

پیش‌بینی جریان رودخانه با استفاده از مدل‌های استوکاستیک

(مطالعه موردی: حوضه آبریز خرم آباد)

روفیا امیدی

دانشجوی کارشناسی ارشد مهندسی منابع آب دانشگاه شهید چمران اهواز Rofiaomidi@gmail.com

فریدون رادمنش

استادیار دانشکده مهندسی علوم آب دانشگاه شهید چمران اهواز Fereidon_radmanesh@yahoo.com

حیدر زارعی

استادیار دانشکده مهندسی علوم آب دانشگاه شهید چمران اهواز zareideidar@gmail.com

چکیده

پیش‌بینی مناسب مقدار آبدهی رودخانه‌ها، نقشی اساسی در برنامه ریزی، مدیریت و بهره برداری بهینه و پایدار از منابع آبی ایفا می‌کند. از آنجا که وقوع سیل در این حوضه بی‌سابقه نبوده است به همین دلیل این تحقیق تلاشی در جهت شناسایی، به منظور ساماندهی در راستای برنامه‌ها و اهداف و آینده‌نگری مخاطرات طبیعی در این حوضه می‌باشد. یکی از روش‌های پیش-بینی آبدهی استفاده از سری‌های زمانی است که ابزاری قدرتمند جهت پیش‌بینی دبی می‌باشند. لذا در تحقیق حاضر داده‌های ۳۰ ساله روزانه دبی، بارش و دما در بازه زمانی (۱۳۹۱-۱۳۶۱) مورد استفاده قرار گرفت. سری‌های زمانی مختلف همچون ARMAX، ARIMA برای پیش‌بینی آبدهی رودخانه بر داده‌ها برازش داده شد. سپس برای تشخیص درستی الگوی برازش داده شده از آزمون بررسی فرض استقلال باقیمانده‌ها استفاده گردید. نتایج نشان داد که مدل ARMAX به دلیل ورود پارامتر-های اقلیمی در مدل دارای نتایج بهتری در مقایسه با مدل آرما بوده است.

واژگان کلیدی: پیش‌بینی، سری زمانی، ARMAX، ARIMA، رودخانه خرم‌آباد

جهت برنامه ریزی صحیح و استفاده بهینه و منطقی از منابع با ارزش آب و همزمان حفظ این منابع برای نسلهای آینده، استفاده از دانش روز هیدرولوژی اجتناب ناپذیر به نظر می‌رسد (مهدوی، ۱۳۸۵). به علت پیچیدگی و عدم دانش کافی در مورد فرایندهای فیزیکی در چرخه هیدرولوژیک، ساخت مدل‌های آماری و گسترش آنها برای بیان این فرایندها همیشه مورد توجه مهندسين بوده است. اساس بسیاری از تصمیم‌گیرها در فرآیندهای هیدرولوژیک و تصمیمات بهره‌برداری از منابع آب بر پایه پیش‌بینی و تحلیل سریهای زمانی می‌باشد. استفاده از سریهای زمانی امروزه به عنوان ابزاری مناسب برای پیش‌بینی‌های مختلف به کار می‌رود. سری زمانی مجموعه‌ای از مشاهدات است که بر حسب زمان مرتب شده‌اند و چنانچه این مشاهدات به طور منظم و در فاصله‌های مساوی ثبت و یا اندازه‌گیری شده باشد، یک سری زمانی گسسته به دست می‌آید (نیرومند و بزرگنیا، ۱۳۸۷). با استفاده از مدل‌های سری زمانی تولید داده‌های مصنوعی هیدرولوژیک، پیش‌بینی وقایع هیدرولوژیک، تشخیص روند و پرش در داده‌ها، تکمیل خلا آماری و تطویل دوره آماری امکانپذیر خواهد بود. مدل‌های سری زمانی از دو بخش یا مولفه اصلی شامل مولفه تصادفی و مولفه مدل تشکیل شده‌اند که مولفه مدل با استفاده از ارقام مشاهداتی و مولفه تصادفی با استفاده از روش‌های مختلف استوکاستیک بدست می‌آید. لذا ساختار مدل‌های سری زمانی می‌تواند با ساختار سریهای هیدرولوژیک در صورت انتخاب درست مدل و محاسبات صحیح آن سازگاری و مطابقت ویژه‌ای داشته باشد مدل‌های مختلف آماری مانند مدل‌های خود همبسته^۱ (AR)، میانگین متحرک^۲ (MA)، خودهمبسته با میانگین متحرک^۳ (ARMA)، آریما^۴ (ARIMA) هر یک شامل مجموعه‌ای از مدل‌ها با پارامترهای گوناگون می‌باشند و می‌توانند به عنوان انتخاب‌های ممکن برای مدلسازی استفاده شوند. اگر یک فرایند تصادفی محض دارای میانگین صفر و واریانس مشخص و دارای یک الگوی رگرسیون چندگانه باشد به آن فرایند اتورگرسیو گفته می‌شود فرایندهای اتورگرسیو میانگین متحرک یک سری زمانی دارای فرایند ایستا است و با نماد اختصاری $ARMA(p,q)$ نشان داده می‌شود که در آن P مرتبه اتورگرسیو و q مرتبه میانگین متحرک است. از طرف دیگر به یک الگوی نایستای همگن الگوی اتورگرسیو تلفیق شده با میانگین متحرک گفته شده و با نماد اختصاری $ARIMA(p,d,q)$ نشان داده می‌شود که در آن d مرتبه تفاضلی است که برای تبدیل سری نایستا به سری ایستا بکار می‌رود (فولادوند، ۱۳۸۹). در سال‌های گذشته مطالعات زیادی در خصوص استفاده از مدلسازی سری زمانی صورت گرفته است و هر کدام از آنها بنابر هدف خاصی از این تکنیک بهره گرفته‌اند. کومورنیک^۵ و همکاران (۲۰۰۶) به مقایسه و پیش‌بینی کارایی مدل‌های هیدرولوژیک سری‌های زمانی در جمهوری چک پرداختند که نتایج آنها کارایی بالای مدل‌های مذکور در پیش‌بینی فرایندهای هیدرولوژیک را بیان می‌کند. دامل و یالکین^۶ (۲۰۰۷) به پیش‌بینی سیلاب با کاربرد سری‌های زمانی در رودخانه می‌سی‌سی‌پی آمریکا پرداختند. نتایج آنها نیز حاکی از توانایی سریهای زمانی در ساخت داده‌های دبی روزانه و صحت پیش‌بینی‌های حاصله است. هاسانا و همکاران^۷ (۲۰۱۳) برای پیش‌بینی سیلاب از مدل تابع انتقال استفاده کردند و به این نتیجه رسیدند که مدل تابع انتقال نسبت به مدل سری زمانی آریما پیشگویی بهتری دارد. بشری (۱۳۸۹) به مقایسه روش‌های مختلف تحلیل سری‌های زمانی در

1. Auto Regressive

2. Moving Average

3. Auto Regressive Moving Average

4. Auto Regressive Integrated Moving Average

5. Komornik

6. Damel and Yalkin

7. Yulianti Hasanah

پیش‌بینی دبی ماهانه حوضه آبریز کرخه پرداخت و به این نتیجه رسید که روش آنالیز روند بهترین پیش‌بینی و پس از آن مدل‌های مختلف خودهمبسته با میانگین متحرک با اختلاف جزئی در اولویت بعدی قرار دارند و نیز بررسی معیار اطلاعاتی آکاییکه در مدل‌های اخیر نشان داد که مرتبه‌های بالاتر، دقت بالاتری در مدلسازی دارند ولی در زمینه پیش‌بینی مقادیر آینده مرتبه‌های پایین‌تر مدل موفق‌تر هستند. دودانگه و همکاران (۱۳۹۱) نیز کاربرد مدل‌های سری زمانی را به منظور تعیین روند پارامترهای اقلیمی از جمله درصد رطوبت، تبخیر، دمای هوا، سرعت باد و تعداد ساعات آفتابی در آینده برای اصفهان مورد ارزیابی قرار دادند، نتایج این مطالعه بیانگر سودمندی مدل‌های سری‌زمانی در مطالعات منابع آب از طریق پیش‌بینی پارامترهای اقلیمی و تعیین روند پارامترهای اقلیمی در آینده است.

هدف از این تحقیق مقایسه کارایی مدل‌های ARIMA و ARMAX در پیش‌بینی فرایندهای هیدرولوژیکی است.

مواد و روش‌ها:

منطقه مورد مطالعه

حوضه آبریز رودخانه خرم‌آباد در استان لرستان و در بخش کوهستانی زاگرس واقع شده است از آنجا که وقوع سیل در این حوضه بی‌سابقه نبوده است به همین دلیل این تحقیق تلاشی در جهت شناسایی، به منظور ساماندهی در راستای برنامه‌ها و اهداف و آینده‌نگری مخاطرات طبیعی در این حوضه می‌باشد. در این بررسی رودخانه خرم‌آباد انتخاب گردید و از ایستگاه آبسنجی چمن‌انجیر در پی اطلاعات مورد نیاز استفاده شد. مشخصات ایستگاه آبسنجی چمن‌انجیر در جدول (۱) آمده است:

جدول (۱): جدول مشخصات ایستگاه آبسنجی چمن‌انجیر

دوره موجود	مساحت حوضه (Km ²)	تجهیزات	ارتفاع (متر)	طول درجه-دقیقه	عرض درجه-دقیقه
۱۳۶۱-۹۱	۲۵۰۶	اشل-پل تلفریک- لیمونوگراف	۱۱۴۰	۴۸-۱۴	۲۶-۳۲

مدل سازی سری زمانی

به منظور تجزیه و تحلیل سری زمانی، بررسی نرمال بودن داده‌ها امری ضروری می‌باشد. نرمال بودن داده‌ها به این دلیل اهمیت دارد که تئوری سری‌های زمانی براساس نرمال بودن داده‌ها توسعه یافته و در صورتی که داده‌ها نرمال نباشند بایستی از روش‌های مختلف چون لگاریتم‌گیری، جذر، نمایی و یا تبدیل توانی Box-Cox آن‌ها را نرمال نمود. در مقاله حاضر به منظور نرمال نمودن داده‌ها از روش لگاریتم‌گیری بر مبنای عدد نپر استفاده شده است. سپس با استفاده از توابع خودهمبستگی ACF^۱ و خودهمبستگی جزئی PACF^۲ مدل‌های سری زمانی مناسب به داده‌ها برآزش داده شد و با استفاده از رفتار این دو تابع خواص ایستایی و فصلی بودن داده‌ها مورد بررسی قرار گرفت. مدلی که به فراوانی استفاده می‌شود مدل ARIMA است. دو شکل کلی مدل‌های ARIMA، عبارتند از ARIMA غیرفصلی (P,d,q) و فصلی-ضربی

^۱ Autocorrelation Function

^۲ Partial Autocorrelation Function

و میانگین متحرک فصلی است که p و q به ترتیب پارامترهای اتورگرسیو و میانگین متحرک غیرفصلی و P و Q پارامترهای اتورگرسیو و میانگین متحرک فصلی است. دو پارامتر دیگر یعنی D و d پارامترهای تفاضلی برای ایستا کردن سری زمانی است. عملگر تفاضلی مورد استفاده برای سری‌های زمانی پویا عبارتند از $\Delta = 1 - B$ (عملگر جهش به عقب است)، این شکل از مدل‌های ARIMA غیرفصلی به شکل زیر نوشته می‌شود:

$$\Phi(B)Z_t = \Phi(B)(1 - B)Z_t = \theta(B)a_t \quad (1)$$

که Z_t سری‌های مشاهده شده، $\Phi(B)$ رتبه چند جمله‌ای p و $\theta(B)$ رتبه چند جمله‌ای q است.

بعد از مشخص شدن مدل، باید برآورد موثر از پارامترها صورت گیرد. پارامترها باید دو شرط ایستایی و معکوس پذیری را برای اتورگرسیو و میانگین متحرک داشته باشند، پارامترها باید از نظر معنی داری مورد آزمون قرار گیرند که در ارتباط با مقادیر خطای برآوردها و برآورد مقادیر t می‌باشد، اگر θ برآورد نقطه‌ای از پارامتر مورد نظر و $S\theta$ خطای برآورد باشد، مقدار t به صورت زیر خواهد بود.

$$t = \frac{\theta}{S\theta} \quad (2)$$

اگر فرض صفر با در نظر گرفتن احتمال خطای برابر یا بیشتر از $\alpha = 0.05$ رد شود در این صورت پارامتر معنی‌دار خواهد بود و در مدل باقی می‌ماند. همچنین اگر مقدار P-value کوچکتر یا مساوی سطح معنی‌دار بودن آزمون یعنی ۰.۰۵ باشد، فرض صفر (H_0) رد می‌شود (دودانگه، ۱۳۹۰).

فرایند ARMAX

مزیت اساسی این روش نسبت به روش‌های دیگر سری زمانی که در آنها فقط از اطلاعات پارامتر مورد پیش‌بینی استفاده می‌شود این است که در این روش از پارامترهای دیگری نیز که با پارامتر مورد بررسی همبستگی دارند استفاده می‌شود.

اگر X_t و Y_t سری‌های درست‌تیدیل شده ایستا باشند در یک سیستم خطی با یک ورودی و یک خروجی، سری خروجی Y_t و سری ورودی X_t تحت یک صافی به صورت:

$$y_t = v_0x_t + v_1x_{t-1} + v_2x_{t-2} + \dots + n_t = v(B)x_t + n_t \quad (3)$$

مرتبطند که در آن $v(B) = \sum_{j=-\infty}^{+\infty} v_j B^j$ توسط باکس و جنکینز تابع انتقال صافی نامیده شده و n_t سری اغتشاش سیستم بوده که مستقل از سری ورودی X_t است. باکس و جنکینز معادله (۳) را الگوی تابع انتقال نامیده‌اند. که این معادله الگوی ARMAX نامیده می‌شود (جهانبخش و همکاران، ۱۳۸۵).

پس از انتخاب مدل برتر باید فرضیه استقلال باقیمانده‌ها مورد آزمون قرار گیرد تا از انتخاب مدل اطمینان حاصل شود.

نتایج و بحث:

در این قسمت ابتدا به بررسی نتایج حاصل از تجزیه و تحلیل آماری اطلاعات جمع‌آوری شده پرداخته شد و سپس مدل‌های مختلف مورد بررسی قرار گرفتند.

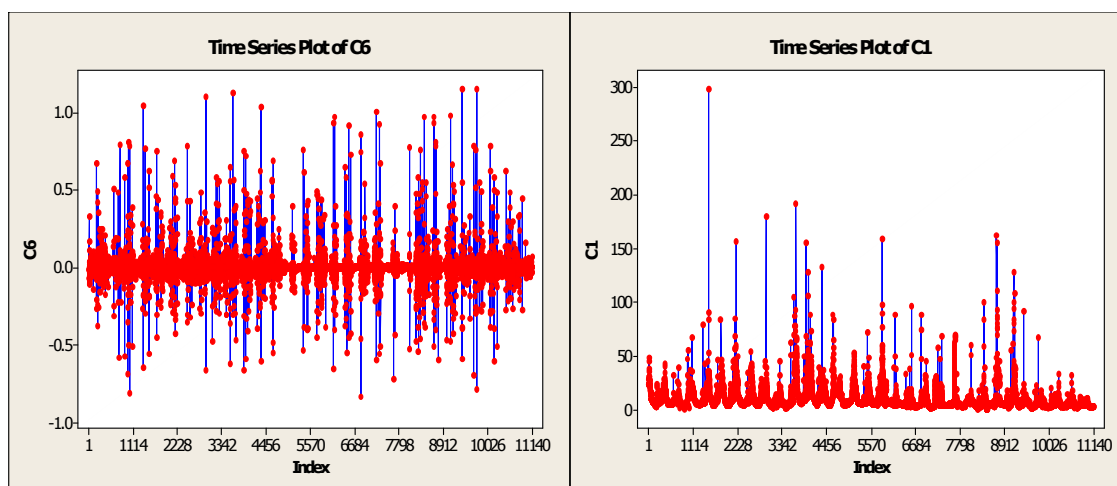
تحلیل آماری متغیرهای تحقیق در جدول ۲ آمده است:

جدول (۲): تحلیل آماری متغیرهای تحقیق

آماره های توصیفی						
تعداد	میانگین	میانه	انحراف معیار	کمترین مقدار	بیشترین مقدار	
۱۱۱۴۳	۹.۵۴	۶.۰۹	۱۰.۹۹	۰.۴	۲۹۷.۸۶	دبی
۱۱۱۴۳	۱.۲۴	۰	۴.۸۴	۰.۰	۷۸	بارش
۱۱۱۴۳	۱۶.۷۲	۱۶.۵	۸.۸۵	-۷	۳۵	دما

برازش الگو

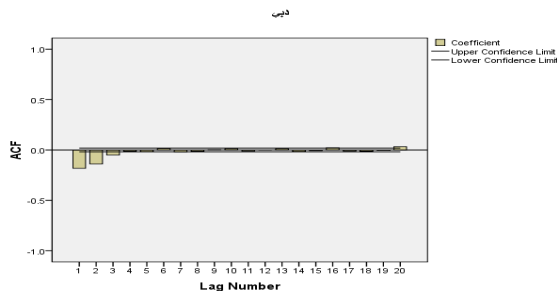
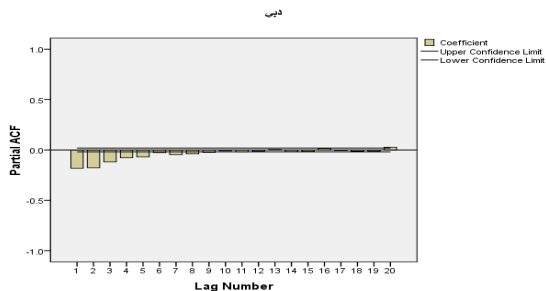
اولین مرحله در آنالیز سری‌های زمانی مشاهده گرافیکی داده‌ها می‌باشد. روند تغییرات دبی حوضه آبریز مورد مطالعه در شکل ۱ ارائه شده است، همانطور که این شکل نشان می‌دهد سری در میانگین و واریانس ناپایستا است. بنابراین یک تبدیل لگاریتمی به منظور پایداری در واریانس و تفاضلی کردن به منظور پایداری در میانگین انجام شد. این موضوع در شکل ۲ نشان داده شده است:



شکل (۲): نمودار سری ایستا شده

شکل (۱): داده‌های اصلی دبی حوضه آبریز رودخانه خرم‌آباد

شکل ۳ و ۴ نیز نمودار توابع خود همبستگی و خود همبستگی جزئی را نشان می‌دهد.



شکل (۳): تابع خود همبستگی داده‌های دبی حوضه آبریز / شکل (۴): تابع خود همبستگی جزئی داده‌های دبی حوضه آبریز

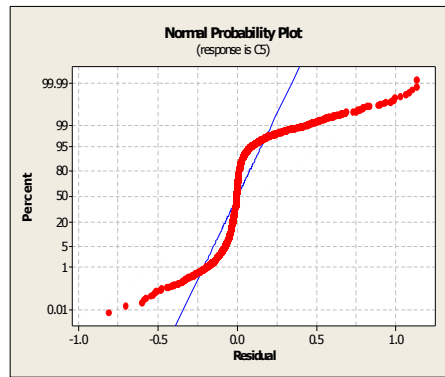
با توجه به نمودار توابع ACF, PACF, رتبه مدل تعیین شد، با توجه به شکل ۳ و ۴ مشاهده می‌شود که نمودار ACF بعد از تاخیر ۲ به سمت صفر میل می‌کند و نمودار PACF بعد از تاخیر ۴ کاهش یافته و در محدوده اطمینان قرار می‌گیرد. که یک مدل ARIMA(2,1,4) را مشخص می‌کند. پارامترهای این مدل در جدول ۳ آمده است:

جدول (۳): پارامترهای مدل ARIMA

	Estimate	SE	t	p-value
Constant	۰.۰۰	۰.۰۰۲	-۰.۲۱۷	۸۲۸
AR Lag 1	-۰.۵۲۳	۰.۱۰۹	-۴.۸۰۸	۰.۰۰۰
Lag 2	۰.۴۰۲	۰.۰۸۱	۴.۹۷۸	۰.۰۰۰
Difference	۱			
MA Lag 1	-۰.۱۷۰	۰.۱۰۹	-۱.۵۵۸	۰.۱۱۹
Lag 2	۰.۷۷۰	۰.۰۵۹	۱۳.۱۱۰	۰.۰۰۰
Lag 3	۰.۰۳۰	۰.۰۵۳	۰.۵۸۰	۰.۵۶۲
Lag 4	-۰.۰۱۲	۰.۰۲۶	-۰.۴۷۶	۰.۶۳۴

با توجه به جدول AR(1) و AR(2) و MA(2) دارای $p\text{-value} < 0.05$ هستند پس معنی دار می‌باشند در حالیکه $p\text{-value}$ مربوط به جمله ثابت و MA (1) و MA (3) و MA (4) معنی دار نیست (بیشتر از 0.05)، بنابراین وجود روند قطعی در مدل مورد تایید نمی‌باشد و همچنین در بین پارامترهای میانگین متحرک α_1 و α_3 و α_4 معنی دار نیستند و می‌توان آن را از مدل حذف کرد.

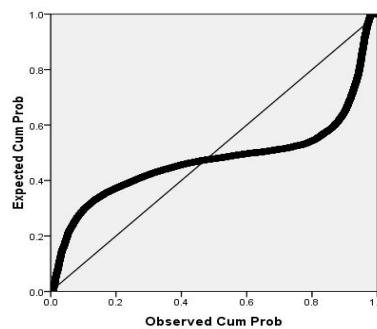
در نهایت جهت پذیرش مدل باید باقیمانده‌های مدل از نظر نرمال بودن و تصادفی بودن مورد آزمون قرار گیرد. نتایج حاصل از این بررسی در شکل ۵ نشان داده شده است.



شکل(۵): بررسی نرمال بودن باقیمانده‌ها در مدل ARIMA

همان طور که دیده می‌شود نمودار احتمال نرمال باقیمانده فرضیه نرمال بودن باقیمانده‌های حاصل از برازش مدل $ARIMA(2,1,4)$ را بخوبی تایید می‌کند. در مرحله بعد پارامترهای اقلیمی، دما و بارش را که بر دبی تاثیر دارند به عنوان ورودی وارد مدل کرده و مدل ARMAX ایجاد می‌شود. باقیمانده‌های مدل نیز در این قسمت مورد بررسی قرار گرفت که شکل ۶ نتیجه حاصل از این بررسی را نشان می‌دهد.

Normal P-P Plot of Noise residual from VAR00002-Model_1



شکل (۶): بررسی نرمال بودن باقیمانده‌های مدل ARMAX

سپس بررسی فرض استقلال باقیمانده‌ها مورد آزمون قرار گرفت که با توجه به نمودار توابع خود همبستگی باقیمانده‌ها در دو مدل، هیچ یک از تاخیرها تفاوت معنی‌داری با صفر نداشتند و فرضیه استقلال باقیمانده‌ها پذیرفته شد. بنابراین هر دو مدل برازش داده شده صحیح می‌باشند. در نهایت به مقایسه آماره‌های برازش دو مدل پرداخته می‌شود که در جدول ۴ نتایج حاصل از برازش این دو مدل ارائه گردیده است.

جدول(۴): نتایج برآورد مدل‌های سری زمانی برازش داده شده

آماره‌های برازش	مدل ARMAX	مدل $ARIMA(2,1,4)$
R-squared	۰.۷	۰.۶۶
RMSE	۶.۰۰۲	۶.۴۶
MAE	۱.۵۵	۱.۶۱
Normalized BIC	۳.۵۹	۳.۷۳

همان طور که دیده می شود ضریب تعیین مدل ARMAX بیشتر، مقدار RMSE و Normalized BIC کمتر می باشد که نشان دهنده ارجحیت و مناسب بودن مدل می باشد.

نتیجه گیری:

پیش بینی جریان رودخانه جهت مدیریت حوضه های آبخیز و به ویژه در حوضه های آبخیز شهری از اهمیت فوق العاده ای برخوردار است. در این مناطق وقوع پدیده های حدی هیدرولوژیک نظیر سیلابها می تواند علاوه بر خسارت های مالی، خسارت های جانی را نیز در پی داشته باشد. بنابراین انجام بررسی و مطالعه در جهت انتخاب مدل هایی که بتوانند پیش بینی های دقیق از پارامتر مورد نظر (دبی جریان) ارائه دهند از نکات کلیدی در آبخیزداری شهری است. در این مطالعه مدل های سری زمانی ARIMA و ARMAX جهت پیش بینی جریان انتخاب گردیدند، نتایج نشان داد که در مدل ARMAX به دلیل اینکه از پارامترهای دیگری نیز که با پارامتر مورد بررسی همبستگی دارند استفاده می شود، بنابراین نسبت به مدل آریمای دقت بیشتری در پیش بینی جریان داشته است. توصیه می شود در مطالعات آینده نتیجه پیش بینی این مدل در سناریوهای مختلف که مربوط به بازه های زمانی مختلف است با پیش بینی مدل های دیگر از جمله روش هوش مصنوعی مورد مقایسه قرار گیرد.

منابع:

- ۱- بشری، م.، وفاخواه، م. ۱۳۸۹. مقایسه روش های مختلف تحلیل سری های زمانی در پیش بینی دبی ماهانه حوزه آبخیز کرخه، فصلنامه علمی پژوهشی مهندسی آب و آبیاری، سال اول، شماره ۲.
- ۲- جهانبخش، س.، ساری صراف، ب.، فاخری فرد، ا.، میر موسوی، ح. ۱۳۸۵. کاربرد مدل های تابع انتقال خطی در بررسی نوسانات پارامترهای اقلیمی (مطالعه موردی ایستگاه تبریز)، مجله پژوهشی دانشگاه اصفهان (علوم انسانی)، جلد بیستم، شماره ۱، صص ۹۲-۷۵
- ۳- دودانگه، ا.، سلطانی کوپائی، س. ۱۳۹۰. بررسی کارایی مدل های سری زمانی وینترز و آریمای در پیش بینی جریان رودخانه (مطالعه موردی: ایستگاه پلدختر)، مجموعه مقالات هفتمین همایش ملی علوم و مهندسی آبخیزداری، دانشکده منابع طبیعی دانشگاه اصفهان.
- ۴- دودانگه، ا.، عابدی کوپائی، ج.، گوهری، ع. ۱۳۹۱. کاربرد مدل های سری زمانی به منظور تعیین روند پارامترهای اقلیمی در آینده در راستای مدیریت منابع آب، مجله علوم و فنون کشاورزی و منابع طبیعی علوم آب و خاک، سال شانزدهم، ۵۹: ۴۷-۵۹.
- ۵- فولادوند، ح. ۱۳۸۹. پیش بینی ماهانه تبخیر-تعرق پتانسیل گیاه مرجع در استان فارس، مجله دانش آب و خاک، جلد ۲۰/۱: ۱۶۹-۱۵۷
- ۶- مهدوی، م. ۱۳۸۵. هیدرولوژی کاربردی، دانشگاه تهران، ۳۴۲ ص
- ۷- نیرومند، ح.، بزرگنیا، ا. ۱۳۸۷. سری های زمانی، انتشارات دانشگاه پیام نور، ۱۶۴ ص

8-Damle, C. and A. Yalcin. 2007. Flood prediction using time series data mining. Journal of Hydrology. 333, (2-4): 305-316.

9- Hasanah, Y., M. Herlina, et al. 2013. "Flood Prediction using Transfer Function Model of Rainfall and Water Discharge Approach in Katulampa Dam." Procedia Environmental Sciences 17(0): 317-326.

10- Komornik, J. Komornikova, M. Mesiar, R. Szokeova, D. and J. Szolgay. 2006. Comparison of forecasting performance of nonlinear models of hydrological time series Physics and Chemistry of the Earth, 31: 1127-11