



A Comparison among the Performance of the Stochastic Models in Generating the Monthly Streamflow and Rainfall Data

M. Montaseri^{1*} and J. Heydari²

Abstract

Synthetic data generation models have been recognized as useful tools to predict and generate alternative time series or long-term series throughout the studies conducted in the domain of water resources management. Accordingly, these models have widely been used by different researchers across the world. In the recent decades, these models have been developed to generate annual, monthly, and daily rainfall or river flow data. Among the synthetic data generated, monthly data are of great importance since they are used in the critical and important studies in the field of water resource systems, such as storage tanks and drought monitoring. Accordingly, the utilization of the monthly synthetic data models leads to more detailed analyses about the real performance of such systems. On the other hand, the theoretical basis of different stochastic models is the generation of diverse monthly data and the performance of these models can remarkably be affected by this fact. Therefore, one can argue that selecting an appropriate model is one of the major concerns of water resources experts. As such, this study made use of the Monte Carlo simulation method to compare and assess the performance of four types of non-parametric Bootstrap models as well as parametric models of Valencia-Schaake, Thomas-Fiering, and Fragment in monthly synthetic data generation. To do this, the monthly flow data of Nazluchay, Shaharchay, and Barandozchay rivers, located at the Western Azerbaijan province in the North West of Iran, were analyzed over a 47-year period. Then, 1000 synthetic time series of monthly flows for these rivers were generated and used for each of the desired seven models over a 47-year period thereof. The results indicated that, compared to other models, the Valencia-Schaake distribution model had a very high performance in terms of all well-known assessment statistical indicators.

Keywords: Synthetic data generation, Non-parametric models, Parametric models, Bootstrap, Valencia and Schaake, Fragment, Thomas-Fiering, Hydrological system, Storage systems, Synthetic time series.

Received: January 13, 2015

Accepted: August 5, 2015

مقایسه عملکرد مدل‌های استوکاستیک در تولید داده‌های ماهیانه جریان و بارندگی

مجید منتصری^{۱*} و جواد حیدری^۲

چکیده

مدل‌های تولید داده مصنوعی به عنوان ابزار مناسبی جهت پیش‌بینی و تولید سری‌های زمانی جانشین یا یک سری بسیار طولانی مدت در مطالعات منابع آب شناخته شده و این مدل‌ها به طور وسیعی در سراسر دنیا توسط محققین مختلف مورد استفاده قرار می‌گیرد. این مدل‌ها برای تولید داده‌های مصنوعی سالیانه، ماهیانه و روزانه بارندگی یا جریان رودخانه‌ای در طی دهه‌های اخیر توسعه داده شده‌اند. در این میان مدل‌های تولید داده مصنوعی ماهیانه به دلیل استفاده در مطالعات مهم و اساسی سیستم‌های منابع آب از جمله سیستم مخازن ذخیره، پایش خشکسالی و غیره دارای اهمیت خاصی بوده و استفاده از آنها تحلیل دقیق‌تر از رفتار عملکرد واقعی چنین سیستم‌هایی را فراهم می‌کند. از طرف دیگر، پایه نظری مدل‌های استوکاستیک مختلف تولید داده‌های ماهیانه متفاوت بوده و این امر می‌تواند اثر قابل توجهی در عملکرد آنها داشته باشد. لذا انتخاب یک مدل مناسب‌تر در مطالعه و تحلیل دقیق و درست یک سیستم منابع آب یکی از دغدغه‌های مهم متخصصین منابع آب می‌باشد. در این مطالعه با استفاده از روش شبیه سازی مونت کارلو، عملکرد چهار تیپ از مدل غیرپارامتریک Bootstrap و مدل‌های پارامتریک Valencia-Schaake، Thomas-Fiering و Fragment در تولید داده‌های ماهیانه مصنوعی مورد مقایسه و ارزیابی قرار گرفته است. بدین منظور داده‌های جریان ماهیانه ۴۷ ساله سه رودخانه نازلوچای، شهرچای و باراندوزچای واقع در استان آذربایجان غربی در شمال غرب ایران مورد استفاده قرار گرفته و ۱۰۰۰ سری زمانی مصنوعی جریان‌های ماهیانه ۴۷ ساله برای هر سه رودخانه مذکور به ازاء هر کدام از مدل‌های هفتگانه مورد مطالعه، تولید و مورد استفاده قرار گرفته است. نتایج مطالعه نشان داد که مدل توزیعی Valencia-Schaake دارای عملکرد بسیار مطلوب نسبت به سایر مدل‌ها به ازاء تمامی آماره‌های مطرح ارزیابی است.

کلمات کلیدی: مدل استوکاستیک، مدل غیر پارامتریک، مدل پارامتریک، Bootstrap، Valencia and Schaake، Fragment، Thomas-Fiering، سیستم هیدرولوژیک، سیستم ذخیره، سری‌های زمانی مصنوعی

تاریخ دریافت مقاله: ۹۳/۱۰/۲۳

تاریخ پذیرش مقاله: ۹۴/۵/۱۴

1- Associate Professor, Department of Water Engineering, Urmia University, Urmia, Iran, Email: montaseri@hotmail.com

2- Department of Water Engineering, Urmia University, Urmia, Iran

*- Corresponding Author

۱- دانشیار گروه مهندسی آب، دانشگاه ارومیه.

۲- کارشناس ارشد سازه های آبی، دانشگاه ارومیه.

*- نویسنده مسئول

بالا و پیچیده را تشکیل می‌دهد که برای تولید مصنوعی سری‌های جریان رودخانه و بارندگی در دهه ۶۰ و به بعد توسعه داده شدند (Brittan, 1961). Barnes (1954) ایده مدل ریاضی یا پارامتریک تولید مصنوعی جریان رودخانه را معرفی کرد که در آن از جدول اعداد تصادفی نرمال برای انتخاب جریان‌های مصنوعی با توزیع نرمال و دارای میانگین و واریانس یکسان با سری‌های تاریخی استفاده کرد و این سری‌ها را برای آنالیز حجم مخزن به کار برد. روش Barnes به سرعت برای دخالت دادن همبستگی بین جریان‌های متوالی سالیانه یا ماهیانه توسعه داده شد و این روش‌های توسعه داده شده تحت عنوان مدل‌های اتورگرسیو مصطلح شدند (Yevjevich, 1972).

با گسترش کامپیوتر از دهه ۶۰ به بعد و با مبنا قرار دادن زنجیر مارکف (۱۹۲۲) مدل‌های پارامتریک از جمله مدل‌های اتورگرسیو و سایر مدل‌های دیگر هرچه بیشتر تکمیل و توسعه پیدا کردند. پارامترهای این قبیل مدل‌ها باید با استفاده از داده‌های تاریخی ثبت شده که منبع اصلی تأثیرات عوامل نامشخص در روش‌های پارامتریک است، برآورد شود (Yevjevich, 1972).

مدل‌های تولید داده مصنوعی در دو کلاس کلی غیر پارامتریک^۲ و پارامتریک^۳ طبقه‌بندی می‌شوند. در روش‌های غیر پارامتریک داده‌های مصنوعی مستقیماً به وسیله نمونه‌گیری مجدد از داده‌های ثبت شده تاریخی تولید می‌شوند (Salas et al., 1980). مدل‌های پارامتریک از یک رابطه ریاضی که معمولاً یک رابطه خطی بوده تبعیت کرده و عمدتاً از مدل‌های غیرپارامتریک بسیار پیچیده‌تر هستند به همین جهت اکثر محققین ترجیح می‌دهند از روش‌های غیر پارامتریک که ساختار تجربی داده‌های مشاهداتی را حفظ می‌کنند، استفاده کنند. در حالی که استفاده از مدل‌های پارامتریک نیازمند فرضیاتی در مورد توزیع احتمالی حاشیه‌ای و تغییرات ساختاری مکانی و زمانی متغیرها می‌باشد (Vogel and Shallcross, 1996).

مدل‌های استوکاستیک به طور وسیع در مطالعات و تحلیل سیستم‌های منابع آب مورد استفاده قرار می‌گیرد. این در حالی است که تحقیقات محدودی جهت مقایسه عملکرد مدل‌های مذکور در دسترس بوده و از طرف دیگر این امر با تعداد کم و از جنبه‌های خاص انجام پذیرفته است که ذیلاً به موارد عمده و مهم آن پرداخته شده است.

منابع آب و محدودیت‌های آن یکی از چالش‌های بزرگ در چشم‌انداز پیش روی بشر است. 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 (Chow et al., 1993). اکثر مطالعات هیدرولوژی بر پایه داده‌های در دسترس هیدرولوژیک از قبیل بارندگی، دما، رطوبت، شدت جریان‌های آبراه‌های و... است. در مدل کردن رفتار سیستم‌های مختلف هیدرولوژی مانند حوضه‌های آبریز، رودخانه‌ها و... هرچه داده‌ها بیشتر باشد مدل به شرایط طبیعی نزدیکتر است و استفاده از آن در مسایل منابع آب منجر به نتایج دقیق‌تر و درست‌تر می‌شود. اما داده‌های ثبت شده (مشاهداتی) برای پدیده‌های هیدرولوژیکی ضمن اینکه معمولاً دارای دوره آماری بسیار کوتاه هستند (۲۰-۳۰ سال)، به عنوان یک نمونه آماری مطرح بوده که این نمونه آماری نمی‌تواند معرفی دقیق برای یک سیستم هیدرولوژیک با عمری چندین هزار ساله باشد (Montaseri and Adeloye, 1999).

مدل‌های تولید داده مصنوعی امکان تولید سری‌های جانشین داده‌های هیدرولوژیک که رخداد آنها در آینده محتمل است را فراهم می‌کنند که منجر به نتایج دقیق و واقعی‌تر در مطالعات منابع آب می‌گردد (Srikanthan and McMahon, 2001). استفاده از تکنیک‌های دوباره جایگزینی^۱ داده‌های تاریخی برای تولید سری‌های مصنوعی به عنوان اولین تلاش‌ها در این زمینه بوده که توسط محققین مختلف مورد استفاده قرار گرفت. به عنوان مثال Hazen (1914) جریان‌های سالیانه استاندارد شده ۱۴ رودخانه را برای ایجاد ۳۰۰ سال داده مصنوعی ترکیب کرد. مطالعه او به عنوان اولین تجربه در هیدرولوژی به منظور کاربرد جریان مصنوعی در سیستم منابع آب مورد توجه است. Solder (1927) از مقادیر پنجاه سال جریان سالیانه رودخانه که در داخل ۵۰ کارت قرار داده شده بود استفاده کرد. این کارت‌ها را به صورت تصادفی و بدون جایگزینی مجدد انتخاب کرد تا اینکه ۲۰ سری جریان سالیانه به طول ۵۰ سال را تولید کرد و سپس این سری‌ها را برای ایجاد داده‌های جریان سالیانه مصنوعی به طول ۱۰۰۰ سال، به منظور برآورد کردن ظرفیت حجم مخزن مورد نیاز، با یکدیگر ترکیب کرد. مشکل اصلی در روش Solder این بود که هر یک از ۲۰ سری محتوی یک مجموعه یکسان از مقادیر جریان می‌باشد، به طوری که این عمل در طبیعت غیرمحتمل است. مشکل دیگر این است که سری‌های متعدد با طول‌های بزرگتر از طول داده‌های تاریخی ثبت شده نمی‌توان تولید کرد. با این وجود، این روش‌های اولیه، اساس بیشتر مدل‌های سطح

آخر سال قبل حفظ نشده و این محدودیت را میتوان با شروع سال هیدرولوژی با ماهی که داده‌های آن ماه دارای کمترین ضریب همبستگی با ماه ماقبل خود است کاهش داد.

Fragment (1996) با استفاده از مدل Maheepala and Perera (1996) سری‌های مصنوعی ماهیانه را تولید و نتایج به دست آمده را با داده‌های تاریخی مقایسه کردند. ایشان اظهار داشتند که مدل مذکور، به غیر از همبستگی داخلی^۶ و بیرونی^{۱۰} با ضریب تأخیر یک بقیه مشخصات آماری داده‌های تاریخی را در حد خوبی حفظ می‌کنند.

مطالعات بسیار متعددی در دهه‌های اخیر با استفاده از مدل‌های مختلف تولید داده‌های ماهیانه مصنوعی در زمینه‌های مختلف مهندسی آب از جمله سیستم مخازن ذخیره و پایش خشکسالی انجام پذیرفته است. تمامی مطالعات انجام یافته با تأکید بر مزیت استفاده از مدل‌های مذکور برای ارزیابی دقیق و واقعی‌تر از پدیده مورد مطالعه بوده است. با این وجود مزیت استفاده از مدل‌های استوکاستیک ماهیانه تابعی از مدل انتخابی می‌باشد. لذا مشخصات تنوریک و عملکرد مدل در حفظ کاملتر و دقیق‌تر پارامتری آماری داده‌های تاریخی نقش بسیار مهمی در تحلیل‌های استوکاستیک منابع آب را دارند. از طرف دیگر تحقیقات محدودی برای مقایسه عملکرد مدل‌های مختلف تولید داده‌های ماهیانه در منابع علمی و پژوهشی معتبر داخلی و خارجی گزارش گردیده و اکثریت این تحقیقات نیز تعداد محدودتری از مدل‌ها را شامل شده و از جنبه‌های خاص و محدود انجام پذیرفته است. لذا در این تحقیق عملکرد هفت مدل تولید داده مصنوعی که بطور گسترده در مطالعات مختلف منابع آب مورد استفاده قرار می‌گیرند، با استفاده از معیارهای آماری مطرح (پارامترهای آماری، ساختمان همبستگی داده‌های ماهیانه، ساختمان همبستگی مابین داده‌های ماهیانه و سالیانه، منحنی فرکانس، متوسط رخدادهای حدی، ضریب Hurst) مورد بررسی و ارزیابی قرار گرفته است. بدین منظور هفت مدل تولید داده‌های ماهیانه مذکور در سه رودخانه اصلی غرب دریاچه ارومیه جهت تولید ۱۰۰۰ سری ۴۷ ساله از داده‌های ماهیانه جریان برای ارزیابی جامع و دقیق‌تر مشخصات تنوریک و ذاتی مدل‌ها مورد استفاده قرار گرفته است.

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

در این مطالعه از داده‌های دبی ماهیانه و سالیانه سه رودخانه نازلوچای، شهرچای و باراندوزچای واقع در غرب دریاچه ارومیه و شمال غرب ایران به ترتیب در ایستگاه‌های هیدرومتری تپیک، میرآباد و هاشم آباد به عنوان ایستگاه‌های مبناء استفاده شده است.

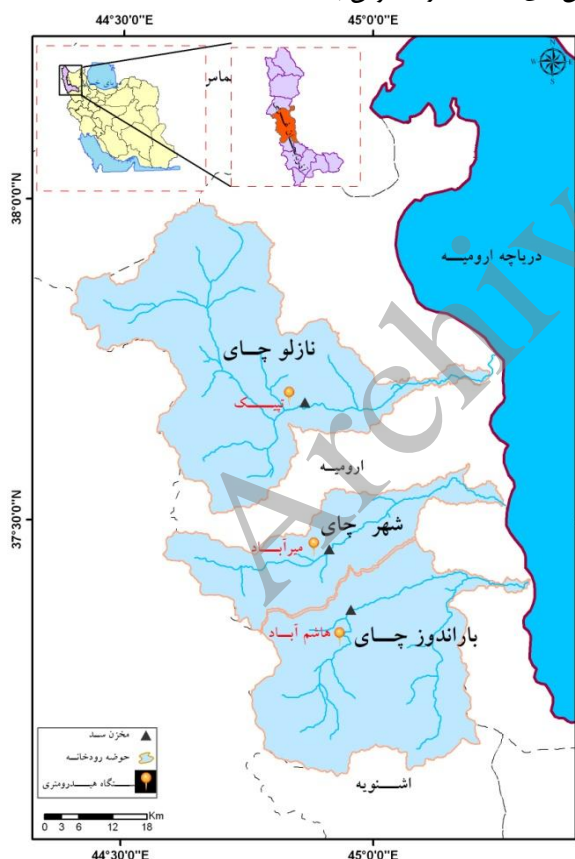
(1996) Vogel and Shallcross به منظور تلاش برای تولید مجدد ضریب همبستگی با تأخیر یک^۴ در سری‌های مصنوعی تولید شده با استفاده از تکنیک Bootstrap مدل Moving Block Bootstrap را آزمایش کردند. ایشان به این نتیجه رسیدند که روش Moving Block Bootstrap با بلوک‌های متحرک ۲ ساله، در مقایسه با روش Bootstrap ساده اریب کمتری دارد، هرچند اریب حاصل از این روش نسبت به نتایج گرفته شده از مدل پارامتریک AR(1) بزرگ بود.

(2001) Srikanthan and McMahon مدل‌هایی که پایداری دراز مدت^۵ و کوتاه مدت^۶ را حفظ می‌کنند را مورد امتحان و آزمایش قرار دادند و به این نتیجه رسیدند که مدل‌های ساده سالیانه از قبیل مدل AR(1) به همراه مدل‌های توزیعی یا یخشی، برای تولید جریان ماهیانه به منظور شبیه سازی مخزن دقیقتر هستند زیرا این مدل‌ها برای سری جریان‌های سالیانه با طول مناسب می‌توانند پایداری کوتاه مدت و دراز مدت را به طور منطقی و مناسب مجدداً تولید کنند.

(1999) Montaseri and Adeloye در مطالعات خود مدل AR(1) را به دو صورت پایه و با اعمال پارامتر عدم قطعیت در برآورد پارامترهای مدل مورد استفاده قرار دادند، و به این نتیجه رسیدند که در حالت دوم ضریب تغییرات داده‌های تولید شده اندکی بیشتر از ضریب تغییرات داده‌های مشاهداتی می‌باشد که این نکته بیانگر عملکرد منطقی و واقعی‌تر این حالت مدل است. (2004) Nurul et al. با استفاده از مدل‌های مستقیم^۷ و توزیعی^۸ اقدام به تولید داده‌های ماهیانه مصنوعی دبی رودخانه و بارندگی کرده و نتایج به دست آمده را با داده‌های تاریخی مقایسه و به این نتیجه رسیدند که مدل Valencia-Schaake به عنوان یکی از بهترین مدل‌های تولید داده مصنوعی می‌باشند. (2007) Debele با استفاده از مدل‌های Disaggregation و MuDRain با تولید داده‌های دبی ماهیانه و داده‌های ۲۴ ساعته بارندگی، دما و باد به این نتیجه رسید که مدل‌های Disaggregation برای تولید داده‌های ماهیانه و مدل MuDRain برای تولید داده‌های روزانه مناسب می‌باشند.

(1985) Srikanthan and McMahon با استفاده از مدل Fragment اقدام به تولید داده‌های ماهیانه بارندگی کردند. در این تحقیق سری داده‌های تولیدی جریان سالیانه توسط روش مذکور به ماه‌های مختلف سال به‌عنوان داده‌های ماهیانه توزیع گردید. آنها به این نتیجه رسیدند که ضریب همبستگی بین ماه اول سال با ماه

تاریخی وجود ندارد. در حالیکه این امر دارای احتمال بسیار ضعیفی بوده و تردیدهای جدی نسبت به نتایج چنین مطالعاتی وجود خواهد داشت (Chow et al., 1993). برای رفع این نقیصه، محققین ضرورت استفاده از روش شبیه سازی مونت کارلو را مطرح نموده‌اند (Montasari and Adeloeye, 2002). در روش شبیه سازی مونت کارلو، سری داده‌های مصنوعی دبی یا بارندگی در نمونه‌های بسیار مثلاً ۱۰۰۰ سری، به عنوان سری داده‌های جانشین با استفاده از یک مدل استوکاستیک مناسب تولید می‌شود. چنین سری‌های تولیدی ویژگی‌های رخدادهای هیدرولوژیکی مانند دبی و بارندگی که احتمال رخداد آنها در آینده وجود دارد، به طور واقعی‌تر توصیف می‌کنند و شبیه‌سازی سیستم‌های مختلف منابع آب براساس چنین سری‌های زمانی مصنوعی منجر به تحلیل درست و واقعی‌تر از عملکرد سیستم مورد نظر خواهد بود (Salas, 1993). بنابراین انتخاب و به کارگیری یک مدل استوکاستیک تولید داده مناسب دارای اهمیت بسیار بالایی بوده و این امر برای موقعی امکان پذیر خواهد بود که بر پایه نتایج یک تحقیق جامع و کامل، ارزیابی خصوصیات تئوریک و عملکرد مدل‌های مختلف در دسترس باشد.



شکل ۱- موقعیت منطقه مورد مطالعه در ایران و استان آذربایجان غربی

محدوده حوضه آبریز و موقعیت ایستگاه‌های هیدرومتری سه گانه مذکور در شکل ۱ نشان داده شده است. همچنین مشخصات عمومی ایستگاه‌های هیدرومتری و پارامترهای آماری اصلی داده‌های سالیانه و ماهیانه به ترتیب در جدول ۱ و ۲ ارائه گردیده است.

داده‌های هیدرولوژی به عنوان متغیرهای تصادفی برای استفاده در تحلیل‌های مختلف سیستم‌های منابع آب بایستی از نظر همگنی، تصادفی و ایستایی بودن مورد بررسی و آزمون قرار گیرند (Adeloeye and Montasari, 2002). بدین منظور در این مطالعه از روش منحنی جرم مضاعف برای آزمون همگنی داده‌ها استفاده گردید، که نتایج نشان دهنده همگنی داده‌ها با ضریب همبستگی ۰/۹۹ می‌باشد (Bars, 1990). برای آزمون ایستایی از روش همبستگی اسپیرمن (Kendall and Stuart, 1976) و برای آزمون تصادفی بودن داده‌ها نیز از روش ران تست استفاده گردید (Adeloeye and Montasari, 2002; McGhee, 1985). نتایج آزمون فوق درستی و صحت فرضیه تصادفی و ایستایی بودن داده‌های مورد مطالعه را تأیید می‌کند. نتایج سه آزمون مذکور به دلیل کمبود جا ارائه نشده است.

یک گام اساسی در تولید داده مصنوعی برای شبیه سازی مونت کارلو، انتخاب توزیع آماری مناسب داده‌های مورد استفاده می‌باشد. بدین منظور در این مطالعه برای تعیین توزیع آماری مناسب داده‌های دبی سالیانه و ماهیانه از روش (PPCC)^{۱۱} به عنوان مناسب‌ترین روش تعیین توزیع احتمالی برتر به ازای پنج توزیع احتمالی نرمال، لاگ نرمال دو پارامتری، لاگ نرمال سه پارامتری، پیرسون تیپ سوم و لاگ پیرسون تیپ سوم استفاده شده است (Vogel and Kroll, 1989). نتایج آزمون PPCC نشان می‌دهد که توزیع پیرسون تیپ سوم به عنوان توزیع برتر داده‌های ماهیانه و سالیانه در رودخانه‌های مورد مطالعه است. برازش گرافیکی داده‌های سالیانه رودخانه‌های سه گانه نازلو، شهر و باراندوز با توزیع پیرسون نوع سوم به عنوان توزیع برتر در شکل ۲ نشان داده شده است.

۲-۲- شبیه سازی مونت کارلو

مطالعات سیستم‌های مختلف منابع آب از جمله سیستم مخازن ذخیره، پایش خشکسالی و بررسی تغییرات اقلیم با براساس داده‌های تاریخی (ثبت شده) و یا با استفاده از شبیه سازی مونت کارلو انجام می‌پذیرد (Douglas, 2010). در حالت اول، تحلیل سیستم آبی مورد نظر با فرض تکرار داده‌های تاریخی در آینده یا دوره بهره‌برداری بوده و احتمال رخداد خشکسالی‌های شدیدتر از خشکسالی‌های دوره

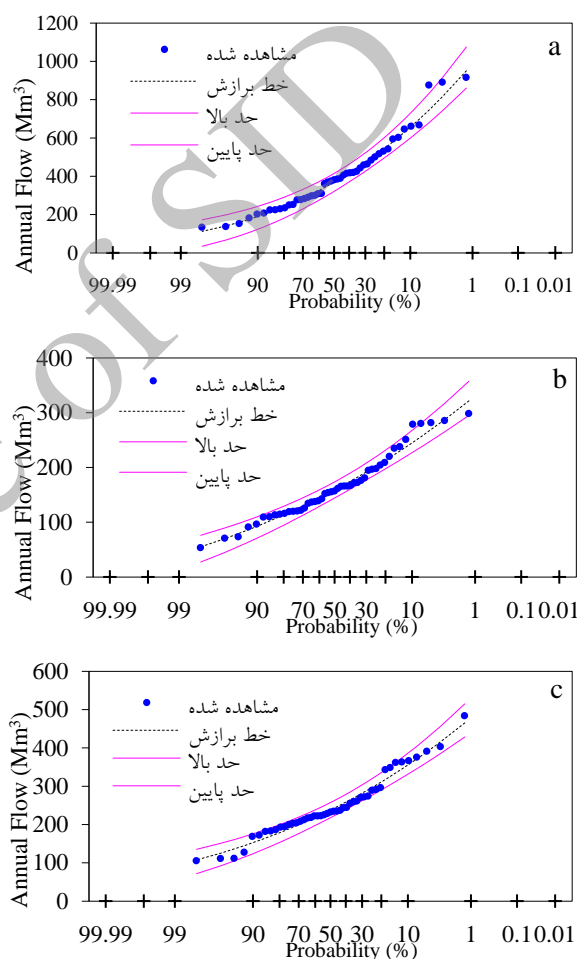
۳-۲- مدل‌های تولید داده مصنوعی

مدل‌های تولید داده‌های مصنوعی در دو گروه عمده پارامتریک^{۱۲} و غیرپارامتریک^{۱۳} طبقه‌بندی می‌شوند. در مدل‌های غیر پارامتریک، داده‌های مصنوعی مستقیماً به وسیله نمونه‌گیری مجدد^{۱۴} از داده‌های ثبت شده تاریخی تولید می‌شوند. اما مدل‌های پارامتریک مشتمل بر رابطه ریاضی بوده که اغلب به وسیله یک رابطه خطی تعریف می‌شوند. اکثریت این مدل‌ها برای داده‌های دارای توزیع نرمال توسعه یافته و داده‌های دارای توزیع‌های آماری دیگر بایستی ابتدا به توزیع نرمال انتقال یافته و سپس در مدل بندی مورد استفاده قرار گیرند.

مدل غیر پارامتریک Bootstrap به صورت تک^{۱۵} و یا دوبل بلوکی^{۱۶} به طور گسترده‌ای برای تولید داده‌های ماهیانه در تحقیقات متعدد مورد استفاده قرار گرفته است (Srinivas and Srinivasan, 2005). این مدل دارای مزیت‌های اساسی از نظر طول دوره آماری، تعداد سری، مستقل بودن سری‌های تولیدی از همدیگر نسبت به مدل Moving window بوده که این امر در تحلیل درست سیستم‌های مختلف منابع آب دارای اهمیت به سزایی است. از طرف دیگر مدل‌های پارامتریک برای تولید داده‌های ماهیانه در دو روش مستقیم^{۱۷} و پخشی^{۱۸} کلاس بندی شده است. در روش اول داده‌های ماهیانه به طور مستقیم توسط مدل تولید می‌شود. در حالی که در روش پخشی، ابتدا داده‌های سالیانه توسط یک مدل مناسب مانند AR(1) تولید و سپس داده‌های سالیانه تولید شده توسط مدل‌های پخشی مانند Fragment, Proration یا Valencia-Schaake به ماه‌های مختلف سال توزیع می‌شوند. در این مطالعه چهار مدل رایج تولید داده‌های مصنوعی ماهیانه، Thomas-Fiering، Valencia-Schaake، Fragment و Bootstrap در چهار تیپ (Single and double block for with and without replacement) که به طور وسیعی در مطالعات استوکاستیک یا شبیه‌سازی مونت کارلو سیستم‌های مختلف منابع آب به کار می‌رود، جهت مقایسه و ارزیابی خصوصیات تئوری، ذاتی و عملکرد آنها در تولید داده‌های ماهیانه جریان یا بارندگی مورد استفاده قرار گرفته است. مشخصات مدل‌های چهارگانه مورد مطالعه ذیلاً به طور اختصار ارائه شده است.

۳-۲-۱- مدل Bootstrap (B.S)

این روش غیر پارامتریک تولید سری‌های زمانی مصنوعی، براساس روش آماری Bootstrap می‌باشد. این روش اغلب برای تست آماری یا برای ارزیابی تغییرپذیری برآورد نقطه‌ای (میانگین، واریانس، ...) یک نمونه، که جمعیت آماری اصلی آن مشخص نیست، استفاده می‌شود.



شکل ۲- برازش گرافیکی داده‌های سالیانه رودخانه‌های نازلو چای (a)، شهرچای (b) و باراندوزچای (c) با توزیع پیرسون تیپ سوم.

لذا تعیین نوع تابع توزیع احتمالی داده‌های تاریخی و نرمالیزه کردن (انتقال داده‌ها از توزیع مربوط به توزیع نرمال) قدم اول به شمار می‌رود. پارامترهای این مدل‌ها با استفاده از داده‌های تاریخی نرمالیزه برآورد می‌شوند (Maheepala and Perera, 1996). مدل‌های

جدول ۱- مشخصات عمومی ایستگاه‌های مبنا و خصوصیات آماری دبی‌های سالیانه تاریخی رودخانه‌های مورد مطالعه.

علامت اختصار	نام رودخانه	نام ایستگاه	مختصات جغرافیایی		پارامترهای آماری داده‌های دبی سالیانه (۱۳۸۶-۱۳۴۰)				
			طول جغرافیایی	عرض جغرافیایی	میانگین (Mm ³)	انحراف معیار	چولگی	ضریب تغییرات	توزیع برتر
S1	نازلو	تیپک	۴۴° ۵۴'	۳۷° ۴۰'	۴۰۲/۵	۱۹۰/۸	۱/۰۲	۰/۴۷	پیرسون تیپ سوم
S2	شهر	میر آباد	۴۴° ۵۲'	۳۷° ۲۶'	۱۶۵/۲	۶۰/۳	۰/۵۷	۰/۳۶	پیرسون تیپ سوم
S3	باراندوز	هاشم آباد	۴۴° ۵۴'	۳۷° ۱۷'	۲۴۶/۸	۷۷/۸	۰/۵۱	۰/۳۲	پیرسون تیپ سوم

جدول ۲- خصوصیات آماری دبی‌های ماهیانه رودخانه‌های مورد مطالعه در ایستگاه‌های مبنا (مترمکعب بر ثانیه)

رودخانه	پارامتر	مهر	آبان	آذر	دی	بهمن	اسفند	فروردین	اردیبهشت	خرداد	تیر	مرداد	شهریور
نازلوچای	میانگین	۹/۶۶	۱۴/۵۷	۱۳/۶۴	۱۳/۱۰	۱۳/۸۲	۲۱/۶۱	۶۵/۹۱	۱۲۱/۵۸	۸۰/۹۸	۲۹/۹۰	۱۰/۷۸	۷/۰۹
	CV	۰/۶۹	۰/۵۶	۰/۴۶	۰/۴۱	۰/۳۹	۰/۴۵	۰/۴۸	۰/۴۸	۰/۶۶	۰/۸۶	۰/۸۷	۰/۸۷
	چولگی	۰/۷۶	۰/۶۴	۰/۸۶	۱/۰۱	۰/۸۴	۱/۰۰	۱/۰۶	۰/۷۶	۱/۱۲	۱/۵۵	۱/۴۰	۱/۴۵
	ماکزیمم	۲۴/۷۶	۳۷/۸۸	۳۱/۹۱	۲۹/۶۵	۲۹/۸۱	۵۴/۷۵	۱۴۷/۳۹	۲۷۳/۶۳	۲۳۵/۶۲	۱۰۱/۳	۳۹/۸۰	۲۶/۱۴
	مینیمم	۱/۰۴	۱/۲۲	۳/۰۰	۴/۸۵	۳/۲۶	۵/۹۹	۲۴/۹۴	۳۵/۱۷	۱۳/۰۴	۱/۶۹	۰/۵۱	۰/۲۷
باراندوزچای	میانگین	۶/۶۶	۱۰/۴۴	۹/۹۶	۸/۵۰	۸/۲۱	۱۲/۶۸	۳۴/۶۴	۶۰/۵۷	۵۱/۵۲	۲۶/۷۴	۱۰/۷۷	۶/۱۸
	CV	۰/۴۱	۰/۶۰	۰/۴۱	۰/۳۷	۰/۳۳	۰/۳۹	۰/۳۸	۰/۳۴	۰/۴۳	۰/۵۶	۰/۵۹	۰/۴۷
	چولگی	۱/۳۶	۱/۶۸	۱/۹۳	۲/۲۳	۱/۳۵	۰/۹۸	۱/۰۴	۰/۶۶	۰/۴۶	۰/۸۷	۱/۱۷	۰/۷۴
	ماکزیمم	۱۵/۱۵	۳۰/۹۵	۲۶/۰۲	۲۰/۷۶	۱۶/۹۵	۲۶/۶۰	۷۶/۱۹	۱۱۲/۹۹	۱۰۴/۰۶	۶۸/۹۴	۳۰/۸۸	۱۲/۹۷
	مینیمم	۲/۹۰	۳/۵۵	۵/۱۳	۴/۳۰	۴/۹۵	۶/۰۶	۱۶/۸۷	۲۲/۳۶	۹/۴۳	۳/۵۹	۰/۷۸	۱/۱۲
شهرچای	میانگین	۲/۳۴	۴/۰۷	۳/۹۲	۳/۷۹	۳/۶۴	۵/۳۷	۲۰/۴۳	۴۷/۱۸	۴۶/۴۴	۱۸/۶۴	۶/۴۲	۲/۹۵
	CV	۰/۵۲	۰/۶۴	۰/۵۲	۰/۳۲	۰/۳۰	۰/۳۸	۰/۴۴	۰/۳۵	۰/۵۱	۰/۷۱	۰/۷۶	۰/۸۸
	چولگی	۰/۹۴	۱/۲۳	۲/۳۳	۰/۸۵	۰/۸۴	۰/۷۵	۱/۰۲	-۰/۰۹	۰/۷۳	۱/۴۴	۱/۳۲	۳/۹۳
	ماکزیمم	۵/۴۲	۱۱/۵۱	۱۳/۱۲	۷/۸۵	۷/۱۵	۱۱/۵۳	۴۷/۷۰	۸۲/۲۳	۱۰۷/۷۰	۵۵/۵۸	۲۱/۴۹	۱۷/۱۸
	مینیمم	۰/۶۷	۱/۰۶	۱/۲۴	۱/۴۰	۱/۶۱	۱/۶۴	۳/۴۹	۱۲/۰۵	۱۰/۴۷	۲/۵۲	۰/۶۴	۰/۴۳

مدل‌های تولید داده مصنوعی	غیر پارامتریک	Bootstrap		
		Moving Window		
	پارامتریک	ماهیهانه	مستقیم	Thomas-Fiering
		سالیانه	غیرمستقیم	Proration Method Fragment Valencia-Schaake
		AR(p) ARMA(p,q) FGN BL		

شکل ۳- تقسیم بندی کلی مدل‌های مختلف تولید داده مصنوعی.

استفاده قرار می‌گیرد. در تک بلوکی داده‌های ماه‌های مختلف هر سال به عنوان بلوک تکی در نظر گرفته شده و در سری داده‌های تولیدی به عنوان یک بلوک تولید شده استفاده می‌شود. در صورتی که بلوک به صورت دو ساله باشد و داده‌های ماهیانه دو سال متوالی را شامل گردد آنرا Bootstrap دویل بلوکی می‌نامند.

روش Bootstrap به وسیله نمونه‌گیری مجدد از داده‌های ثبت شده مشاهداتی، با دوباره جایگزینی^{۱۹} و یا بدون جایگزینی^{۲۰}، برای تولید تعداد زیاد نمونه‌های مستقل (N_s) هریک به اندازه n_u سال (n_u اندازه داده‌های اصلی است) به کار می‌رود (Vogel and Shallcross, 1996; Salas, 1993). این مدل به صورت تک و دویل بلوکی برای تولید سری داده‌های ماهیانه مورد

۲-۳-۲- مدل (V.S) Valencia and Schaaque

در این روش ابتدا داده‌های سالیانه با یک مدل مناسب تولید و سپس مقادیر سالیانه در ماه‌های مختلف سال توزیع می‌شود که در این تحقیق از مدل اتورگرسیو با تاخیر یک ساله AR(1) برای تولید داده‌های سالیانه استفاده شده است. این مدل به دلیل سادگی و عملکرد بسیار مطلوب به طور وسیع در سیستم‌های منابع آب برای تولید جریان یا بارندگی سالیانه از سال ۱۹۶۰ به کار می‌رود (McMahon and Mein, 1986). رابطه کاربردی این مدل به صورت رابطه ۱ می‌باشد:

$$Q_{i+1} = \bar{Q} + \rho(Q_i - \bar{Q}) + v_i s \sqrt{1 - \rho^2} \quad [1]$$

که در آن Q_i و Q_{i+1} به ترتیب جریان سالیانه تولید شده مربوط به سال i و $i+1$ ، \bar{Q} میانگین جریان‌های سالیانه مشاهداتی، v_i متغیر تصادفی از توزیع نرمال استاندارد با میانگین صفر و واریانس واحد، ρ ضریب خود همبستگی با تاخیر یک و s انحراف معیار داده‌های مشاهداتی سالیانه می‌باشد.

سپس این داده‌های تولید شده توسط مدل AR(1) با استفاده از مدل والنسیا - شاکي به داده‌های ماهیانه تبدیل می‌گردد. مدل والنسیا - شاکي برای توزیع جریان‌های سالیانه به ماهیانه به صورت رابطه ۲ می‌باشد:

$$X_i = AZ_i + BV_i \quad [2]$$

که در آن X_i یک بردار (12×1) داده‌های ماهیانه، V_i یک بردار (12×1) متغیرهای تصادفی از توزیع نرمال استاندارد می‌باشد که از Z_i مستقل است، داده سالیانه تولید شده، A و B ماتریس ضرایب می‌باشند که ابعاد آنها به ترتیب (12×1) و (12×12) می‌باشد. مقادیر ماتریس ضرایب A و B با استفاده از داده‌های تاریخی ماهیانه و سالیانه برآورد می‌شود (Lin and Lee, 1992).

۲-۳-۳- مدل (Frag) Fragment

این مدل نیز جزء مدل‌های پخشی در تولید داده ماهیانه می‌باشد که توسط Harm and Kambel (1967) ارائه گردید است (Fairing, 1967). سپس Srikanthan and McMahon (1982). این روش را اصلاح و به عنوان روش "Method of Fragment" نامیدند. Porter and Pink (1991) نیز این روش را توسعه داده و تحت عنوان "Method of Synthetic Fragment" مورد استفاده قرار دادند.

در کارکرد این مدل، ابتدا براساس داده‌های تاریخی نسبت داده‌های ماهیانه به جریان سالیانه در همان سال محاسبه می‌شود و سپس مقادیر جدید به دست آمده از تقسیم حجم جریان‌های ماهیانه به جریان سالیانه در همان سال به‌عنوان Fragments شناخته می‌شود. در مرحله بعد داده‌های سالیانه به طور صعودی مرتب شده و Fragment‌های هر سال در این ترتیب جدید حفظ می‌گردد. سپس برای داده‌های سالیانه مشاهداتی کلاس‌بندی در نظر گرفته می‌شود که اولین کلاس، حداقل صفر و آخرین کلاس حداکثرش بی‌نهایت است و محدوده ما بین کلاس‌ها به وسیله متوسط ارقام سالیانه مرتب شده به دست می‌آید. در مرحله بعد، داده‌های جریان سالیانه با استفاده از یک مدل مناسب تولید می‌شود. سپس کلاس داده سالیانه تولیدی به ازای کلاس بندی ذکر شده تعیین و با استفاده از Fragment‌های همان کلاس در ماه‌های مختلف سال توزیع می‌شود. این مدل مشخصه‌های آماری ماهیانه را حفظ کرده و همچنین در این مدل نیازی به تعیین توزیع آماری داده‌های ماهیانه نیست (Porter and Pink, 1991).

۲-۳-۴- مدل (T.F) Thomas and Fiering

Thomas and Fiering (1962) با استفاده از زنجیره مارکوف فرمولی برای تولید جریان‌های ماهیانه به طور مستقیم ارائه کردند که رابطه کاربردی آن به صورت رابطه ۳ می‌باشد:

$$Q_{i+1} = \bar{Q}_{j+1} + b_j(Q_i - \bar{Q}_j) + e_i S_{j+1} \sqrt{1 - r_j^2} \quad [3]$$

که در آن Q_i و Q_{i+1} عبارتست از جریان‌های ماهیانه تولید شده در طی پرپود $i, i+1$ ، \bar{Q}_j و \bar{Q}_{j+1} متوسط جریان ماهیانه در ماه‌های $j, j+1$ و b ضریب همبستگی حداقل مربعات برای محاسبه جریان ماه $j+1$ از جریان در ماه j که از رابطه ۴ به دست می‌آید:

$$b_j = r_j \left[\frac{S_{j+1}}{S_j} \right] \quad [4]$$

که در این رابطه r_j ضریب همبستگی بین جریان‌های ماه‌های $j+1$ و j ، S_j ، S_{j+1} انحراف معیار برای ماه‌های $j, j+1$ و e_i متغیر تصادفی از توزیع نرمال استاندارد با میانگین صفر و انحراف معیار یک می‌باشد (Thomas and Fiering, 1962).

۲-۳-۵- مراحل تحقیق

در این مطالعه داده‌های دبی ماهیانه و سالیانه ۴۷ ساله (از سال ۱۳۴۰ الی ۱۳۸۶) رودخانه‌های نازلوچای، شهرچای و باراندوزچای در ایستگاه‌های هیدرومتری مینا استفاده شده است. در اولین قدم

- ساختمان همبستگی داده‌های ماهیانه یعنی همبستگی بین ماه‌های مختلف داده‌های تاریخی و تولیدی؛

- همبستگی مابین داده‌های ماهیانه در ماه‌های مختلف با داده‌های سالیانه داده‌های تولیدی و تاریخی؛

- منحنی فرکانس^{۳۳} براساس داده‌های تولیدی و تاریخی؛

- متوسط رخدادهای حدی حداقل مانند مقادیر دبی با احتمالات وقوع ۹۰ و ۹۸ درصد داده‌های تاریخی و تولیدی؛

- ضریب Hurst به عنوان شاخص گرایش بلند مدت به ازاء داده‌های تولیدی و تاریخی؛

- ساختمان همبستگی با تاخیر صفر و یکساله ماهیانه و سالیانه مابین ایستگاه‌ها، به ازاء داده‌های تولیدی و تاریخی در تحلیل سیستم‌های چندگانه منابع آب، که در این تحقیق مطرح نمی‌باشد.

در مطالعه حاضر تمامی معیارهای ارزیابی مذکور به استثناء معیار انتهایی که برای تحلیل سیستم‌های چندگانه دارای اهمیت بسیار بالایی می‌باشد، مورد ملاحظه قرار گرفته و ذیلاً برای داده‌های سالیانه و ماهیانه به تفکیک تنظیم و ارائه شده است.

۳-۱- خصوصیات آماری داده‌های سالیانه تولیدی و تاریخی

در این مطالعه داده‌های سالیانه به واسطه مدل $AR(1)$ در روش‌های Fragment و Valencia-Schaake، روش دوباره جاگذاری داده‌های تاریخی در مدل Bootstrap تک و دوبل بلوکی و استنتاج از داده‌های ماهیانه تولیدی در روش Thomas-Fiering حاصل شده است. نتایج حاصل از روش Bootstrap تک بلوکی کاملاً مشابه داده‌های تاریخی به استثناء خود همبستگی با تاخیر یکساله می‌باشد. لذا در این بخش عملکرد سه مدل $AR(1)$ ، Thomas-Fiering و Bootstrap دوبل بلوکی در سطح داده‌های سالیانه مورد ارزیابی قرار گرفته است.

۳-۲- پارامترهای آماری پایه دبی‌های سالیانه

مشخصات مهم آماری داده‌های دبی سالیانه تاریخی (His) و تولیدی سه مدل مذکور در جدول ۳ نشان داده شده است. براساس نتایج مذکور هر سه مدل میانگین، انحراف معیار و ضریب چولگی داده‌های سالیانه را به خوبی حفظ می‌کنند. خود همبستگی با تأخیر یک ساله داده‌های سالیانه نقش اصلی را در تعیین رفتار برون سالی سیستم‌های منابع آب مانند مخازن ذخیره و پایش خشکسالی داشته و لذا حفظ این پارامتر توسط مدل‌های تولید داده دارای اهمیت به‌سزایی می‌باشد. براساس نتایج حاصل در جدول ۳ تنها مدل $AR(1)$ پارامتر خود همبستگی با تأخیر یک ساله داده‌های تاریخی را به نحو مطلوبی حفظ یا دوباره تولید می‌کند. دو مدل بعدی بسیار

همگنی، تصادفی و ایستایی بودن داده‌ها تست و تأیید شده و سپس توزیع احتمالی برتر داده‌های دبی در ایستگاه‌های هیدرومتری مینا یعنی توزیع پیرسون نوع سوم تعیین شده است. در مرحله بعد تمامی داده‌های ماهیانه و سالیانه در هر سه ایستگاه مینا با استفاده از رابطه Wilson-Hilferty از توزیع پیرسون نوع سوم به توزیع نرمال انتقال یافته است. سپس پارامترهای سه مدل پارامتریک با استفاده از داده‌های انتقال یافته به توزیع نرمال برآورد گردیده است. در قدم بعدی با استفاده از چهارمدل $Bootstrap(i, j)$ (که در آن i بیانگر بلوک یکساله (۱) و یا بلوک دوساله (۲) و j بیانگر با جایگزینی (۱) و بدون جایگزینی (۰) می‌باشد)، Thomas-Fiering، Fragment و Valencia-Schaake، ۱۰۰۰ سری داده‌های ماهیانه برابر با ۴۷ سال در توزیع نرمال تولید و سپس به توزیع اصلی خود انتقال داده می‌شوند. در نهایت سری داده‌های ماهیانه تولید شده مدل‌های چهارگانه برای هر سه رودخانه مورد مطالعه برای تعیین مشخصات آماری موثر جهت ارزیابی خصوصیات تئوری، ذاتی مدل‌ها و عملکرد آنها در حفظ مشخصات آماری دبی‌های ماهیانه و سالیانه مورد استفاده قرار گرفته است.

۳- بحث و نتایج

تکنیک‌های استوکاستیک تولید داده (جریان و بارندگی) عموماً براساس فرض ایستایی داده‌ها توسعه یافته‌اند (Salas, 1993). بنابراین تکنیک یا مدل‌های مذکور بایستی قادر به دوباره تولید یا حفظ خصوصیات مهم آماری داده‌های تاریخی باشند. نهایتاً مهمترین فاکتورها در انتخاب مدل استوکاستیک، توانایی تولید داده‌های مورد نظر و حفظ خصوصیات داده‌های تاریخی در داده‌های تولیدی است. سیستم‌های منابع آب از جمله مخازن ذخیره و رخدادهای خشکسالی دارای رفتار همزمان درون سالی^{۳۱} و برون سالی^{۳۲} می‌باشند. مطالعات چنین سیستم‌هایی یکی از کاربردهای خیلی مهم مدل‌های تولید داده محسوب می‌شوند. در این تحقیقات، مدل تولید داده بایستی قادر باشد که به طور دقیق خصوصیات آماری داده‌های جریان یا بارندگی را در هر دو سطح سالیانه و ماهیانه حفظ کند. نتیجتاً پروسه ارزیابی برای مدل‌های تولید داده‌های ماهیانه برای تحلیل و مطالعه سیستم‌های منابع آب مذکور، ترجیحاً مقایسه‌های ذیل را شامل می‌شود (Maheepala and Perera, 1996; Savic et al., 1989; McMahon and Mein, 1986; Montaseri and Adeloye, 2002)

- پارامترهای آماری پایه یعنی میانگین، انحراف معیار، ضریب چولگی و خود همبستگی با تاخیر یکساله داده‌های تولیدی و تاریخی در دو سطح سالیانه و ماهیانه

ضعیف عمل کرده، مخصوصاً مدل T.F پارامتر مذکور را کاملاً تخریب^{۲۴} می‌کند.

۳-۳- ضریب Hurst

ضریب Hurst به عنوان پارامتر گرایش بلند مدت^{۲۵} نیز تنها توسط مدل AR(1) به خوبی حفظ شده و این در حالی است که مدل AR(1) یک مدل گرایش کوتاه مدت^{۲۶} بوده و حفظ ضریب Hurst در آن لحاظ نشده است. بنابراین مدل AR(1) به طور تجربی توان حفظ گرایش بلند مدت داده‌های تاریخی را داشته و این امر به عنوان یک مزیت نسبی مهم این مدل نسبت به مدل‌های دیگر مخصوصاً مدل‌های با گرایش بلندمدت مانند Fractional Gaussian Noise و Broken line محسوب می‌شود و این نتیجه در مطالعات متعدد دیگری نیز تأیید شده است (Srikanthan and McMahon, 1979; Montaseri and Adeloye, 2004).

از طرف دیگر منحنی توزیع احتمالی داده‌ها از مدل Bootstrap از یک منحنی هموار تبعیت نکرده و دارای بیشترین اختلاف از منحنی توزیع نظری داده‌های تاریخی است.

۳-۴- توزیع تجربی^{۲۷} داده‌های سالیانه دبی

همچنین دبی‌های حداقلی با احتمال وقوع ۹۰ و ۹۸ درصد از نمودارهای منحنی فرکانس استخراج و در جدول ۳ ارائه شده است. مقادیر دبی با احتمال‌های وقوع ۹۰ و ۹۸ درصد به ازاء مدل‌های مختلف بیانگر عملکرد بهتر مدل AR(1) در مقایسه مدل‌های دیگر در دوباره تولید سری دبی‌های حداقلی پایین‌تر از رخدادهای تاریخی است.

۳-۶- خصوصیات آماری داده‌های ماهیانه تولیدی و تاریخی

در این مطالعه داده‌های ماهیانه با استفاده از مدل‌های ماهیانه Bootstrap (چهار تیپ)، Thomas-Fiering، Fragment و Valencia-Schaake تولید شده است. لذا در این بخش عملکرد مدل‌های فوق در حفظ یا دوباره تولید خصوصیات داده‌های تاریخی در سطح ماهیانه مورد بررسی و مقایسه قرار گرفته است.

در این مقایسه پارامترهای آماری، همبستگی داده‌های ماهیانه در ماه‌های مختلف با داده‌های سالیانه، ساختمان همبستگی داده‌های ماهیانه مابین ماه‌های مختلف و توزیع تجربی داده‌های ماهیانه مورد ارزیابی قرار گرفته است. لازم به ذکر است که پارامترهای آماری داده‌های تولیدی با متوسط‌گیری از ۱۰۰۰ سری برای سه مورد اول و مورد چهارم از داده‌های تمام سری‌های تولیدی تعیین شده است.

۳-۷- پارامترهای آماری پایه دبی‌های ماهیانه

پارامترهای مهم آماری یعنی میانگین، انحراف معیار و ضریب چولگی

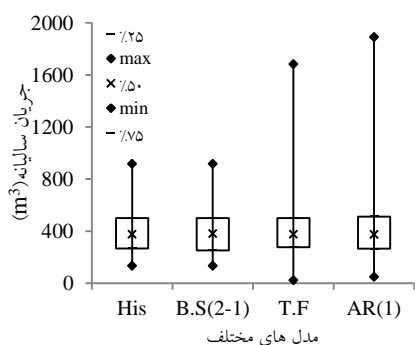
برای نمایش توزیع تجربی داده‌های سالیانه تاریخی و تولیدی از سه مدل مورد مطالعه از نمودار جعبه‌ای (Box Plot) در شکل ۴ استفاده شده است. نمودار جعبه‌ای در شکل ۴ مقادیر حداکثر، حداقل، ۲۵، ۵۰ و ۷۵ درصد داده‌های سالیانه تولیدی از ۱۰۰۰ سری هر سه مدل مذکور را در مقایسه با داده‌های تاریخی نشان می‌دهد. مقادیر حداکثر و حداقل داده‌های سالیانه تولیدی و تاریخی به ازاء جدول ۳ و شکل ۴ نشان می‌دهد که دامنه تغییرات داده‌های سالیانه تولیدی از مدل Bootstrap کاملاً محدود به داده‌های سالیانه تاریخی بوده ولی این دامنه تغییرات برای مدل‌های T.F و AR(1) برای هر سه رودخانه مذکور بسیار بیشتر از داده‌های تاریخی است و در این میان مدل AR(1) دارای عملکرد بهتری نسبت به مدل T.F دارد. این امر برای رخدادهای مورد انتظار یا آتی هیدرولوژیکی از جمله دبی بسیار محتمل بوده و بایستی مدل استوکاستیک توان تولید خشکسالی و ترسالی‌های شدیدتر از رویدادهای تاریخی را داشته باشد. همچنین نمودارهای جعبه‌ای مذکور بیانگر توزیع داده‌های سالیانه با چولگی به طرف راست یعنی ضریب چولگی مثبت است.

۳-۵- منحنی‌های فرکانس دبی و آماره‌های مقادیر حدی حداقل

طرح، مدیریت و بهره‌برداری سیستم‌های منابع آب از جمله مخازن ذخیره و تحلیل پدیده خشکسالی به‌طور خاصی متأثر از دبی‌های (یا بارندگی‌های) حداقلی^{۲۸} می‌باشند. لذا مقایسه توزیع احتمالی داده‌های تولیدی و تاریخی با تأکید بر محدوده مقادیر حداقلی دارای اهمیت

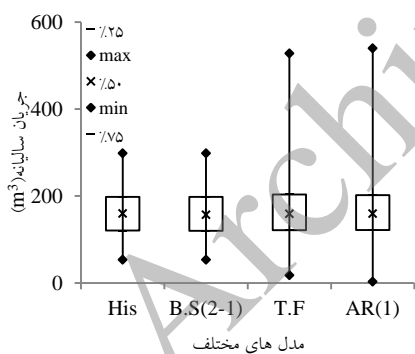
جدول ۳- مقایسه مشخصات آماری دبی سالیانه داده‌های تاریخی و تولیدی با مدل‌های مختلف در رودخانه‌های باراندوز، نازلو و شهرچای

رودخانه	Model	Min (m ³ /s)	Max (m ³ /s)	Mean (m ³ /s)	Cv	Skew	Auto Correlation lag-1	Hurst	Q90% (m ³ /s)	Q98% (m ³ /s)
باراندوز	His	۱۰۵/۲	۴۳۳/۰	۲۴۶/۹	۰/۳۲	۰/۵۱	۰/۳۸	۰/۶۲	۱۷۲/۱	۱۱۱/۲
	AR(1)	۲۲/۰	۷۰۴/۰	۲۴۷/۶	۰/۳۰	۰/۴۶	۰/۳۶	۰/۷۳	۱۵۳/۳	۱۰۹/۵
	T.F	۶۳/۱	۶۹۹/۴	۲۴۷/۷	۰/۲۹	۰/۵۹	۰/۰۲	۰/۶۴	۱۶۰/۷	۱۲۳/۲
	B.S(2-1)	۱۰۵/۲	۴۳۳/۰	۲۴۶/۶	۰/۳۱	۰/۵۶	۰/۱۶	۰/۶۷	۱۶۹/۱۰	۱۰۵/۲
نازلو	His	۱۳۴/۰	۹۱۶/۰	۴۰۲/۴	۰/۴۷	۱/۰۲	۰/۳۶	۰/۷۲	۲۰۷/۱	۱۳۷/۲
	AR(1)	۴۸/۶	۱۸۹۰/۰	۴۰۵/۸	۰/۴۶	۰/۹۵	۰/۳۲	۰/۷۲	۱۹۲/۳	۱۲۳/۹
	T.F	۲۲/۸	۱۶۸۲/۶	۴۰۲/۱	۰/۴۲	۰/۸۷	۰/۰۶	۰/۶۵	۲۰۲/۹	۱۳۳/۱
	B.S(2-1)	۱۳۳/۹	۹۱۶/۰	۴۰۱/۹	۰/۴۶	۱/۰۹	۰/۱۷	۰/۶۷	۲۰۲/۴	۱۳۴/۰
شهر	His	۵۳/۶	۲۹۸/۴	۱۶۵/۲	۰/۳۶	۰/۵۷	۰/۲۲	۰/۷۴	۱۰۹/۳	۷۰/۷
	AR(1)	۲/۹	۵۳۹/۸	۱۶۵/۰	۰/۳۵	۰/۵۳	۰/۲۵	۰/۷۰	۹۲/۷	۶۰/۳
	T.F	۱۷/۸	۶۴۲/۱	۱۶۶/۵	۰/۳۸	۰/۶۹	۰/۰۱	۰/۶۴	۹۱/۲	۵۸/۸
	B.S(2-1)	۵۳/۶	۲۹۶/۴	۱۶۵/۰	۰/۳۶	۰/۶۲	۰/۱۰	۰/۶۶	۹۶/۵	۵۳/۶

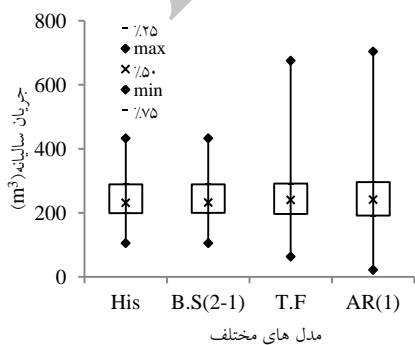


داده‌های استوکاستیک مذکور برای سه رودخانه نازلو، شهر و باراندوز بررسی و مورد مقایسه قرار گرفته است. شکل ۶ نمودار تغییرات پارامترهای آماری تولیدی مذکور برای رودخانه‌های شهرچای و باراندوز چای را به عنوان نمونه نشان می‌دهد. نتایج حاصل بیانگر عملکرد مناسب تمامی مدل‌های هفت گانه در حفظ پارامترهای آماری پایه داده‌های ماهیانه رودخانه مورد مطالعه می‌باشد.

۳-۸- همبستگی داده‌های ماهیانه در ماه‌های مختلف با داده‌های سالیانه



این پارامتر به عنوان یکی از پارامترهای مهم بیان کننده همگرایی رفتار پدیده‌های هیدرولوژیکی مانند دبی و بارندگی در دو سطح سالیانه و ماهیانه است (Maheepala and Perera, 1996). نمودار تغییرات ضریب همبستگی داده‌های دبی ماهیانه در ماه‌های مختلف با داده‌های سالیانه به ازاء مدل‌های Thomas-Fiering, Fragment و Valencia-Schaake برای هر سه رودخانه در شکل ۷ ارائه شده است. نتایج به دست آمده نشان می‌دهد که عملکرد مدل Thomas-Fiering در حفظ پارامتر مذکور کاملاً مایوس کننده بوده ولی مدل Fragment و Valencia-Schaake به ترتیب دارای عملکرد بسیار مطلوب و رضایت بخش می‌باشند.



در این مقایسه از نتایج مدل Bootstrap بلوکی که داده‌های ماهیانه و سالیانه در هر سال به عنوان یک بلوک مورد استفاده قرار می‌گیرد و در نتیجه پارامتر مذکور به طور سیستماتیک در داده‌های تولیدی حفظ می‌شود، استفاده نگردیده است.

شکل ۴- مقایسه نمودار جعبه‌ای جریان‌های سالیانه داده‌های تاریخی و تولیدی (از بالا به پائین: نازلو، شهر و باراندوز).

۳-۹- ساختمان همبستگی داده‌های ماهیانه مابین ماه‌های مختلف

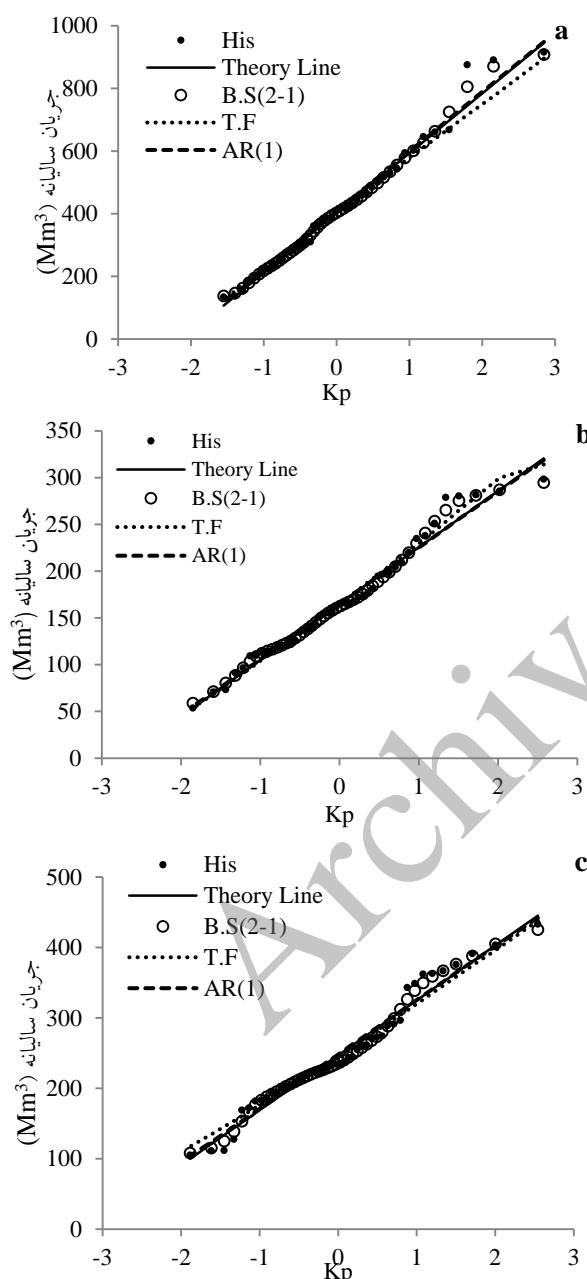
رخداد‌های هیدرولوژیکی مانند بارندگی و جریان در پریودهای کوتاه مدت ماهیانه متأثر از شرایط هیدرولوژیکی پیوسته و طولانی مدت یک یا چند ساله بوده لذا چنین رخداد‌های هیدرولوژیکی کوتاه مدت (ماهیانه) دارای همبستگی بالایی مابین ماه‌های مختلف سال است (Montaseri and Adeloey, 2002). در جدول ۴ مقادیر ضریب همبستگی داده‌های ماهیانه تولیدی و تاریخی برای رودخانه باراندوز چای به ازاء مدل‌های استوکاستیک Thomas-Fiering و Valencia-Schaake ارائه شده است. در این بخش نتایج مدل‌های استوکاستیک Bootstrap و Fragment به دلیل حفظ سیستماتیک ساختمان همبستگی ماهیانه در این مدل‌ها ارائه نشده است. براساس جداول مذکور مدل Thomas-Fiering به استثنای همبستگی دو ماه متوالی دارای عملکرد ضعیف در حفظ همبستگی داده‌های ماهیانه مابین ماه‌های مختلف سال است و بر عکس مدل Valencia-Schaake عملکرد بسیار خوبی در حفظ ساختمان همبستگی داده‌های ماهیانه را ارائه نموده است.

۳-۱۰- توزیع تجربی داده‌های ماهیانه

برای نمایش توزیع تجربی داده‌های ماهیانه از داده‌های تولیدی و تاریخی ماه‌های فروردین و شهریور به عنوان معرفی از دوره پرآبی و کم آبی با استفاده از نمودار جعبه‌ای (Box Plot) استفاده شده است. شکل ۸ نمودار جعبه‌ای داده‌های تولیدی و تاریخی ماه‌های فروردین و شهریور برای رودخانه نازلو چای به عنوان نمونه را به ازاء هفت مدل استوکاستیک نشان می‌دهد. نمودارهای مذکور بیان می‌کند که توزیع تجربی داده‌های تولیدی (چهارتیب) مدل Bootstrap به طور کامل مشابه داده‌های تاریخی بوده یعنی در این مدل‌ها داده‌های تولیدی در محدوده داده‌های تاریخی تغییر کرده و رخداد‌های مورد انتظار و محتمل کمتر یا بیشتر از محدوده داده‌های تاریخی دارای هیچ شانس نمی‌باشد. در حالیکه توزیع تجربی داده‌های ماهیانه به ازاء مدل‌های Thomas-Fiering، Fragment و Valencia-Schaake دارای محدوده تغییرات بسیار بزرگتری بوده که رخداد‌های محتمل مورد انتظار حتمی کوچکتر و بزرگتر از محدوده داده‌های تاریخی در آینده را شامل شده است و این امر به عنوان یک مزیت مهم مدل‌های پارامتریک نسبت به مدل‌های غیرپارامتریک محسوب می‌شود. همچنین در میان سه مدل Thomas-Fiering، Valencia-Schaake و Fragment Valencia-Schaake مدل Valencia-Schaake عملکرد نسبی بهتری داشته است.

۴- نتیجه‌گیری

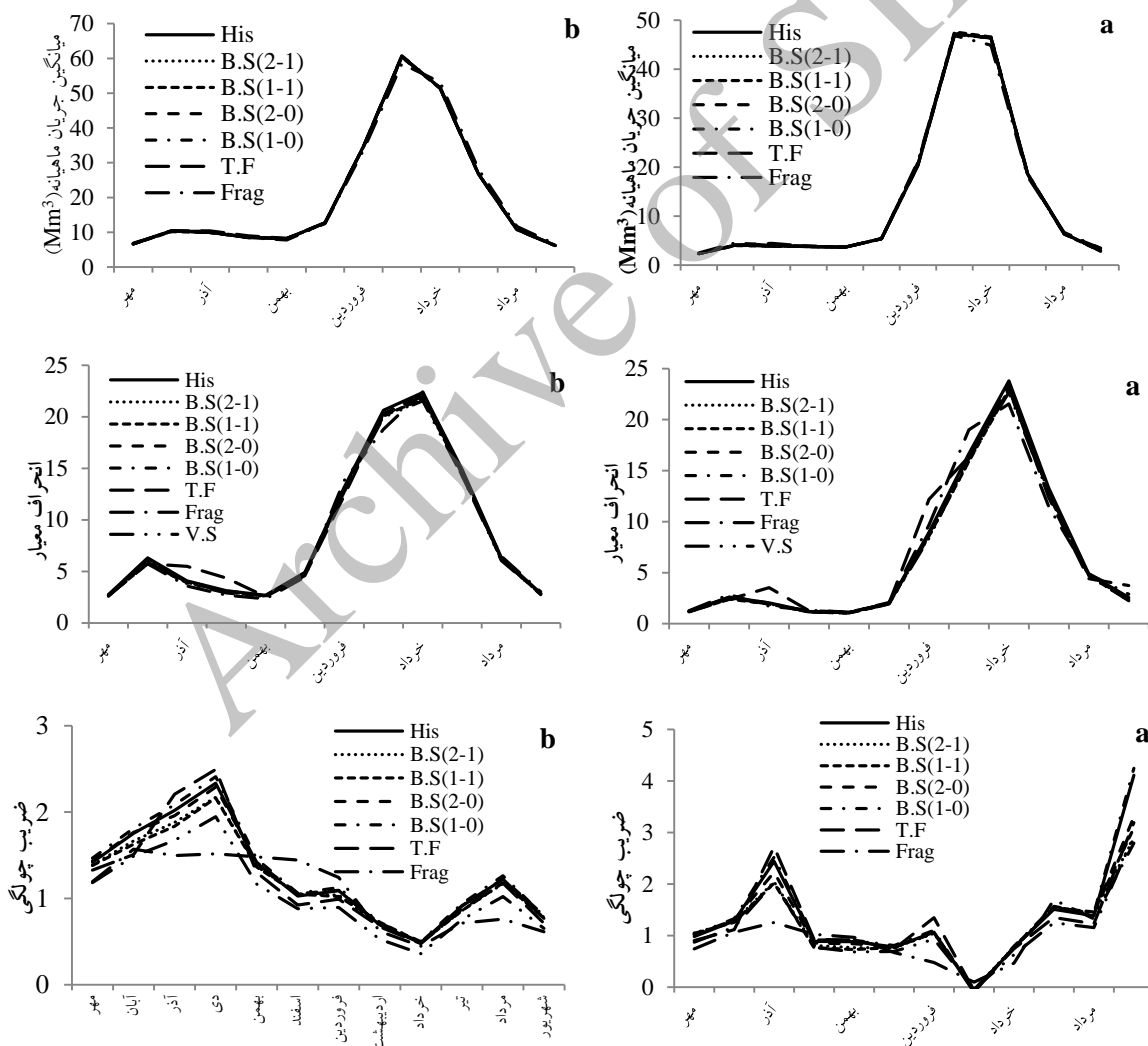
نتایج این مطالعه نشان می‌دهد ترکیب مدل Valencia-Schaake با مدل AR(1) برای تولید داده‌های جریان و بارندگی مطمئن‌ترین روش می‌باشد و می‌توان از این مدل جهت تحلیل سیستم مخازن، پایش خشکسالی و بررسی تغییرات احتمالی اقلیم در آینده با اطمینان بیشتر استفاده کرد.



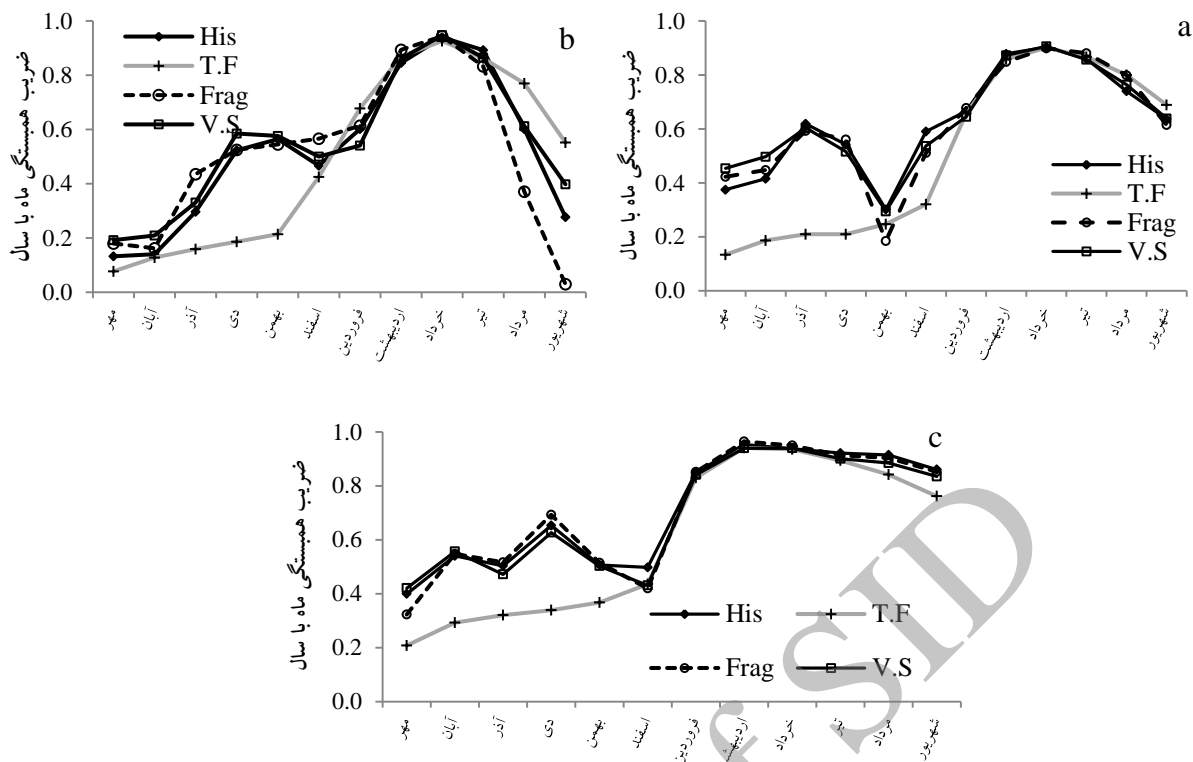
شکل ۵- برازش گرافیکی داده‌های تاریخی و تولیدی با توزیع پیرسون نوع سوم، نازلو چای (a)، شهرچای (b) و باراندوزچای (c)

مدل Thomas-Fiering نیز می‌تواند در تولید داده مصنوعی به کار گرفته شود اما با توجه به نتایج به دست آمده این مدل در حفظ ضرایب همبستگی و دامنه تغییرات داده‌ها دارای محدودیت جدی نسبت به مدل Valencia-Schaake است. نتایج حاصل این مطالعه بر اساس داده‌های جریان برای یک اقلیم نیمه خشک در شمال غرب ایران بوده و لذا مقایسه مدل‌های استوکستیک با به‌کارگیری سایر داده‌های هیدرولوژیکی مانند بارندگی و ارزیابی عملکرد مدل‌ها در اقلیم‌های مختلف و استفاده از سایر مدل‌های تولید داده مصنوعی برای مطالعات آتی می‌تواند مطرح باشد.

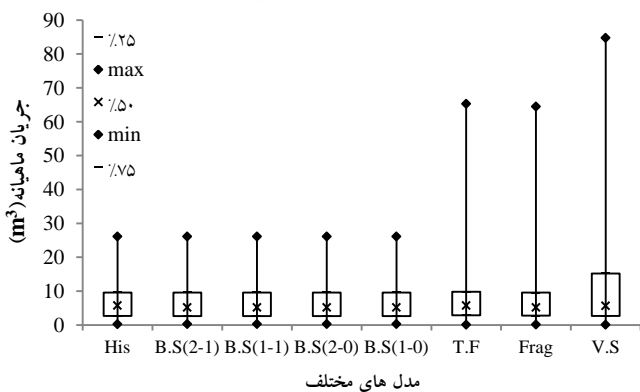
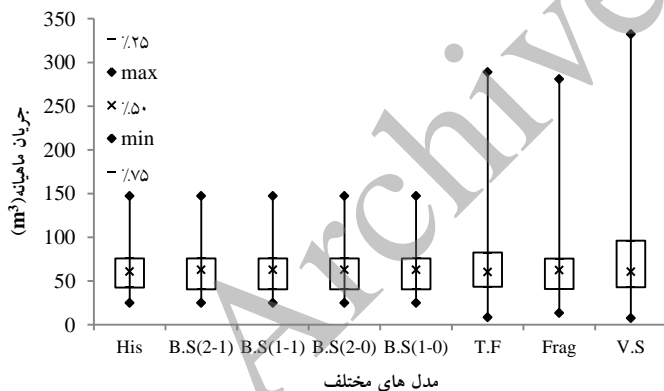
مدل Bootstrap با اینکه نتایج موفق‌تری در حفظ پارامترهای آماری دارد اما این داده‌ها در واقع همان داده‌های تاریخی ثبت شده می‌باشند که فقط ترتیب وقوع آنها عوض گردیده و داده‌های بزرگتر و کوچکتر از داده‌های تاریخی تولید نمی‌کند که این مغایر با احتمال مورد انتظار در طبیعت می‌باشد. مدل Fragment نیز همانند مدل Valencia-Schaake به خوبی پارامترهای آماری داده‌ها را حفظ می‌کند اما چون در این مدل توزیع داده‌های سالیانه به ماهیانه به صورت قطعی انجام می‌پذیرد نتایج غیرواقعی نسبت به مدل Valencia-Schaake داشته و تفاوت این مدل را می‌توان در دامنه تغییرات و برازش با توزیع‌های آماری مشاهده کرد.



شکل ۶- مقایسه میانگین، انحراف معیار و ضریب چولگی داده‌های تاریخی و تولیدی مدل‌های مختلف، شهرچای (a) و باراندوزچای (b).



شکل ۷- ضریب همبستگی ماه با سال داده‌های تاریخی و تولیدی مدل‌های مختلف، نازلو چای (a)، شهرچای (b) و باراندوزچای (c).



شکل ۸- مقایسه نمودار جعبه‌ای جریان‌های ماهیانه تاریخی و مصنوعی در ماه فروردین (پرآب) و ماه شهریور (کم آب) - رودخانه نازلو چای.

جدول ۴- همبستگی داده‌های ماهیانه تاریخی و تولیدی با مدل‌های مختلف مابین ماه‌های مختلف در رودخانه‌های باراندوز چای.

شهریور	مرداد	تیر	خرداد	اردیبهشت	فروردین	اسفند	بهمن	دی	آذر	آبان	مهر	مدل	
-	-	-	-	-	-	-	-	-	-	-	۱/۰۰	His	مهر
-	-	-	-	-	-	-	-	-	-	-	۱/۰۰	T.F	
-	-	-	-	-	-	-	-	-	-	-	۱/۰۰	V.S	
-	-	-	-	-	-	-	-	-	-	۱/۰۰	۰/۶۴	His	آبان
-	-	-	-	-	-	-	-	-	-	۱/۰۰	۰/۶۱	T.F	
-	-	-	-	-	-	-	-	-	-	۱/۰۰	۰/۶۰	V.S	
-	-	-	-	-	-	-	-	-	۱/۰۰	۰/۵۹	۰/۵۱	His	آذر
-	-	-	-	-	-	-	-	-	۱/۰۰	۰/۵۴	۰/۳۴	T.F	
-	-	-	-	-	-	-	-	-	۱/۰۰	۰/۵۴	۰/۴۵	V.S	
-	-	-	-	-	-	-	-	۱/۰۰	۰/۵۲	۰/۶۰	۰/۵۴	His	دی
-	-	-	-	-	-	-	-	۱/۰۰	۰/۴۵	۰/۲۶	۰/۱۶	T.F	
-	-	-	-	-	-	-	-	۱/۰۰	۰/۴۵	۰/۵۵	۰/۴۸	V.S	
-	-	-	-	-	-	-	۱/۰۰	۰/۵۴	۰/۳۶	۰/۴۸	۰/۳۸	His	بهمن
-	-	-	-	-	-	-	۱/۰۰	۰/۴۸	۰/۲۴	۰/۱۶	۰/۱۰	T.F	
-	-	-	-	-	-	-	۱/۰۰	۰/۴۸	۰/۳۲	۰/۴۲	۰/۳۴	V.S	
-	-	-	-	-	-	۱/۰۰	۰/۵۹	۰/۴۷	۰/۴۵	۰/۵۰	۰/۴۴	His	اسفند
-	-	-	-	-	-	۱/۰۰	۰/۵۷	۰/۲۸	۰/۱۵	۰/۱۰	۰/۰۷	T.F	
-	-	-	-	-	-	۱/۰۰	۰/۵۶	۰/۴۲	۰/۴۰	۰/۴۶	۰/۴۰	V.S	
-	-	-	-	-	۱/۰۰	۰/۳۳	۰/۱۸	۰/۴۰	۰/۳۰	۰/۲۸	۰/۱۶	His	فروردین
-	-	-	-	-	۱/۰۰	۰/۳۱	۰/۱۸	۰/۰۸	۰/۰۴	۰/۰۳	۰/۰۲	T.F	
-	-	-	-	-	۱/۰۰	۰/۳۲	۰/۱۷	۰/۳۴	۰/۲۶	۰/۲۵	۰/۱۳	V.S	
-	-	-	-	۱/۰۰	۰/۶۳	۰/۴۷	۰/۲۴	۰/۴۸	۰/۵۱	۰/۴۹	۰/۳۶	His	اردیبهشت
-	-	-	-	۱/۰۰	۰/۶۲	۰/۲۰	۰/۱۱	۰/۰۵	۰/۰۳	۰/۰۲	۰/۰۱	T.F	
-	-	-	-	۱/۰۰	۰/۶۱	۰/۴۵	۰/۳۲	۰/۴۲	۰/۴۵	۰/۴۵	۰/۳۲	V.S	
-	-	-	۱/۰۰	۰/۷۶	۰/۴۸	۰/۳۲	۰/۰۲	۰/۳۸	۰/۵۳	۰/۳۴	۰/۳۷	His	خرداد
-	-	-	۱/۰۰	۰/۷۵	۰/۴۶	۰/۱۵	۰/۰۸	۰/۰۴	۰/۰۲	۰/۰۱	۰/۰۰	T.F	
-	-	-	۱/۰۰	۰/۷۴	۰/۴۶	۰/۳۱	۰/۰۱	۰/۳۲	۰/۴۷	۰/۳۱	۰/۳۳	V.S	
-	-	۱/۰۰	۰/۹۰	۰/۶۲	۰/۴۵	۰/۴۱	۰/۰۷	۰/۳۶	۰/۵۶	۰/۲۱	۰/۳۱	His	تیر
-	-	۱/۰۰	۰/۸۹	۰/۶۶	۰/۴۰	۰/۱۳	۰/۰۷	۰/۰۴	۰/۰۲	۰/۰۱	۰/۰۰	T.F	
-	-	۱/۰۰	۰/۸۹	۰/۵۹	۰/۴۳	۰/۳۸	۰/۰۶	۰/۳۱	۰/۵۱	۰/۱۹	۰/۲۸	V.S	
-	۱/۰۰	۰/۹۱	۰/۸۱	۰/۵۱	۰/۳۱	۰/۳۷	۰/۱۳	۰/۳۷	۰/۵۲	۰/۲۱	۰/۴۰	His	مرداد
-	۱/۰۰	۰/۹۰	۰/۸۰	۰/۵۹	۰/۳۷	۰/۱۲	۰/۰۶	۰/۰۳	۰/۰۲	۰/۰۰	۰/۰۰	T.F	
-	۱/۰۰	۰/۹۰	۰/۷۹	۰/۴۸	۰/۲۸	۰/۳۴	۰/۱۱	۰/۳۰	۰/۴۶	۰/۱۸	۰/۳۵	V.S	
۱/۰۰	۰/۸۴	۰/۶۸	۰/۶۴	۰/۵۰	۰/۱۹	۰/۲۵	۰/۲۸	۰/۳۲	۰/۴۰	۰/۲۱	۰/۴۱	His	شهریور
۱/۰۰	۰/۸۳	۰/۷۶	۰/۶۸	۰/۵۱	۰/۳۱	۰/۱۰	۰/۰۶	۰/۰۳	۰/۰۲	۰/۰۰	۰/۰۱	T.F	
۱/۰۰	۰/۸۲	۰/۶۵	۰/۶۳	۰/۴۷	۰/۱۷	۰/۲۲	۰/۲۶	۰/۲۷	۰/۳۵	۰/۲۲	۰/۳۷	V.S	

5- Long term persistence

6- Short term persistence

7- Direct

8- Disaggregation

9- Autocorrelation lag one

10- Cross correlation lag one

پی‌نوشت‌ها

1- Resample

2- Parametric

3- Non- parametric

4- Lag one serial correlation coefficient

- Kendall MG, Stuart A (1976) The advanced theory of statistics. Charles Griffin & Company, London, High Wycombe, 400,401.
- Lin GF, Lee FC (1992) An aggregation-disaggregation approach for hydrologic time series modeling, *Journal of Hydrology*, 138:543-557.
- Maheepala S, Perera BJC (1996) Monthly hydrologic data generation by disaggregation, *Journal of Hydrology*, 178:277-291.
- McGhee JW (1985) Introductory statistics, West Publishing Co., New York, USA.
- McMahon TA, Mein RG (1986) Water and reservoir yield. Water Resources Publication, Littleton, Colorado.
- Montaseri M, Adeloje AJ (1999) Critical period of reservoir systems for planning purpose, *Journal of Hydrology*, 224(4-3):136-115.
- Montaseri M, Adeloje AJ (2002) Effects of integrated planning on capacity-yield-performance functions, *Journal of Water Resources Planning and Management*, 128(6):456-461.
- Montaseri M, Adeloje AJ (2004) A graphical rule for volumetric evaporation loss correction in reservoir capacity-yield-performance planning in Urmia region, Iran, *Water Resources Management*, 18:55-74.
- Nurul AI (2004) Synthetic simulation of stream flow and rainfall data using disaggregation and aggregation models, *Jurnal Kejuruteraan Awam* 16(2):56-65.
- Porter JW, and Pink BJ, 1991. A method of synthetic fragments for disaggregation in stochastic data generation, Pp. 187-191. *Proceedings In Hydrology and Water Resources Symposium*. Australia.
- Salas JD (1993) Analysis and modeling of hydrologic time series, In *Handbook of Hydrology*, Edited by Maidment, McGraw – Hill book CO., New York.
- Salas JD, Deller JW, Yevjevich V, and Lane WL (1980) Applied modeling of hydrologic time series, Water Resources Publications, Littleton, Colorado.
- Savic DA, Burn DH, and Zrinji Z (1989) A comparison of streamflow generation models for reservoir capacity-yield analysis, *Water Resources Bulletin*, 25(5): 977-983.
- Srikanthan R, and McMahon TA (1982) Stochastic generation of monthly streamflows, *Journal of the Hydraulics Division, ASCE*, 108(HY3), 419-441.
- Srikanthan R, McMahon TA (1985) Stochastic generation of rainfall and evaporation data AWRC Technical Paper No.84, 301pp.
- 11- Probability Plot Correlation Coefficient
- 12- Parametric time series model
- 13- Non parametric time series model
- 14- Resampling
- 15- Single block
- 16- Double block
- 17- Direct
- 18- Disaggregation
- 19- With replacement
- 20- Without replacement
- 21- Within year
- 22- Over year
- 23- Frequency curve
- 24- Destroy
- 25- Long term persistence
- 26- Short term persistence
- 27- Empirical
- 28- Low flow

۵- مراجع

Adeloje AJ, and Montaseri M (2002) Preliminary stream flow data analyses prior to water resources planning study, *Hydrological Sciences Journal*, 47(5):679-692.

Barnes FB (1954) Storage required for a city water supply, *Journal of the Institute of Engineers Australia*, 26(9): 198-203.

Bars RL (1990) *Hydrology: an introduction to hydrologic science*. Addison-Wesley Publishing Co., New York, USA.

Brittan MR (1961) Probability analysis to the development of a synthetic hydrology for the Colorado River, in part IV of past and probable future variations in stream flow in the upper Colorado River, University of Colorado.

Chow VT, Maidment, DR, and Mays LW (1993) *Applied hydrology*, Mc Grave-Hill book Co., New York.

Debele B (2007) Accuracy evaluation of weather data generation and disaggregation methods at finer timescales, *Advance water Resource*, 30(5):1286-1300.

Douglas H (2010) *How to measure anything: finding the value of intangibles in business*. John Wiley and Sons, Inc, Canada.

Fiering MB (1967) *Stream flow synthesis*, Harvard Univ. Press, Cambridge, Mass.

Hazen A (1914) Storage to be provided in impounding reservoirs for municipal water supply, *Trans., ASCE*, 77: 1539-1540.

- G. Fair (Eds.), Harvard University Press, Cambridge, Massachusetts USA.
- Vogel RM, Kroll CN (1989) Low flow frequency analysis using probability plot correlation coefficients. *Journal of Water Resources Planning and Management*, 115(3):338-357.
- Vogel RM, Shallcross AL (1996) The moving blocks bootstrap versus parametric time series models, *Water Resources Research*, 32(6):1875-1882.
- Yevjevich V (1972) Probability and statistics in hydrology. Water Resources Publications Fort Collins, Colorado, philosophy.
- Srikanthan R, McMahon TA (2001) Stochastic generation of annual, monthly and daily climate data. *Hydrology and Earth System Sciences*, 5(4):653-670.
- Srinivas VV, Srinivasan K (2005) Hybrid moving block bootstrap stochastic simulation of multi-site multi-season for stream flows. *Journal of Hydrology*, 302:307-330.
- Thomas HA, Fiering MP (1962) Mathematical synthesis of stream flow sequences for the analysis of river basins by simulation. Chapter 12 in: Design of water resources systems A. Mass, S. Marglin and

Archive of SID