

مقایسه روش‌های کاپیولا و خودرگرسیون میانگین متحرک فصلی در مدل‌سازی جریان رودخانه هیرمند

احمدعلی کیخا^۱ - مهدیه مسنن مظفری^{۲*} - محمود صبوحی^۳ - غلامرضا سلطانی^۴

تاریخ دریافت: ۱۳۹۲/۹/۱۲

تاریخ پذیرش: ۱۳۹۳/۴/۲

چکیده

مدل‌سازی جریان رودخانه در برنامه‌ریزی و مدیریت منابع آب از اهمیت ویژه‌ای برخوردار است. با این وجود، ایجاد داده‌هایی که بتواند شباهت زیادی با داده‌های حقیقی داشته باشد، به دلیل اندک بودن داده‌های جریان رودخانه و وابستگی و همبستگی ماهانه و سالانه آنها، کاری مشکل و پیچیده است. در این مطالعه از آمار ۵۰ ساله و دو روش خودرگرسیون میانگین متحرک فصلی و کاپیولاهای (مفصل‌های) فرانک و کلیتون که به ترتیب روش‌های پیش‌بینی و شبیه‌سازی در مدل‌سازی جریان رودخانه هستند، برای ایجاد داده‌های جریان رودخانه هیرمند استفاده شد. نتایج نشان داد، داده‌های بازسازی شده در روش خودرگرسیون میانگین متحرک مقادیر کم جریان رودخانه را به خوبی پیش‌بینی می‌کنند، ولی از همبستگی داده‌های تاریخی برخوردار نبوده و حداکثر جریان رودخانه را کمتر از حد واقعی نشان می‌دهند. از طرف دیگر، کاپیولا همبستگی داده‌های جریان رودخانه را حفظ کرده و مقادیری مشابه داده‌های واقعی ایجاد می‌کند. همچنین مقدار ریشه میانگین مربع خطا در روش کاپیولا ۰/۳ بوده که کمتر از روش میانگین متحرک (۰/۴) است. لذا می‌توان گفت، کاپیولا روش مناسبی برای مدل‌سازی جریان رودخانه هیرمند است.

واژه‌های کلیدی: شبیه‌سازی، پیش‌بینی، مدیریت منابع آب، همبستگی

مقدمه

مدل‌سازی جریان رودخانه، اولین قدم در مدیریت منابع آب یک منطقه است. مدل مناسب جریان رودخانه، مدلی است که بتواند داده‌هایی با توزیع آماری مشابه جریان واقعی رودخانه ایجاد نماید. با استفاده از روش‌های نوین شبیه‌سازی و نرم‌افزارهای موجود، علاوه بر افزایش سرعت اجرای محاسبات، ویژگی‌های آماری داده‌های موجود که مسئله‌ای مهم در مدل‌سازی است، حفظ شده و در نتیجه داده‌هایی با شباهت بیشتر به جریان واقعی رودخانه ایجاد می‌شود (۵). روش‌های مختلفی برای شبیه‌سازی جریان رودخانه توسط محققین کشورهای دنیا استفاده شده است که ون درلی (۱۸) آن‌ها را به دو روش پیش‌بینی و شبیه‌سازی گروه‌بندی نمود. در هر دو روش مطالعات مختلفی انجام شده است. در روش

پیش‌بینی، وانگ (۱۹) با استفاده از تکنیک سری‌های زمانی خودرگرسیون میانگین متحرک متداول و دوره‌ای به پیش‌بینی ماهانه جریان رودخانه زرد^۵ در چین پرداخت. دوگان و همکاران (۸) برای پیش‌بینی روزانه جریان رودخانه حوضه ساکاریا^۶ از دو روش شبکه‌های عصبی مصنوعی و خودرگرسیون استفاده کردند. طرازکار و صدق‌آمیز (۵) دبی جریان رودخانه کرخه را با استفاده از روش‌های خودرگرسیون برداری و شبکه عصبی مصنوعی پیش‌بینی کرده و با استفاده از معیارهای مختلف آن‌ها را باهم مقایسه کردند. ولی در سال‌های اخیر ساخت متغیرهای تصادفی جریان رودخانه با استفاده از مدل‌های شبیه‌سازی کاربرد بیشتری داشته است.

هارداکر و همکاران (۱۱) دلیل استفاده از روش‌های شبیه‌سازی را حفظ همبستگی بین داده‌ها در ماه‌ها و یا سال‌های مختلف، می‌دانند. هوپ و بارت (۱۲) طی مطالعه خود در ۴۱ حوضه آبریز کالیفرنیا برای به دست آوردن دبی با فراوانی ۸۰ درصد و مدت زمانی که رودخانه خشک است، از منحنی‌های تداوم جریان استفاده کردند. آنها به این

۱- دانشیار گروه اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه زابل
۲ و ۳- دانشجوی دکتری و دانشیار گروه اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه فردوسی مشهد

*- نویسنده مسئول: (Email: mmosannan@gmail.com)

۴- استاد گروه اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه شیراز

سالانه جریان ورودی رودخانه هیرمند طی سال‌های ۱۳۳۸-۱۳۸۸ وجود داشته است. به طوری که میانگین جریان رودخانه بین ۱۷ میلیون مترمکعب در سال ۱۳۸۱ و ۶۲۵ میلیون مترمکعب در سال ۱۳۷۱ نوسان داشته است. در کل ورود آب به رودخانه روندی نزولی داشته و میانگین ورودی آب در ۵۰ سال گذشته، سه برابر میانگین آن در ۱۰ سال اخیر (۱۳۷۹-۱۳۸۸) بوده است (۷).

آب رودخانه هیرمند در ماه‌های مختلف سال نیز نوسانات زیادی دارد. به طوری که در اردیبهشت سال ۱۳۸۱ حدود ۲ میلیارد مترمکعب آب از این رودخانه وارد ایران شد ولی در مهر و آبان ماه سال ۱۳۸۱ هیچ آبی در رودخانه جاری نشده است. میانگین ماهانه آب این رودخانه در ۵۰ سال مورد مطالعه، در شکل ۲ نشان داده شده است (۷).

با توجه به شکل فوق حداکثر آب ورودی در اردیبهشت ماه اتفاق می‌افتد به طوری که ۵۵ درصد میانگین حجم آب ورودی در فصل بهار بوده و تنها ۴۵ درصد آب در فصول دیگر سال از رودخانه هیرمند وارد ایران می‌شود (۷).

با توجه به تغییرات ماهانه و سالانه جریان رودخانه هیرمند در ورود به ایران و همبستگی جریان آب طی سال‌های مورد مطالعه، ساختن مدلی که بتواند همه ویژگی‌های ورودی آب رودخانه را حفظ کند امری پیچیده و بسیار مهم می‌باشد. به همین دلیل در این تحقیق از روش‌هایی استفاده شد که علاوه بر کاربرد آن توسط سایر محققین، بتواند شرایط این رودخانه را مدل‌سازی کند. در ادامه این روش‌ها تشریح می‌شوند.

مواد و روش‌ها

از آنجا که در هر سری زمانی بررسی ایستایی داده‌ها اولین قدم در کاربرد آن‌ها است، ابتدا ایستایی و سپس مدل‌های مورد استفاده در این مطالعه شامل مدل خودرگرسیون میانگین متحرک و کاپیولا (مفصل) تشریح خواهند شد.

فرآیند تصادفی هنگامی ایستا نامیده می‌شود که میانگین و واریانس آن طی زمان ثابت بوده و مقدار کواریانس بین دو دوره زمانی تنها به فاصله یا وقفه بین دو دوره بستگی داشته باشد. آزمونی که اغلب برای بررسی ایستایی مورد استفاده قرار می‌گیرد و از دقت بالاتری نسبت به سایر روش‌ها برخوردار است، آزمون ریشه واحد و استفاده از آزمون دیکی فولر تعمیم‌یافته است که به صورت معادله ۱ می‌باشد (۶).

$$\Delta y_t = (\rho - 1)y_{t-1} + u_t = \delta y_{t-1} + u_t \quad (1)$$

در این معادله y_t داده‌های سری زمانی، u_t جملات اخلال و δ پارامتر ایستایی هستند. لذا اگر فرض صفر $\delta = 0$ رد شود، سری زمانی ایستا خواهد بود.

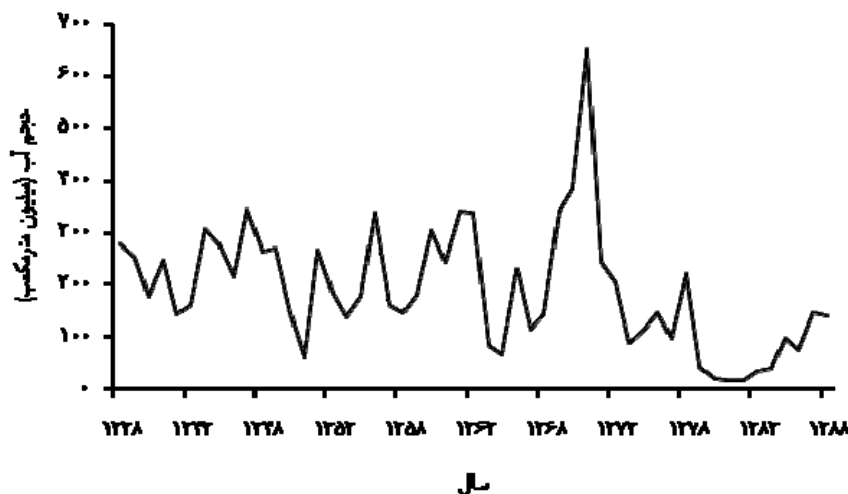
نتیجه رسیدند که منحنی تداوم جریان با زمین‌شناسی منطقه رابطه‌ی مستقیمی داشته و یکی از مزایای آن نسبت به سایر روش‌ها استفاده از آن در صورت وجود داده‌های کم جریان رودخانه است. دلیری و خلقی (۳) با استفاده از روش اعداد تصادفی مستقل با توزیع یکنواخت که شرایط هیدرولوژیکی را با اعداد تصادفی ترکیب می‌کند، متوسط جریان آب رودخانه اریه استان خراسان رضوی را به صورت تصادفی شبیه‌سازی کردند. نتایج آن‌ها نشان داد مقدار خطای پیش‌بینی بسیار کم بوده و این روش برای شبیه‌سازی جریان تصادفی این رودخانه مناسب است. از آنجا که همبستگی بین متغیرهای تصادفی در شبیه‌سازی داده‌ها امری مهم می‌باشد، بررسی این رابطه با استفاده از توابعی خاص مانند کاپیولا، صورت می‌گیرد. موضوع کاپیولا در سال ۱۹۵۹ توسط اسکالر معرفی شد و پس از آن در تحقیقات فراوانی مورد استفاده قرار گرفت. می‌توان گفت، کاپیولاها توابعی هستند که توابع توزیع چند متغیره را به توابع توزیع حاشیه‌ای تک بعدی متصل می‌کنند (۱۶). با توجه به اهمیت پیش‌بینی در بازارهای بورس، کاربرد این روش بیشتر در بازارهای مالی بوده و برای پیش‌بینی جریان رودخانه استفاده کمی داشته است. برای مثال کیخا (۱۵) با استفاده از این روش جریان رودخانه سیستان را شبیه‌سازی کرده و برای مدل خود از آن استفاده نمود. فاور و همکاران (۹) اولین بار از کاپیولا در مدل‌سازی متغیرهای هیدرولوژیک دو حوضه آبریز در کبک^۱ کانادا استفاده کردند. آن‌ها بیان کردند همبستگی بین متغیرهای موجود در مباحث هیدرولوژی را می‌توان با استفاده از این روش مدل‌سازی کرد و نتایج بهتری نسبت به روش‌های پیش‌بینی داشت.

لذا در این مطالعه جریان رودخانه هیرمند با استفاده از روش‌های پیش‌بینی (خودرگرسیون میانگین متحرک) و شبیه‌سازی (کاپیولا)، مدل‌سازی و باهم مقایسه خواهند شد. با توجه به اهمیت شناخت خصوصیات رودخانه در ایجاد داده‌های آن، قبل از ساختن مدل مناسبی از جریان رودخانه، ویژگی‌های جریان ورودی به رودخانه هیرمند تشریح خواهد شد.

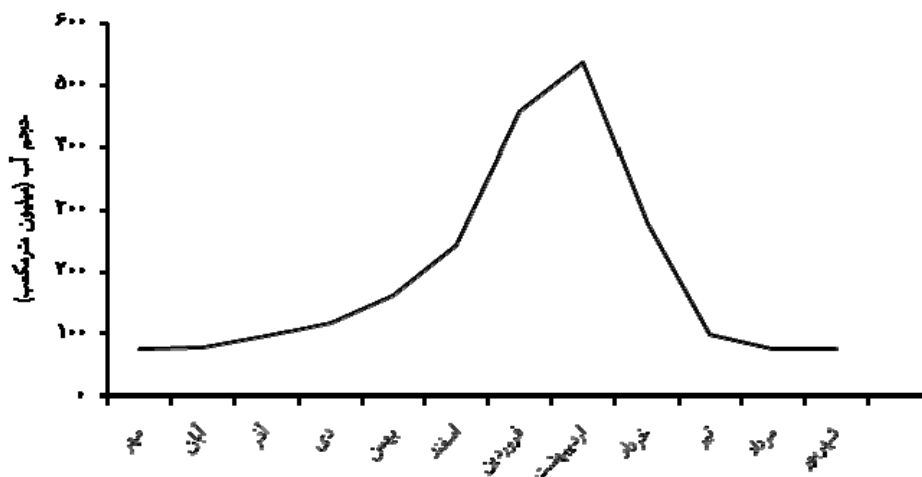
رودخانه مورد مطالعه

دشت سیستان با مساحتی افزون بر ۸۰۰۰ کیلومتر مربع در شمال استان سیستان و بلوچستان از رسوبات رودخانه هیرمند در طول هزاران سال به وجود آمده است. این رودخانه با حوضه آبریزی به وسعت ۳۵ هزار کیلومتر مربع بزرگ‌ترین رود فلات ایران و کشور افغانستان است (۲).

رودخانه هیرمند تنها ورودی آب به منطقه سیستان است که در ورود به ایران به دو رودخانه سیستان و پریان تقسیم می‌شود. همان‌طور که در شکل ۱ مشاهده می‌شود، نوسانات زیادی در میانگین



شکل ۱- میانگین سالانه آب ورودی به رودخانه هیرمند طی سال‌های مورد مطالعه



شکل ۲- میانگین ماهانه آب ورودی به رودخانه هیرمند طی سال‌های مورد مطالعه

فصلی^۲ تبدیل خواهد شد (۱).

اغلب برای ایجاد الگوی خودرگرسیون میانگین متحرک انباشته از روش باکس جنکینز^۳ استفاده می‌شود. بر اساس این روش تشکیل الگوی خودرگرسیون میانگین متحرک انباشته دارای سه مرحله تشخیص مقدار وقفه سری زمانی، تخمین مدل مناسب و پیش‌بینی و انتخاب بهترین مدل بر اساس اصل صرفه‌جویی است. در مرحله اول مقدار درجه‌های خودرگرسیون (p)، میانگین متحرک (q) و انباشتگی (I) تعیین می‌شود. برای تخمین پارامترهای مدل، با استفاده از

پس از بررسی ایستایی، روند و اختلال‌های موجود در سری زمانی، مدل مناسب ساخته خواهد شد. یکی از مدل‌های به‌کاربرده شده در این مطالعه روش خودرگرسیون میانگین متحرک انباشته^۱ $ARIMA(p, I, q)$ است که با استفاده از داده‌های ایستا شده به صورت معادله ۲ نوشته می‌شود:

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \sum_{i=0}^q \beta_i \varepsilon_{t-i} \quad (2)$$

در آن y_{t-i} جز خودرگرسیون از درجه p و ε_{t-i} جز میانگین متحرک از درجه q هستند. اگر در معادله فوق تغییرات فصلی نیز وارد شوند، این مدل به الگوی خودرگرسیون میانگین متحرک انباشته

2- SARIMA
3- Box Jenkins

1- ARIMA

اگر تابع ρ به صورت زیر باشد کاپیولایی با نام فرانک^۵ خواهیم داشت:

$$\rho(t) = \log\left(\frac{e^{\alpha t} - 1}{e^{\alpha} - 1}\right) \quad (۷)$$

و اگر ρ به شکل زیر باشد کاپیولا از نوع گامبل-هوگارد^۶ خواهد بود:

$$\rho(t) = (-\log(t))^{\alpha} \quad (۸)$$

در توابع فوق α ارتباط بین متغیرها را بیان می کند (۱۴).

از آنجا که ارتباط و همبستگی بین متغیرها در توابع کاپیولا بسیار مهم است، برای بیان این ارتباط از ماتریس خودهمبستگی کندال^۷ تاو^۷ استفاده می شود. برای اعداد تصادفی ماتریس $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ کندال تاو به صورت زیر است:

$$\tau_N = \binom{N}{2}^{-1} \sum_{i < j} \text{sign}[(x_i - x_j)(y_i - y_j)] \quad (۹)$$

که در آن N تعداد مشاهدات، sign برابر با یک است اگر $x_i \leq x_j$ و $y_i \leq y_j$ و در غیر این صورت sign برابر -1 خواهد شد و $i, j = 1, 2, \dots, N$ هستند (۱۷).

رابطه بین داده ها در کاپیولای فرانک، کلیتون و داده های حقیقی در شکل ۳ آورده شده اند. همانطور که در شکل ۳ مشاهده می شود، داده های حقیقی دارای همبستگی در مقادیر کم دارند که شباهت زیادی با کاپیولای کلیتون دارد. این در صورتی است که کاپیولای گامبل دارای همبستگی از بالا است. از طرف دیگر چون کاپیولای فرانک مقادیر بالا و پایین همبستگی را نشان می دهد در این مطالعه از این روش نیز برای بررسی داده ها استفاده شد (۱۳).

برای مقایسه بین داده های حقیقی و شبیه سازی شده علاوه بر استفاده از میانگین، حداقل و حداکثر از آزمون ریشه میانگین مربع خطا^۸ نیز استفاده شد. این آزمون به صورت زیر می باشد:

$$RMSE = \sqrt{\frac{\sum (\hat{y}_t - y_t)^2}{n}} \quad (۱۰)$$

که در آن \hat{y}_t اعداد محاسبه شده، y_t داده های حقیقی و n تعداد مشاهدات می باشد (۲۰).

برای تخمین توابع کاپیولا مراحل زیر انجام می شود: ابتدا داده ها ایستا می شوند.

در مرحله دوم ماتریس کندال تاو تخمین زده می شود.

در نهایت با استفاده از توابع کاپیولای مختلف داده ها شبیه سازی می شوند.

داده های مورد استفاده در این مطالعه از شرکت مدیریت منابع آب سیستان طی سال های ۱۳۳۸-۱۳۸۸ جمع آوری شد. به منظور

آماره های R^2 ، میانگین مجذور مربعات خطا، آکاییک و شوارتز بهترین الگو از بین سایر الگوها مشخص شده و در پیش بینی از آن استفاده می شود (۴). پیش بینی آخرین مرحله این روش است. در این قسمت مدلی که بهترین پیش بینی را نسبت به سایر مدل ها داشته باشد، انتخاب خواهد شد. برای بررسی این موضوع بهتر است با استفاده از مدل انتخابی، قسمتی از داده های پیش بینی شده با داده های واقعی مقایسه شوند. در نهایت مدلی بهترین پیش بینی را خواهد داشت که خطای کمتری نسبت به داده های واقعی داشته باشد (۱).

دومین مدل مورد استفاده، کاپیولا است. کاپیولا توابعی هستند که توابع توزیع احتمال چند متغیره را به توابع توزیع احتمال حاشیه ای تک بعدی متصل می کنند (۱۶). فرض کنید $(X^{(1)}, \dots, X^{(d)})$ برداری از اعداد تصادفی با توزیع احتمال تاو زیر است:

$$F(x_1, \dots, x_d) = \Pr(\{X^{(1)} \leq x_1\} \cap \dots \cap \{X^{(d)} \leq x_d\}) \quad (۳)$$

تابع کاپیولا به صورت زیر تعریف می شود:

$$c: [0; 1]^d \rightarrow [0; 1] \text{ and } F(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \quad (۴)$$

تابع c برای مدل سازی ارتباط بین متغیرها استفاده می شود. متغیرها با توزیع های مختلف با استفاده از توابع کاپیولا شبیه سازی می شوند. دو نوع از انواع توابع کاپیولا که استفاده وسیعی از آن ها شده، کاپیولای ارشمیدسی^۱ و کاپیولای بیضی شکل^۲ هستند که به دلیل استفاده از کاپیولای ارشمیدسی در این مطالعه، در ادامه روش ساخت و انواع آن تشریح خواهد شد (۲۰). به این دلیل از کاپیولای ارشمیدسی استفاده شد که در مدیریت منابع آب کاربرد بیشتری داشته، به راحتی می توان آنها را ایجاد کرد، مدل های زیادی در این روش وجود دارد و انواع وابستگی بین متغیرها را در نظر می گیرد (۱۰).

فرض کنید $\rho: [0, 1] \rightarrow [0, +\infty]$ تابعی پیوسته و کاهشی است که در آن $\rho(1) = 0$ و $(-1)^k d^k \rho^{-1}(t) \geq 0$ و همه t ها و $k=1, \dots, d$ که d تعداد بعد را نشان می دهد. تابع زیر کاپیولای ارشمیدسی است (۱۴).

$$c: [0; 1]^d \rightarrow [0; 1], c(x_1, \dots, x_d) = \begin{cases} \rho^{-1}(\sum_{i=1}^d \rho(F_i(x_i))) & \text{if } \sum_{i=1}^d \rho(F_i(x_i)) \leq \rho(0), \\ 0 & \text{در غیر این صورت} \end{cases} \quad (۵)$$

حال اگر تابع مولد^۳ ρ به صورت زیر باشد کاپیولایی با نام کلیتون^۴ خواهیم داشت:

$$\rho(t) = t^{-\alpha} - 1 \quad (۶)$$

5- Frank
6- Gumbel-Hougaard
7- Kendall's tau
8- RMSE (Root Mean Square Error)

1- Archimedean
2- Elliptical
3- Generator
4- Clayton

می‌باشد تخمین زده شد. هر کدام از مدل‌ها، با استفاده از آماره‌های R^2 ، میانگین مجذور مربعات خطا، آکاییک و شوارتز مقایسه و در نهایت بهترین مدل تخمین زده شد. این مدل و پارامترهای آن در جدول ۲ مشاهده می‌شود.

جدول ۲- مقادیر p و q در روش خود رگرسیون میانگین متحرک

متغیرها	ضرایب	آماره t
$AR(2)$	-0.37	$6/12$
$SAR(12)$	-0.28	$7/0.5$
$MA(1)$	-0.51	$13/0.2$
	R^2	0.4
	آکاییک	$13/1$
	شوارتز	$13/0.2$

در جدول ۲، $AR(p)$ نشان‌دهنده تعداد وقفه‌های خود رگرسیونی، $SAR(p)$ مقدار وقفه الگوی خودرگرسیونی فصلی و $MA(q)$ مقدار میانگین متحرک می‌باشند که ضرایب همه آن‌ها معنادار است. در مرحله آخر که هدف پیش‌بینی جریان رودخانه است، با استفاده از این مدل پیش‌بینی انجام شد. سپس آماره‌های مختلف از قبیل میانگین، حداقل، حداکثر و انحراف معیار برای آن‌ها محاسبه و با داده‌های واقعی برای دو دوره مقایسه شد. این نتایج طی دو دوره که دوره اول از سال ۱۳۷۵ تا ۱۳۸۸ و دوره دوم ۱۳۸۲ تا ۱۳۸۸ می‌باشد، در جدول ۳ مشاهده می‌شود.

با استفاده از این داده‌ها نمودارهای متعلق به دو دوره ترسیم و باهم مقایسه شدند. همان‌طور که در شکل‌های ۴ و ۵ مشاهده می‌شود، داده‌های پیش‌بینی شده نوسانات و مقادیر کم جریان رودخانه را به خوبی نشان می‌دهند، ولی حداکثر جریان داده‌های بازسازی‌شده با جریان حقیقی رودخانه شباهت زیادی ندارد. به همین دلیل میانگین محاسبه شده بیشتر از میانگین داده‌های واقعی هستند. مقدار ریشه میانگین مربعات خطا در کل دوره 0.4 می‌باشد که نشان‌دهنده خوبی برازش است ولی طبق محاسبات انجام‌شده همبستگی داده‌های ایجادشده با همبستگی داده‌های تاریخی مطابقت ندارد. لذا این داده‌ها مجدداً با استفاده از روش کاپیولا بازسازی خواهند شد.

برای شبیه‌سازی داده‌ها با استفاده از روش کاپیولا از داده‌های ماهانه هر سال استفاده می‌شود. بدین دلیل بررسی تغییرات روند، دوره‌ای بودن^۳ و ایستایی آن‌ها جداگانه مورد بررسی قرار گرفت. با توجه به نتایج، اکثر ماه‌ها دارای روند زمانی، دوره تناوب سالانه و نوسان شدید هستند. پس از حذف این موارد، ایستایی داده‌ها بررسی شد.

مدل‌سازی جریان رودخانه هیرمند، مقادیر موجود ابتدا توسط نرم‌افزار Eviews7 با استفاده از روش خودرگرسیون میانگین متحرک مدل-سازی و سپس با روش کاپیولا توسط نرم‌افزار^۱ شبیه‌سازی شد. در نهایت این دو روش باهم مقایسه و بهترین روش برای ایجاد داده‌های جریان رودخانه انتخاب شد.

نتایج و بحث

به منظور مدل‌سازی جریان رودخانه هیرمند، ابتدا وجود یا عدم وجود روند و چرخه در داده‌ها مشخص و سپس با استفاده از آزمون دیکی فولر تعمیم‌یافته، ایستایی داده‌های جریان رودخانه به صورت ماهانه و سری زمانی سالانه بررسی شد.

بدین منظور، از سال به عنوان روند، سینوس و کسینوس برای بررسی سیکل و فصلی بودن^۲ داده‌ها و متغیرهای موهومی برای از بین بردن نوسانات شدید جریان رودخانه در سری زمانی ماهانه استفاده شد. نتایج نشان داد، داده‌های جریان رودخانه علاوه بر داشتن روند، دارای سیکل سالانه و ماهانه و نوسانات زیاد در ماه‌های فروردین و اردیبهشت بوده که پس از حذف این موارد ایستایی آنها بررسی شد. با توجه به جدول ۱ از آنجا که فرض صفر عدم ایستایی داده‌ها است، t محاسبه شده از t جدول در هر سطحی بالاتر بوده و در نتیجه داده‌های ماهانه ایستا هستند. چون داده‌ها ایستا هستند انباشتگی وجود ندارد.

جدول ۱- بررسی ایستایی داده‌های رودخانه هیرمند

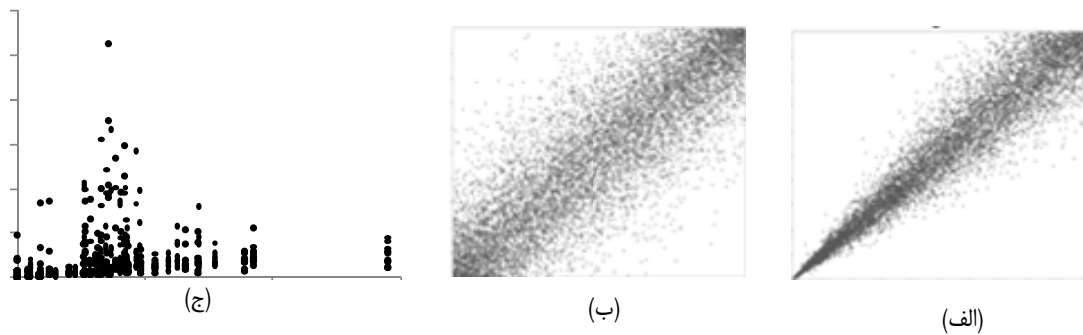
آماره آزمون	$-13/18$
مقدار بحرانی در سطح ۵ درصد	$-1/94$

سپس با استفاده از مدل باکس جنکینز داده‌های سالانه جریان رودخانه طی مراحل زیر شبیه‌سازی شد:

در مرحله تشخیص، با استفاده از نمودار خود همبستگی و خود همبستگی جزئی تعداد وقفه‌های موجود در مدل بررسی و مقادیر p و q متناسب با آن مشخص شد. با توجه به این که هر دو نمودار دارای روند سینوسی بوده و در وقفه‌های یک تا سه دارای بیرون آمدگی‌هایی هستند، این وقفه‌ها آزمون شد. از طرف دیگر به دلیل مشاهده روند فصلی در داده‌ها، تغییرات فصلی مدل خود رگرسیون میانگین متحرک نیز بررسی شد.

در مرحله تخمین، مدل‌هایی که دارای p و q های متفاوت و مقادیر مختلف فصلی برای جز خود رگرسیون و میانگین متحرک

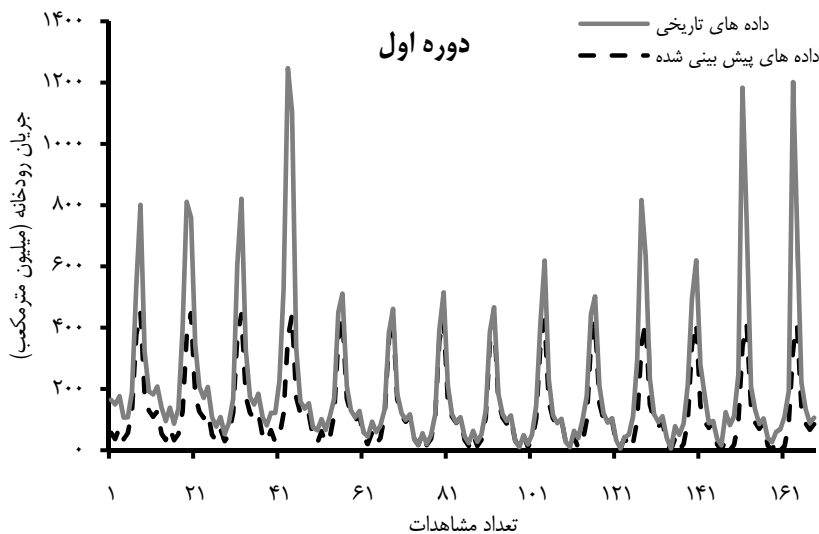
- 1- Simetar
- 2- Seasonality



شکل ۳- همبستگی داده ها در کاپیولای کلیتون (الف) کاپیولای فرانک (ب) و داده های حقیقی (ج)

جدول ۳- مقایسه داده های پیش بینی شده و تاریخی جریان رودخانه هیرمند طی دو دوره

آماره ها	دوره اول (۱۳۷۵-۱۳۸۸)		دوره دوم (۱۳۸۲-۱۳۸۸)	
	داده های تاریخی	داده های پیش بینی شده	داده های تاریخی	داده های پیش بینی شده
میانگین	۸۷/۰۲	۱۲۶/۰۳	۷۹/۱	۸۳/۶
حداقل
حداکثر	۸۸۰/۲	۴۵۰/۲	۸۶۶/۸	۸۳۱/۶
انحراف معیار	۱۴۳/۳	۱۲۸/۵	۱۴۵/۴	۱۲۸/۱
ریشه میانگین مربعات خطا	۱/۳۳	۱/۳۳	۰/۹۸	۰/۹۸
تعداد مشاهدات	۱۶۸	۱۶۸	۸۴	۸۴



شکل ۴- مقادیر واقعی و پیش بینی شده جریان رودخانه هیرمند در دوره اول

همگامی^۱ دارد.

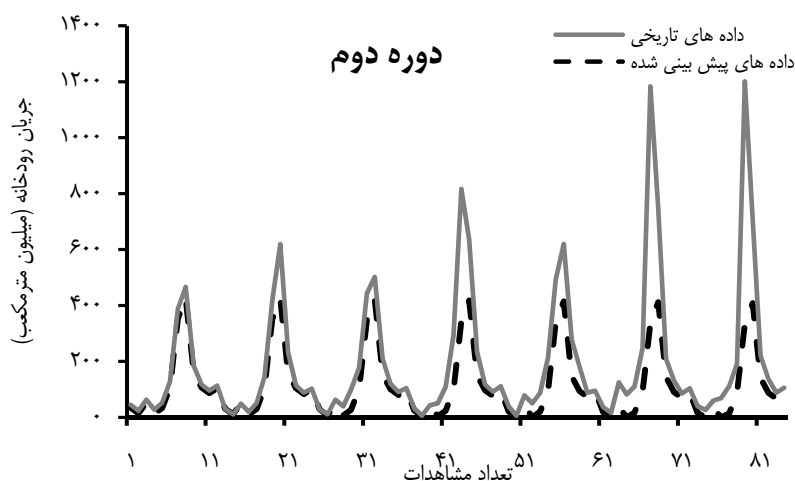
با استفاده از داده های ایجاد شده در قسمت های قبل، از کاپیولاهای کلیتون و فرانک برای مدل سازی استفاده شد. سپس داده ها، ۱۲۰ بار شبیه سازی شده و نتایج با داده های واقعی مقایسه شدند. این مقایسه در جدول ۵ گزارش شده است.

نتایج نشان داد، همه داده ها ایستا بوده و با استفاده از داده های روند زدایی شده و با حذف دوره تناوب و نوسانات شدید از آن ها می توان جریان رودخانه را شبیه سازی نمود.

با استفاده از ماتریس کندال - تاو، ارتباط و همبستگی بین داده ها مشخص شد. این ماتریس در جدول ۴ نشان داده شده است.

با توجه به این جدول در اکثر ماه ها جریان رودخانه با سایر ماه ها

1- Concordance



شکل ۵- مقادیر واقعی و پیش‌بینی شده جریان رودخانه هیرمند در دوره دوم

جدول ۴- ماتریس همگامی کندال-تاو جریان ماهانه رودخانه هیرمند

	مهر	آبان	آذر	دی	بهمن	اسفند	فروردین	اردیبهشت	خرداد	تیر	مرداد	شهریور
مهر	۱	۰/۸ (۰)	۰/۶۶ (۰)	۰/۵۱ (۰)	۰/۳۷ (۰/۰۳)	۰/۳۱ (۰/۳۷)	۰/۱۲ (۰/۶۹)	۰/۲۱ (۰/۳۸)	۰/۳۳ (۰/۱۵)	۰/۴۱ (۰/۰۰۲)	۰/۴۷ (۰)	۰/۴۸ (۰)
آبان		۱	۰/۷۴	۰/۶ (۰)	۰/۴۲ (۰)	۰/۲۸ (۰/۲۹)	۰/۱۴ (۰/۹۲)	۰/۲۳ (۰/۱۷)	۰/۳۵ (۰/۰۳)	۰/۴۲ (۰)	۰/۵۱ (۰)	۰/۴۶ (۰)
آذر			۱	۰/۷۲ (۰)	۰/۴۶ (۰/۰۰۴)	۰/۳۲ (۰/۳۵)	۰/۱۶ (۰/۸۳)	۰/۲۶ (۰/۲۶)	۰/۳۴ (۰/۲۲)	۰/۴۴ (۰/۰۰۲)	۰/۵۲ (۰)	۰/۴۵ (۰)
دی				۱	۰/۵۸ (۰)	۰/۳۳ (۰/۵۵)	۰/۱۸ (۰/۶۵)	۰/۲۵ (۰/۳۲)	۰/۳۴ (۰/۱۶)	۰/۴۵ (۰/۰۰۶)	۰/۴۸ (۰)	۰/۴۲ (۰/۰۰۴)
بهمن					۱	۰/۵۱ (۰)	۰/۳۴ (۰)	۰/۳۱ (۰/۲۶)	۰/۳۸ (۰/۲۲)	۰/۴۵ (۰/۰۰۳)	۰/۳۸ (۰/۹۵)	۰/۳۹ (۰/۴۴)
اسفند						۱	۰/۵۸ (۰)	۰/۵۳ (۰/۰۰۰۲)	۰/۵ (۰/۰۰۰۳)	۰/۴۷ (۰/۰۲)	۰/۳۷ (۰/۰۸)	۰/۴ (۰)
فروردین							۱	۰/۶۵ (۰)	۰/۴۹ (۰)	۰/۳۶ (۰)	۰/۲۵ (۰/۹۵)	۰/۲۵ (۰/۴۴)
اردیبهشت								۱	۰/۶۷ (۰)	۰/۵ (۰)	۰/۳۸ (۰/۰۲)	۰/۴ (۰)
خرداد									۱	۰/۷۱ (۰)	۰/۵۷ (۰)	۰/۵۷ (۰)
تیر										۱	۰/۶۸ (۰)	۰/۶۳ (۰)
مرداد											۱	۰/۷۹ (۰)
شهریور												۱

سلول‌های هاشور خورده به معنای معناداری ارتباط بین متغیرها می‌باشد. مقادیر داخل پرانتز احتمال (p-value) هستند.

با توجه به جداول ۵ و ۶، میانگین داده‌های شبیه‌سازی شده شباهت زیادی به داده‌های واقعی دارند.

از آنجا که نتایج داده‌های شبیه‌سازی شده در دو روش کاپیولای فرانک و کلیتون شباهت زیادی باهم دارند، در ادامه تنها از نتایج کاپیولای کلیتون استفاده می‌شود.

جدول ۵- مقایسه داده‌های شبیه‌سازی شده با کاپیولای کلیتون و فرانک و جریان رودخانه هیرمند طی ماه‌های مختلف

	کاپیولای کلیتون			کاپیولای فرانک			داده‌های واقعی		
	میانگین	حداکثر	حداقل	میانگین	حداکثر	حداقل	میانگین	حداکثر	حداقل
مهر	۷۵/۳	۱۷۴/۵	۰	۷۵/۶	۱۷۴/۴	۰	۷۶/۳	۲۹۰/۲	۰
آبان	۷۶/۱	۱۸۱/۳	۰	۷۵/۴	۱۸۱/۳	۰	۷۸/۵	۱۹۹/۷	۰
آذر	۱۱۱/۳	۷۳۳/۱	۰/۶	۱۰۸/۵	۷۳۳/۱	۰/۶	۹۷/۴	۴۵۰/۲	۰/۲
دی	۱۲۹/۱	۵۱۹/۸	۸/۷	۱۲۸/۱	۵۱۹/۸	۸/۷	۱۱۷/۱	۴۰۵/۱	۳/۳
بهمن	۱۶۰/۲	۱۱۲۰/۵	۱۸/۹	۱۶۴/۹	۱۱۲۰/۵	۱۸/۹	۱۶۱/۶	۴۹۷/۱	۱۲/۳
اسفند	۲۴۲/۵	۱۱۰۰/۶	۸/۲	۲۳۸/۸	۱۱۰۰/۶	۸/۲	۲۴۳/۱	۱۰۳۶/۴	۷/۱
فروردین	۴۵۶/۱	۳۴۹۸/۸	۲۱/۸	۴۵۱/۱	۳۴۹۸/۸	۲۱/۸	۴۵۶/۹	۲۶۲۷/۸	۲۴/۲
اردیبهشت	۵۳۳/۳	۱۹۲۳/۷	۳۳/۳	۵۳۱/۱	۱۹۲۳/۷	۳۳/۳	۵۳۶/۸	۱۷۶۱/۸	۲۷/۳
خرداد	۲۷۱/۱	۱۱۲۷/۶	۳۰/۶	۲۷۰/۱	۱۰۹۶/۹	۳۰/۶	۲۷۷/۴	۱۰۱۱/۹	۱۷/۱
تیر	۱۱۰/۹	۵۱۰/۴	۱۹/۸	۱۰۸/۶	۵۱۰/۴	۱۹/۸	۹۸/۵	۳۹۶/۱	۷/۷
مرداد	۷۹/۴	۴۱۸/۵	۰	۷۸/۴	۴۱۸/۵	۰	۷۶/۲	۲۲۸/۳	۰/۹
شهریور	۷۵/۶	۲۲۲/۵	۰/۲	۷۵/۶	۲۲۲/۵	۰/۲	۷۴/۶	۱۹۹/۲	۰/۱

توسط هوپ و بارت (۱۱)، دلیری و خلقی (۳)، کیخا (۱۳) و فاوور و همکاران (۹) مورد استفاده قرار گرفته، برای مدل‌سازی جریان رودخانه مناسب تر بوده و علاوه بر ایجاد داده‌هایی با شباهت بیشتر به داده‌های واقعی، همبستگی بین آنها را نیز حفظ می‌کند.

می‌توان گفت، روش خودرگرسیون قادر به پیش‌بینی مقادیر کم جریان رودخانه می‌باشد ولی حداکثر جریان رودخانه را بیشتر از حد واقعی محاسبه می‌کند که این امر منجر به افزایش مقادیر میانگین می‌شود.

همچنین پس از مقایسه واریانس و همبستگی داده‌های اولیه با داده‌های ایجادشده، مشاهده شد (جدول ۶ و ۷) روش کاپیولا همبستگی، وابستگی، میانگین و واریانس داده‌ها را به خوبی حفظ می‌کند. می‌توان گفت این روش، در مدل‌سازی جریان ماهانه رودخانه هیرمند نتایج بهتری ارائه داده و کاربرد آن مناسب‌تر از روش اول خواهد بود.

با وجودی که وانگ (۱۷)، دوگان و همکاران (۸) و طرازکار و صدق‌آمیز (۵) روش‌های پیش‌بینی را برای مدل‌سازی جریان رودخانه مناسب دانستند در این مطالعه مشاهده شد روش شبیه‌سازی که

جدول ۶- مقایسه میانگین و واریانس داده‌های ایجادشده با کاپیولا و واقعی

مقایسه میانگین (آزمون t)	مقایسه واریانس (آزمون F)	
-۰/۲۲	۱/۱۶	مهر
-۰/۱۵	۱/۱۲	آبان
-۱/۱۸	۱/۱۵	آذر
-۰/۹۹	۱/۲۱	دی
۰/۰۷	۱/۰۶	بهمن
-۰/۰۲	۱/۰۵	اسفند
۰/۰۱	۱/۱	فروردین
-۰/۰۵	۱/۰۱	اردیبهشت
-۰/۱۸	۱/۳۳	خرداد
-۱/۱	۱/۱۲	تیر
-۰/۳۵	۱/۰۱	مرداد
-۰/۱۳	۱/۱۸	شهریور
۲/۳	۱/۴۶	مقادیر بحرانی

جدول ۷- معناداری همبستگی بین داده‌های ماهانه شبیه‌سازی شده با کاپیولا و جریان واقعی رودخانه هیرمند

	شهریور	مرداد	تیر	خرداد	اردیبهشت	فروردین	اسفند	بهمن	دی	آذر	آبان
مهر	۳/۲۸	۳/۶۴	۴/۰۲	-۰/۵۲	-۰/۹۴	-۰/۴۵	-۰/۴۳	۱/۳۹	۲/۷	۱/۸۱	۱/۹۱
آبان	۴/۲۸	۴/۶۵	۴/۶۸	-۰/۶۲	-۰/۸۳	-۰/۶۷	-۰/۴۴	۱/۴	۳/۱	۲/۲۶	
آذر	۳/۶۲	۳/۹۳	۴/۲۷	۲/۶	-۰/۹۱	-۰/۴۴	-۰/۵۹	۱/۴۳	۱/۴۷		
دی	۴/۳۴	۴/۶۵	۴/۷۷	۳/۰۳	-۰/۹۳	-۰/۶	-۰/۴۹	۱/۲۱			
بهمن	۱/۲۱	۱/۸۲	۱/۳۹	-۰/۶۸	-۰/۲۴	-۰/۷	۰				
اسفند	-۰/۰۸	-۰/۵۲	-۰/۱۲	۰	-۰/۰۶	-۰/۲					
فروردین	۱/۰۹	-۰/۷۱	-۰/۸۷	-۰/۷۷	-۰/۳۶						
اردیبهشت	-۰/۱۵	-۰/۳۹	-۰/۲۴	-۰/۰۲							
خرداد	۱/۴۴	۱/۶۱	-۰/۵۵								
تیر	۲/۷۱	۲/۱۳									
مرداد	۱/۷۷										

سلول‌های هاشور خورده به معنای معناداری همبستگی بین متغیرها می‌باشد (مقدار بحرانی با احتمال ۹۹ درصد = ۳/۳۸)

با توجه به نتایج به دست آمده می‌توان گفت، روش خودرگرسیون میانگین متحرک فصلی گرچه با واردکردن عامل تصادفی، داده‌هایی مانند داده‌های اصلی ایجاد می‌کند ولی نمی‌تواند همبستگی بین داده‌ها در ماه‌های مختلف سال را حفظ نماید. در روش کاپیولا این مشکل رفع شده و داده‌های شبیه‌سازی شده علاوه بر داشتن میانگینی مانند جریان واقعی رودخانه، همبستگی و وابستگی مشابه جریان تاریخی دارند. لذا می‌توان گفت کاپیولا روشی مناسب برای شبیه‌سازی جریان تصادفی رودخانه هیرمند است؛ که پیشنهاد می‌شود از داده‌های ایجادشده در روش کاپیولا برای مدل‌سازی سیستم منابع آب منطقه و در نتیجه تصمیمات آینده استفاده شود. همچنین از این روش برای شبیه‌سازی جریان سایر رودخانه‌ها استفاده‌شده و نتایج آن با روش‌های دیگر مقایسه شود. از طرف دیگر، کاربرد سایر روش‌های کاپیولا برای مدل‌سازی جریان رودخانه می‌تواند موضوع تحقیقات آتی باشد.

همچنین این روش همبستگی بین داده‌های جریان رودخانه را حفظ نمی‌کند. پس این روش برای مدل‌سازی جریان رودخانه هیرمند مناسب نخواهد بود.

از طرف دیگر، آزمون‌های مختلف نشان داد، روش کاپیولا جریان ماهانه رودخانه را مشابه داده‌های واقعی نشان می‌دهد و مقادیر حداکثر و حداقل با جریان واقعی تفاوت معناداری ندارند. همچنین در این روش همبستگی بین داده‌ها حفظ‌شده و داده‌هایی مشابه داده‌های حقیقی ایجاد می‌شود. کاپیولا دارای میانگین مربع خطای کمتری بوده که این عوامل منجر به بهتر بودن کاپیولا نسبت به روش خودرگرسیون در مدل‌سازی جریان رودخانه هیرمند می‌باشد.

نتیجه‌گیری

در این مطالعه به منظور مدل‌سازی جریان رودخانه هیرمند، از روش‌های خودرگرسیون میانگین متحرک فصلی و کاپیولا استفاده شد و این روش‌ها برای انتخاب بهترین مدل باهم مقایسه شدند.

منابع

- ۱- ابریشمی ح. ۱۳۸۳. مبانی اقتصادسنجی (ترجمه). چاپ سوم، انتشارات دانشگاه تهران، صفحات ۹۰۷ تا ۹۲۵.
- ۲- خلقی اشکلک م. ۱۳۸۶. اندازه‌گیری میزان رادون و رادیوم آب منطقه سیستان. پایان‌نامه کارشناسی ارشد. گروه فیزیک. دانشکده علوم دانشگاه پیام نور، واحد مشهد.
- ۳- دلیری ف.، و خلقی م. ۱۳۸۸. اصلاح روش IUDRN به منظور شبیه‌سازی استوکاستیکی دبی سالانه رودخانه‌ها (مطالعه موردی: رودخانه اریه استان خراسان رضوی). علوم و مهندسی آبخیزداری ایران. ۳ (۶)، ۸-۱.
- ۴- اندرس و. ۱۳۸۶. اقتصادسنجی سری‌های زمانی با رویکرد کاربردی (صادقی م.، و شوال پور س.). چاپ اول، انتشارات دانشگاه امام صادق، ۵۲۶ ص.
- ۵- طرازکار م.، و صدق آمیز ع. ۱۳۸۷. مقایسه پیش‌بینی دبی جریان ماهانه رودخانه کرخه با استفاده از روش‌های سری زمانی و هوش مصنوعی. پژوهش و سازندگی در منابع طبیعی. ۸۰، ۵۱-۵۸.

- ۶- نوفرستی م. ۱۳۷۸. ریشه واحد و همجمعی در اقتصادسنجی. چاپ اول، موسسه خدمات فرهنگی رسا، ۱۸۳ ص.
- ۷- وزارت نیرو. ۱۳۹۰. گزارش برنامه‌ریزی منابع آب رودخانه و مخازن چاه نیمه‌های سیستان. جلد دوم. شرکت سهامی آب منطقه‌ای استان سیستان و بلوچستان، زابل.
- 8- Dogan I., Toluk T., and Sandalci M. 2007. Daily stream flow forecasting using artificial neural networks, International Congress River Flood Management, 448-459.
- 9- Favre A.C., El Adlouni S., Perreault L., Thiemonge N., and Bobee B. 2004. Multivariate hydrological frequency analysis using copulas, Water Resource Research, 40:1-12.
- 10- Genest C., and Mackay L. 1986. The joy of copulas: Bivariate distributions with uniform marginals, The American Statistician, 40 (4), 280-283.
- 11- Hardaker J.B., Hurine R.B.M., Anderson J. R., and Lein G. 2004. Coping with Risk in Agriculture. 2nd den, CABI, Wallingford, Oxford, U.K.
- 12- Hope A., and Bart R. 2012. Evaluation of a regionalization approach for daily flow duration curves in central and southern California watersheds, Journal of the American Water Resources Association (JAWRA), 48(1): 123-133.
- 13- <http://www.vosesoftware.com/index.php>
- 14- Kao S.C., and Govindaraju R.S. 2010. A copula-based joint deficit index for droughts, Journal of Hydrology, 380: 121-134.
- 15- Kehkha A.A. 2005. Modeling water resources management in the Sistan region of Iran, Thesis for the degree of Doctor of Philosophy, University of New England.
- 16- Rémillard B., Papageorgiou N., and Soustra F. 2012. Copula-based semi parametric models for multivariate time series, Journal of Multivariate Analysis, 110:30-42.
- 17- Renard B., and Lang M. 2007. Use of a Gaussian copula for multivariate extreme value analysis: Some case studies in hydrology, Advances in Water Resources, 30: 897-912.
- 18- Van der Lee J.J. 2002. A decision support framework for participative and integrated river basin management: an application of the triple bottom line, Ph.D. thesis, University of New England, Australia.
- 19- Wang W. 2005. Testing and modeling autoregressive conditional heteroskedasticity of stream flow process, Nonlinear processes in Geophysics, 12: 55-66.
- 20- Zhang L., and Singh V.P. 2007. Bivariate rainfall frequency distributions using Archimedean copulas, Journal of Hydrology, 332, 93-109.

Comparison Copula and Seasonal Auto Regressive Moving Average Methods in Helmand River Flow Modeling

A.A. Keikha¹- M. Mosannan Mozafari^{2*}- M. Sabouhi³- Gh. Soltani⁴

Received:03-12-2013

Accepted:24-06-2014

Abstract

River flow modeling has special importance in water resources management. Since the actual river flow data are often low and they correlate and depend yearly and monthly, making the data similar to historical data is so difficult and complex. In this study, 50 year data and Seasonal Auto Regressive Moving Average (SARMA) and Clayton and Frank Copulas which are the prediction and simulation methods of the river flow molding, were used to generate random flow data of Helmand River. Results show, SARMA model forecasts minimum river flow data very good, but the generated data hasn't correlation of historical data and usually the maximum river flow is greater than real data. Otherwise, Copula preserved concordance of real data and make the data that are similar to real river flow. Also Root Mean Square Error of Copula method was 0.3 that is was less than SARMA method (0.4). Therefore Copulas are good methods for Helmand river flow modeling.

Keywords: Simulation, Prediction, Water Resources Management, Correlation

1- Associate Professor of Agricultural Economic Department, Faculty of Agriculture, University of Zabol
2,3- PhD Graduated and Associate Professor of Agricultural Economic Department, Ferdwosi University of Mashhad
(*- Corresponding Author Email: mmosannan@gmail.com)
4- Professor of Agricultural Economic Department, University of Shiraz