

مدل‌سازی و پیش‌بینی رواناب ماهانه در دامنه‌ی زمان (مطالعه موردی: حوضه آبریز قره‌سو)

حامد قزل‌سغلی: دانش‌آموخته کارشناسی ارشد آبخیزداری، گروه مرتع و آبخیزداری، دانشگاه گنبد کاووس، گنبد کاووس

نادر جندقی*: استادیار گروه مرتع و آبخیزداری، دانشگاه گنبد کاووس، گنبد کاووس

مجتبی قره‌محمودلو: استادیار گروه مرتع و آبخیزداری، دانشگاه گنبد کاووس، گنبد کاووس

مجید عظیم‌محسنی: دانشیار گروه آمار، دانشگاه گلستان، گلستان

سیدمرتضی سیدیان: استادیار گروه مرتع و آبخیزداری، دانشگاه گنبد کاووس، گنبد کاووس

نوع مقاله: پژوهشی

تاریخ پذیرش: ۱۴۰۰/۱۲/۱۸

تاریخچه مقاله (تاریخ دریافت): ۱۴۰۰/۱۰/۴



[20.1001.1.22517812.1401.12.3.6.5](https://doi.org/10.22517/12.1401.12.3.6.5)

چکیده

در پژوهش حاضر برای مدل‌سازی رواناب ماهانه، از داده‌های چهار ایستگاه هیدرومتری پل توسکاستان، نهارخوران، غازمحل و سیاه‌آب در حوضه آبریز قره‌سو در یک دوره آماری ۳۶ ساله استفاده شد. سپس بررسی همگنی سری داده‌ها با استفاده از آزمون چاو انجام شد. پس از مرتب‌سازی داده‌ها، برای مدل‌سازی مقادیر رواناب ماهانه از چهار روش باکس و جنکینز (SARIMA)، شبکه عصبی مصنوعی (ANN)، شبکه عصبی مصنوعی - فازی (ANFIS) و الگوریتم ژنتیک (GA) در ایستگاه‌های هیدرومتری منتخب استفاده شد. پس از مدل‌سازی مقادیر رواناب ماهانه با استفاده از چهار مدل فوق، به پیش‌بینی تغییرات رواناب ماهانه در ایستگاه‌های هیدرومتری منتخب برای دوازده ماه آینده پرداخته شد و این امر با کمک نرم‌افزارهای Minitab، R و SPSS صورت گرفت. با توجه به نوع پراکنش مقادیر رواناب و وجود داده صفر، برای تثبیت واریانس از تبدیل $\log(1+Y_t)$ در مدل استفاده شد. در مرحله بعد، اعتبارسنجی مقادیر پیش‌بینی شده توسط مدل‌ها با استفاده از شاخص‌های MAD، RMSE و MAPE ارزیابی شد. نتایج نشان داد که در اکثر ایستگاه‌های هیدرومتری منتخب، مدل شبکه عصبی مصنوعی بهترین عملکرد را در بین چهار مدل مورد استفاده داشت. بعد از شبکه عصبی مصنوعی، شبکه عصبی مصنوعی - فازی دارای مناسب‌ترین عملکرد بود. روش باکس و جنکینز نیز با وجود اینکه در تشخیص روند تغییرات به صورت مناسب عمل کرده بود، در بین چهار مدل مورد استفاده عملکرد ضعیف‌تری را در پیش‌بینی مقادیر رواناب داشت.

واژگان کلیدی: رواناب ماهانه، شبکه عصبی مصنوعی، الگوریتم ژنتیک، باکس و جنکینز، حوضه آبریز قره-

سو.

۱- مقدمه

امروزه با افزایش تقاضا برای مصرف آب، اغلب مناطق دنیا و به خصوص مناطقی مانند خاورمیانه با محدودیت منابع آب سطحی و زیرزمینی مواجه شده‌است؛ بنابراین، مدیریت بهینه و بهره‌برداری مناسب از منابع آب در این مناطق امری ضروری است (Farzin and Alizadeh Sanami, 2017). رشد روزافزون جمعیت، توسعه کشاورزی در مناطق خشک و نیمه‌خشک، صنعتی شدن جوامع و غیره باعث شده‌است تأمین آب مطمئن، به یکی از چالش‌های مهم تبدیل شود. به‌طور معمول برای مدیریت منابع آب کشور و استفاده بهینه از آن، نیاز است برخی از پارامترها از جمله جریان رودخانه‌ها پیش‌بینی و برآورد شود. بسته به نوع هدف مطالعه، این پیش‌بینی ممکن است از نظر زمانی کوتاه‌مدت یا بلندمدت باشد. با در نظر گرفتن این واقعیت که مدل‌های پیش‌بینی کننده نمی‌توانند به‌طور کامل با داده‌های مشاهداتی انطباق داشته باشند، سعی می‌شود تا با بررسی خطای پیش‌بینی، مدل مناسب انتخاب شود. امروزه از سیستم‌های هوشمند برای پیش‌بینی پدیده‌های غیرخطی به‌طور گسترده استفاده می‌شود. روش شبکه عصبی مصنوعی، الگوریتم ژنتیک و شبکه عصبی مصنوعی - فازی و غیره از جمله این مدل‌ها است که برای پیش‌بینی مقادیر رواناب از سوی محققان بسیار قابل توجه قرار گرفته‌است. چنین پیش‌بینی‌هایی از جریان رودخانه با فراهم کردن هشدارهای لازم در مورد وقوع سیل و کمک به تنظیم جریان خروجی منابع آب در هنگام پایین بودن تراز جریان رودخانه، به مدیریت منابع آب کمک می‌کند (Akhtar et al, 2010 & Besaw et al, 2009). با توجه به اهمیت پیش‌بینی جریان آب در رودخانه‌ها که از دیدگاه‌های مختلف اهمیت دارد، یافتن مدلی مناسب در این خصوص ضروری به نظر می‌رسد (Farzin et al, 2020).

Hosseini (1999) برای پیش‌بینی هیدروگراف سیل، از شبکه عصبی مصنوعی و شبکه عصبی - فازی استفاده کرد. نتایج نشان داد که این دو روش به میزان قابل قبولی می‌تواند به تخمین رواناب حاصل منجر شود. Kamali و همکاران (2007) در مطالعه خود با استفاده از آمار ایستگاه هیدرومتری پل شالو در یک دوره آماری ۴۷ ساله، جریان ماهانه ورودی به مخزن سد را با استفاده از سری‌های زمانی مدل تهیه کردند. مقایسه نتایج این مدل با داده‌های اندازه‌گیری شده، بیانگر کارایی قابل قبول مدل برازش شده بود. Najafi و همکاران (2014)، به شبیه‌سازی دبی حداکثر لحظه‌ای در چند ایستگاه هیدرومتری تهران با استفاده از شبکه عصبی مصنوعی پرداختند و نشان دادند که در همه ایستگاه‌های مورد بررسی، استفاده از شبکه عصبی کارایی بالایی دارد. Solgi و همکاران (2015) در تحقیقی با استفاده از شبکه عصبی مصنوعی، به مدل‌سازی هوشمند سری زمانی جریان ماهانه بر رودخانه شور قروه پرداختند. در این بررسی برای بهینه‌سازی وزن‌های اتصال در بین لایه‌های مختلف شبکه عصبی مصنوعی، از الگوریتم ژنتیک استفاده شد. همچنین در این تحقیق، دقت نسبتاً بالای شبکه عصبی مصنوعی در برآورد جریان ماهانه تأیید شد. Baba Ali و Dehghini (2019) برای برآورد رواناب روزانه در ایستگاه بادآور نورآباد، از مدل شبکه عصبی موجک استفاده و نتایج آن را با روش‌های شبکه عصبی مصنوعی مقایسه کردند. نتایج نشان داد که هر دو مدل برای تخمین رواناب روزانه قابلیت خوبی دارد، ولی مدل شبکه عصبی موجک عملکرد بهتری داشت. در پژوهشی دیگر توسط Jandaghi و Azimmohseni (2021)، برای مدل‌سازی و پیش‌بینی تراز آب زیرزمینی در استان گلستان از تحلیل سری زمانی در یک دوره آماری ۳۳ ساله استفاده شد. نتایج نشان داد که

تراز آب زیرزمینی پیش بینی شده توسط مدل، در مقایسه با داده های مشاهداتی دقت مناسبی داشت. Jandaghi و همکاران (2021) نیز برای مدل سازی فرایند بارش - رواناب، از مدل سری زمانی و تابع انتقال استفاده کردند. نتایج نشان داد که از دو دیدگاه پیش بینی و برآزش مدل، تابع انتقال نسبت به مدل سری زمانی دقت بالاتری داشت.

در تحقیقی دیگر، Dastorani (2007) برای شبیه سازی جریان در کشور انگلستان از شبکه عصبی مصنوعی استفاده کرد. براساس نتایج به دست آمده، هر چند شبکه های عصبی مصنوعی به صورت عمومی و کلی کارایی مناسبی را در شبیه سازی و پیش بینی رواناب جریان از خود نشان دادند، ولی نوع شبکه عصبی مصنوعی و خصوصیات داده های ورودی مدل به خصوص داده های آموزشی، پارامترهای بسیار مهمی است که بر روی خروجی های مدل تأثیر عمده ای دارد. Kurtulus و Razask (2010) برای پیش بینی دبی روزانه در جنوب شرقی فرانسه، از دو روش شبکه عصبی مصنوعی و شبکه عصبی مصنوعی - فازی (ANFIS) استفاده کردند. نتیجه بررسی آنها نشان داد که مدل ANFIS نسبت به شبکه عصبی مصنوعی عملکرد مناسب تری دارد. Krishna و همکاران (2011) با استفاده از شبکه عصبی موجک، به پیش بینی دبی های روزانه رودخانه مالارابها هند پرداختند و به این نتیجه رسیدند که مدل شبکه عصبی موجک نسبت به شبکه عصبی مصنوعی و مدل های خود همبسته AR عملکرد بهتری دارد. Elsafi (2014) برای پیش بینی وقوع سیلاب در قسمتی از رودخانه نیل واقع در کشور سودان، از شبکه عصبی مصنوعی (ANN) استفاده کرد. نتایج تحقیق نشان داد که استفاده از مدل (ANN)، روشی قابل اطمینان برای تشخیص خطر سیل در رودخانه نیل است. Kasiviswanathana و همکاران (2016)، کارایی شبکه عصبی موجک را برای پیش بینی دبی روزانه رودخانه آلبرتا در کشور کانادا بررسی کردند. نتایج نشان داد که مدل شبکه عصبی موجک نسبت به روش شبکه عصبی مصنوعی دقت بالاتری داشت. Tsakiri و همکاران (2018) در تحقیقی برای پیش بینی سیل در رودخانه موهاک نیویورک، از روش شبکه عصبی مصنوعی (ANN) و رگرسیون خطی چندگانه (MLR) استفاده کردند. نتایج نشان داد که شبکه عصبی مصنوعی مدل در مقایسه با مدل رگرسیون خطی چندگانه، پیش بینی کارآمدتری ارائه می دهد. Mirza و همکاران (2021) برای پیش بینی تغییرات سطح آب در رودخانه های مستعد وقوع سیلاب، از یک مدل یادگیری ماشین ترکیبی از شبکه عصبی مصنوعی - فازی (ANFIS) و الگوریتم ژنتیک استفاده کردند. نتایج این بررسی نشان داد که روش ANFIS بهینه شده با الگوریتم ژنتیک در مقایسه با ANFIS معمولی دقت بیشتری دارد. Park و Park (2021) برای پیش بینی بارش در کره جنوبی، از یک سیستم واسط بین الگوریتم میکروژنتیک (μ -GA) و مدل WRF استفاده کردند. نتایج این مطالعه نشان داد که بهینه سازی ترکیبی طرح های فیزیکی در مدل WRF، یکی از راه حل های ممکن برای افزایش دقت پیش بینی بارش است. Malekpour و همکاران (2021) برای مدل سازی جریان رواناب در رودخانه الن بروک استرالیا، از شبکه ترکیبی موجک - عصبی بر اساس استخراج ویژگی استفاده کردند. آنها ابتدا داده ها را با روش موجک تجزیه و تحلیل کردند، سپس تعداد ورودی های ANN را با استخراج ویژگی بر اساس انرژی، انحراف استاندارد و حداکثر مقادیر داده های تحلیل شده تعیین کردند. نتایج نشان داد که مدل استخراج ویژگی پیشنهادی، به خصوص برای پیش بینی طولانی مدت بهتر از هم تایان خود عمل می کند. Kazemi و همکاران (2021) نیز برای پیش بینی سری زمانی دمای هوا از مدل ترکیبی هوشمند الگوریتم ژنتیک و شبکه عصبی استفاده کردند. در این تحقیق

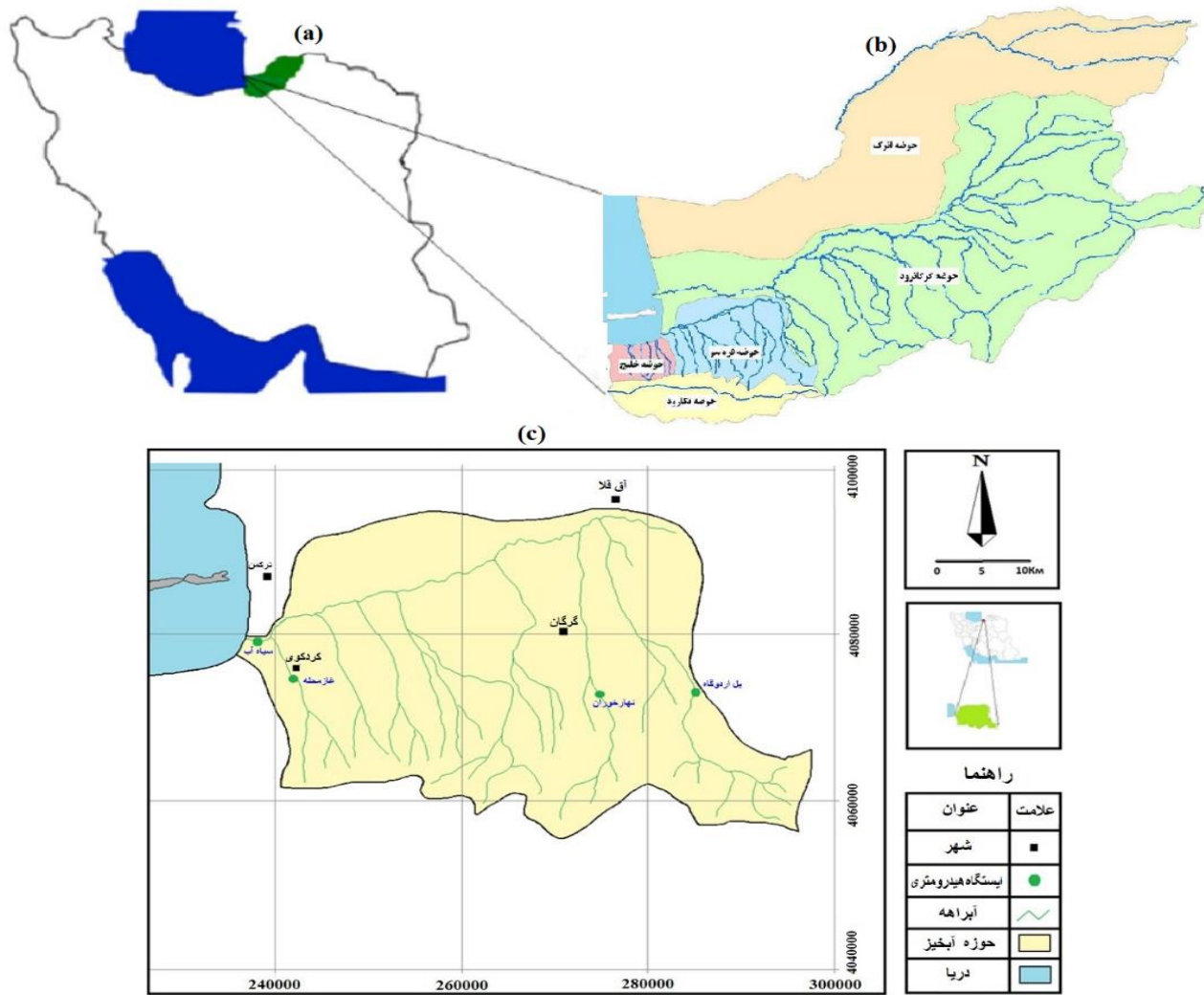
برای انتخاب ویژگی، بر اساس الگوریتم ژنتیک (GA) برای تعیین متغیرهای ورودی مدل و از قدرت پیش‌بینی بالای شبکه عصبی استفاده شد. در مدل پیشنهادی از ساختار مدل سری زمانی خود رگرسیون استفاده شد. نتایج بیانگر دقت بالای مدل توسعه‌یافته در مقایسه با مدل‌های دیگر بود و نشان داد که انتخاب ویژگی در پیش‌بینی سری‌های زمانی اهمیت زیادی دارد. با توجه به اهمیت پیش‌بینی هر چه دقیق‌تر مقدار رواناب رودخانه‌ها و بررسی تغییرات آن در طول سال - که اطلاع داشتن از آن برای مدیریت منابع آب‌های سطحی، مخازن سدها و غیره ضروری است - در پژوهش حاضر با بهره‌گیری از روش‌های هوش مصنوعی شامل شبکه عصبی مصنوعی، شبکه عصبی مصنوعی - فازی، الگوریتم ژنتیک و مدل باکس و جنکینز به پیش‌بینی جریان رواناب ماهانه در ایستگاه‌های هیدرومتری واقع در حوضه آبریز قره‌سو در استان گلستان پرداخته شد تا مدل مناسب برای مدیریت کارآمد منابع آب معرفی شود.

۲- منطقه مورد مطالعه

حوضه آبریز قره‌سو با مساحت تقریبی ۱۶۲۴ کیلومترمربع (۸٪ از مساحت استان)، در غرب استان گلستان با مختصات UTM طول ۲۳۵۷۵۴ تا ۲۹۷۵۴۴ شرقی و عرض ۴۰۵۶۳۴۷ تا ۴۰۹۶۲۴۲ شمالی واقع شده است (شکل ۱). این حوضه از قسمت شمال و شرق با حوضه آبریز گرگانرود، از سمت غرب به حوضه آبریز خلیج گرگان و دریای خزر و از جنوب به حوضه آبریز نکارود منتهی می‌شود. شاخه اصلی رودخانه قره‌سو با طول ۲۰۵ کیلومتر از جنوب شرقی گرگان سرچشمه گرفته است و در روستای قره‌سو به خلیج گرگان وارد می‌شود. تقریباً هشت درصد از منابع آب سطحی استان گلستان معادل با صد میلیون مترمکعب در این حوضه جریان دارد. از زیرحوضه‌های حوضه آبریز قره‌سو می‌توان به گرمابدشت و نومل، زیارت، توشن، شصت کلاته، شמושک، میان‌دره، شش‌دانگ، غازمحل و زواردشت اشاره کرد.

۳- مواد و روش

پس از تعیین موقعیت جغرافیایی حوضه آبریز قره‌سو، مقادیر رواناب ماهانه ایستگاه‌های هیدرومتری موجود در داخل حوضه از آغاز تأسیس جمع‌آوری شد. در انتخاب ایستگاه‌های هیدرومتری نیز سعی شد ایستگاهی انتخاب شود که به بازسازی آماری نیاز نداشته باشد. در نهایت، پس از بررسی آمار تمام ایستگاه‌های هیدرومتری موجود در این حوضه آبریز، چهار ایستگاه هیدرومتری پل توسکاستان، نهارخوران، غازمحل و سیاه‌آب با توجه به داده‌های آماری مناسب برای این تحقیق انتخاب شد (شکل ۱C) که اطلاعات آنها در جدول ۱ نمایش داده شده است. در تحقیق حاضر، دوره آماری مشترک برای ایستگاه‌های هیدرومتری منتخب ۳۶ ساله و از سال آبی ۱۳۶۱-۱۳۶۲ تا ۱۳۹۷-۱۳۹۶ انتخاب و مقادیر رواناب ایستگاه‌ها به صورت ماهانه جمع‌آوری شد (شرکت آب منطقه‌ای استان گلستان).



شکل ۱: موقعیت جغرافیایی حوزه آبریز قره سو در ایران (a) و استان گلستان (b)

جدول ۱: مشخصات ایستگاه های هیدرومتری منتخب در حوزه آبریز قره سو

شماره ایستگاه	نام ایستگاه	کد ایستگاه	مختصات UTM		ارتفاع (متر)
			عرض	طول	
۱	پل توسکاستان	۰۸۵-۱۲	۴۰۷۳۴۸۵	۲۸۴۲۷۲	۴۶۵
۲	نهارخوران	۰۴۳-۱۲	۴۰۷۱۷۷۱	۲۷۴۶۲۲	۵۰۰
۳	غازمچه	۰۴۹-۱۲	۴۰۶۹۱۲۶	۲۴۲۵۴۴	۵/۵
۴	سیاه آب	۰۹۷-۱۲	۴۰۷۹۶۷۲	۲۳۷۲۹۱	-۲۶

با توجه به اینکه مقادیر رواناب به زمان وابسته است، در ابتدا این داده‌ها به صورت سری زمانی مرتب شد. در این تحقیق مقادیر رواناب ماهانه ایستگاه‌های منتخب در یک دوره زمانی ۳۶ ساله، به صورت سری زمانی مرتب شد. سپس بررسی همگنی سری داده‌ها با استفاده از آزمون چاو^۱ صورت گرفت (Ceylan and Ceyda, 2016). در پژوهش حاضر پس از مرتب‌سازی داده‌ها، برای مدل‌سازی مقادیر رواناب ماهانه از چهار روش باکس و جنکینز (SARIMA)، شبکه عصبی مصنوعی (ANN)، شبکه عصبی مصنوعی - فازی (ANFIS) و الگوریتم ژنتیک (GA) در ایستگاه‌های هیدرومتری منتخب استفاده شد. پس از مدل‌سازی مقادیر رواناب ماهانه با استفاده از چهار مدل فوق، به پیش-بینی تغییرات رواناب ماهانه در ایستگاه‌های هیدرومتری منتخب برای دوازده ماه آینده صورت گرفت. با توجه به نوع پراکنش مقادیر رواناب و وجود مقادیر صفر، برای تثبیت واریانس از تبدیل یک به یک $\log(1+Y_t)$ در مدل استفاده شد. توضیح این امر لازم است که در پژوهش حاضر، برای اجرای محاسبات آماری از نرم‌افزارهای SPSS، MINITAB و R استفاده شد.

۱-۳- مدل‌های باکس و جنکینز (SARIMA)

فرم کلی یک مدل SARIMA به صورت رابطه ۱ ارائه شده است:

$$\nabla^d \nabla_T^D w(B)W(B^T)X_t = \theta(B)\Theta(B^T)Z_t \quad \text{رابطه ۱}$$

که در آن، Z_t سری زمانی نوفه سفید، $w(B)$ و $\theta(B)$ چند جمله‌ای‌ها از عملگر پس‌رو و $W(B^T)$ و $\Theta(B^T)$ چند-جمله‌ای‌هایی از عملگر پس‌رو فصلی می‌باشند (Box et al, 2015). هر مدل SARIMA با شش مشخصه تعیین می‌شود که به صورت رابطه ۲ نمایش داده شده است (Cramer, 1986):

$$(p, d, q) \times (P, D, Q)_T \quad \text{رابطه ۲}$$

که در آن، d و D به ترتیب مرتبه تفاضلی کردن معمولی و فصلی سری زمانی، p و q به ترتیب مرتبه خودبازگشتی و میانگین متحرک معمولی و P و Q نیز به ترتیب مرتبه خودبازگشتی و میانگین متحرک فصلی با دوره تناوب T مدل را تعیین می‌کند. مدل‌سازی سری‌های زمانی بر اساس مدل‌های SARIMA بر فرضیاتی استوار است که ممکن است سری‌های زمانی همواره از این فرضیات پیروی نکنند. شیوه کلی مدل‌سازی SARIMA در گام نخست، حذف روند و روند فصلی و سپس مدل‌سازی بر مقادیر باقیمانده مدل است. فرض این مدل‌سازی این است که روند و روند فصلی در کل دوره زمانی تغییر نکرده و قابل تفکیک شدن است، اما در همه سری‌های زمانی این امکان وجود ندارد. برای بررسی نکویی برازش یک مدل SARIMA، معمولاً از آماره Q باکس - پیرس و فرم‌های اصلاح‌شده آن استفاده می‌شود. این آماره به صورت رابطه ۳ محاسبه می‌شود (Bowerman and O'Connel, 1993):

¹ Chow

$$Q = n \sum_{i=1}^k r_i^2(Z_t, Z_t) \quad \text{رابطه ۳}$$

که در آن، $r_i(Z_t, Z_t)$ تابع خودهمبستگی سری زمانی باقیمانده را در گام یا تأخیر i ام نشان می‌دهد. آماره Q دارای توزیع مجانبی $\chi^2(k)$ است. این آماره بر این اساس شکل گرفته است که سری زمانی باقیمانده تا چه میزان تبدیل به نوفه سفید شده که بیانگر برازش مناسب مدل است. این بررسی در محیط نرم افزار MINITAB انجام شد (Box et al, 2015).

۲-۳- مدل شبکه عصبی مصنوعی (ANN)

مدل‌های شبکه‌های عصبی مصنوعی به دلیل نوع طراحی آنها، انعطاف پذیری بسیار خوبی دارد که در وضعیت‌های مختلف دارای عملکرد مناسبی است. این مدل‌ها بر هیچ فرضیه خاصی مبتنی نیست که این امر از مزیت‌های این روش محسوب می‌شود. همچنین طراحی آنها به گونه‌ای است که با وجود هر روندی چه خطی چه غیرخطی، چه فصلی و چه غیرفصلی قابل استفاده می‌باشد. در صورت تغییرات در الگوی‌ها در دوره زمانی، شبکه‌های عصبی به خوبی آنها را شناسایی می‌کند و نیازی به حذف روند و روند فصلی از مجموعه داده‌ها نیست. اگر چه انواع مختلفی از شبکه‌های عصبی مصنوعی ارائه شده است، اما در سری‌های زمانی معمولاً از شبکه‌های عصبی پیش‌رو استفاده می‌شود (Khan and Khan, 2019). این مدل می‌تواند به صورت رابطه ۴ نوشته شود:

$$y_t = \alpha_0 + \sum_{j=1}^n \alpha_j f(\sum_{i=1}^m \beta_{ij} y_{t-j} + \beta_{0j}) + \varepsilon_t \quad \text{رابطه ۴}$$

که در آن، m تعداد گره‌های ورودی، n تعداد گره‌های پنهان و f یک تابع انتقال‌دهنده از ورودی به خروجی است که معمولاً به صورت سیگموئیدی در نظر گرفته شده و فرم معمول آن تابع لجستیک است (رابطه ۵). α_j وزن مسیرهای گره-های پنهان تا لایه خروجی و β_{ij} وزن مسیرهای بین لایه ورودی و گره‌های پنهان را مشخص می‌کند (Zhang and Qi, 2005).

$$f(x) = \frac{1}{1+e^{-x}} \quad \text{رابطه ۵}$$

اگرچه از شبکه‌های عصبی مصنوعی در پیش‌بینی مدل‌های رگرسیونی به طور گسترده‌ای استفاده می‌شود، در مورد سری‌های زمانی کمتر به آنها پرداخته شده است. برای ساخت یک مدل مناسب شبکه عصبی مصنوعی، داده‌ها به سه دسته تقسیم می‌شود: قسمت اول، مشاهدات آموزشی است که از آن برای برآورد پارامترهای مدل استفاده می‌شود. قسمت دوم، مشاهداتی که از آن برای اعتبارسنجی درونی مدل و قسمت سوم، برای اعتبارسنجی بیرونی یا پیش‌بینی مدل استفاده می‌شود. در سری‌های زمانی متناوب، مشاهدات به صورت مضربی از دوره تناوب تقسیم‌بندی می‌شود. در سری‌های زمانی متناوب، استفاده از چند دوره زمانی قبل مسأله‌ای مهم است که با استفاده از شبکه‌های عصبی می‌توان به آن پاسخ داد. در این پژوهش، مراحل زیر برای مدل‌سازی شبکه‌های عصبی بر اساس فرایند پیش‌رو طی شد:

گام اول: سال قبل به عنوان متغیر پیش بینی کننده در نظر گرفته شد، سپس مقادیر پیش بینی تعیین و شاخص RMSE برای این مقادیر محاسبه شد.

گام دوم: دومین سال به مدل اضافه شد سپس با دو متغیر، مقادیر پیش بینی تعیین و مقدار RMSE محاسبه شد.
گام سوم: اضافه کردن سال های گذشته در پیش بینی تا مرحله ای انجام شد که دیگر تغییر محسوسی در نتایج شاخص RMSE مربوط به مقادیر پیش بینی ایجاد نشود. در نهایت، مناسب ترین مدل شبکه عصبی برای سری داده ها انتخاب شد. این بررسی در محیط نرم افزار SPSS صورت گرفت.

۳-۳- مدل شبکه عصبی مصنوعی - فازی (ANFIS)

در این روش، از ترکیب شبکه های عصبی مصنوعی و منطق فازی در پیش بینی سری زمانی استفاده می شود. نحوه مدل سازی آن شبیه شبکه های عصبی مصنوعی است؛ با این تفاوت که انتخاب مسیر بر اساس منطق فازی و یک تابع عضویت که توسط داده ها برآورد می شود، صورت می گیرد؛ به عبارت دیگر، هر مسیر از لایه ورودی تا لایه های پنهان و لایه های پنهان تا لایه خروجی با یک درجه عضویت می تواند به پیش بینی سری زمانی منجر شود. این روش علاوه بر بالا بردن انعطاف پذیری شبکه های عصبی، ریسک خطا را نیز کاهش می دهد. تابع عضویت معمولاً به صورت تابعی زنگوله ای شکل در نظر گرفته می شود (رابطه ۶):

$$\mu(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \quad \text{رابطه ۶}$$

که در آن، پارامترهای a_i ، b_i و c_i توسط داده ها برآورد یا آموزش داده می شود (Jin et al, 2013). این بررسی در محیط نرم افزار R انجام خواهد شد.

۴-۳- مدل الگوریتم ژنتیک (GA)

از آنجا که مدل سازی، یک مسأله بهینه سازی محسوب می شود؛ می توان از الگوریتم ژنتیک به عنوان یک روش بهینه سازی استفاده کرد. در حالت کلی، الگوریتم ژنتیک شامل مراحل زیر است:
گام اول: همه حالت های ممکن برای پارامترها، برای شروع الگوریتم در نظر گرفته می شود.
گام دوم: به ازای همه مقادیر ممکن، تابع هدف محاسبه می شود.
گام سوم: انتخاب مقادیری که تابع هدف در آنها بهینه است.
گام چهارم: با مجموعه مقادیر جدید الگوریتم ژنتیک تکرار می شود.
گام پنجم: این مراحل آن قدر ادامه می یابد که بهترین جواب حاصل شود.
 الگوریتم ژنتیک در عین سادگی، از لحاظ ریاضی بسیار قدرتمند عمل می کند. در این روش، تابعی که بر اساس آن بهینه سازی صورت می گیرد، هیچ محدودیت شکلی ندارد و می تواند شکل های مختلفی داشته باشد که این خود از مزیت های این روش به شمار می رود. حتی می توان از توابع غیر خطی، چند جمله ای ها و غیره نیز در آن استفاده کرد. همچنین می توان با

اعمال محدودیت های مختلف، بهینه سازی را انجام داد. مسأله پیش بینی در سری های زمانی را می توان به زبان ریاضی به صورت رابطه ۷ بیان کرد:

$$X_t = f(X_{t-1}, \dots, X_{t-p}) \quad \text{رابطه ۷}$$

که در آن، f رابطه بین زمان های گذشته یک سری زمانی با زمان آینده آن را مشخص می کند و معمولاً به صورت چند جمله ای در نظر گرفته می شود. یک چند جمله ای از مرتبه q به صورت رابطه ۸ است:

$$f(X_{t-1}, \dots, X_{t-p}) = \sum_{j=1}^p \sum_{k=1}^q c_{jk} X_{t-j}^k \quad \text{رابطه ۸}$$

تابع هدفی که بر اساس آن بهینه سازی انجام می شود، به صورت میانگین مربعات خطا است (رابطه ۹):

$$Fitness = \frac{1}{n} \sum_{t=1}^n (X_t - \hat{X}_t)^2 \quad \text{رابطه ۹}$$

که در آن، \hat{X}_t مقادیر درونیابی سری زمانی است. با کمینه کردن تابع $Fitness$ ، پارامترهای تابع f برآورد می شود و بر اساس آنها می توان به پیش بینی سری زمانی پرداخت (Khadka, 2010). این بررسی در محیط نرم افزار R و به کمک بسته GA انجام شد. پس از برازش مدل های مختلف می توان مقادیر پیش بینی شده را محاسبه کرد. توضیح این امر لازم است که چنانچه در سری زمانی خروجی تبدیلی (مانند تبدیل لگاریتمی) انجام سپس مدل سازی شده باشد، برای محاسبه مقادیر برازش شده یا پیش بینی واقعی باید از تبدیل معکوس (مانند آنتی لوگ) استفاده کرد.

۵-۳- اعتبارسنجی مدل

در مرحله بعد به اعتبارسنجی^۱ مدل های مورد استفاده پرداخته شد. اعتبارسنجی هر مدل سری زمانی می تواند به دو شکل درونی و بیرونی انجام شود. در اعتبارسنجی درونی، مقادیر برازش شده مدل محاسبه و با مقادیر واقعی سری زمانی مقایسه می شود. در اعتبارسنجی بیرونی، معمولاً یک یا چند دوره از سری زمانی در مدل سازی شرکت داده نمی شود و پس از برازش مدل، مقادیر پیش بینی برای دوره های کنار گذاشته شده محاسبه و با مقادیر واقعی مقایسه می شود. تاکنون شاخص های استاندارد مختلفی برای مقایسه مقادیر حاصل از مدل و مقادیر واقعی ارائه شده است که در زیر به بیان مهم ترین آنها (روابط ۱۰ تا ۱۲) پرداخته می شود (Moriasi, 2007):

¹ Validation

الف) شاخص میانگین قدرمطلق فاصله^۱

$$MAD = \frac{1}{m} \sum_{i=1}^m |Y_t - \hat{Y}_t|$$

رابطه ۱۰

ب) شاخص مربع میانگین خطا^۲

$$RMSE = \sqrt{\frac{\sum_{i=1}^m (Y_t - \hat{Y}_t)^2}{m}}$$

رابطه ۱۱

ج) شاخص میانگین قدرمطلق درصد خطا^۳

$$MAPE = \frac{\sum_{i=1}^m \frac{|Y_t - \hat{Y}_t|}{Y_t}}{m}$$

رابطه ۱۲

در روابط فوق، \hat{Y}_t مقادیر برازش شده یا پیش‌بینی شده توسط مدل، Y_t مقدار سری زمانی مشاهده شده در زمان t و m تعداد مقادیر برازش شده یا مقادیر پیش‌بینی است. علاوه بر محاسبه شاخص‌های ذکر شده، برای مقایسه روش‌های مختلف از لحاظ پیش‌بینی از دیاگرام تیلور نیز استفاده شد. این بررسی در محیط نرم‌افزار R صورت گرفت.

۴- یافته‌ها (نتایج)

۴-۱- برازش مدل باکس و جنکینز

بر اساس نمودارهای سری زمانی مشخص شد که برای مقادیر رواناب ماهانه دو ایستگاه هیدرومتری غازمحل و سیاه‌آب روند وجود ندارد، اما در سری داده‌های دو ایستگاه هیدرومتری پل توسکاستان و نهارخوران به‌خصوص در سال‌های اخیر روند نزولی مشاهده شده است. برای ایستگاه‌هایی که روند داشتند، حذف روند انجام شد.

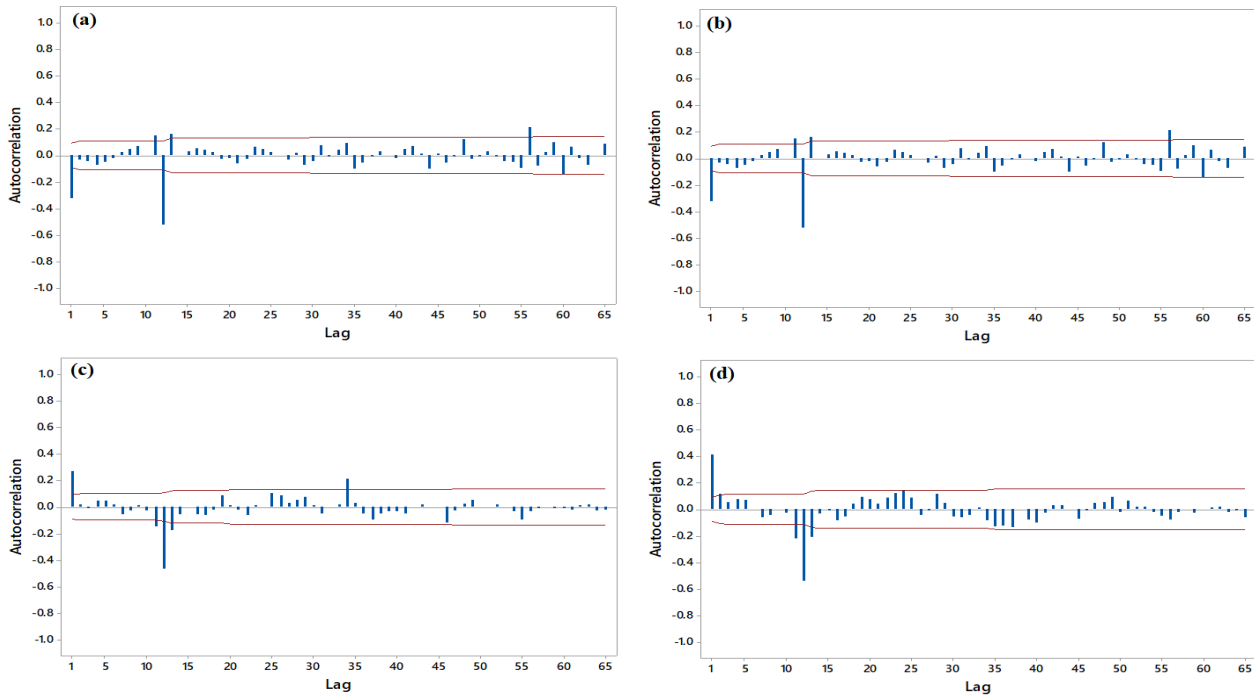
شکل ۲، نمودار خودهمبستگی‌نگار مربوط به مقادیر رواناب ماهانه را در ایستگاه‌های هیدرومتری منتخب نشان می‌دهد. با توجه به نمودار خودهمبستگی‌نگار (شکل ۲)، مشخص شد که سری‌های زمانی در چهار ایستگاه هیدرومتری مورد بررسی دارای روند فصلی با دوره تناوب دوازده ماهه است. با توجه به وجود روند فصلی از تفاضل‌گیری فصلی برای ایستاسازی مشاهدات استفاده شد.

بهترین مدل برازش شده به داده‌های رواناب ماهانه در ایستگاه‌های هیدرومتری منتخب به همراه مقادیر معنی‌داری در جدول ۲ معرفی شده است. در این جدول چنانچه مقدار معنی‌داری بزرگ‌تر از ۰/۰۵ شود، مدل پذیرفته خواهد شد. مقادیر معنی‌داری در جدول ۲ نشان می‌دهد که در سری‌های زمانی باقی‌مانده مدل‌ها، همبستگی وجود ندارد. در نتیجه، مدل‌ها به-خوبی برازش داده شده‌اند (شکل ۳).

¹ Mean Absolute Distance

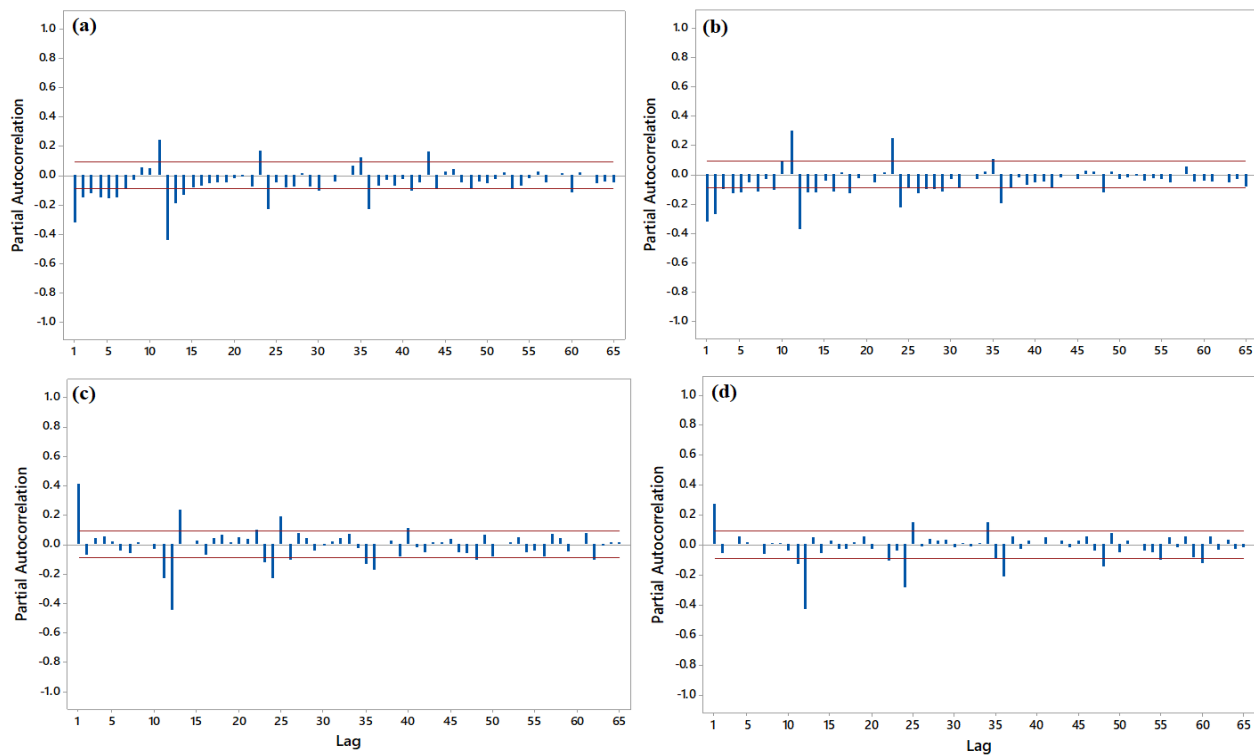
² Root Mean Squares of Errors

³ Mean Absolute Percentage Error



شکل ۲: نمودار خودهمبستگی نگار مربوط به مقادیر رواناب ماهانه در ایستگاه‌های هیدرومتری پل توسکاستان (a)، نهارخوران (b)، غازمجله (c) و سیاه‌آب

(d)



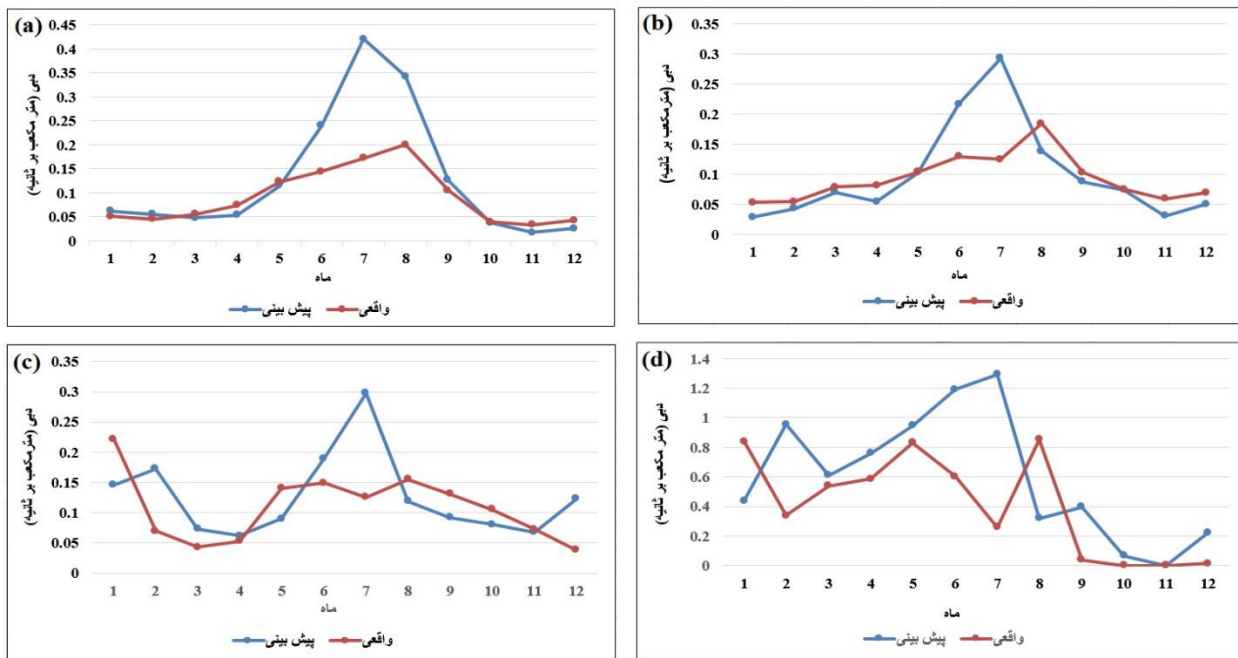
شکل ۳: نمودار خودهمبستگی جزئی سری زمانی مقادیر رواناب ماهانه در ایستگاه‌های هیدرومتری پل توسکاستان (a)، نهارخوران (b)، غازمجله (c) و

سیاه‌آب (d)

جدول ۲: نتایج مدل سازی باکس و جنکینز برای سری زمانی رواناب ماهانه در ایستگاه های هیدرومتری منتخب

ایستگاه هیدرومتری	مدل باکس و جنکینز	مقدار معنی داری
پل توسکاستان	$(1,1,3) \times (1,1,1)_{12}$	۰/۷۶۵
نهارخوران	$(1,1,1) \times (0,1,2)_{12}$	۰/۷۰۲
غازمحله	$(1,0,2) \times (1,1,1)_{12}$	۰/۷۰۰
سیاه آب	$(1,0,4) \times (2,1,1)_{12}$	۰/۴۷۶

سپس مقادیر رواناب ماهانه پیش بینی شده توسط مدل باکس و جنکینز در ایستگاه های هیدرومتری منتخب محاسبه شد. مطابق شکل ۴، تغییرات روند مقادیر در هر چهار ایستگاه هیدرومتری مورد بررسی، واقعی و پیش بینی توسط مدل باکس و جنکینز منطبق بر یکدیگر بود؛ با این وجود، مقادیر پیش بینی نسبت به مقادیر واقعی در ماه های اسفند و فروردین (ماه های ۶ و ۷ روی شکل ۴) بیش برآوردی داشت. در سایر ماه ها، مقدار پیش بینی توسط این مدل مناسب بود.

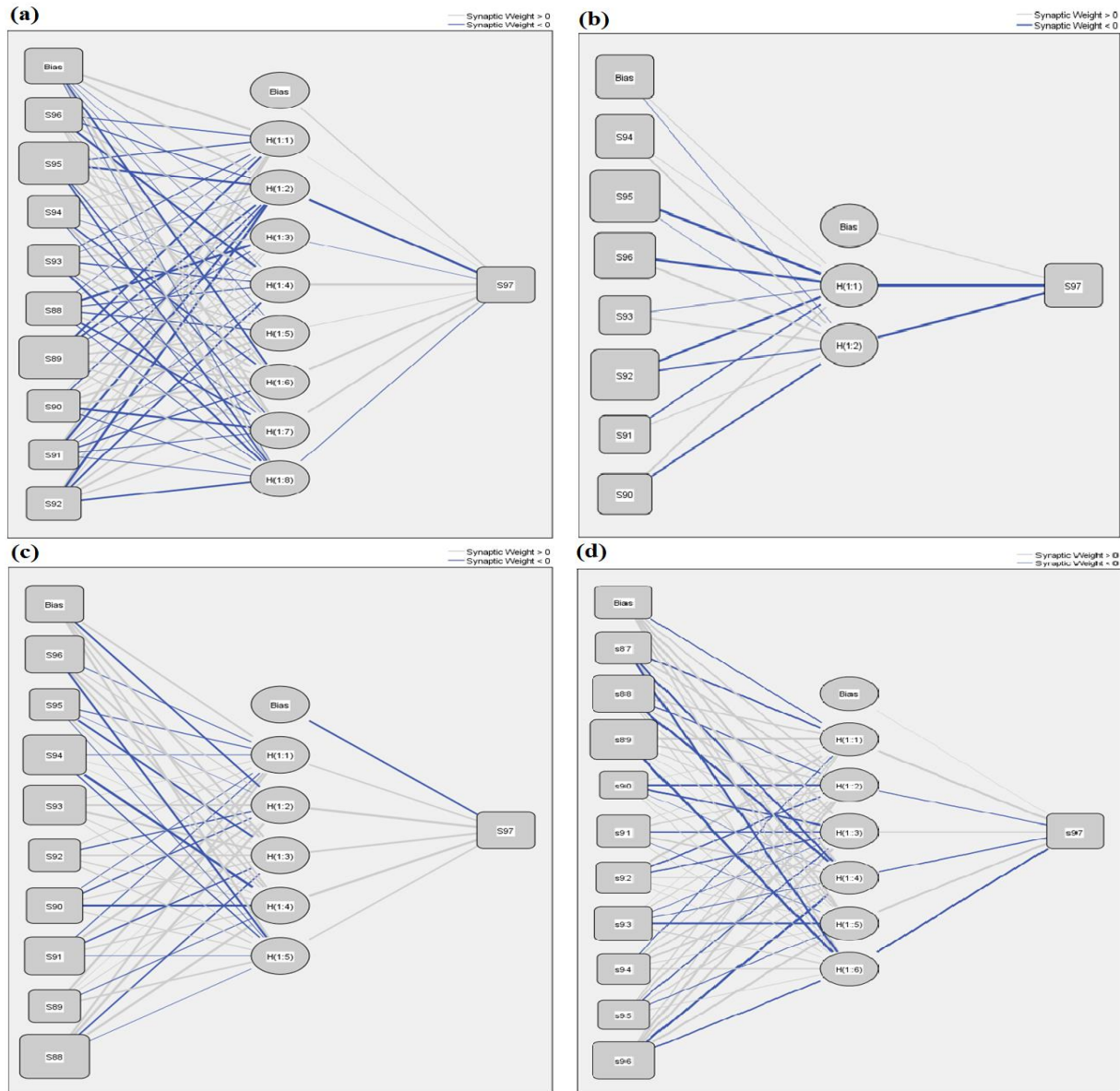


شکل ۴: نمودار مقایسه مقادیر رواناب ماهانه واقعی و پیش بینی شده توسط مدل باکس و جنکینز در ایستگاه های هیدرومتری پل توسکاستان (a)، نهارخوران (b)، غازمحله (c) و سیاه آب (d)

۲-۲-۴- برازش مدل شبکه عصبی مصنوعی (ANN)

به دلیل وجود دوره تناوب دوازده ماهه در سری های زمانی مقادیر رواناب، مشاهدات ماهانه به صورت سری زمانی چندگانه با بعد دوازده ماهه دسته بندی شد و مشاهدات سال های گذشته، به عنوان متغیرهای پیش بینی کننده در مدل شبکه عصبی در نظر گرفته شد. از آنجا که سال های نزدیک تر به وضعیت حال، اطلاعات دقیق تری از وضعیت این زمان دارند سال ها به صورت فرایند پیش رو در شبکه های عصبی در نظر گرفته شد.

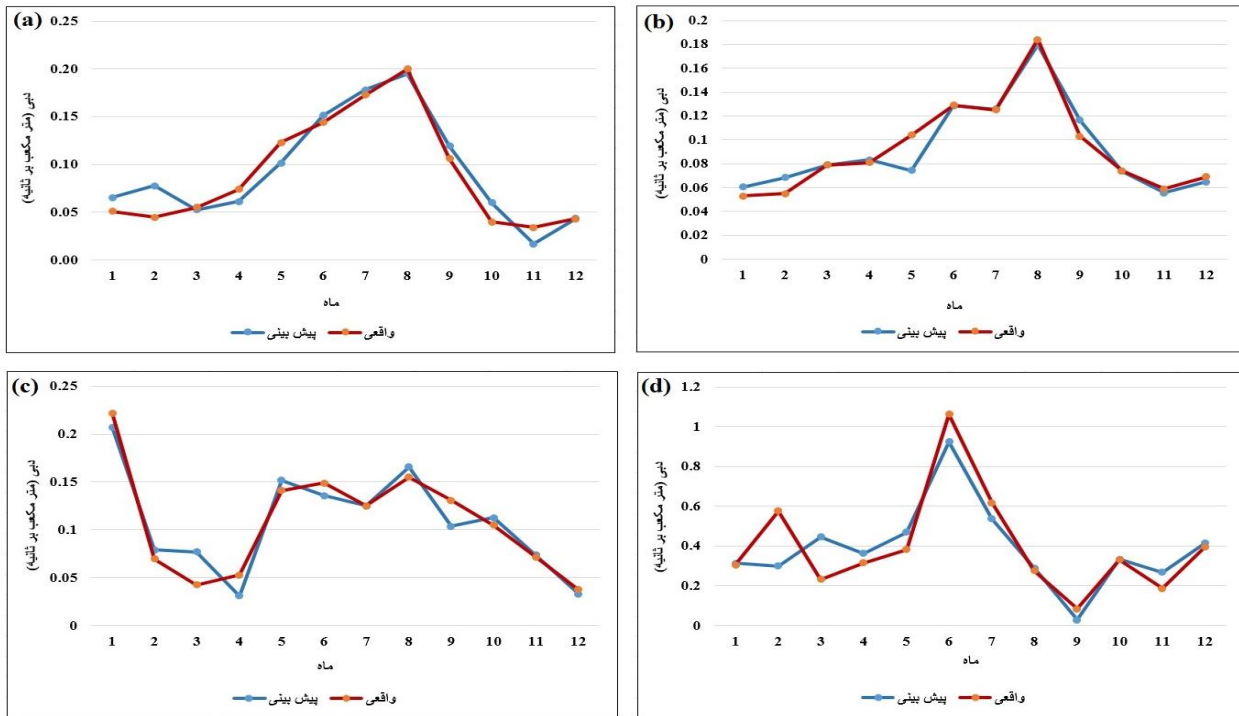
از آنجا که هدف، پیش بینی سری زمانی برای مقادیر آینده است؛ می توان مراحل فوق را با مقادیر سال آخر به عنوان متغیر پاسخ و سال های قبل از آن را به عنوان متغیرهای پیش بینی کننده انجام و شبکه را آموزش داد. سپس می توان بر اساس آن شبکه، مقادیر پیش بینی را محاسبه کرد. همان طور که در شکل (۵، قسمت a) در ایستگاه هیدرومتری پل توسکاستان مشخص است، مقادیر رواناب ماهانه ۹ سال گذشته این ایستگاه به عنوان متغیرهای پیش بینی کننده به مدل وارد شد و بهترین عملکرد را داشت که این کار مطابق با الگوریتم فوق انجام شد. همچنین بررسی ساختار شبکه عصبی مصنوعی در سه ایستگاه هیدرومتری نهارخوران، غازمحل و سیاه آب نشان داد که به ترتیب مقادیر رواناب ماهانه ۷، ۹ و ۱۰ سال گذشته این ایستگاه ها به عنوان متغیرهای پیش بینی کننده به مدل وارد شد و بهترین عملکرد را داشت.



شکل ۵: ساختار شبکه عصبی مصنوعی در ایستگاه‌های هیدرومتری پل توسکاستان (a)، نهارخوران (b)، غازمجله (c) و سیاه‌آب (d)

شکل ۶، نمودار مقایسهٔ مقادیر رواناب ماهانه واقعی و پیش‌بینی شده توسط مدل شبکه عصبی مصنوعی را در ایستگاه‌های هیدرومتری منتخب نشان می‌دهد. مطابق شکل ۶ می‌توان قضاوت کرد که مدل شبکه‌های عصبی در برآزش مقادیر رواناب ماهانه در هر چهار ایستگاه هیدرومتری منتخب، نسبت به روش باکس و جنکینز عملکردی بهتر داشت و بسیار مناسب برآزش داده شد.

هیدرومتری منتخب در نظر گرفته شد. تابع مجموع مربعات خطا $SSE(b)$ ، به عنوان تابع هدف برای برآورد ضرایب ترکیب خطی در الگوریتم ژنتیک به کار گرفته شد.



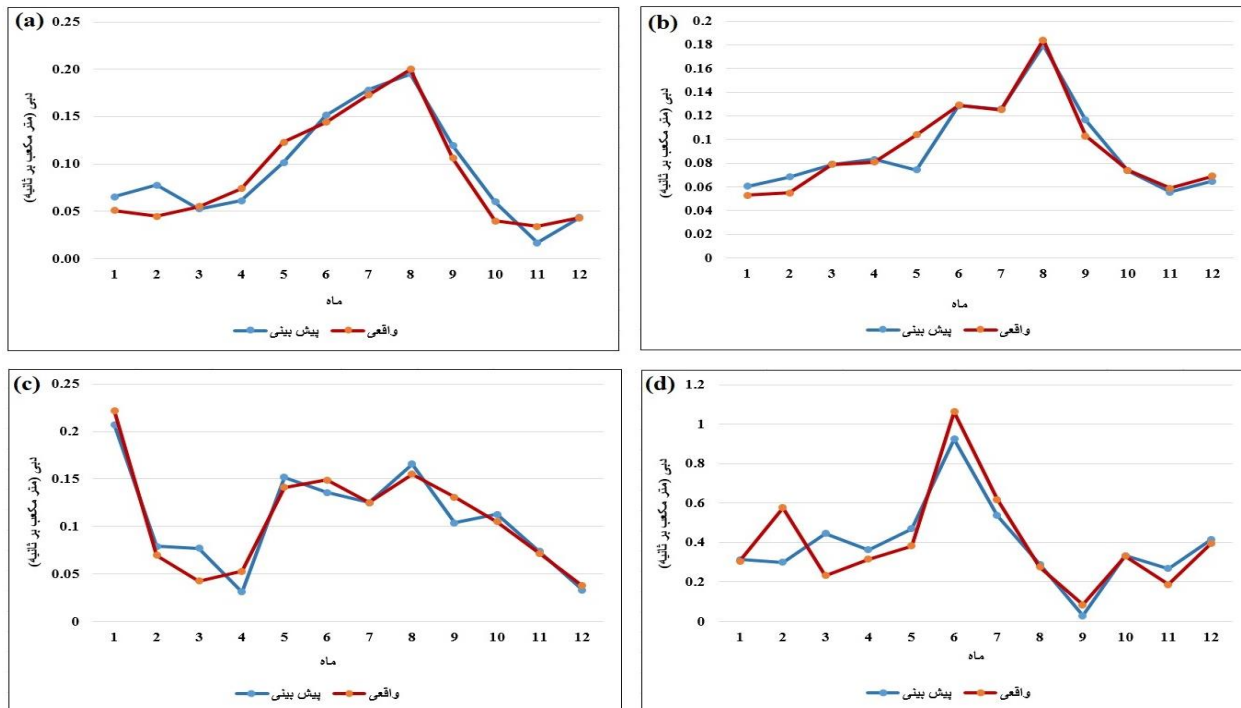
شکل ۷: نمودار مقایسه مقادیر رواناب ماهانه واقعی و پیش‌بینی شده توسط مدل شبکه عصبی مصنوعی - فازی در ایستگاه‌های هیدرومتری پل توسکاستان (a)، نهارخوران (b)، غازمحل (c) و سیاه‌آب (d)

ضرایب با پانصد تکرار در الگوریتم ژنتیک برآورد شد، سپس پیش‌بینی مقادیر رواناب ماهانه بر اساس ترکیب خطی برآورد شده صورت گرفت. برای این منظور، از بسته GA در نرم‌افزار R استفاده شد (Scrucca, 2013). شکل ۸، نمودار مقایسه مقادیر رواناب ماهانه واقعی و پیش‌بینی شده توسط روش الگوریتم ژنتیک را در ایستگاه‌های هیدرومتری منتخب نشان می‌دهد. مطابق شکل ۹ می‌توان قضاوت کرد که روش الگوریتم ژنتیک در برازش مقادیر رواناب ماهانه در هر چهار ایستگاه هیدرومتری منتخب، عملکرد مناسبی داشت و توانست روند تغییرات داده‌ها را تشخیص دهد.

۴-۵- نتایج اعتبارسنجی مدل‌های مورد استفاده

بر اساس نتایج اعتبارسنجی (جدول ۳) مشخص شد که برای پیش‌بینی مقادیر رواناب ماهانه در ایستگاه هیدرومتری پل توسکاستان، مدل شبکه عصبی مصنوعی بهترین عملکرد را در بین چهار مدل مورد استفاده داشت. بعد از شبکه عصبی مصنوعی، شبکه عصبی مصنوعی - فازی نیز مناسب‌ترین عملکرد را داشت. همچنین روش باکس و جنکینز با وجود اینکه در تشخیص روند تغییرات به صورت مناسب عمل کرده بودند، ولی در بین چهار مدل مورد استفاده عملکرد ضعیف‌تری در پیش‌بینی مقادیر رواناب داشت. در ایستگاه هیدرومتری نهارخوران نیز مدل‌های شبکه عصبی مصنوعی و شبکه عصبی

مصنوعی - فازی، مناسب ترین عملکرد را داشتند و مدل الگوریتم ژنتیک و باکس و جنکینز عملکرد ضعیف تری داشتند. توضیح این امر لازم است که دو ایستگاه هیدرومتری نهارخوران و پل توسکاستان، در سرشاخه های غربی حوضه آبریز قره سو واقع شده اند.



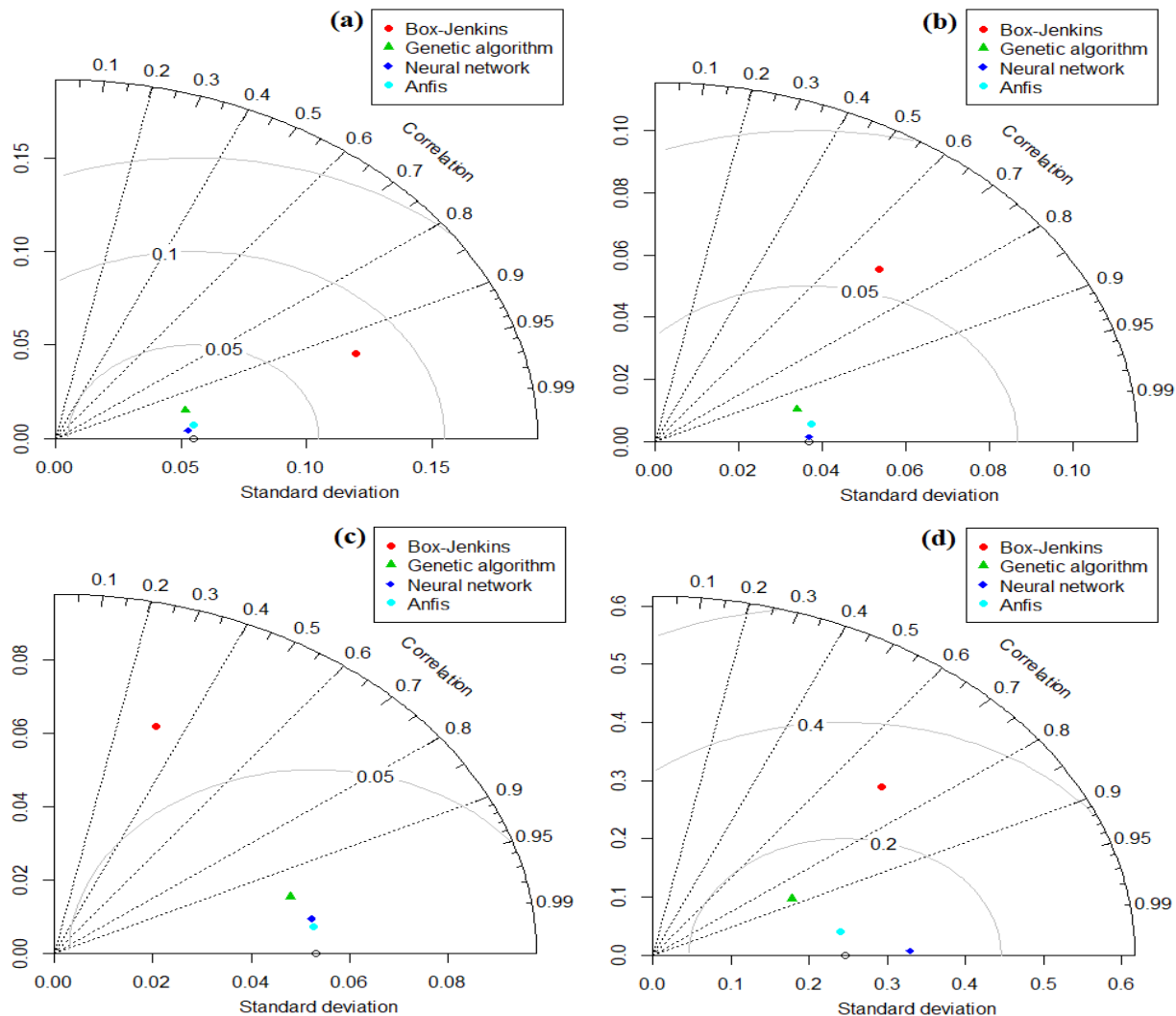
شکل ۸: نمودار مقایسه مقادیر رواناب ماهانه واقعی و پیش بینی شده توسط روش الگوریتم ژنتیک در ایستگاه های هیدرومتری پل توسکاستان (a)، نهارخوران (b)، گازمعه (c) و سیاه آب (d)

جدول ۳: نتایج اعتبارسنجی مدل های مورد بررسی در ایستگاه های هیدرومتری منتخب

مدل مورد استفاده				شاخص (معیار)	ایستگاه هیدرومتری
الگوریتم ژنتیک	شبکه عصبی مصنوعی - فازی	شبکه عصبی مصنوعی	باکس و جنکینز (SARIMA)		
۰/۰۱۲۷	۰/۰۰۶۱	۰/۰۰۳۷	۰/۰۴۹۹	MAD	پل توسکاستان
۰/۰۱۵۵	۰/۰۰۷۴	۰/۰۰۴۷	۰/۰۸۷۹	RSME	
۰/۰۱۱۸	۰/۰۰۵۷	۰/۰۰۳۴	۰/۰۴۳۳	MAPE	
۰/۰۰۶۶	۰/۰۰۵۱	۰/۰۰۱۱	۰/۰۳۶۳	MAD	نهارخوران
۰/۰۱۰۶	۰/۰۰۶۲	۰/۰۰۱۳	۰/۰۵۸۳	RSME	
۰/۰۰۶۱	۰/۰۰۴۶	۰/۰۰۱۰	۰/۰۳۲۶	MAPE	
۰/۰۱۳۱	۰/۰۰۵۶	۰/۰۰۸۰	۰/۰۵۵۹	MAD	غازمحله
۰/۰۱۶۳	۰/۰۰۷۳	۰/۰۰۹۹	۰/۰۷۱۹	RSME	
۰/۰۱۱۸	۰/۰۰۵۱	۰/۰۰۷۳	۰/۰۵۰۲	MAPE	
۰/۰۸۵۷	۰/۰۳۴۰	۰/۰۰۶۱	۰/۳۴۷۰	MAD	سیاه آب
۰/۱۱۸۶	۰/۰۴۱۵	۰/۰۰۸۱	۰/۴۵۲۹	RSME	
۰/۰۶۰۰	۰/۰۲۴۵	۰/۰۰۰۳	۰/۲۴۹۰	MAPE	

نتایج اعتبارسنجی مدل های مورد استفاده در ایستگاه هیدرومتری غازمحله - که در سرشاخه های شرقی حوضه آبریز قره-سو واقع شده است - با دو ایستگاه دیگر کمی متفاوت بود. در این ایستگاه مدل شبکه عصبی مصنوعی - فازی، مناسب ترین عملکرد و پس از آن، مدل شبکه عصبی مصنوعی بهترین عملکرد را داشت. همچنین مدل باکس و جنکینز نیز دارای ضعیف ترین عملکرد در بین چهار مدل مورد استفاده بود. نتایج اعتبارسنجی مدل های مورد استفاده برای پیش بینی مقادیر رواناب ماهانه در ایستگاه هیدرومتری سیاه آب واقع در خروجی حوضه آبریز قره سو نیز کاملاً مشابه با نتایج ایستگاه های هیدرومتری پل توسکاستان و نهارخوران بود. در این ایستگاه نیز مدل های شبکه عصبی مصنوعی و شبکه عصبی مصنوعی - فازی، مناسب ترین عملکرد را داشت و مدل الگوریتم ژنتیک و باکس و جنکینز عملکرد ضعیف تری داشت. همچنین در این تحقیق برای اعتبارسنجی چهار مدل مورد استفاده، از دیاگرام تیلور استفاده شد که نتایج شاخص های MAD، RMSE و MAPE را در انتخاب مناسب ترین مدل تأیید کرد (شکل ۹).

در این بررسی همچنین مشخص شد که سه مدل باکس و جنکینز، شبکه عصبی مصنوعی - فازی و الگوریتم ژنتیک بر اساس معیار MAPE، در ایستگاه هیدرومتری نهارخوران مناسب ترین عملکرد را بین ایستگاه های هیدرومتری منتخب داشت (جدول ۳). بر این اساس، مدل شبکه عصبی مصنوعی در ایستگاه هیدرومتری سیاه آب، مناسب ترین عملکرد را در بین چهار ایستگاه هیدرومتری مورد بررسی داشت.



شکل ۹: نتایج اعتبارسنجی مدل‌های مورد بررسی با استفاده از دیاگرام تیلور در ایستگاه‌های هیدرومتری پل توسکاستان (a)، نهارخوران (b)، غازمحله (c) و سیاه‌آب (d)

۵- بحث و نتیجه‌گیری

در پژوهش حاضر با استفاده از آمار ۳۶ ساله چهار ایستگاه هیدرومتری واقع در حوضه آبریز قره‌سو، با استفاده از چهار مدل باکس و جنکینز (SARIMA)، شبکه عصبی مصنوعی، شبکه عصبی مصنوعی - فازی و الگوریتم ژنتیک به برآزش مدل پرداختند و مقادیر رواناب‌های متوسط ماهانه برای دوازده ماه آینده با کمک نرم‌افزارهای SPSS، Minitab و R پیش‌بینی شد. سپس مقایسه و اعتبارسنجی مدل‌های برآزش داده شده با استفاده از چند شاخص استاندارد ارزیابی شد. با توجه به نمودار خودهمبستگی‌نگار مشخص شد که سری‌های زمانی رواناب ماهانه در هر چهار ایستگاه هیدرومتری مورد بررسی، روند فصلی با دوره تناوب دوازده ماهه داشت. در این پژوهش با توجه به اینکه سال‌های نزدیک‌تر به وضعیت حال، اطلاعات دقیق‌تری از وضعیت این زمان دارند، سال‌ها به صورت فرایند پیش‌رو در شبکه‌های عصبی در نظر گرفته شد.

در نهایت پس از تعیین تعداد سال‌های مؤثر، مناسب‌ترین مدل شبکه عصبی برای سری داده‌ها انتخاب و بر سایر مدل‌ها نیز تعمیم داده شد. به‌طور کلی با توجه به نتایج حاصل از برازش مدل‌ها، پیش‌بینی و اعتبارسنجی هر یک از آنها با استفاده از سه شاخص MAD ، $RMSE$ و $MAPE$ می‌توان چنین اظهار کرد که مدل شبکه عصبی مصنوعی در سه ایستگاه هیدرومتری نهارخوران، پل توسکاستان و سیاه‌آب، بهترین عملکرد را در بین چهار مدل مورد استفاده داشت و پس از آن، مدل شبکه عصبی مصنوعی - فازی به‌عنوان مناسب‌ترین مدل انتخاب شد. در این بررسی همچنین مشخص شد که در ایستگاه هیدرومتری غازمحل به ترتیب دو مدل شبکه عصبی مصنوعی - فازی و شبکه عصبی مصنوعی، بهترین عملکرد را در بین چهار مدل مورد استفاده داشت. نتایج این بررسی با مطالعات Hosseini (1999)، Najafi و همکاران (2014)، Solgi و همکاران (2015)، Dehghani and Baba Ali (2019)، Dastorani (2007)، Elsafi (2014) و Tsakiri و همکاران (2018) که همگی آنها معتقدند شبکه عصبی مصنوعی برای برازش و پیش‌بینی دقت نسبتاً بالایی دارد، همسو است. اما با مطالعه Razask and Kurtulus (2010) که معتقدند مدل ANFIS عملکرد مناسب‌تری نسبت به شبکه عصبی مصنوعی دارد، چندان همسویی ندارد. در این پژوهش اگرچه عملکرد دو روش شبکه عصبی مصنوعی و ANFIS با هم مشابه بود، ولی مدل ANFIS تنها در ایستگاه غازمحل بهترین عملکرد را داشت و در سایر ایستگاه‌ها، شبکه عصبی مصنوعی به‌عنوان مناسب‌ترین مدل انتخاب شد.

در این بررسی همچنین مشخص شد که در تمام ایستگاه‌های منتخب، مدل الگوریتم ژنتیک و باکس و جنکینز پس از دو مدل شبکه عصبی مصنوعی و ANFIS عملکرد ضعیف‌تری داشت. توضیح این امر لازم است که این دو مدل و به خصوص الگوریتم ژنتیک، در پیش‌بینی و تشخیص روند تغییرات رواناب ماهانه عملکرد بسیار مناسبی داشت، اما عملکرد آنها نسبت به مدل‌های شبکه عصبی مصنوعی و ANFIS ضعیف‌تر بود. نتایج این بررسی با مطالعات انجام شده توسط Kamali و همکاران (2007)، Krishna و همکاران (2011) و Jandaghi و Azimmohseni (2021) که استفاده از مدل سری زمانی را برای پیش‌بینی مناسب ارزیابی کردند، ولی عملکرد آن را ضعیف‌تر از روش شبکه عصبی مصنوعی دانسته‌اند، همسویی دارد.

همچنین نتایج این تحقیق نشان داد که به دلیل وجود تغییرات فیزیکی و اقلیمی در بالادست رودخانه‌ها، سری‌های زمانی مقادیر رواناب در ایستگاه‌ها دچار تغییرات ساختاری شده‌است. بنابراین ممکن است زمان‌های دورتر سری‌های زمانی نسبت به زمان حال، به ایجاد انحراف در مدل‌سازی و در نتیجه پیش‌بینی نتایج منجر شود. در نتیجه، استفاده از الگوریتمی برای انتخاب تعداد سال‌های مؤثر در مدل‌سازی و پیش‌بینی می‌تواند مفید باشد. شبکه عصبی مصنوعی به دلیل دقت زیاد در پیش‌یابی و پیش‌بینی، معیاری مناسب برای انتخاب سال‌های مؤثر است. بر اساس سال‌های مؤثر تعیین شده با این روش، می‌توان مدل‌های دیگر پیش‌بینی را اصلاح کرد و به عنوان یک ناظر، مقادیر ورودی مناسب‌تری برای خروجی سایر مدل‌های پیش‌بینی فراهم کرد.

منابع

1. Akhtar, M. K.; Corzo, G. A.; Van Andel, S. J.; & A. Jonoski, 2009. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information, case study of the Ganges River Basin, *Hydrology and Earth System Sciences*, 9, 1607-1618.
2. Baba Ali, H. R., & R. Dehghani., (2019). Evaluating the Performance of Wavelet Neural Network Models in Estimation of Daily Discharge. *Irrigation Sciences and Engineering*. 42(3), 105-116. (In Persian).
3. Besaw L. E.; Rizzo, D. M.; Bierman, P. R.; & W. R. Hackett, 2010. Advances in ungauged stream flow prediction using artificial neural networks, *Journal of Hydrology*, 1(4), 27-37.
4. Bowerman, B. L., & R. O'Connel, (1993). Forecasting and time series: An applied approach, Third edition. Amazon Publication. 722 p.
5. Box, G. E.; Jenkins, G. M.; & G. C. Reinsel, 2015. Time series analysis: forecasting and control, 4th Edition, New Jersey: Wiley, 734 p.
6. Ceylan, Y., & Y. Ceyda, (2016). Comparison of homogeneity tests for temperature using a simulation study. *International Journal of Climatilyg*. 36, 62-81.
7. Crayer, J., 1986. Time series analysis, PWP Publication, Boston, 286 p.
8. Dastorani, M. T., 2007. Evaluation of the Application of Artificial Intelligence Model for Simulation and Real – Time Prediction of Flood Flow, *Journal of Water and Soil Science*, 11(40), 27-37. (In Persian).
9. Elsafi, H., 2014. Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan, *Alexandria Engineering Journal*, 53(2), 655-662.
10. Farzin, S., & F. Alizadeh Sanami., (2017). Modeling and Analyzing of Hydrological Regime Trend of Rivers (Case Study: Gorganrood River, Tamar Hydrometric Stations). *Iran-Watershed Management Science & Engineering*. 11(37), 35-46. (In Persian).
11. Farzin, S.; Mirhashemi, H.; Abbasi, H.; Maryanaji, Z.; & P. Khosravinia, 2020. Assessing memory signal of time-series and simulation of rainfall-runoff process, using neural networks and wavelet-neural hybrid models, *Watershed Engineering and Management*, 11(4), 1059-1074. (In Persian).
12. Hosseini, H., 1999. Artificial Neural Network Model for Determining Flood Hydrograph. Master's Thesis in Civil Engineering, Iran University of Science and Technology, 103 p. (In Persian).
13. Jandaghi, N.; Azimmohseni, M.; & M. Ghareh Mahmoodlu, 2021. Rainfall-runoff process modeling using time series transfer function, *Environmental Erosion Research Journal*, 11(2), 111-128. (In Persian).
14. Jandaghi, N., & M. Azimmohseni., (2021). Modeling and Predicting of Groundwater Level Using Time Series Analysis (Case Study: Ghareghach Well, Golestan Province). 39 National Congress & 4th International Congress of Earth Sciences. Tehran. 1-7. (In Persian).
15. Jin, X. B.; Wang, J. F.; Zhang, H. Y.; & L. H. Cao, 2013. ANFIS model for time series prediction. In Applied Mechanics and Materials, Vol. 385, 1411-1414. Trans Tech Publications Ltd.
16. Kamali, A.; Mahmoudian Shushtari, M.; & N. Kamali, 2007. Predicting the monthly inflow to the reservoir of Shahid Abbaspour Dam using Box-Jenkins time series models, 7th International River Engineering Conference, Ahwaz, P: 300. (In Persian).

17. Kasiviswanathana, K. S.; Jianxun, H.; Sudheerb, K. P.; & T. Joo-Hwa, 2016. Potential application of wavelet neural network ensemble to forecast streamflow for flood management, *Journal of Hydrology*, 536(4), 161-173.
18. Khadka, M. S.; Popp, B.; George, K. M.; & N. Park, 2010. A New Approach for Time Series Forecasting based on Genetic Algorithm. 23rd International Conference on Computer Applications in Industry and Engineering 2010, Caine 2010, 226-231.
19. Khan, M. Z., & M. F. Khan., (2019). Application of ANFIS, ANN and fuzzy time series models to CO2 emission from the energy sector and global temperature increase, *International Journal of Climate Change Strategies and Management*, 11(5), 622-642.
20. Krishna, B.; Satyaji Rao, Y. R.; & P. C. Nayak, 2011. Time Series Modeling of River Flow Using Wavelet Neural Networks, *Journal of Water Resource and Protection*, 3(3), 350-59.
21. Kurtulus, B., & M. Razask., (2010). Modeling daily discharge responses of a large karstic aquifer using soft computing methods: artificial neural network and neuro-fuzzy, *Journal of Hydrology*, 381, 101-111.
22. Kazemi, S. M.; Saffarian, M.; & V. Babaiyan, 2021. Time series forecasting of air temperature using an intelligent hybrid model of genetic algorithm and neural network, *Journal of Industrial and Systems Engineering*, 13(3), 1-15.
23. Malekpour Heydari, S.; Aris, T. N. M.; Yaakob, R.; & H. Hamdan, 2021. Data-Driven Forecasting and Modeling of Runoff Flow to Reduce Flood Risk Using a Novel Hybrid Wavelet-Neural Network Based on Feature Extraction, *Sustainability*, 13(20), 11537.
24. Mirza, O.; Shaikh, A. P.; Mohamm, R.; & S. R. Kishan, 2021. Forecasting water level of Glacial fed Perennial River using a genetically optimized hybrid Machine learning mode, *Materials Today: Proceedings*, 46(20), 11113-11119.
25. Moriasi, D. N.; Arnold, J. G.; Van Liew, M. W.; Bingner, R. L.; Harmel, R. D.; & T. L. Veith, 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Transactions of the ASABE*, 50(3), 885-900.
26. Najafi, E.; Safari, A.; Ghanavati, A.; & A. Karam, 2014. Simulation and analysis of maximum instantaneous discharges using artificial neural network (case study: hydrometric stations, Sulaghan, Golak, and Maghdebik stations in Tehran metropolis), *Journal of Quantitative Geomorphology Research*, 4(1), 90-103. (In Persian).
27. Park, S., & S. K. Park., (2021). A micro-genetic algorithm for combinatorial optimization of physics parameterizations in weather research and forecasting model for quantitative precipitation forecast in Korea. *Geoscientific Model Development Discussions*. pp,1-28.
28. Scrucca, L., 2013. GA: a package for genetic algorithms in R, *Journal of Statistical Software*, 53(1), 1-37.
29. Solgi, A.; Radmanesh, F.; & K. Soltan, 2015. Intelligent modeling time series of monthly flow of basins Shor River of Qorveh with Artificial Neural Network, *Journal of Water and Soil Conservation*, 22 (1), 309-318. (In Persian).
30. Tsakiri, K. G.; Marsellos, A.; & S. Kapetanakis, 2018. Artificial neural network and multiple linear regression for flood prediction in Mohawk River, *New York. Water*, 10(9), 1-20.
31. Wang, L. X., & J. M. Mendel., (1992). Generating fuzzy rules by learning from examples. *IEEE Transactions on systems, man, and cybernetics*. 22(6), 1414-1427.
32. Zhang, G. P., & M. Qi., (2005). Neural network forecasting for seasonal and trend time series. *European journal of operational research*. 160(2), 501-514.

Modeling and Forecasting of Monthly Runoff in the Time Domain (Case Study: Gharasou River Basin)

Hamed Ghezelsefla: *Msc Student in Watershed Management, Rangeland and Watershed Management Department, Faculty of Agriculture & Natural Resources, University of Gonbad Kavous*

Nader Jandaghi¹: *Assistant Professor in Engineering Hydrology, Rangeland and Watershed Management Department, Faculty of Agriculture & Natural Resources, University of Gonbad Kavous*

Mojtaba Ghareh Mahmoodlu: *Assistant Professor in Environmental Hydrogeology, Rangeland and Watershed Management Department, Faculty of Agriculture & Natural Resources, University of Gonbad Kavous*

Majid Azimmohseni: *Associate Professor in Statistics, Department of Statistics, Faculty of Science, Golestan University*

Seyed Morteza Seyedian: *Assistant Professor in Hydraulic Structures, Rangeland and Watershed Management Department, Faculty of Agriculture & Natural Resources, University of Gonbad Kavous*

Article History (Received: 2021/12/25

Accepted: 2022/03/9)



[20.1001.1.22517812.1401.12.3.6.5](https://doi.org/10.1001.1.22517812.1401.12.3.6.5)

Extended Abstract

1- Introduction

Nowadays, demand for water is increasing especially in arid and semi-arid regions (e.g., Iran) due to population growth, economic development, higher standard of living, and changes in consumption patterns. Hence, optimal management of water resources in these areas is essential. Furthermore, climate change and increasingly extreme weather events have caused a surge in natural disasters (e.g., floods) over the past 50 years in arid and semi-arid regions. Thus, forecasting and modeling of runoff data is extremely necessary for planning and managing of water resources. Water flow forecasting plays a key role in flood reduction, reservoir optimization, and reservoir management. These models are mostly developed and applied for simulation and prediction. Therefore, different types of forecasting methods have been proposed over the decades including: Box and Jenkins (SARIMA), Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Genetic Algorithm (GA) models. Forecasting hydrological reactions invariably involves uncertainty. So far, numerous studies have been performed to improve the reliability and accuracy of hydrological forecasts, resulting in reduced risk error. Therefore, the main objective of current research was to use artificial intelligence methods consisting of ANN, ANFIS, GA, and SARIMA models to predict monthly runoff data and also select the best model for the efficient management of water resources in the Gharasou River basin.

2- Methodology

Gharasou river basin with an area of 1624 square kilometers is located in the west of Golestan province and has an important role in providing water resources required in this province. In this research, to model and forecast the runoff process, the monthly runoff time series of 4 hydrometric stations of Pol-Tuskestan, Naharkhoran, Ghazmahale, and Siah-ab over Gharasou River basin were used for a period of 36 years (1982-2018). The time series homogeneity was examined using the Chow's method. Runoff data are time dependent, so initially these data were arranged in time series. After sorting the data, four models consisting of Box and Jenkins (SARIMA), Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Genetic Algorithm (GA) models were used to forecast monthly runoff. To

¹ Corresponding Author: nader.jandaghi@gmail.com

increase the prediction accuracy of other methods, the far time series of monthly runoff were first ignored based on to neural network method and then the number of effective years for modeling was determined. Later, the monthly runoff was forecasted for the next 12 months using four models consisting of SARIMA, ANN, ANFIS, and GA. Lastly, based on the forecasted values and using MAD, RMSE and MAPE indices, the accuracy and precision of SARIMA, ANN, ANFIS and GA models were compared. Modeling and forecasting were done using Minitab, R and SPSS software packages.

3- Results

Based on the type of distribution of monthly runoff and the presence of zero data, $\log(I+Y_t)$ conversion was used in the models to stabilize the variance. The results according to the autocorrelation diagrams revealed that the time series in all stations have seasonal trend with a period of 12 months. Then, the monthly runoff of the next 12 months was forecasted using four models including SARIMA, ANN, ANFIS and GA. Model validation results using three indicators of MAD, RMSE and MAPE revealed that the ANN model in the three hydrometric stations of Naharkhoran, Pol-Tuskestan and Siah-ab had the best performance. In these three hydrometric stations, after ANN model, the ANFIS model has been selected as the most suitable model. However, the performance of these two models has been very similar. In the Ghazmahaleh hydrometric station, two models of ANFIS and ANN had the best performance, respectively. In this study, it was also found that in four selected hydrometric stations, the GA model had a good performance after the two models of ANN and ANFIS. Although SARIMA model performed very well in identifying the trend of monthly runoff changes, it had the weakest performance among the methods. The forecast data using SARIMA model were overestimated compared to the actual data for March and April, but in other months, the forecast data using this model were relatively appropriate.

4- Discussion & Conclusions

In this research, to model and predict the monthly runoff process, four models including SARIMA, ANN, ANFIS, and GA models were used for four selected hydrometric stations in Gharasou River basin. The results of model validation using three indicators of MAD, RMSE and MAPE showed that the ANN and ANFIS models had the best performance among the four models used. It was also found that time series of runoff data in hydrometric stations have undergone structural changes due to the physical and climatic alterations in the upstream of rivers. Therefore, the distant past of time series may cause deviations in modeling and forecasting results. To overcome this problem, making use of an algorithm to select the number of effective years in modeling and forecasting can be useful. Artificial neural network provides a suitable criterion for selecting the number of effective years due to its high accuracy in modeling and forecasting. Based on the effective years identified by this model, other models can be modified and provide more appropriate input data for forecasting from other models.

Keywords: Monthly runoff, Artificial Neural Network, Genetic Algorithm, Box and Jenkins, Gharasou River Basin