.
4. Dastorani, M. and H. Afkhami. 2011. Application of artificial neural networks on drought prediction in Yazd (Central Iran). Desert, 16(1): 39-48.
5. Farmani, R., H.J. Henriksen, D. Savic and D. Butler. 2012. An evolutionary Bayesian belief network methodology for participatory decision making under uncertainty: An application to groundwater management. Integrated environmental assessment and management, 8(3): 456-461.
6. Ghorbani, M.A., J. Shiri and H. Kazemi. 2010. Estimation of maximum, mean and minimum mir temperature in Tabriz City using Artificial Intelligent methods. Water and Soil Science, 20(3): 84-107 (In Persian).
7. Habibipoor, H., M.T. Dastorani, M.R. Ekhtesasi and H. Afkhami. 2012. Evaluation of the Effects of Data range Modification on Efficiency of Regression Decision Tree and Artificial Neural Networks for Drought Prediction, Journal of Watershed Management Research, 2(3): 63-79 (In Persian).
8. Haykin, S. 1994. Neural Networks: A Comprehensive Foundation: Prentice Hall PTR.
9. Hosseini-Moghari, S.M. and S. Araghinejad. 2016. Application of Statistical, Fuzzy and Perceptron Neural Networks in Drought Forecasting (Case Study: Gonbad-e Kavous Station). Water and Soil (Agricultural Sciences and Technology), 30(1): 247-259 (In Persian).
10. Hosking, J.R. 2009. L-Moments Wiley StatsRef: Statistics Reference Online.
11. Kempes, C., O. Myers, D. Breshears and J. Ebersole. 2008. Comparing response of Pinus edulis tree-ring growth to five alternate moisture indices using historic meteorological data Journal of Arid Environments, 72(4): 350-357.
12. Koza, J.R. and R. Poli. 2005. Genetic programming. In: Search Methodologies. Springer, pp: 127-164.
13. Maca, P. and P. Pech. 2016. Forecasting SPEI and SPI drought indices using the integrated artificial neural networks. Computational Intelligence and Neuro science, 2016, 1-17.
14. Madadgar, S. and H. Moradkhani. 2014. Spatio-temporal drought forecasting within Bayesian networks. Journal of Hydrology, 512: 134-146.
15. Mehdizadeh, S., J. Behmanesh and H. Saadatnejad Gharahassanlou. 2016. Evaluation of gene expression programming and Bayesian networks methods in predicting daily air temperature. Journal of Agricultural Meteorology, 4(2): 1-13 (In Persian).
16. Miller, G.T. and S. Spoolman. 2012. Environmental science: Cengage Learning.
17. Mishra, A. and V. Desai. 2005. Drought forecasting using stochastic models. Stochastic Environmental Research and Risk Assessment, 19(5): 326-339.
18. Mishra, A. and V. Desai. 2006. Drought forecasting using feed-forward recursive neural network. ecological modelling, 198(1-2): 127-138.
19. Morid, S., V. Smakhtin and K. Bagherzadeh. 2007. Drought forecasting using artificial neural networks and time series of drought indices. International Journal of Climatology, 27(15): 2103-2111.
20. Thornthwaite, C.W. 1948. An approach toward a rational classification of climate. Geographical Review, 38(1): 55-94.

21. Vicente-Serrano, S.M., S. Begueria and J.I. Lopez-Moreno. 2010. A multi scalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, 23(7): 1696-1718.
22. Wilhite, D. A. 2000. Drought as a Natural Hazard: Concepts and Definitions, in Wilhite, D. (ed.), *Drought: A Global Assessment*, Routledge, London & New York, 1: 3–18.
23. Zabihi, M., R. Mostafazadeh and M. Sharari. 2017. Analysis of Wet and Dry Spells Intensity and Duration Using Precipitation-Based and Evapotranspiration Influenced Indices. *Journal of Watershed Management Research*, 8(15): 125-136 (In Persian).
24. Zare Amini, F., M.A. Ghorbani and S. Darbandi. 2014. Evaluation of Genetic Programming in Estimation of Soil Temperature. *Geographical Space*, 47(4): 19-38 (In Persian).
25. Zanetti, S., E. Sousa, V. Oliveira, F.Almeida and S. Bernardo. 2007. Estimating evapotranspiration using artificial neural network and minimum climatological data. *Journal of Irrigation and Drainage Engineering*, 133(2): 83-89.

Comparison of Artificial Neural Networks, Bayesian Network and Gene Expression Programming in Drought Prediction (Case Study: Maragheh Synoptic Station)

Abbas Abbasi¹, Keivan Khalili², Javad Behmanesh³ and Akbar Shirzad⁴

1- Ph.D. In Water Resources Engineering Urmia University

2- Assistant Professor Urmia University, (Corresponding author: khalili2006@gmail.com s)

3- Professor Urmia University

4- Assistant Professor Urmia University of Technology

Received: October 1, 2018

Accepted: May 26, 2019

Abstract

Drought is an inseparable part of any climate that has significant effects on different parts of the community and it increases the stress on water resources. Therefore, predicting its future status can help planners and decision makers in different sectors. In this study, for predicting drought in different time scales of the SPEI drought index, from 5 different inputs, including SPEI values with a lag of 1 to 5 months, then three intelligent methods including Gene Expression Programming (GEP), Bayesian Network (BN) and Artificial Neural Networks (ANNs) were used to predict future values. The results showed that all three methods in the short-term time-scale of the SPEI index are not appropriate so that the best performance in the one-month time scale is related to the Bayesian network model with a correlation coefficient of 0.142 and in the 3-month time-scale is related to the ANN model with correlation coefficient of 0.704. The results also showed that predictive accuracy of the model has a direct correlation with the SPEI calculation scale and, with increasing SPEI time scale, predictive accuracy increases. Also, all three methods have good performance in long-term time-scales.

Keywords: Drought, Intelligent Model, Prediction, SPEI, Maragheh