

مقایسه کارایی پیش‌بینی دبی ماهانه با استفاده از روش‌های شبکه عصبی مصنوعی و سری‌های زمانی

مجید خزایی^۱ و محمدرضا میرزایی^{۲*}

^۱ دانشجوی دکتری، دانشکده کشاورزی و منابع طبیعی، دانشگاه هرمزگان و ^۲ استادیار، دانشکده کشاورزی، دانشگاه یاسوج

تاریخ دریافت: ۹۱/۰۷/۲۶ تاریخ پذیرش: ۹۲/۰۲/۱۴

چکیده

پیش‌بینی در هیدرولوژی به معنی تخمین شرایط هیدرولوژیکی و هواشناسی در یک بازه زمانی خاص می‌باشد. در همین راستا، فهم رابطه بین بارش و رواناب یکی از ضروری‌ترین مسائل برای مدیریت منابع آب می‌باشد. پژوهش حاضر با هدف مقایسه بین مدل‌های مختلف شبکه عصبی مصنوعی (MLP و RBF) و سری‌های زمانی آرما (ARMA) در برآورد دبی ماهانه در حوزه آبخیز طالقان برای یک دوره ۳۰ ساله بین سال‌های ۱۳۵۶ تا ۱۳۸۶ پی‌ریزی شد. در روش شبکه عصبی مصنوعی از توابع محرک سیگموئیدی و پارامترهای تعداد تکرار، ضریب یادگیری، تعداد نرون مخفی و خطای هدف که با استفاده از آزمون و خطا به دست آمده، استفاده شد. همچنین، در روش آرما از بین مدل‌های مختلف روشی که دارای کمترین میزان خطا و معیار سنجش آکائیک (AIC) بود به عنوان مدل بهینه انتخاب شد. نتایج مدل‌سازی سری‌های زمانی با استفاده از مدل‌های آنالیز روند، حالت وینترز و باکس-جنکینز (آرما) حاکی از دقت بیشتر مدل‌های آرما (۲ و ۳) ($R=0/77$) و حالت وینترز ($R=0/72$) بوده است. در مقایسه بین مدل‌های شبکه عصبی مصنوعی، مدل MLP با میانگین ضریب هم‌بستگی $0/83$ نسبت به مدل RBF با میانگین ضریب هم‌بستگی $0/81$ دقت بیشتری در پیش‌بینی دبی نشان داده است. در مجموع دقت‌سنجی مدل‌ها براساس آماره‌های ریشه میانگین مربعات خطا و ضریب هم‌بستگی حاکی از دقت بیشتر شبکه عصبی مصنوعی (ANN) نسبت به مدل‌های سری زمانی (ARMA) می‌باشد. همچنین، ارزیابی دقت در مدل‌های مختلف حاکی از دقت بیشتر مدل یک ($R=0/86$) و $6/45$ (RMSE=) با ورودی‌های دبی یک ماه تا چهار ماه قبل بوده است. بهترین معماری در روش شبکه عصبی مصنوعی نوع MLP، مدل شماره ۱ با آرایش ۴-۲۰-۱ به ترتیب با چهار نرون در لایه ورودی، ۲۰ نرون در لایه مخفی و یک نرون در لایه خروجی شناخته شد.

واژه‌های کلیدی: حوزه آبخیز طالقان، دقت، ضریب هم‌بستگی، مدیریت منابع آب، میانگین متحرک

مقدمه

مدیریت آبخیز و مدیریت منابع آب، بسیار مؤثر باشد (Nouraei و همکاران، ۲۰۰۹). روش‌های پیش‌بینی مختلف سعی در ارتباط بین متغیرهای مستقل و وابسته داشته‌اند (Karamouz و Araghinejad، ۲۰۰۵). در همین راستا رابطه بین متغیرهای چرخه

پیش‌بینی در هیدرولوژی به معنی تخمین شرایط هیدرولوژیکی و هواشناسی در یک بازه زمانی خاص می‌باشد. پیش‌بینی جریان رودخانه در منابع آب از اهمیت زیادی برخوردار است و می‌تواند در بهینه‌سازی

* نویسنده مسئول: mmirzaei@yu.ac.ir

پیش‌بینی جریان رودخانه‌ها استفاده می‌شوند. یکی از رایج‌ترین روش‌های مورد استفاده برای پیش‌بینی توسط هیدرولوژیست‌ها روش پرسپترون چندلایه MLP^3 می‌باشد. در شبکه‌های عصبی چندلایه، هر نرون حاصل ضرب خروجی نرون‌های لایه پیشین در وزن میانی آن نرون و نرون‌های لایه پیشین را به‌عنوان ورودی خام دریافت کرده، مجموع آن‌ها را به‌عنوان ورودی مؤثر برای قرارگیری در تابع فعال‌سازی محاسبه می‌کنند (Farjamnia و همکاران، ۲۰۰۷). این روش (پرسپترون چندلایه) که در سال ۱۹۶۸ توسط روملمهارت و مک‌کلیلاند پیشنهاد شد به معنای این می‌باشد که خطاها به‌سمت عقب در شبکه تغذیه می‌شوند تا وزن‌ها را اصلاح کنند و پس از آن مجدداً ورودی مسیر پیش‌سوی خود تا خروجی را تکرار کنند (Asghari Moghaddam و همکاران، ۲۰۰۹).

روش دیگری که اخیراً برای پیش‌بینی در هیدرولوژی به‌کار برده می‌شود، سری زمانی است. منظور از یک سری زمانی مجموعه‌ای از داده‌های آماری است که در فواصل زمانی مساوی و منظمی جمع‌آوری شده باشند و روش‌های آماری که این گونه داده‌های آماری را مورد استفاده قرار می‌دهد، تحلیل سری‌های زمانی نامیده می‌شود. مهم‌ترین نوع سری زمانی مدل ترکیبی اتورگرسیو و میانگین متحرک^۴ می‌باشد. این روش ابتدا توسط Box و Jenkins (۱۹۷۶) ارائه شده است.

اگر یک سری زمانی دارای میانگین ثابت نبوده اما تغییرات یا تفاضل تغییرات این پارامتر ایستا باشد، می‌توان با استفاده از عمل‌گر تفاضل آن سری زمانی را ایستا نمود. استفاده از d مین تفاضل سری و مدل‌سازی آن با مدل خود هم‌بسته میانگین متحرک منجر به پیدایش سری جدیدی از مدل‌های آماری موسوم به آرما می‌شود (Araghinejad و Karamouz، ۲۰۰۵). مدل آرما به‌طور موفقیت‌آمیزی برای شبیه‌سازی سری‌های زمانی هیدرولوژیکی برای زمان‌های طولانی به‌کار می‌رود (Zhou و همکاران، ۲۰۰۸). مدل اتورگرسیو و خود هم‌بستگی میانگین متحرک (ARMA) نوعی فرآیند تولید داده در

هیدرولوژیک و رواناب مسئله مهمی در هیدرولوژی سطحی می‌باشد و معمولاً چالش اصلی برای هیدرولوژیست‌هاست (Antar و همکاران، ۲۰۰۶).

از جمله مدل‌های پیش‌بینی رواناب مدل‌های جعبه سیاه می‌باشند که صرفاً مبتنی بر اطلاعات ورودی و خروجی بازیابی شده از حوزه آبخیز می‌باشد که از آن جمله مدل شبکه عصبی مصنوعی^۱ (Dawson و Wilby، ۲۰۰۱) و سری‌های زمانی^۲ (Abrahart و See، ۲۰۰۰) می‌باشد.

شبکه عصبی مصنوعی (ANN) در واقع مدل ساده شده‌ای از مغز انسان می‌باشد که قادر به نشان دادن ترکیبات غیرخطی برای ارتباط بین ورودی‌ها و خروجی‌های هر سیستمی است. این شبکه طی مراحل یادگیری آموزش دیده و برای پیش‌بینی در آینده مورد استفاده قرار می‌گیرد (Nouraei و همکاران، ۲۰۰۹). گسترش وسیع استفاده از شبکه‌های عصبی مصنوعی به‌عنوان مدل تجربی و کارآمد در علوم مختلف از جمله هیدرولوژی نشان‌دهنده ارزش بالای مطالعه این مدل‌هاست (Asghari Moghaddam و همکاران، ۲۰۰۹).

ساختار عادی شبکه عصبی مصنوعی از سه لایه متفاوت تشکیل شده است (شکل ۱) که شامل لایه ورودی، لایه‌های میانی (مخفی)، و لایه خروجی است (Jahangeer و همکاران، ۲۰۰۸). در داخل هر لایه تعدادی نرون قرار دارد که با اتصالات وزن‌دار به هم مربوط می‌شوند که تعداد این نرون‌ها بستگی به تعداد متغیرهای ورودی و خروجی مدل دارد ولی انتخاب تعداد نرون‌های لایه میانی به‌صورت سعی و خطا تعیین می‌شود (Rahimi Khoob و همکاران، ۲۰۰۸). آخرین لایه یا لایه خروجی شامل مقادیر پیش‌بینی - شده به‌وسیله شبکه می‌باشد و خروجی مدل را معرفی می‌کند. لایه‌های میانی که از گره‌های پردازش‌گر تشکیل شده، محل پردازش داده‌ها است (Asghari Moghaddam و همکاران، ۲۰۰۹).

از اواخر دهه ۱۹۹۰ تاکنون شبکه‌های عصبی مصنوعی به‌طور گسترده در مسائل هیدرولوژی و هیدروژئولوژی مثل مدل‌سازی بارش-رواناب و

³ Multilayer Perceptron

⁴ Autoregressive Integrated Moving Average

¹ Artificial Neural Network

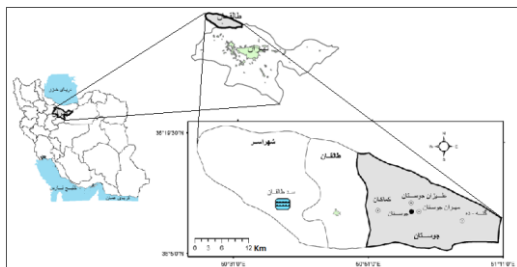
² Time Series

مصنوعی و روش خود هم‌بسته و میانگین متحرک در ایستگاه جویستان حوزه آبخیز طالقان می‌باشد.

مواد و روش‌ها

حوزه آبخیز طالقان یکی از شاخه‌های حوضه سفید رود است که در دامنه جنوبی رشته‌کوه‌های البرز در شمال شرق تهران واقع شده است. این حوزه آبخیز در ۹۰ کیلومتری تهران و با مختصات جغرافیایی به طول $39^{\circ} 39' 50''$ الی $51^{\circ} 11' 26''$ شرقی و عرض $36^{\circ} 20' 17''$ الی $36^{\circ} 20' 45''$ شمالی قرار دارد. مساحت این حوزه آبخیز $94028/82$ کیلومتر مربع و محیط آن $168/3$ کیلومتر است که مساحت و محیط زیرحوزه جویستان به ترتیب 6548 هکتار و $8/47$ کیلومتر می‌باشد. متوسط ارتفاع حوضه مورد مطالعه 2816 متر و شیب متوسط آن 26 درجه است. براساس طبقه‌بندی اقلیمی دومارتن، حوزه آبخیز طالقان دارای اقلیم‌های مدیترانه‌ای، نیمه‌مرطوب، مرطوب و خیلی مرطوب می‌باشد. شکل ۱ شهرستان طالقان و حوزه آبخیز طالقان را در ایران و در استان تهران نشان می‌دهد. کاربری غالب در این منطقه، مراتع است که به‌طور متوسط 70 درصد سطح اراضی را در بر گرفته و از جمله گیاهان مرتعی غالب در آن، انواع گونه‌های گون می‌باشد (Integrated Studies of Taleghan river basin، ۱۹۹۳).

روش پژوهش: برای انجام این پژوهش آمار دبی ماهانه ایستگاه جویستان در حوزه آبخیز طالقان از سال ۱۳۵۶ تا ۱۳۸۶ از سازمان تحقیقات منابع آب تهیه شد.



شکل ۱- نمایی از شهرستان طالقان، ایستگاه‌های هیدرومتری و محل سد طالقان در حوزه آبخیز طالقان

برای مدل‌سازی دبی با استفاده از روش شبکه عصبی مصنوعی از نرم‌افزار MATLAB ۲۰۰۹ استفاده

سری‌های زمانی می‌باشد. در این نظریه فرآیند تولید داده به رابطه آماری بین متغیرهای گذشته و حال می‌پردازد (Zhou و همکاران، ۲۰۰۸).

پژوهش‌های متفاوتی در ایران و جهان در رابطه با این موضوع انجام شده است، به‌طوری که Méndez و Castellano و همکاران (۲۰۰۴) با مدل‌سازی ماهانه و روزانه رواناب در رودخانه Xallas در کشور آمریکا با استفاده از روش‌های شبکه عصبی مصنوعی و مدل باکس-جنکینز^۱ به این نتیجه رسیدند که شبکه عصبی مصنوعی نسبت به مدل باکس-جنکینز برای مدل‌سازی روابط بارش و رواناب قابلیت بیشتری دارد. Zhou و همکاران (۲۰۰۸) با مقایسه مدل‌های ANN، ARMA و شبکه عصبی مصنوعی موجکی^۲ Wavelet به دقت بیشتر روش Wavelet در پیش‌بینی دبی پی بردند. Kisi (۲۰۰۹) با مقایسه شبکه عصبی مصنوعی با مدل ترکیبی شبکه عصبی مصنوعی موجکی نشان داد که مدل‌های ترکیبی باعث افزایش دقت پیش‌بینی از طریق کاهش خطای ریشه میانگین مربعات خطا به میزان ۶۵ درصد و افزایش ضریب تبیین به میزان ۴۰ درصد می‌شود.

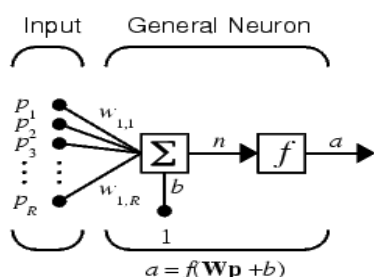
Abrihamchi و همکاران (۲۰۰۵) مدل‌های استوکاستیک^۲ منطقه‌ای جریان سالانه در حوزه‌های آبخیز کرخه دز و کارون را بر پایه مدل اتورگرسیو و کاربرد تحلیل رگرسیون چند متغیره ارائه کرده‌اند. در حوزه‌های آبخیز یاد شده مدل استوکاستیک مناسب جریان سالانه، مدل اتورگرسیو مرتبه اول تشخیص داده شده است. Ghanbarpour و همکاران (۲۰۰۸) با ارزیابی مدل‌های سری زمانی و برازش آن‌ها بر داده‌های دبی ماهانه در حوزه آبخیز کرخه به این نتیجه رسیدند که به‌ترتیب مدل‌های آرما و آرما از دقت بیشتری در پیش‌بینی دبی ماهانه نسبت به مدل‌های توماس-فیرینگ برخوردار است.

از پژوهش‌های متعدد نتیجه گرفته می‌شود که روش شبکه عصبی مصنوعی روش مناسبی برای پیش‌بینی دبی است. براین اساس پژوهش حاضر با در نظر گرفتن این پیش‌فرض، در پی ارزیابی روش‌های پیش‌بینی دبی با استفاده از روش‌های شبکه عصبی

¹ Box-Jenkins Model

² Stochastic Models

تبیین ادامه یافت. همچنین، بعد از انتخاب ساختار شبکه به‌منظور تعیین مقادیر تعداد تکرار، خطای هدف و ضریب یادگیری برای هر مدل نیز با استفاده از سعی و خطا حاصل شد. تکنیک دوم به‌کار رفته در این پژوهش برای استفاده شبکه عصبی مصنوعی RBF^4 می‌باشد. آموزش این شبکه دلالت بر یافتن مجموعه‌ای از گره‌ها و وزن‌های اولیه دارد. بنابراین، فرآیند یادگیری به‌منظور پیدا کردن بهترین خط در داده‌های آموزشی می‌باشد (شکل ۳). در شبکه RBF مقادیر خطای هدف، تعداد نرون و ضریب SPREAD نیز با استفاده از سعی و خطا حاصل شد. در نهایت خروجی مدل‌های منتخب با مقادیر مشاهداتی مقایسه و مدل‌های بهینه انتخاب شدند (شکل ۴).



شکل ۲- ساختار پرسپترون چندلایه با تابع f Hagan. (۱۹۹۴)

مدل‌سازی سری زمانی: پیش‌بینی دبی با استفاده از روش آرما در نرم‌افزار Minitab14 صورت گرفت. در مدل‌های باکس-جنکینز که به نام ARMA معروف هستند، مراحل زیر انجام گرفت؛ با آماده‌سازی داده‌ها برای مدل‌سازی پس از اطمینان از عدم وجود روند و عدم وجود داده پرت با روش آرما، داده‌ها با کمک تبدیل‌گر باکس-کاکس^۵ نرمال و با استفاده از تفاضل‌گیری داده‌ها ایستا شدند. سپس ضرایب و نمودارهای خود هم‌بستگی و خود هم‌بستگی جزئی بررسی شد.

مراحل مدل‌سازی با این روش در پژوهش حاضر شامل شناسایی، تعیین مرتبه، تعیین پارامترها، اعتبارسنجی و مدل‌سازی بود. در یک مدل آرما هر دو فرآیند اتورگرسیو و میانگین متحرک مورد استفاده

شد. چگونگی انجام مدل‌سازی با روش شبکه عصبی مصنوعی به‌طور خلاصه در شکل ۳ ارائه شده است. براین اساس داده‌های موجود طوری به دو دسته تقسیم شدند که از مجموع داده‌های مورد استفاده برای آموزش و یادگیری همه فصول و ماه‌های خشک و مرطوب را شامل شد. به‌طوری‌که ۸۰ درصد داده‌ها (۲۹۰ ماه) برای آموزش^۱ و ۲۰ درصد (۱۵۰ ماه) برای آزمایش^۲ جدا شد، سپس برای وارد کردن داده‌ها به نرم‌افزار MATLAB، به‌منظور استانداردسازی و یکسان‌سازی واحدها از رابطه (۱) استفاده شد (Rajurkar و همکاران، ۲۰۰۴).

$$X_s = 0.8 \times \left[\frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \right] + 0.1 \quad (1)$$

که در آن، X_s متغیر استاندارد شده، X_i داده‌های اصلی، X_{\max} متغیر حداکثر و X_{\min} متغیر حداقل در بین کل داده‌ها می‌باشد. الگوی مورد استفاده در شبکه عصبی پرسپترون چندلایه (MLP) الگوی لونیبرگ-مارکوورت^۳ (LM) می‌باشد که به‌منظور شناسایی بهترین روش با بازدهی بالا برای آموزش شبکه استفاده شده است (Asghari Moghaddam و همکاران، ۲۰۰۹).

در پژوهش کنونی مراحل آموزش به‌کمک این الگوریتم شامل اختصاص ماتریس وزن تصادفی به هر یک از اتصالات، انتخاب بردار ورودی و خروجی، محاسبه خروجی نرون‌ها در لایه خروجی، به‌هنگام‌سازی وزن‌ها به‌روش انتشار خطای شبکه به لایه‌های قبل، ارزیابی عملکرد شبکه و سرانجام برگشت به قسمت پایان آموزش بوده است. پس از تکمیل آموزش، شبکه برای سری اطلاعات معلوم امتحان شد. در این شبکه، تابع انتقال سیگموئیدی برای لایه اول و تابع انتقال خطی برای لایه دوم به‌عنوان بهترین معماری شبکه استفاده شد (شکل ۲).

برای تعیین تعداد نرون‌های مطلوب در لایه مخفی، شبکه با قراردادن مقادیر مختلف در لایه مخفی مورد آموزش قرار گرفت که این اقدام تا دستیابی به مقدار حداقل خطا و مقدار بهینه ضریب

¹ Training

² Testing

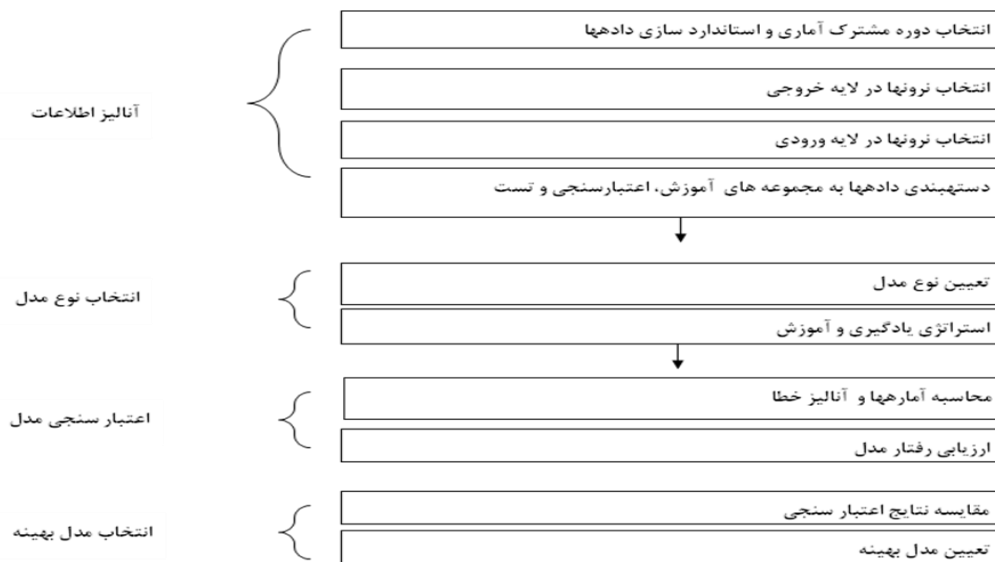
³ Levenberg-Marquardt Back propagation

⁴ Radial basis function

⁵ Box-Cox

تفاضل گیری یا استفاده از مدل های فصلی ایستا شده و با استفاده از شکل های تابع خود هم بستگی (ACF) و تابع خود هم بستگی جزئی (PACF) فرآیند ARMA شناسایی شد (Sadeghi و Rahmani، ۲۰۰۴).

قرار گرفت. مدل ARMA (p,q) کلی ترین مدل سری زمانی است که با توجه به ماهیت داده ها و مدل شناسایی شده برای سری، مقادیر عددی p و q محاسبه و ارائه می شود. در این بخش ابتدا داده ها با



شکل ۳- مراحل مدل سازی با شبکه عصبی مصنوعی

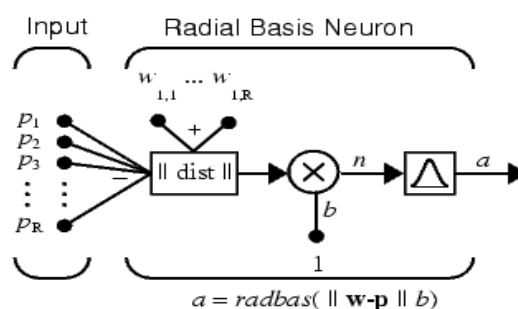
تصادفی و غیرقطعی مسائل مهندسی آب و مدیریت منابع آب، به عنوان یکی از روش های توصیه شده در پیش بینی های پدیده های تصادفی نظیر سیلاب ها می باشد. بعد از این که مدل های مختلف بر سری زمانی آرما عبور داده شد از معیار آکائیک (AIC)^۳ برای تعیین بهترین مدل موجود استفاده شد. چگونگی محاسبه معیار آکائیک در رابطه (۲) مشخص است.

$$AIC(p, q) = N \ln(\delta_{\Sigma}^2) + 2(p + q) \quad (2)$$

که در آن، N تعداد اطلاعات سری زمانی و δ_{Σ}^2 واریانس خطاهای باقی مانده است که بر این اساس مدلی که آکائیک کمتری داشته باشد، مدل بهتری می باشد. همچنین، با ترسیم مقادیر مشاهده ای و برآوردی در یک نمودار و به صورت پراکنش نقاط عملکرد مدل ها مورد ارزیابی قرار گرفت.

نتایج و بحث

نتایج حاصل از روش شبکه عصبی مصنوعی با الگوریتم پس انتشار خطا و تابع LM و روش RBF در ایستگاه جویستان در جدول ۱ ارائه شده است. در این



شکل ۴- ساختار مدل RBF با تابع radbas (Hagan، ۱۹۹۴)

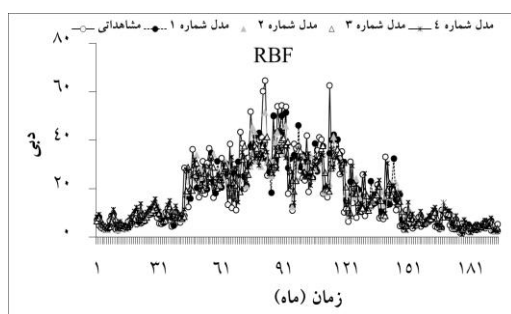
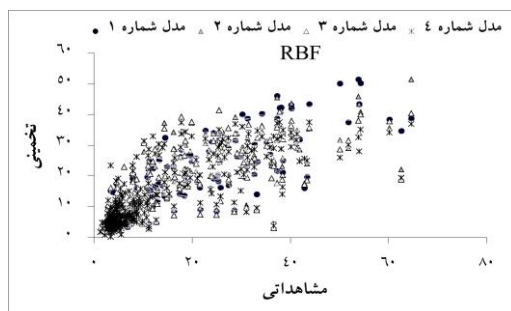
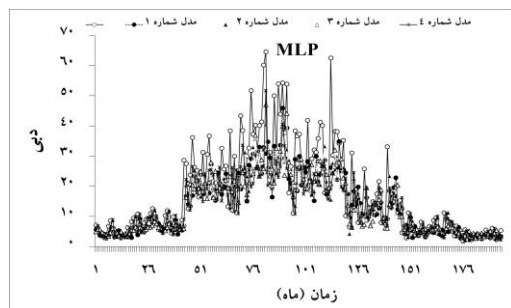
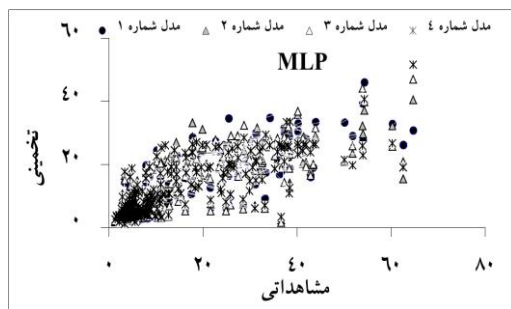
برای تعیین و شناسایی نوع و مرتبه مدل باید ابتدا بر روی نمودارهای ACF^۱ و PACF^۲ قضاوت نمود. به طور معمول مدل های ARMA، خود هم بسته (AR) و میانگین متحرک (MA) قابل انتخاب می باشند. دلیل استفاده گسترده از این مدل ها را می توان مربوط به توانایی آن ها در ایجاد هم بستگی بین مقادیر زمان حال با زمان های پیشین و همچنین، سادگی ساختار این مدل ها دانست (Smith و Salas، ۱۹۸۱). استفاده از مدل های تصادفی و سری های زمانی با توجه به ماهیت

^۱ Autocorrelation Function

^۲ Partial Autocorrelation Function

^۳ Akaike

حساسیت ورودی‌های مدل در شبکه عصبی مصنوعی در شکل ۶ ارائه شده است



شکل ۵- مقایسه بین مقادیر مشاهداتی و محاسباتی

مدل ۲ با دبی یک ماه قبل تا دبی سه ماه قبل، مدل ۳ با دبی یک ماه قبل تا دبی دو ماه قبل و مدل ۱ با دبی یک ماه قبل، به ترتیب در اولویت‌های بعدی قرار دارند. به طوری که نتایج جدول ۱ حاکی از افزایش دقت و کاهش خطای مدل‌ها با افزایش تعداد

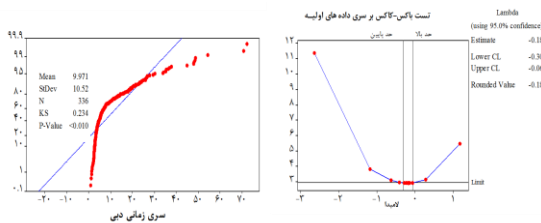
جدول مقادیر R و RMSE برای چهار مدل ارائه شده است. شکل ۵ نیز مقایسه بین مقادیر مشاهداتی با مقادیر محاسباتی و پلات‌های پراکنش نقاط در روش شبکه عصبی مصنوعی را در ایستگاه جویستان و در مدل‌های مختلف نشان می‌دهد.

جدول ۱- مقادیر R و RMSE برای شبکه MLP و RBF

| نام مدل | MLP | | RBF | | متغیر ورودی (دبی متوسط ماهیانه) |
|----------------------------|------|------|------|------|---------------------------------|
| | RMSE | R | RMSE | R | |
| دبی یک تا چهار مرتبه تأخیر | ۷/۴۰ | ۰/۸۵ | ۶/۴۵ | ۰/۸۶ | دبی یک تا چهار مرتبه تأخیر |
| دبی یک تا سه مرتبه تأخیر | ۷/۵۵ | ۰/۸۴ | ۷/۰۷ | ۰/۸۳ | دبی یک تا سه مرتبه تأخیر |
| دبی یک تا دو مرتبه تأخیر | ۷/۷۱ | ۰/۸۳ | ۷/۵۹ | ۰/۸۰ | دبی یک تا دو مرتبه تأخیر |
| دبی یک تا مرتبه تأخیر | ۸/۰۸ | ۰/۸۰ | ۷/۹۹ | ۰/۷۷ | دبی یک تا مرتبه تأخیر |

نتایج جدول ۱ نشان می‌دهد که مدل ۱ با دبی یک ماه تا دبی چهار ماه قبل (به عنوان داده ورودی) برای شبکه MLP و RBF بهترین کارایی را ارائه داده است. ساختار شبکه MLP در مدل ۱ شامل چهار نرون در لایه ورودی، ۲۰ نرون در لایه میانی و یک نرون در لایه خروجی (دبی مشاهداتی) می‌باشد؛ پارامترهای شبکه MLP که با سعی و خطا به دست آمد شامل تعداد تکرار، ضریب یادگیری، تعداد نرون مخفی و خطای هدف است که میزان مقادیر به دست آمده در مدل ۱ به ترتیب، ۵۰۰، ۰/۲۲، ۲۰ و ۰/۰۰۵ می‌باشد. ساختار شبکه RBF در مدل ۲ شامل چهار نرون در لایه ورودی و یک نرون در لایه خروجی (دبی مشاهداتی) می‌باشد. در شبکه نوع RBF مهم‌ترین پارامترها شاخص SPREAD و خطای هدف می‌باشد که مقدار آن‌ها با سعی و خطا به دست آمد. با تغییر این پارامترها تغییر قابل توجهی در نتایج مقادیر تخمینی صورت می‌گیرد به طوری که شاخص SPREAD معادل یک و میزان خطای هدف معادل صفر بهترین کارایی را ارائه داده است. آزمون آنالیز

کاکس بهترین روش نرمال سازی داده‌ها مشخص شد. نمودار آزمون نرمالیت و آزمون باکس-کاکس برای سری‌های زمانی دبی ماهانه ایستگاه جویستان در شکل ۷ ارائه شده است.

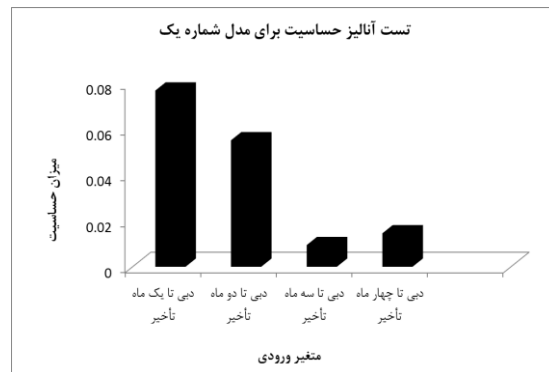


شکل ۷- آزمون نرمالیت و آزمون باکس- کاکس برای سری‌های زمانی دبی ماهانه ایستگاه جویستان

مرحله بعد تعیین مؤلفه‌های روند و حذف آن از سری زمانی و برازش مدل ایستا بر سری زمانی می‌باشد. توابع خود هم‌بستگی (ACF) و خود هم‌بستگی جزئی (PACF) برای داده‌ها به ازای ۶۰ و ... و $K=1$ ترسیم شد. پس از مقایسه با مقادیر توابع، مدل‌های خودهم‌بستگی (AR)، میانگین متحر (MA) و ترکیبی از مدل خودهم‌بستگی و میانگین متحر (ARMA)، مرتبه مدل شناسایی و مدل انتخاب شد. شکل ۸ تابع خود هم‌بستگی و خود هم‌بستگی جزئی را برای داده‌های فصلی و غیرفصلی نشان می‌دهد. یکی از نیازهای اصلی مدل آماری باکس-جنکینز شناسایی مؤلفه‌های سری زمانی برای تعیین ضرایب مورد نیاز مدل می‌باشد. ضرایب p و q از روی نمودارهای خود هم‌بستگی و خود هم‌بستگی جزئی حاصل می‌شود. با ترسیم مقادیر ضریب هم‌بستگی در مقابل تاخیر زمانی، نمودار خود هم‌بستگی ایجاد می‌شود که برای تفسیر مجموعه ضرایب هم‌بستگی داده‌ها مورد استفاده قرار می‌گیرد. در مدل آرما نمودارهای خود هم‌بستگی و خود هم‌بستگی جزئی به صورت سینوسی و نمایی نوسان می‌کنند (Farjamnia و همکاران، ۲۰۰۷).

برای صحت فرآیند تفاضل‌گیری در حذف الگوی روند از داده‌ها، نمودارهای خود هم‌بستگی و خود هم‌بستگی جزئی بر داده‌های تفاضل‌گیری شده محاسبه و براساس آن‌ها نمودار خود هم‌بستگی ترسیم شد که می‌توان از طریق آن‌ها ضرایب مناسبی برای p و q استخراج کرد (شکل ۸).

تاخیرهای دبی می‌باشد، که این نتایج با نتایج پژوهش Antar و همکاران (۲۰۰۶) و Kisi (۲۰۰۹) مبنی بر این‌که اگر شبکه با داده‌های ورودی مناسبی آموزش داده شود قابلیت پیش‌بینی بهتری در رابطه بین بارش و رواناب نسبت به مدل‌های دیگر دارد، مطابقت دارد.



شکل ۶- آزمون آنالیز حساسیت ورودی‌های مدل در شبکه عصبی مصنوعی

نتایج نشان داد که به‌طور کلی در مقایسه بین مدل‌ها، مدل MLP با میانگین ضریب هم‌بستگی $0/83$ نسبت به مدل RBF با میانگین ضریب هم‌بستگی $0/81$ دقت بیشتری در پیش‌بینی دبی نشان داده است که در این زمینه محققان دیگری نیز نتایج مشابهی را به‌دست آورده‌اند (Fatehi Marj و Borhani Dariane، ۲۰۰۸؛ Alp و Kerem، ۲۰۰۹). همان‌طور که ملاحظه می‌شود در برازش ترسیمی (شکل ۵) نیز بیشترین مطابقت در مدل ۱ بین داده‌های دبی مشاهداتی و محاسباتی مشاهده می‌شود. با این وصف مدل ۱ به‌عنوان بهترین مدل برای این منطقه پیشنهاد می‌شود. آزمون آنالیز حساسیت ورودی‌های مدل در شبکه عصبی مصنوعی (شکل ۶) نیز نشان‌دهنده کاهش حساسیت مدل نسبت به متغیرهای ورودی با افزایش میزان تاخیرها است.

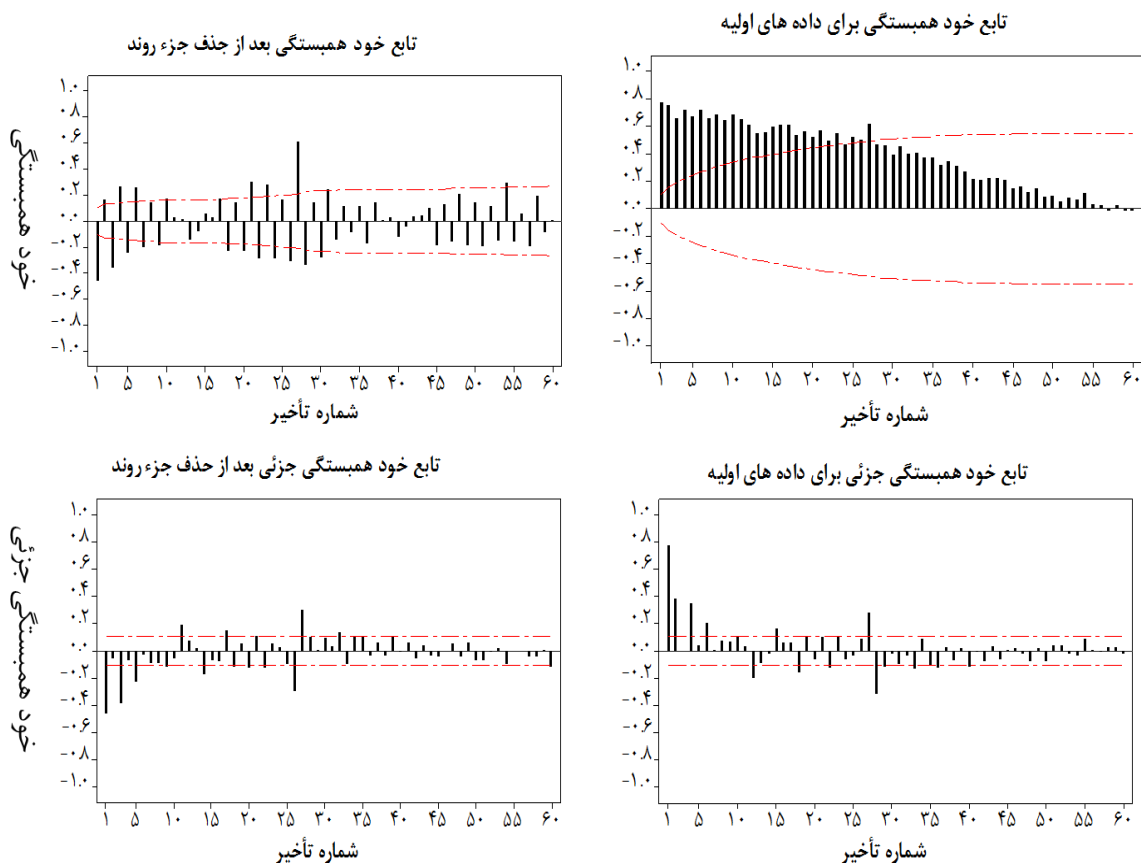
نتایج مدل سازی سری زمانی: در این پژوهش سری زمانی ۳۰ ساله مربوط به دبی متوسط ماهانه زیرحوضه جویستان بررسی شده است. برای مدل سازی سری زمانی ابتدا داده‌ها را به صورت سری زمانی مرتب کرده و سپس با استفاده از آزمون‌های نرمالیت، نرمال بودن داده‌ها بررسی شد. با توجه به نرمال نبودن داده‌های دبی ایستگاه جویستان با استفاده از آزمون باکس-

داده‌ها در روش باکس-جنکینز با توجه به توابع ACF و PACF از مدل‌های آرما تبعیت کردند به طوری که در این روش چهار ساختار آرما (۱و۰)، (۱و۱)، (۱و۲) و (۲ و ۲) مدل‌سازی شد. نتایج مدل‌سازی سری زمانی نیز در شکل‌های ۹ و ۱۰ و جدول ۲ ارائه شده است.

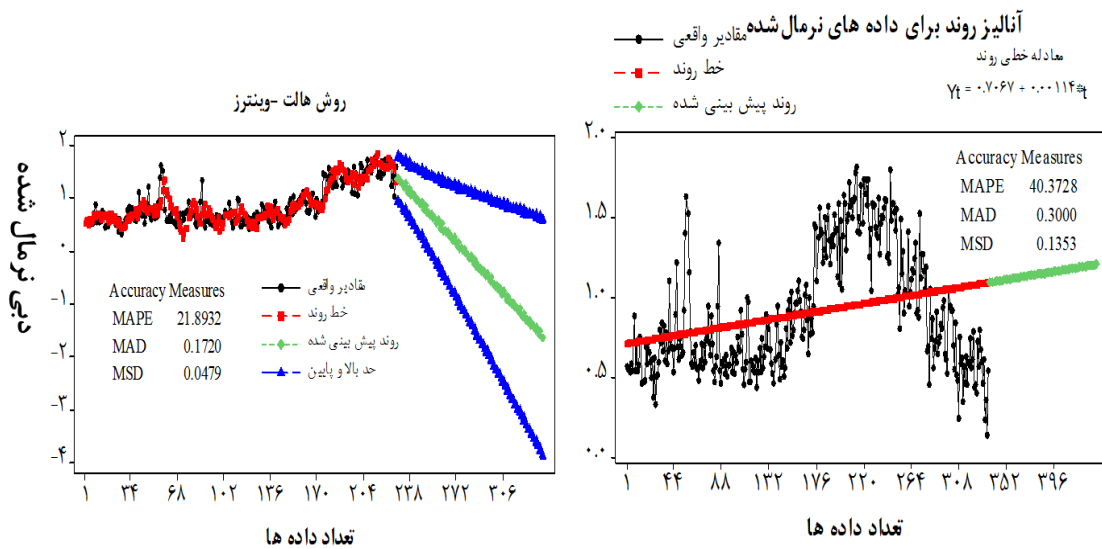
همان‌گونه که در شکل‌های ۹ و ۱۰ ملاحظه می‌شود روند پیش‌بینی مدل‌های سری زمانی به کار رفته در پژوهش حاضر نشان داد که روش‌های باکس-جنکینز و حالت وینترز داده‌های دبی آینده را به صورت کاهشی پیش‌بینی نموده است در حالی که روش آنالیز روند دبی را به صورت صعودی پیش‌بینی نموده است. در بین مدل‌های مختلف باکس-جنکینز بهترین پیش‌بینی را مدل آرما نوع چهار از مقادیر دبی ارائه کرده است. مدل آرما نوع چهار با کمترین میزان خطا برترین مدل را در بین مدل‌های مختلف ارائه نموده است.

همان‌گونه که از نمودار خود همبستگی مشخص است مقدار ACF بازه اطمینان را قطع کرده و نزول کرده است که به دلیل وجود روند است. ولی نمودار خود همبستگی پس از تفاضل‌گیری، عامل روند را از خود نشان نداده است و نمودار فقط میل به تناوب دارد و مقدار همبستگی در برخی تأخیرها خط اطمینان را قطع کرده است که مبین نوسانات فصلی می‌باشد. برای حذف این عامل از سری، از روش تفاضل‌گیری فصلی استفاده شد. در نهایت بررسی توابع خود همبستگی و خود همبستگی جزئی نشان می‌دهد که داده‌های موجود در هر دو ایستگاه مطالعاتی از مدل‌های آرما تبعیت می‌کنند.

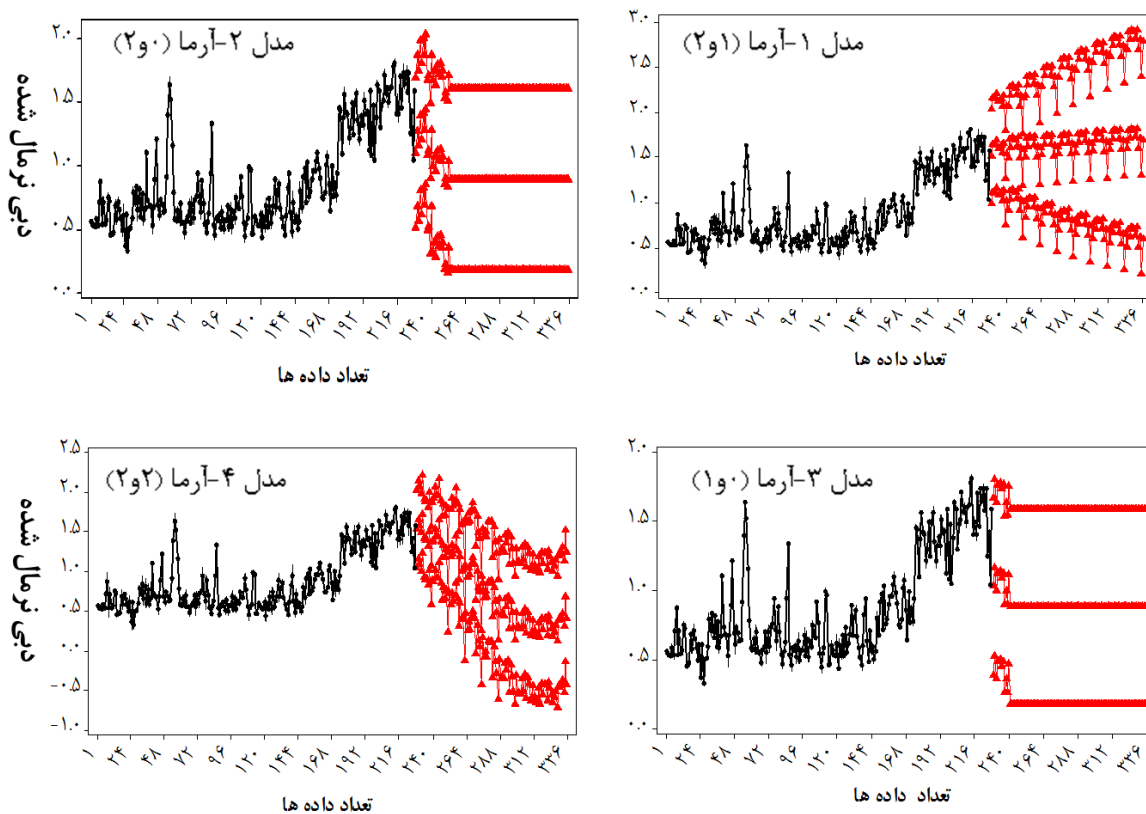
برای مدل‌سازی دو سوم داده‌ها برای مدل‌سازی و یک سوم داده‌ها به منظور ارزیابی صحت پیش‌بینی تفکیک شدند. در نهایت روش‌های آنالیز روند، حالت-وینترز و باکس-جنکینز بر داده‌ها برازش داده شد.



شکل ۸- نمودارهای توابع خود همبستگی و خود همبستگی جزئی برای داده‌های اصلی و تفاضل‌گیری شده



شکل ۹- نتایج روش های آنالیز روند، حالت وینترز در ایستگاه جوستان



شکل ۱۰- مدل آرما بر سری زمانی ماهانه دبی ایستگاه جوستان

به طوری که مشاهده می شود در بین مدل های مختلف مدل ۴ با مقدار p و q به ترتیب معادل ۲ و ۲ کمترین آکائیک را دارا می باشند. مقایسه مقادیر ضریب تبیین و ریشه مجذور مربعات خطا و معیار آکائیک در بین روش های مختلف سری زمانی حاکی از دقت بیشتر مدل ۴ با ساختار (۲ و ۲) را نشان می دهد.

آزمون و معیار آکائیک (AIC) یکی از روش های مقایسه مدل های مختلف می باشد. این روش بر این مبنا استوار است که از بین مدل های مناسب، مدلی که دارای کمترین مقدار ضریب آکائیک باشد، به عنوان بهترین مدل انتخاب می شود. مقادیر آکائیک به دست آمده از مدل های مختلف در جدول ۲ ارائه شده است.

روی هم رفته به عقیده محققان مختلف شبکه عصبی مصنوعی به دلایل عمده از جمله برخورداری از توانایی تشخیص الگو، رابطه خوب بین ورودی و خروجی، حساسیت کمتر نسبت به وجود خطا در اطلاعات ورودی، فرآیند پردازش کاملاً موازی (Soltani, ۲۰۰۲ و Ashrafzadeh, ۱۹۹۹)، نیاز به تعداد داده‌های ورودی کمتر و روند صحت‌سنجی و واسنجی سریع‌تر و آسان‌تر (Borhani Dariane و Fatehi Marj, ۲۰۰۸) و توانایی در کشف و پیش‌بینی ارتباط پارامترهای حوزه آبخیز، (Souza Filho و Lall, ۲۰۰۳ و Sveinsson Oli و همکاران، ۲۰۰۳) نسبت به دیگر روش‌های مدل‌سازی نتیجه قابل قبول‌تری ارائه می‌دهد.

با این وجود پیشنهاد می‌شود که از روش‌های دیگر هوش مصنوعی از جمله نرون فازی، مدل‌های هیبرید، شبکه‌های مختلف عصبی مصنوعی، شبکه عصبی موجکی، برنامه‌ریزی ژنتیک و رگرسیون چندمتغیره برای ارزیابی دقت آن‌ها در پیش‌بینی دبی در حوزه آبخیز طالقان استفاده شود.

جدول ۲- ضرایب خطاسنجی مدل‌های مختلف سری زمانی در

| | ایستگاه جوستان | | |
|------------------|----------------|-------|------|
| | معیار آکائیکه | RMSE | R |
| آرما-مدل ۱ (۲۰۱) | -۲۰۱/۷۱ | ۳۶/۸۰ | ۰/۲۸ |
| آرما-مدل ۲ (۲۰۰) | -۱۷۵/۷۳ | ۱۲/۰۷ | ۰/۵۶ |
| آرما-مدل ۳ (۱۰۰) | -۱۶۰/۷۳ | ۱۲/۵۹ | ۰/۴۷ |
| آرما-مدل ۴ (۲۰۲) | -۲۱۳/۱۵ | ۹/۰۱ | ۰/۷۴ |
| آنالیز روند | - | ۳۳/۳۲ | ۰/۶۶ |
| هالت وینترز | - | ۱۳/۱۲ | ۰/۷۲ |

همچنین، با مقایسه مقادیر ضریب تبیین و ریشه مجذور مربعات خطا در روش سری زمانی آرما با مقادیر مشابه در روش شبکه عصبی مصنوعی دقت بیشتر روش شبکه عصبی را نشان می‌دهد که با یافته Méndez Castellano و همکاران (۲۰۰۴) و Antar و همکاران (۲۰۰۶) مبنی بر این‌که شبکه عصبی مصنوعی نسبت به مدل باکس-جنکینز برای مدل‌سازی دبی قابلیت بیشتری دارد، مطابقت دارد.

منابع مورد استفاده

1. Abrahart, R.J. and L. See. 2000. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Journal of Hydrological Processes*, 14: 2157-2172.
2. Abrichamchi, A., M. Tajrishy and B. Chehreneghar. 2005. Regional stochastic models of annual Streamflow for West Basin of Iran (Karkheh, Dez and Karoon). *Iran-Water Resources Research*, 1(1): 48-57 (in Persian).
3. Alp, M. and H. Kerem. 2009. Suspended sediment load simulation by two artificial neural network methods using hydro-meteorological data. *Environmental Modelling and Software*, 22: 2-13.
4. Antar, A., I. Ellassiouti and M.N. Allam. 2006. Rainfall-runoff modelling using artificial neural networks technique: a Blue Nile catchment case study. *Hydrological Processes Journal*, 20: 1201-1216.
5. Asghari Moghaddam, A., V. Nourani and A. Nadiri. 2009. Modeling of Tabriz plain rainfall using artificial neural networks. *Journal of Agricultural Science*, 18(1): 1-15 (in Persian).
6. Ashrafzadeh, A. 1999. Application of artificial neural networks for prediction of evaporation from evaporative ponds, MSc Thesis, Tehran University, Tehran, 131 pages (in Persian).
7. Borhani Dariane, A. and A. Fatehi Marj. 2008. Application of artificial neural network in stream flow forecasting using climatic indices, Case study: Nazloochoy river basin. *Journal of Faculty of Engineering*, 35(3): 51-63 (in Persian).
8. Box, G.E.P. and G.M. Jenkins. 1976. *Time series analysis: forecasting and control*. San Francisco, Holden-Day, 550 pages.
9. Dawson, C.W., R.L. Wilby. 2001. Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, 25(1): 80-108.
10. Farjamnia, I., M. Naseri and S.M.M Ahmadi. 2007. Oil price forecasting; a comparison between ARIMA and ANN models. *Iranian Economic Research*, 32:161-183 (in Persian).

11. Ghanbarpour, M.R., M. Amiri and S.A. Gholami. 2008. Evaluation of monthly discharge forecasting models based on time series analysis (Case study: Karkheh basin). *Journal of the Iranian Natural Resources*, 61(1): 43-56 (in Persian).
12. Hagan, M.T. 1994. Training feed forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 6: 989-993.
13. Integrated Studies of Taleghan river basin. 1993. Faculty of Agricultural, Tehran University (in Persian).
14. Jahangeer, A.R., M. Raeini and M.Z. Ahmadi. 2008. Comparison of artificial neural networks (ANN) simulation of rainfall-runoff process with HEC-HMS model in Kardeh watershed, *Journal of Water and Soil*, 22: 72-84 (in Persian).
15. Karamouz, M. and S. Araghinejad. 2005. *Advanced Hydrology*. Amir Kabir University Press, 460 pages (in Persian).
16. Kisi, O. 2009. Neural networks and wavelet conjunction model for intermittent stream flow forecasting. *Journal Hydrologic Engineering*, 14(8): 773-782.
17. Méndez Castellano, M., W. González Manteiga, M. Febrero Bande, J. Manuel, P. Sánchez and R. Lozano Calderón. 2004. Modelling of the monthly and daily behavior of the runoff of the Xallas river using Box-Jenkins and neural networks methods. *Journal of Hydrology*, 296: 38-58.
18. Nouraei, V., K. Salehi, N. Abulvasset and M. Salehi. 2009. Rainfall-runoff modeling using adaptive wavelet neuro-fuzzy technique and comparison results with wavelet neural network and adaptive neuro-fuzzy modeling. 8th International Congress on Civil Engineering, Shiraz University, Shiraz, Iran, 8 Pages (in Persian).
19. Rahimi Khoob A, S. Behbahani, M. Nazarifar. 2008. Estimating maximum air temperature in Khoozestan province using NOAA satellite images data and artificial neural network. *JWSS-Isfahan University of Technology*, 11: 357-364 (in Persian).
20. Rahmani, A. and M. Sadeghi. 2004. Predication of groundwater level changes in the Plain of Hamedan-Bahar using time series model. *Journal of Water and Wastewater*, 51: 42-49 (in Persian).
21. Salas, J.D. and R.A. Smith. 1981. Physical basis of stochastic models of annual flows. *Water Resources Research*, 17: 428-430.
22. Soltani, S. 2002. Comparison of conceptual and neural network models for rainfall-runoff modeling. MSc Thesis, Tarbiat Modares University, Tehran, 98 pages (in Persian).
23. Souza Filho, F.A. and U. Lall. 2003. Seasonal to inter-annual ensemble stream flow forecasting for Ceara Brazil. *Water Resources Research*, 39(11): 1307-1320.
24. Sveinsson Oli, G.B., J.D. Salas, D.C. Boes and R.A. Pielke. 2003. Modelling the dynamics of long-term variability of hydro-climatic processes. *Journal of Hydrology*, 7: 489-501.
25. Zhou, H.C., Y. Peng and G.H. Liang. 2008. The research of monthly discharge predictor-corrector model based on wavelet decomposition. *Water Resource Management*, 22: 217-227.

Comparison prediction performance of monthly discharge using ANN and time series

Majid Khazaei¹ and Mohammad Reza Mirzaei^{*2}

¹ Ph.D. Student, Faculty of Agricultural and Natural Resources, Hormozgan University, Iran, ² Assistant Professor, Faculty of Agriculture, Yasouj University, Iran

Received: 17 October 2012 Accepted: 04 May 2013

Abstract

Prediction in hydrology is as estimation of hydrological and meteorological conditions in a specific interval time. In this regard, understanding the relationship between precipitation and runoff is necessary for water resources optimal management. The purpose of this study was to compare different models of artificial neural networks (two type of ANNs: RBF and MLP) and time series models (ARMA) to discharge estimation in a part of the Taleghan watershed, using monthly flow discharge data for a period of 30 years between 1977 and 2007. Among the different ARMA models, a model with a lowest error and akaike (AIC) criterion was selected as an optimal model. Using trial and error method, ANNs were designed by specifying the number of hidden layers and neurons in each layer, sigmoidal transfer function, training function, weight/bias learning function and performance function. Using trend analysis, Halt-Winters and Box-Jenkins (ARMA) methods, time series analysis showed that ARMA (2, 2) ($R=0.77$) and Halt-Winters ($R=0.72$) presented more accurate results. In general, it could be concluded that ANNs models produced more accurate predictions of flow discharge than time series approaches. Also, the results revealed that the MLP model (average $R=0.83$) produced more accurate predictions of flow discharge than RBF model (average $R=0.81$). Assessment of accuracy of all models based on RMSE and R showed that the model 1 (with $RMSE=6.45$ and $R=0.86$) obtained with a network architecture of 4-20-1 configuration. Model 1 used the input vector consisting of antecedent monthly discharge with one to four time lag.

Key words: Accuracy, Moving average, Regression coefficient, Taleghan watershed, Water resources management

* Corresponding author: mmirzaei@mail.yu.ac.ir