

## تحلیل تغییرات دقت برازش سری‌های متعامد نرمال به اندازه نمونه دبی اوج سالانه

میثم سالاری جزی<sup>1</sup>

تاریخ دریافت: 1395/3/19 تاریخ پذیرش: 1395/9/3

### چکیده

مطالعات تحلیل ریسک سیلاب بر اساس برآورد تابع چگالی احتمال دبی اوج سیلاب صورت می‌گیرد. روش‌های مرسوم برای تخمین تابع چگالی احتمال توزیع‌های پارامتری هستند که دارای محدودیت‌های مختلف از جمله حساسیت در دقت برازش نسبت به کاهش اندازه نمونه می‌باشند. روش‌های ناپارامتری برآورد تابع چگالی احتمال دارای محدودیت‌های کمتری نسبت به توزیع‌های پارامتری می‌باشند. در این مطالعه از توزیع‌های پارامتری نمایی، لوگ نرمال، گاما و گامبل و روش سری‌های متعامد نرمال به عنوان یک روش نوین ناپارامتری و همچنین معیارهای دقت برازش معیار اطلاعات آکائیک (AIC)، معیار اطلاعات بیزی (BIC) و میانگین مربعات خطا (MSE) استفاده شد تا دقت این روش‌ها و نیز حساسیت آن‌ها به کاهش اندازه نمونه دبی اوج سالانه سیلاب بررسی گردد. بر اساس داده‌های ثبت شده پنج ایستگاه هیدرومتری در استان گلستان (با طول دوره آماری 38 تا 48 سال) دقت و تغییرات دقت برازش توزیع‌های پارامتری و ناپارامتری به کاهش اندازه نمونه‌ها بررسی شد. نتایج نشان داد دقت روش ناپارامتری سری‌های متعامد نرمال از توزیع‌های پارامتری به شکل محسوسی بالاتر می‌باشد. به علاوه روش سری‌های متعامد نرمال در مقایسه با روش‌های پارامتری دارای حساسیت کمتر به کاهش اندازه نمونه می‌باشد که این ویژگی موجب می‌شود بتواند به عنوان یک گزینه مناسب در مواردی که داده‌های ثبت شده کوتاه مدت است به حساب آید.

**واژه‌های کلیدی:** اندازه نمونه، تحلیل حساسیت، توزیع‌های پارامتری، سری‌های متعامد نرمال

### مقدمه

هیدرولوژیکی وجود ندارد (Smakhtin., 2001; Karmakar and Simonovic., 2008; Adamowski., 1985) بر اساس آزمون‌های دقت برازش ممکن است چندین توزیع به خوبی بر داده‌ها منطبق شوند اما در اغلب موارد توزیع‌های مختلف منجر به برآوردهای متفاوت می‌گردند. مزیت روش‌های ناپارامتری عدم نیاز به فرضیات مربوط به توزیع جامعه مورد نظر می‌باشد. روش‌های ناپارامتری برای داده‌های برگرفته از هر جامعه آماری با هر توزیع احتمال معتبر هستند (Haghighat-Jou., 2008). برآوردهای ناپارامتری تابع، با استفاده از میانگین‌های متحرک وزنی داده‌ها در یک همسایگی کوچک نقطه‌ای که برآورد آن انجام می‌شود، این مزیت را دارد که همیشه ویژگی‌های بیان شده توسط نمونه را باز تولید می‌کند (Kim., 2003; Lall., 1995). مطالعات انجام شده در زمینه کاربرد روش‌های ناپارامتری در هیدرولوژی محدود است. این مطالعات نشان داده است که توابع چگالی ناپارامتری انعطاف‌پذیری بالایی دارند و با توزیع تجربی متغیرهای تصادفی در طبیعت می‌توانند به خوبی تطابق نمایند و به ویژه برای توزیع‌های چند نمایی مناسب می‌باشند (Quintela-Del-Río., 2011). مطالعات در زمینه کاربرد روش‌های ناپارامتری در تحلیل سیلاب عمدتاً به کاربرد روش برآورد تابع چگالی احتمال کرنل

رویداد سیلاب یکی از پدیده‌های هیدرولوژیکی است که از جهات متفاوت در مهندسی منابع آب مورد بررسی قرار می‌گیرد (Adib et al., 2010; Adib et al., 2011; Eidipour et al., 2016). این پدیده می‌تواند اثرات متفاوت زیست محیطی در حوضه داشته باشد و از این رو تحلیل‌های مبتنی بر ریسک سیلاب اهمیت می‌یابد (Adib et al., 2010). در این زمینه تحلیل‌های مبتنی بر ریسک عمدتاً بر تحلیل ریسک دبی اوج سیلاب مبتنی می‌باشد (Adamowski., 2000). عمده تلاش‌ها در این زمینه بر کاربرد توزیع‌های پارامتری در تحلیل ریسک دبی اوج سیلاب متمرکز شده و تعداد زیادی از توزیع‌های پارامتری برای تحلیل فراوانی داده‌های سیلاب مورد استفاده قرار گرفته است (Kidson and Richards., 2005). نتیجه بسیاری از مطالعات بیان‌کننده این نکته است که یک تابع توزیع احتمال یا یک فرایند برازش پذیرفته شده عمومی برای متغیرهای

1- استادیار گروه مهندسی آب، دانشکده مهندسی آب و خاک، دانشگاه علوم کشاورزی و منابع طبیعی گرگان

Email: meysam.salarijazi@gau.ac.ir

سری‌های متعامد نرمال و توزیع‌های مرسوم پارامتری را به کاهش اندازه نمونه داده‌های ثبت شده دبی اوج سیلاب بررسی و مقایسه نماید.

### مواد و روش‌ها

**داده‌های مورد استفاده:** بررسی کارایی روش‌های ناپارامتری و پارامتری مورد استفاده در تحقیق به تغییرات اندازه نمونه داده‌های دبی اوج، نیازمند آنست که از مجموعه‌ای از داده‌ها که دارای دوره آماری نسبتاً بلند مدت هستند استفاده گردد. در این تحقیق از داده‌های ثبت شده پنج ایستگاه هیدرومتری استفاده شده است. ایستگاه‌های مورد استفاده در این تحقیق در حوضه آبریز گرگانود در استان گلستان در شمال ایران واقع شده‌اند. داده‌های ثبت شده دبی جریان رودخانه ایستگاه‌های تمر، رامیان، زرین گل، سرمو و لزوره که در این تحقیق مورد استفاده قرار گرفته‌اند به ترتیب دارای دوره آماری ثبت شده 46 سال، 48 سال، 38 سال، 39 سال و 47 سال می‌باشند. داده‌های مورد استفاده از ابتدای دوره ثبت داده‌ها تا سال 1394 می‌باشد که دارای خلا آماری نبوده است.

**توزیع‌های پارامتری:** توزیع‌های پارامتری متفاوتی در هیدرولوژی مورد بررسی و کاربرد قرار گرفته است (Moazed et al., 1999; Hamed and Rao., 2012). در این مطالعه با در نظر گرفتن متغیر مورد بررسی یعنی دبی اوج توزیع‌های نمایی، گاما، گامبل یا حدی نوع 1 و لوگ نرمال مورد بررسی قرار گرفته است. توزیع‌های مورد بررسی و جزییات آن‌ها در جدول 1 ارائه گردیده است.

مرتبط می‌شود (Adamowski and Feluch., 1990; Shabri., 2012; Reddy and Ganguli., 2002). روش سری‌های متعامد نرمال نیز یکی از روش‌های ناپارامتری است که کاربرد آن در مباحث هیدرولوژی و محیط زیست به تازگی مورد توجه قرار گرفته است. کاربرد روش سری‌های متعامد نرمال در هیدرولوژی با کار گوش و ماژومدار شروع شده است که در تحقیقات مرتبط با خشکسالی از این روش استفاده کرده‌اند (Ghosh and Majumdar., 2007). همچنین کارماکار و سیمونویچ با بررسی روش سری‌های متعامد نرمال این روش را به عنوان یک روش منعطف در تخمین تابع چگالی احتمال سیلاب توصیف نموده‌اند (Karmakar and Simonovic., 2008). همچنین در تحقیقی دیگر ویتالروش ناپارامتری سری‌های متعامد نرمال را در تحلیل سیلاب نسبت به روش برآورد چگالی احتمال کرنل ارجح دانسته‌اند (Vittal et al., 2015). بررسی منابع مختلف در این زمینه نشان می‌دهد روش ناپارامتری سری‌های متعامد نرمال از نظر تئوری در مواردی که اندازه نمونه کوچک است نسبت به سایر روش‌هایی تواند به تخمین مناسب‌تری از تابع چگالی احتمال دست یابد (Bowman and Azzalini., 1997; Ghosh and Mujumdar., 2007; Karmakar and Simonovic., 2008). با توجه به اینکه در کشورهای در حال توسعه در اکثر موارد ایستگاه‌های هیدرومتری دارای دوره آماری بلند مدت ثبت شده نیستند استفاده از روشی که به طور نسبی کمتر تحت‌تأثیر اندازه‌های کوچک‌تر نمونه‌ها قرار بگیرد دارای اهمیت می‌باشد. با توجه به اینکه تاکنون مطالعه‌ای در زمینه تحلیل حساسیت دقت روش ناپارامتری سری‌های متعامد نرمال به اندازه نمونه داده‌های دبی اوج سیلاب انجام نشده است هدف این مطالعه آن است که با استفاده از داده‌های ثبت شده در پنج ایستگاه هیدرومتری، تغییرات دقت روش ناپارامتری

جدول 1- توابع توزیع پارامتری مورد استفاده در مطالعه

نام توزیع	پارامترها	تابع چگالی احتمال
نمایی	$\eta$	$f_X(x) = \frac{1}{\eta} e^{-\frac{x}{\eta}}; x > 0$
گاما	$\lambda, \beta$	$f_X(x) = \frac{\lambda^\beta x^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)}; x \geq 0, \lambda > 0, \beta > 0, \Gamma(\beta) = \int_0^\infty u^{\beta-1} e^{-u} du$
گامبل (مقادیر حدی نوع 1)	$\varepsilon, \alpha$	$f_X(x) = \frac{1}{\alpha} \exp[-\frac{x-\varepsilon}{\alpha} - \exp(-\frac{x-\varepsilon}{\alpha})]$
لوگ نرمال	$\mu_y, \sigma_y$	$f_X(x) = \frac{1}{x\sqrt{2\pi\sigma_y}} \exp[-\frac{(y-\mu_y)^2}{2\sigma_y^2}]; y = \log x, x > 0, -\infty < \mu_y < +\infty, \sigma_y > 0$

واقعی نرمال است که ممکن است این فرض صحیح نباشد. برای غلبه بر مشکل برآورد چگالی کرنل روش سری‌های متعامد نرمال برای برآورد تابع چگالی احتمال ناپارامتری مورد استفاده قرار

**روش ناپارامتری سری‌های متعامد نرمال:** وقتی نمونه کوچک باشد این امکان وجود دارد برآورد کرنل کارایی پایین داشته باشد. همچنین عرض باند با این فرض برآورد می‌گردد که چگالی

(Karmakar and Simonovic., 2008).

گام دوم: برآورد ضرایب سری‌های متعامد نرمال: با استفاده از روابط 1 و 2 می‌توان توابع مناسب در روش سری‌های متعامد نرمال را به‌دست آورد. ضرایب  $\theta_j$  ذکر شده در رابطه 3 را می‌توان با رابطه 6 به‌دست آورد که در آن  $f(y)$  تابع چگالی احتمال متغیر تصادفی تغییر مقیاس داده شده  $Y$  می‌باشد.

$$\theta_j = \int_{-\infty}^{+\infty} f(y)\Phi_j(y)dy \quad (6)$$

به عبارت دیگر  $\theta_j$  امیدریاضی  $\Phi_j(y)$  می‌باشد، یعنی  $\theta_j = E[\Phi_j(y)]$  که از یک نمونه به اندازه  $n$  (یعنی  $y_j, j=0,1,2,\dots,n$ ) به صورت زیر برآورد می‌شود.

$$\theta_j = \frac{1}{n} \sum_{l=1}^n \Phi_j(y_l) \quad (7)$$

گام سوم: برآورد برش  $J$ : تعیین برش  $J$  در رابطه 3 در روش سری‌های متعامد نرمال ضروری است.  $J$  را به صورت زیر محاسبه می‌شود (Efromovich., 1999).

$$J = \operatorname{argmin}_{0 \leq J \leq J_n} \sum_{j=0}^J (2\delta_j n^{-1} - \theta_j^2) \quad (8)$$

در رابطه فوق  $\delta_j = \theta_0 + \theta_{2j} 2^{-1/2} - \theta_j^2$  و  $J_n = [C_{J0} + C_{J1} \ln(n)]$  است. ضرایب فوریه  $\theta_j$  با افزایش  $J$  به

طور نسبتاً سریع کاهش می‌یابند بنابراین انتخاب  $\hat{\delta} = \hat{\theta}_0$  به عنوان یک برآورد برای تمامی  $\delta_j$  ها در نظر گرفته می‌شود. مقادیر پیش فرض  $C_{J0}$  و  $C_{J1}$  به ترتیب برابر با 4 و 0/5 پیشنهاد شده است (Efromovich., 1999). رابطه 8 نشان می‌دهد مقداری از  $\hat{J}$  که به

ازای آن عبارت  $\sum_{j=0}^{\hat{J}} (2\delta_j n^{-1} - \theta_j^2)$  به حداقل مقدار خود برسد به عنوان برش  $J$  در نظر گرفته می‌شود.

گام چهارم: هموار سازی تابع چگالی احتمال برآورد شده: در بسیاری از موارد هموار سازی ضرایب فوریه با ضرب آن‌ها در بعضی مقادیر ثابت که در بازه  $[0-1]$  قرار دارند مورد توجه هستند. پس از هموار سازی، یک تابع چگالی احتمال اصلاح شده به صورت رابطه 9 به‌دست می‌آید.

$$\hat{f}_J(y) = \sum_{j=0}^J w_j \theta_j \Phi_j(y); \quad 0 \leq y \leq 1, \quad 0 \leq w_j \leq 1 \quad (9)$$

که در آن وزن‌های  $w_j$  برای هموار سازی ضرایب فوریه به صورت معادلات 10 و 11 تعیین می‌گردند.

می‌گیرد (Efromovich., 1999; Bowman and Azzalini., 1997; Ghosh and Mujumdar., 2007; Karmakar and Simonovic., 2008). تابع چگالی احتمال یک نمونه کوچک را می‌توان با استفاده از روش‌سری‌های متعامد نرمال که لزوماً یک سری توابع متعامد به‌دست آمده از نمونه است، برآورد نمود. یک سری متعامد نرمال یک دنباله از توابع متعامد  $\Phi_j(x)$  و  $\Phi_s(x)$  است که در روابط زیر صدق می‌کند.

$$\int \Phi_s(x)\Phi_j(x)dx = 0 \quad \forall s \neq j \quad (1)$$

$$\int \{\Phi_j(x)\}^2 dx = 1 \quad \forall j \quad (2)$$

یک تابع چگالی تک متغیره تصادفی  $X$  را می‌توان به خوبی با سری متعامد نرمال  $\hat{f}_J(x)$  تقریب زد.

$$\hat{f}_J(x) = \sum_{j=0}^J \theta_j \Phi_j(x) \quad (3)$$

که  $J$  برش نامیده می‌شود.  $\Phi_j(x)$  ها توابع سری متعامد نرمال،  $\theta_j$  ها ضرایب متناظر با هر تابع می‌باشند. در این مطالعه توابع کسینوس از زیر مجموعه سری‌های فوریه به عنوان سری‌های متعامد نرمال انتخاب شده‌اند.

$$\Phi_0(x) = 1 \quad (4)$$

$$\Phi_j(x) = \sqrt{2} \cos(\pi j x), \quad j = 1, 2, 3, \dots$$

تعیین تابع چگالی احتمال بر مبنای سری‌های متعامد نرمال شامل گام‌های زیر است (Ghosh and Mujumdar., 2007; Karmakar and Simonovic., 2008):

گام اول: تعیین تکیه‌گاه و مقیاس‌بندی داده‌ها: روش ارایه شده برای تعیین تابع چگالی احتمال با استفاده از سری‌های متعامد نرمال در حالتی معتبر است که دامنه متغیر تصادفی  $[0-1]$  باشد. متغیر تصادفی مورد بررسی ممکن است دامنه دیگری مانند  $[a-b]$  داشته باشد و بنابراین نیاز است که با تبدیل متغیر  $y$ ، دامنه آن به بازه  $[0-1]$  تغییر کند. متغیر مقیاس‌بندی شده  $y$  می‌تواند به صورت  $y = [(x-a)/(b-a)]$  در نظر گرفته شود. با توجه به اینکه  $a, b$  مقادیر حداقل و حداکثر مجموعه داده می‌باشند تضمینی وجود ندارد که داده‌های غیر مشاهده‌ای در این محدوده باشند بنابراین برای این داده‌ها تکیه‌گاه قابل تعریف است. بر اساس روش ارایه شده اگر داده‌های مشاهداتی  $x_1, x_2, \dots, x_n$  مرتب شده باشند  $(x_1 \leq x_2 \leq \dots \leq x_n)$  آنگاه:

$$a = (x_1 - \Psi_1) \quad , \quad b = (x_n + \Psi_2) \quad (5)$$

در رابطه فوق  $\Psi_1 = (x_{1+s} - x_1)/s$  و  $\Psi_2 = (x_n - x_{n-s})/s$  است و با فرض اینکه چگالی نزدیک مرزهای تکیه‌گاه هموار است  $S$  یک مقدار صحیح مثبت می‌باشد. مقدار پیش فرض  $S$  برابر یک می‌باشد

بعضی از نقاط دامنه مقادیر منفی به خود بگیرد. در چنین شرایطی فقط بخش مثبت  $\hat{f}_J(y)$  در نظر گرفته می‌شود.

$$\hat{f}_J(y) = [\hat{f}_J(y) - c_1]_+ \quad (15)$$

که  $\hat{f}_J(y) = [\hat{f}_J(y) - c_1]_+ = \max\{0, [\hat{f}_J(y) - c_1]_+\}$  مساحت زیر نمودار دوباره محاسبه می‌شود و اگر برابر با واحد نباشد دوباره باید به گام ج رفته و روند را ادامه داد تا اینکه مساحت زیر تابع چگالی احتمال واحد گردد.

گام ششم: برآورد تابع چگالی احتمال برای داده‌های مقیاس بندی نشده: داده‌های مقیاس بندی شده مطابق با چگالی  $f_Y(y)$  توزیع می‌شود که  $y \in [0 - 1]$  می‌باشد. بنابراین تابع چگالی احتمال برآورد شده متناظر با داده‌های مقیاس بندی شده  $y$  در بازه  $[0 - 1]$  می‌باشد. برآورد  $f_X(x)$  داده‌های متغیر اصلی  $x$  در بازه  $[a - b]$  با توجه به رابطه زیر قابل دستیابی است.

$$f_X(x) = (b-a)^{-1} f_Y(y); \quad x \in [a-b] \quad (16)$$

پس از برآورد تابع چگالی احتمال متغیرهای سیلاب با استفاده از روش ناپارامتری سری‌های متعامد نرمال می‌توان با استفاده از انتگرال عددی تابع توزیع تجمعی هر متغیر را به‌دست آورد. یکی از مزیت‌های سری‌های متعامد نرمال نسبت به برآورد چگالی کرنل تولید تابع مناسب می‌باشد. هنگامی که  $\theta_j$  محاسبه شد برآورد تابع چگالی احتمال بر روی تمام مقادیر ممکن قابل تعریف می‌باشد.

#### معیارهای دقت برازش (Precision of Fit Criteria) با

استفاده از روش‌های پارامتری و ناپارامتری توابع چگالی مناسب بر متغیرهای مورد بررسی برازش داده می‌شوند. در گام بعدی توابع توزیع تجمعی منتخب با احتمال عدم وقوع تجربی که از روابط ترسیم موقعیت به‌دستی‌آید مقایسه می‌گردند. روابط ترسیم موقعیت فراوانی توسط محققان مختلف معرفی شده است (Hamed and Rao., 1999). رابطه گرینگورتن یکی از این روابط رایج است که به صورت زیر بیان می‌گردد.

$$P_k = \frac{k - 0.44}{N + 0.12} \quad (17)$$

$k$  ردیف مشاهدات در یک ترتیب صعودی،  $N$  تعداد مشاهدات و  $P_k$  فراوانی تجمعی (احتمال اینکه یک مقدار مشخص از  $k$  امین مشاهده مرتب شده به صورت صعودی کمتر است) می‌باشد.

در این مطالعه میانگین مربعات خطا ( $RMSE$ )، معیار اطلاعات آکائیک ( $AIC$ ) و معیار اطلاعات بیزی ( $BIC$ ) یا معیار شوارز به عنوان معیار دقت برازش برای توابع توزیع احتمال پارامتری و ناپارامتری متغیرهای مورد بررسی استفاده شده است. ریشه میانگین مربعات خطا با استفاده از رابطه زیر تعیین می‌گردد (Akaike., 1974;

$$w_0 = 1 \quad (10)$$

$$w_j = \left[ 1 - \frac{\hat{\delta}}{(n\theta_j^2)} \right]_+, \quad \forall j > 0 \quad (11)$$

$[1 - \hat{\delta}/(n\theta_j^2)]_+ = \max\{0, [1 - \hat{\delta}/(n\theta_j^2)]\}$  یعنی بخش مثبت  $[1 - \hat{\delta}/(n\theta_j^2)]$  به غیر از تعداد  $j$  تای اول ضرایب فوری، یک تابع چگالی نیازمند تعداد نسبتاً زیادی از ضرایب برای تخمین مناسب می‌باشد. بنابراین جمله‌هایی با بسامد بالا اضافه خواهند شد. بعد از اضافه کردن جمله‌ها در رابطه 9 برآورد چگالی احتمال به صورت زیر به‌دستی‌آید.

$$\hat{f}_J(y) = \sum_{j=0}^J w_j \theta_j \Phi_j(y) + \sum_{j=J+1}^{C_{JM} \times J_n} I_{(\theta_j^2 > C_T \hat{\sigma} \ln(n)/n)} \theta_j \Phi_j(y); \quad 0 \leq y \leq 1, \quad 0 \leq w_j \leq 1 \quad (12)$$

که  $C_T$  و  $C_{JM}$  پارامترهای بخش الحاقی می‌باشند که تعداد حداکثر اجزای شامل شده در برآورد تابع چگالی احتمال را تعریف می‌کند. مقادیر پیش فرض 6 و 4 می‌باشند (Efromovich., 1999).  $I_{(A)}$  یک تابع نشانگر است به طوری که اگر  $A$  درست باشد برابر با یک و اگر درست نباشد برابر با صفر می‌باشد.

گام پنجم: تصحیح مساحت زیر تابع چگالی احتمال و مقادیر منفی: پس از برآورد تابع چگالی احتمال این امکان وجود دارد که تابع چگالی احتمال در بعضی نواحی مقداری منفی گردد، در این صورت برای اصلاح تابع چگالی احتمال از الگوریتم پیشنهادی گوش و ماژومدار که شامل مراحل زیر است استفاده می‌گردد (Ghosh and Mujumdar., 2007).

الف - اگر مقادیر منفی در بعضی نقاط وجود دارد حداکثر این مقادیر محاسبه می‌شوند.

ب - حداکثر مقدار منفی یافت شده به  $\hat{f}_J(y)$  اضافه می‌شود. این عمل موجب می‌شود مقادیر تابع در همه دامنه مثبت باشد.

ج - مساحت زیر نمودار تعیین می‌گردد. اگر این مقدار کمتر از واحد باشد مقدار  $C$  به نحوی که رابطه زیر برقرار باشد با استفاده از روش‌های عددی محاسبه شود. مقدار  $\hat{f}_J(y)$  با اضافه کردن مقدار  $C$  به‌دست آمده از رابطه زیر تصحیح می‌گردد.

$$\int_{-\infty}^{+\infty} [\hat{f}_J(y) + c] dy = 1 \quad (13)$$

د - اگر مساحت زیر نمودار کمتر از واحد باشد مقدار  $c_1$  با استفاده از همین فرآیند مشخص می‌گردد.

$$\int_{-\infty}^{+\infty} [\hat{f}_J(y) + c_1] dy = 1 \quad (14)$$

با کم کردن  $c_1$  از  $\hat{f}_J(y)$  ممکن است تابع چگالی احتمال در

(Schwarz., 1978).

(18)

$$MSE = \sqrt{E(X_c - X_o)^2} = \left\{ \frac{1}{N-p} \sum_{i=1}^N [x_c(i) - x_o(i)]^2 \right\}$$

در این رابطه  $E(*)$  امید ریاضی  $(*)$  و  $x_c(i)$  و  $x_o(i)$  مقادیر محاسبه شده و مشاهده شده  $i$ ام می‌باشند و  $p$  تعداد پارامترهایی می‌باشد که در برآورد مقادیر محاسباتی استفاده شده است همچنین  $N$  تعداد مشاهدات است. بهترین تابع توزیع برازش شده تابعی است که حداقل ریشه میانگین مربعات خطا را داشته باشد. معیار اطلاعات آکائیک توسط آکائیک (1974) توسعه داده شد. این معیار برای شناسایی مناسب‌ترین تابع توزیع در بین توزیع‌های مورد بررسی مورد استفاده قرار می‌گیرد. معیار آکائیک شامل دو بخش است (Akaike., 1974). الف-بخشی که عدم برازش را در نظر می‌گیرد ب-بخشی که عدم اطمینان ناشی شده از تعداد پارامترهای مدل را در نظر می‌گیرد و به صورت زیر بیان می‌شود.

$$AIC = -2 \log(\text{ml of model}) + 2(\text{no. parameters}) \quad (19)$$

یا

$$AIC = N \log(MSE) + 2(\text{no. parameters}) \quad (20)$$

$$MSE = E(X_c - X_o)^2 = \left\{ \frac{1}{N-p} \sum_{i=1}^N [x_c(i) - x_o(i)]^2 \right\} \quad (21)$$

معیار اطلاعات بیزی می‌تواند به صورت رابطه زیر بیان گردد (Schwarz., 1978).

(22)

$$BIC = N \log(MSE) + [(no. parameters) \times \log(N)]$$

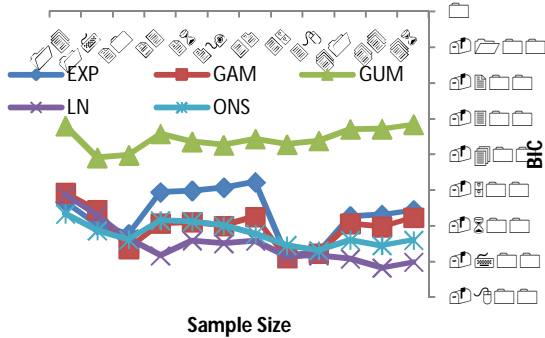
بهترین تابع توزیع احتمال توزیعی است که حداقل معیار اطلاعات آکائیک، معیار اطلاعات بیزی و میانگین مربعات خطا را دارد.

## نتایج و بحث

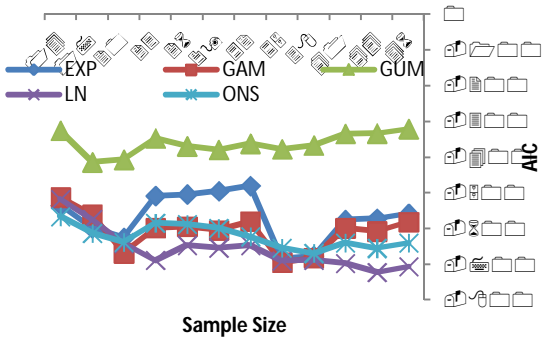
در ابتدا سری اصلی داده‌های دبی اوج سیلاب سالانه برای هر ایستگاه مد نظر قرار گرفت و بر این اساس با در نظر گرفتن گام‌های زمانی سه ساله، اندازه سری تا چهارده سال کاهش یافت. بر اساس سری‌های مختلف تشکیل شده برای هر ایستگاه توزیع‌های پارامتری و روش سری‌های متعامد نرمال بر داده‌ها اعمال شد و نتایج معیارهای دقت برازش ثبت گردید. بررسی شکل 1 نشان می‌دهد که برای داده‌های ایستگاه تمر توزیع پارامتری گامبل دارای کم‌ترین دقت در میان توزیع‌های پارامتری و ناپارامتری مورد بررسی بوده است و همچنین با تغییرات اندازه نمونه، دقت این توزیع نیز نوسان زیادی داشته است. توزیع پارامتری نمایی نیز الگویی تقریبی را توزیع پارامتری گامبل نشان می‌دهد. در بین توزیع‌های پارامتری توزیع لوگ نرمال دقت

کاملاً بهتری را نسبت به سایر توزیع‌های پارامتری نشان می‌دهد و همچنین مشخص است که با تغییر اندازه نمونه، دقت این توزیع در بین توزیع‌های پارامتری کمتر تحت‌تأثیر قرار می‌گیرد. روش ناپارامتری سری‌های متعامد نرمال نیز در این سری داده دارای دقت کاملاً مناسب است و از طرفی با تغییرات در اندازه نمونه نوسان کمی نیز از خود نشان می‌دهد. دقت توزیع لوگ نرمال به عنوان بهترین توزیع پارامتری مورد بررسی در این سری داده در اندازه نمونه‌های بزرگ‌تر (بیشتر از 20 نمونه) از توزیع سری‌های متعامد نرمال کمی بالاتر است اما با کاهش تعداد نمونه این الگو معکوس می‌گردد. همچنین مشخص است که در انتهای نمودار (کاهش تعداد نمونه به 17 و 14 مقدار ثبت شده) همه توزیع‌های پارامتری و ناپارامتری مورد بررسی دارای کاهش محسوس دقت برازش می‌گردند اما تغییرات برای روش ناپارامتری از توزیع‌های پارامتری کمتر می‌باشد. بررسی شکل 2 نشان می‌دهد توزیع‌های پارامتری نمایی و گامبل دارای دقت کمتری نسبت به سایر توزیع‌های مورد بررسی هستند. به طور مشخص هر دو توزیع پارامتری نمایی و گامبل دارای حساسیت بالاتر در دقت برازش نسبت به سایر توزیع‌های مورد بررسی نیز می‌باشند. به طور مشخص توزیع نمایی با کاهش اندازه نمونه به صورت چشمگیری دچار کاهش دقت برازش می‌گردد که از این نظر دارای رفتار متفاوت با دیگر توزیع‌ها در این سری داده است. در بین توزیع‌های پارامتری توزیع‌های لوگ نرمال و گاما دارای دقت مناسب و نیز پایداری نسبی نسبت به تغییرات اندازه نمونه می‌باشند که دقت توزیع لوگ نرمال نسبت به گاما کمی بهتر می‌باشد. مقایسه مناسب‌ترین توزیع پارامتری از لحاظ دقت و نیز پایداری به کاهش تعداد نمونه نشان می‌دهد که وقتی تعداد نمونه بیشتر از 32 باشد کارایی توزیع لوگ نرمال کمی بهتر از توزیع سری‌های متعامد نرمال است اما وقتی تعداد نمونه از این مقدار کاهش می‌یابد توزیع ناپارامتری به مقدار چشمگیری دارای دقت بالاتر از توزیع لوگ نرمال می‌باشد. در این سری داده به طور مشخص، کاهش دقت همه توزیع‌ها، در برازش تابع چگالی احتمال، به واسطه کاهش اندازه نمونه از تعداد بالاتر نمونه‌ها (نسبت به داده‌های ایستگاه تمر یعنی تعداد 38 نمونه) شروع می‌شود اما نرخ کاهش دقت در روش ناپارامتری نسبت به توزیع‌های پارامتری کمتر می‌باشد. بررسی شکل 3، که در برگزیده نتایج به دست آمده برای ایستگاه زرینگل است، نیز نشان دهنده دقت کمتر توزیع‌های گامبل و نمایی نسبت به سایر توزیع‌های پارامتری است. همچنین مشخص است که توزیع نمایی به کاهش تعداد نمونه حساسیت بالایی دارد. در این سری داده نیز توزیع پارامتری لوگ نرمال دارای بهترین کارایی، با در نظر گرفتن دقت برازش و نیز پایداری این دقت نسبت به کاهش اندازه نمونه، در بین روش‌های پارامتری است. همچنین این شکل نشان می‌دهد که توزیع سری‌های متعامد نرمال و توزیع لوگ نرمال دارای دقت بسیار نزدیک به یکدیگر می‌باشند.

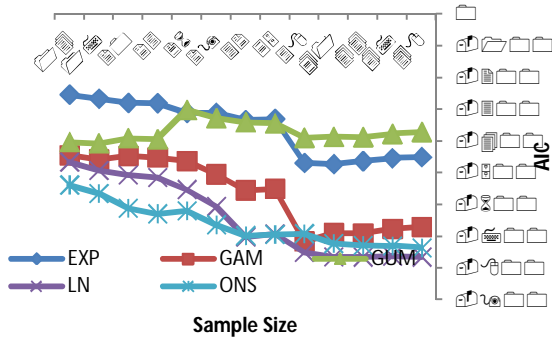
شکل 2-ب- نمودار تغییرات BIC با اندازه نمونه در ایستگاه تمر



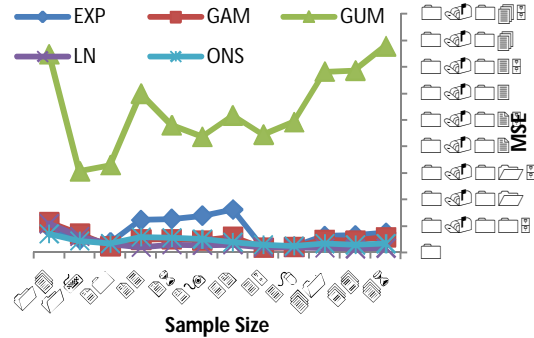
شکل 1-الف- نمودار تغییرات AIC با اندازه نمونه در ایستگاه تمر



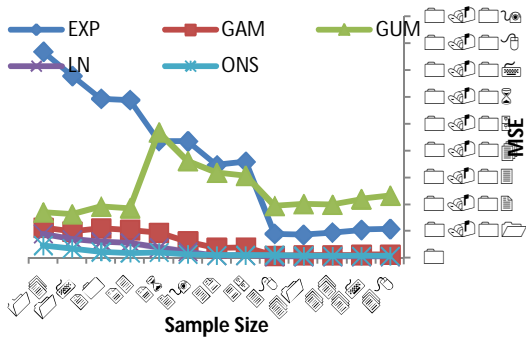
شکل 2-الف- نمودار تغییرات AIC با اندازه نمونه در ایستگاه رامیان



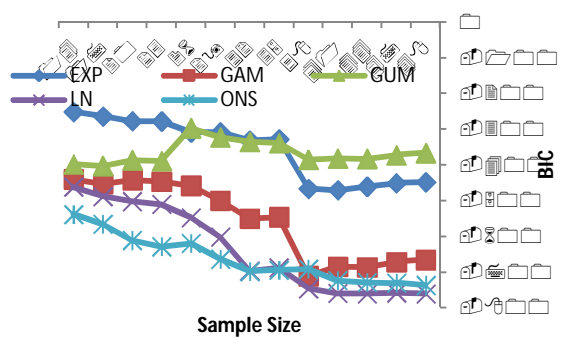
شکل 1-ج- نمودار تغییرات MSE با اندازه نمونه در ایستگاه تمر



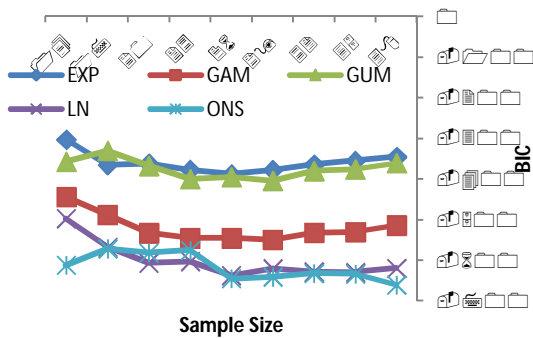
شکل 2-ج- نمودار تغییرات MSE با اندازه نمونه در ایستگاه رامیان



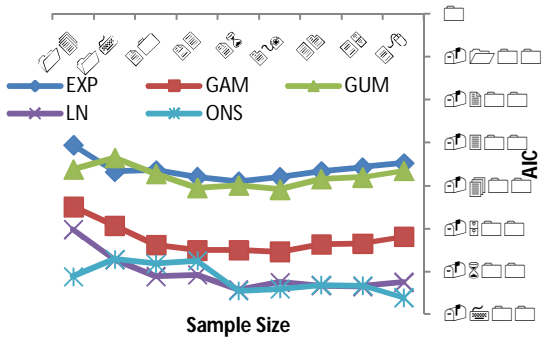
شکل 2-ب- نمودار تغییرات BIC با اندازه نمونه در ایستگاه رامیان



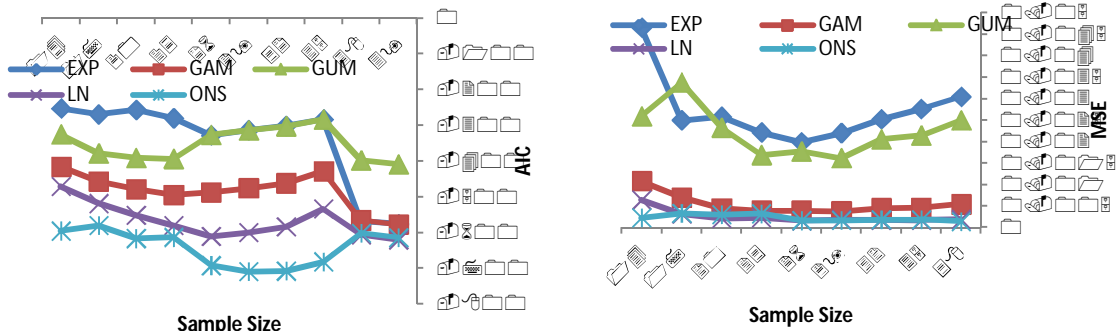
شکل 3-ب- نمودار تغییرات BIC با اندازه نمونه در ایستگاه زرین گل



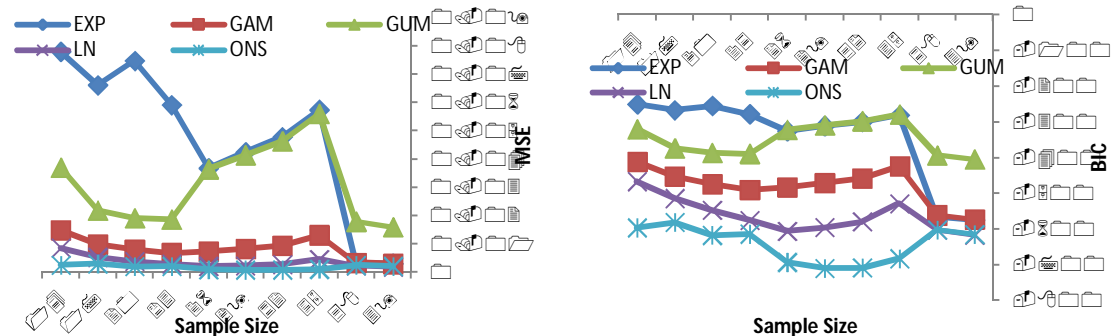
شکل 3-الف- نمودار تغییرات AIC با اندازه نمونه در ایستگاه زرین گل



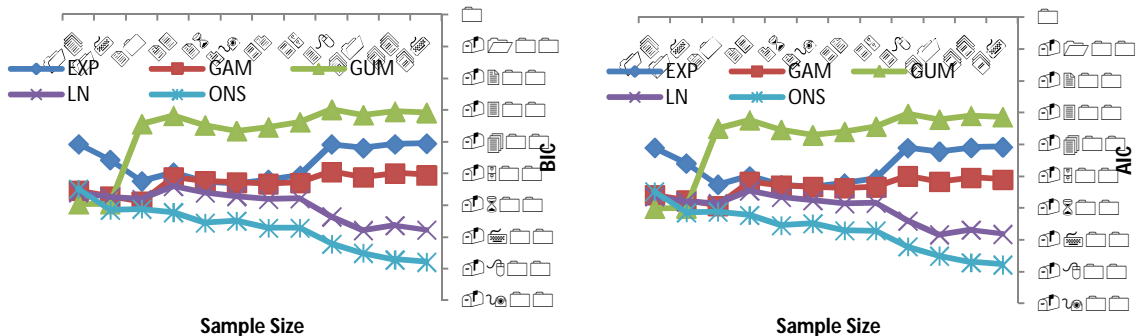
شکل 3-ج - نمودار تغییرات MSE با اندازه نمونه در ایستگاه زرین گل



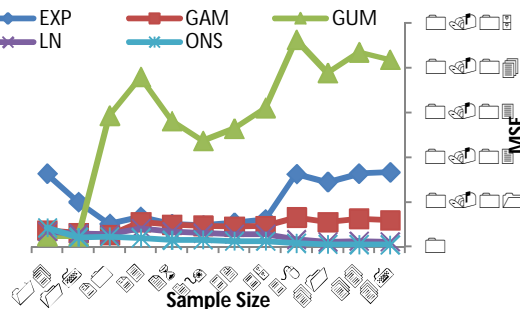
شکل 4-ب - نمودار تغییرات BIC با اندازه نمونه در ایستگاه سرمو



شکل 5-الف - نمودار تغییرات AIC با اندازه نمونه در ایستگاه لزوره



شکل 5-ج - نمودار تغییرات MSE با اندازه نمونه در ایستگاه لزوره



برای داده‌های ایستگاه هیدرومتری سرمو در شکل 4 ارایه شده است. بررسی این شکل نشان می‌دهد که در این سری داده نیز توزیع‌های پارامتری نمایی و گامبل دارای کمترین دقت و بیشترین حساسیت در کاهش تعداد نمونه، در مقایسه با سایر توزیع‌ها بوده‌اند. همچنین توزیع

با کاهش تعداد نمونه به تعداد 26 و مقادیر کمتر از آن یک روند ضعیف کاهش دقت در توزیع‌های پارامتری مشاهده می‌شود در حالیکه توزیع سری‌های متعامد نرمال یک رفتار تقریباً متناوب از خود نسبت به تغییرات اندازه نمونه نشان می‌دهد. معیارهای دقت برآزش

نرمال بر توزیع لوگ نرمال با کاهش دقت نمونه‌ها کاهش می‌یابد که رفتاری متفاوت با سری‌های مورد بررسی قبلی می‌باشد. همان‌طور که مشخص است برای هر سری از داده‌های ایستگاه-های مورد بررسی تعداد چهار توزیع پارامتری و نیز توزیع ناپارامتری سری‌های متعامد نرمال مورد بررسی قرار گرفته است. نمودارهای ارایه شده امکان مقایسه گرافیکی دقت برازش توزیع‌های مختلف مورد بررسی را نشان می‌دهد اما در مواردی که نمودارها دارای فواصل کم نسبت به یکدیگرند نیاز به بررسی عددی مقادیر می‌باشد. برای هر کدام از سری داده‌های مورد بررسی، با در نظر گرفتن سه معیار دقت برازش، دو مشخصه میانگین و انحراف معیار محاسبه و در جدول 2 ارایه شده است. مقدار میانگین می‌تواند نشان‌دهنده این مورد باشد که در هر سری داده مورد بررسی کدام توزیع به شکل متوسط بالاترین دقت در برازش تابع چگالی احتمال را داشته است در حالیکه معیار انحراف معیار بیانگر این نکته است که تغییرات در اندازه نمونه بر هر توزیع تا چه میزان تأثیرگذار بوده و به عبارتی کدام توزیع دارای رفتار پایدارتری در دقت برازش نسبت به کاهش اندازه نمونه بوده است. مقادیر محاسبه شده میانگین و انحراف معیار به ازای اندازه‌های مختلف نمونه‌های کاهش یافته برای ایستگاه‌های متفاوت مورد بررسی در جدول 2 ارایه شده است.

لوگ نرمال مانند سری‌های مورد بررسی قبلی دارای بهترین کارایی در بین توزیع‌های پارامتری بوده است. در این سری داده مشخص است که سری متعامد نرمال در همه مقادیر تعداد نمونه‌های بررسی شده دارای دقت بهتری نسبت به توزیع لوگ نرمال بوده است. با کاهش تعداد نمونه‌ها در این سری به کمتر از 26 مقدار ثبت شده همه توزیع‌های مورد بررسی، کاهش دقت برازش را نشان داده‌اند اما با این وجود شدت روند کاهش دقت در سری‌های متعامد نرمال نسبت به توزیع‌های پارامتری به مراتب کمتر بوده است. در شکل 5 معیارهای دقت برازش برای داده‌های ایستگاه لزوره به نمایش درآمده است. مطابق تمام سری داده‌های بررسی شده قبلی در این مرحله نیز توزیع‌های گامبل و نمایی دارای کمترین دقت و بیشترین حساسیت در تغییرات تعداد نمونه در مقایسه با سایر توزیع‌های باشند. در شرایطی که تعداد نمونه کاهش می‌یابد دقت برازش توزیع گاما و گامبل بر خلاف سایر توزیع‌های مورد بررسی مقادیر افزایشی را نشان می‌دهد که رفتاری غیر معمول می‌باشد. توزیع لوگ نرمال در این سری داده نیز به عنوان بهترین توزیع پارامتری انتخاب می‌شود. دقت سری‌های متعامد نرمال در تمام اندازه نمونه‌های بررسی شده نسبت به توزیع لوگ نرمال بالاتر می‌باشد. نکته مهم دیگری که در بررسی شکل 5 مشخص می‌شود آن است که برتری کاملاً مشخص سری‌های متعامد

جدول 2- مقادیر میانگین و واریانس معیارهای دقت برازش در اندازه‌های مختلف نمونه بررسی شده

تمر		رامیان		زرین گل		سرمو		لزوره		
Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
-567.22	68.57	-360.2	87.22	-361.74	24.89	-343.68	123.41	-469.32	54.42	EXP-AIC
-564.61	68.57	-357.6	87.22	-359.14	24.89	-341.08	123.41	-466.72	54.42	EXP-BIC
0.0041	0.0024	0.0361	0.0247	0.0271	0.0079	0.0468	0.0265	0.0102	0.0053	EXP-MSE
-606.35	52.13	-563.34	109.77	-524.73	34.09	-484.93	52.16	-533.74	28.74	GAM-AIC
-601.14	52.13	-558.13	109.77	-519.52	34.09	-479.72	52.16	-528.53	28.74	GAM-BIC
0.0025	0.0013	0.0054	0.0044	0.0054	0.0023	0.0084	0.0038	0.0048	0.0012	GAM-MSE
-360.93	30.38	-371.54	32.39	-379.03	23.58	-352.79	46.24	-378.87	105.62	GUM-AIC
-355.72	30.38	-366.33	32.39	-373.82	23.58	-347.58	46.24	-373.66	105.62	GUM-BIC
0.0271	0.0078	0.0247	0.0091	0.0223	0.0055	0.031	0.0142	0.0293	0.0145	GUM-MSE
-655.51	58.68	-642.21	120.63	-605.59	44.13	-568.1	47.89	-604.86	49.42	LN-AIC
-650.3	58.68	-637	120.63	-600.38	44.13	-562.89	47.89	-599.65	49.42	LN-BIC
0.0017	0.0013	0.0029	0.003	0.0025	0.0015	0.0037	0.0021	0.0025	0.001	LN-MSE
-621.66	32.43	-663.94	65.62	-616.26	33.23	-641.74	50.27	-671.74	71.02	ONS-AIC
-621.66	32.43	-663.94	65.62	-616.26	33.23	-641.74	50.27	-671.74	71.02	ONS-BIC
0.0021	0.0007	0.0016	0.0012	0.0022	0.0008	0.0018	0.0008	0.0015	0.0011	ONS-MSE

توزیع‌های سری‌های متعامد نرمال و لوگ نرمال، و از نظر پایداری نسبت به تغییرات اندازه نمونه، توزیع سری‌های متعامد نرمال دارای بهترین نتیجه بوده‌اند. در سری داده‌های ایستگاه زرینگل بهترین دقت در برازش مدل به ترتیب متعلق به سری‌های متعامد نرمال و توزیع لوگ نرمال می‌باشد. بر اساس معیارهای AIC و BIC به ترتیب توزیع‌های گامبل، نمایی و سری‌های متعامد نرمال و براساس معیار RMSE به ترتیب توزیع‌های سری‌های متعامد نرمال، لوگ نرمال و گاما دارای بالاترین پایداری نسبت به تغییرات اندازه نمونه بوده‌اند که به نظر می‌رسد بتوان در مجموع توزیع سری متعامد نرمال را به عنوان

بررسی مقادیر میانگین معیارهای AIC، BIC و MSE در سری داده‌های ایستگاه تمر نشان می‌دهد که در مجموع توزیع لوگ نرمال دارای بالاترین دقت برازش بوده است در حالیکه روش ناپارامتری سری‌های متعامد نرمال در جایگاه دوم و با دقتی بالاتر از سایر توزیع‌های پارامتری قرار می‌گیرد. بررسی مقدار انحراف معیار معیارهای دقت برازش نشان می‌دهد از نظر پایداری دقت برازش در تغییرات اندازه نمونه توزیع سری‌های متعامد نرمال دارای بهترین نتیجه بوده است. بررسی مشخصه‌های آماری معیارهای دقت برازش برای داده‌های ایستگاه رامیان نشان می‌دهد از نظر دقت به ترتیب



نشده است. در این مطالعه داده‌های ثبت شده حداکثر دبی اوج سیلاب سالانه پنج ایستگاه هیدرومتری تمر، رامیان، زرین گل، سرمو و لزورهدر استان گلستان انتخاب شده است تا اینکه با استفاده از این داده‌ها تغییرات در دقت تخمین تابع چگالی احتمال دبی اوج سیلاب با استفاده از توزیع‌های مهم پارامتری و توزیع ناپارامتری سری‌های متعامد نرمال بررسی گردد. نتایج بررسی‌ها با در نظر گرفتن میانگین معیارهای دقت برازش در اندازه نمونه‌های مختلف در ایستگاه‌های مورد بررسی، نشان می‌دهد که در مجموع در چهار ایستگاه (رامیان، زرین گل، سرمو و لزوره) بالاترین دقت برازش با استفاده از روش سری‌های متعامد نرمال به دست آمده است در حالیکه در یک ایستگاه باقیمانده (تمر)، روش مذکور پس از توزیع پارامتری لوگ نرمال دومین رتبه را در دقت برازش کسب کرده است. نکته بسیار مهم دیگری که در نتایج ارایه شده وجود دارد آن است که با کاهش اندازه نمونه در داده‌های مورد بررسی دقت و صحت روش سری‌های متعامد نرمال به طور مشخصی از روش‌های پارامتری، در هر پنج ایستگاه مورد بررسی، بالاتر می‌باشد. همچنین با در نظر گرفتن انحراف معیار معیارهای دقت برازش برای داده‌های هر یک از ایستگاه‌ها با در نظر گرفتن تمام اندازه نمونه‌های مختلف بررسی شده مشخص می‌شود که در چهار ایستگاه (تمر، رامیان، زرین گل، سرمو) روش سری‌های متعامد نرمال دارای کمترین حساسیت به تغییرات اندازه نمونه بوده است در حالیکه در ایستگاه لزوره توزیع لوگ نرمال دارای نتایج مناسب‌تر از سری‌های متعامد نرمال در این زمینه بوده است. همچنین بررسی نمودارهای تغییرات معیارهای دقت برازش به اندازه نمونه مشخص می‌کند که هنگامی که تعداد نمونه‌ها کاهش می‌یابد تغییرات معیارهای دقت برازش در سری‌های متعامد نرمال در همه ایستگاه‌ها از توزیع‌های پارامتری کمتر می‌باشد. نتایج به دست آمده این تحقیق، مطالعاتی که در آن به صورت نظری به حساسیت کمتر سری‌های متعامد نرمال به اندازه نمونه در مقایسه با توزیع‌های پارامتری اشاره شده بود را تایید می‌کند (Vittal et al., 2015; Ghosh and Mujumdar., 2007; Karmakar and Simonovic., 2008). به نظر می‌رسد روش سری‌های متعامد نرمال باید به شکل جامع‌تری برای کاربردهای آن در هیدرولوژی مورد بررسی قرار بگیرد.

## منابع

- حقیقت جو، پرویز. (1387). ارزیابی مقایسه ای آمار پارامتری و ناپارامتری در تحلیل فراوانی بارش. رساله دکتری هیدرولوژی. دانشگاه شهید چمران اهواز. 115 صفحه.
- Adamowski, K. 1985. Nonparametric kernel estimation of flood frequencies. *Water Resources Research*. 21.11:1585-1590.

انتخاب نهایی معرفی کرد. مقدار میانگین معیارهای دقت برازش تایید می‌کند که در داده‌های ایستگاه سرمو توزیع‌های سری‌های متعامد نرمال و لوگ نرمال به ترتیب بهترین دقت را داشته‌اند. بررسی مقدار انحراف معیار معیارهای AIC و BIC مبین آن است که توزیع‌های گامبل، لوگ نرمال و سری‌های متعامد نرمال دارای بیشترین پایداری به تغییرات اندازه نمونه بوده‌اند در حالیکه انحراف معیار RMSE به ترتیب توزیع‌های سری‌های متعامد نرمال، لوگ نرمال و گاما را به عنوان برترین توزیع‌ها در این زمینه معرفی می‌کند. با توجه به این موارد به نظر می‌رسد بتوان در مجموع به ترتیب توزیع‌های سری‌های متعامد نرمال و لوگ نرمال را به عنوان برترین توزیع‌ها در پایداری دقت نتایج به تغییرات اندازه نمونه در سری داده‌های ایستگاه سرمو معرفی نمود. با در نظر گرفتن مقادیر میانگین هر سه معیار دقت برازش مشخص می‌شود که برای داده‌های ایستگاه لزوره به ترتیب توزیع‌های ناپارامتری سری‌های متعامد نرمال و لوگ نرمال دارای بهترین صحت برازش می‌باشند. بر اساس انحراف معیار شاخص‌های AIC و BIC به ترتیب توزیع‌های گاما، لوگ نرمال و نمایی و بر اساس انحراف معیار شاخص MSE به ترتیب توزیع‌های لوگ نرمال، سری‌های متعامد نرمال و گاما دارای کمترین حساسیت نسبت به تغییرات اندازه نمونه بوده‌اند که بر این اساس می‌توان توزیع لوگ نرمال را بهترین توزیع در این زمینه با در نظر گرفتن داده‌های ایستگاه لزوره به حساب آورد.

## نتیجه گیری

به طور معمول در مطالعات مرتبط با کارایی توزیع‌های آماری در مبحث هیدرولوژی این موضوع مد نظر قرار می‌گیرد که کاربرد هر توزیع متفاوت چگونه می‌تواند موجب بهبود کارایی برآورد تابع چگالی احتمال متغیرهای مورد بررسی شود. در این زمینه دو رویکرد توزیع‌های پارامتری و روش‌های ناپارامتری مورد استفاده قرار می‌گیرد. تخمین‌های ناپارامتری تابع چگالی احتمال نیاز به فرض اولیه‌ای در مورد داده‌ها ندارد و از این جهت دارای برتری بر توزیع‌های پارامتری است. مطالعات پیشین در حوزه روش‌های ناپارامتری برآورد تابع چگالی احتمال بر کاربرد برآوردگر چگالی احتمال کرنل متمرکز بوده است که این روش علی‌رغم نتایج مناسب تا حدی به کاهش نمونه حساس می‌باشد (Efromovich., 1999; Bowman and Azzalini., 1997; Ghosh and Mujumdar., 2007; Karmakar and Simonovic., 2008). روش ناپارامتری سری‌های متعامد نرمال که مدت زمان زیادی از معرفی آن در هیدرولوژی نمی‌گذرد برای رفع نسبی این محدودیت در مقایسه با روش‌های پارامتری و نیز روش ناپارامتری کرنل ارایه شده است. تاکنون در زمینه تحلیل حساسیت این روش به تغییر اندازه نمونه تحقیقی انجام

- analysis. CRC press.
- Karmakar, S., Simonovic, S.P. 2008. Bivariate flood frequency analysis: Part 1. Determination of marginals by parametric and nonparametric techniques. *Journal of Flood Risk Management*. 1.4: 190-200.
- Kidson, R., Richards, K.S. 2005. Flood frequency analysis: Assumptions and alternatives. *Progress in Physical Geography*. 29.3: 392-410.
- Kim, T.-W., Valdés, J.B., Yoo, C. 2003. Nonparametric approach for estimating return periods of droughts in arid regions. *Journal of Hydrologic Engineering*. 8.5: 237-246.
- Lall, U. 1995. Nonparametric function estimation: Recent hydrologic contributions. *Reviews of Geophysics, Contributions in Hydrology, U.S. National Report to the IUGG 1991-1994*. 1093-1099.
- Moazed, H., Moradzadeh, M., Veysi, S., Kaboli, H.S., Salarijazi, M. 2012. Analysis of drought return periods in Khuzestan province, southwest of Iran. *Journal of Food, Agriculture and Environment*. 10.1: 642-645.
- Quintela-Del-Río, A. 2011. On bandwidth selection for nonparametric estimation in flood frequency analysis. *Hydrological Processes*. 25.5: 671-678.
- Reddy, M.J., Ganguli, P. 2012. Bivariate flood frequency analysis of upper Godavari river flows using Archimedean copulas. *Water resources management*. 26.14: 3995-4018.
- Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics*. 6.2: 461-464.
- Shabri, A. 2002. Nonparametric kernel estimation of annual maximum stream flow quantiles. *Matematika*. 18:99-107.
- Smakhtin, V.U. 2001. Low flow hydrology: A review. *Journal of Hydrology*. 240.3-4: 147-186.
- Vittal, H., Singh, J., Kumar, P., Karmakar, S. 2015. A framework for multivariate data-based at-site flood frequency analysis: Essentiality of the conjugal application of parametric and nonparametric approaches. *Journal of Hydrology*. 525: 658-675.
- Adamowski, K. 2000. Regional analysis of annual maximum and partial duration flood data by nonparametric and L-moment methods. *Journal of Hydrology*. 229.3:219-231.
- Adamowski, K., Feluch, W. 1990. Nonparametric flood-frequency analysis with historical information. *Journal of Hydraulic Engineering*. 116.8. 1035-1047.
- Adib, A., Salarijazi, M., Najafpour, K. 2010. Evaluation of synthetic outlet runoff assessment models. *Journal of Applied Sciences and Environmental Management*. 14.3: 13-18.
- Adib, A., Salarijazi, M., Shooshtari, M.M., Akhondali, A.M. 2011. Comparison between characteristics of geomorphoclimatic instantaneous unit hydrograph be produced by GcIUH based Clark Model and Clark IUH model. *Journal of Marine Science and Technology*. 19.2: 201-209.
- Adib, A., Salarijazi, M., Vaghefi, M., Shooshtari, M.M., Akhondali, A.M. 2010. Comparison between GcIUH-Clark, GIUH-Nash, Clark-IUH, and Nash-IUH models. *Turkish Journal of Engineering and Environmental Sciences*. 34.2: 91-104.
- Akaike, Hirotugu. 1974. New look at the statistical model identification. *IEEE Transactions on Automatic Control*. 19.6: 716-723.
- Bowman, A., Azzalini, A. 1997. *Applied smoothing techniques for data analysis: the Kernel approach with S-plus illustrations*. Oxford University Press, New York. 198p.
- Efromovich, S. 1999. *Nonparametric curve estimation: methods, theory and applications*. Springer-Verlag, New York. 411p.
- Eidipour, A., Akhondali, A.M., Zarei, H., Salarijazi, M. 2016. Flood Hydrograph Estimation Using GIUH Model in Ungauged Karst Basins (Case Study: Abolabbas Basin). *TUEXENIA*. 36.36: 26-33.
- Ghosh, S., Mujumdar, P.P. 2007. Nonparametric methods for modeling GCM and scenario uncertainty in drought assessment. *Water Resources Research*. 43.7:1-19.
- Hamed, K., Rao, A.R. 1999. Flood frequency

## Analysis of Precision Changes of Ortho-Normal Series to Annual Maximum Discharge Sample Size

M. Salarijazi<sup>1</sup>

Received: Jun.08, 2016

Accepted: Nov.23, 2016

### Abstract

Flood risk analysis studies carry out based on estimation of the probability density function (PDF) of flood peak. Conventional approaches to estimate the probability density function distributions are parametric methods that have constraints such as sensitivity of fitness precision to reduction in sample size. Non-parametric methods to estimate the PDF have fewer restrictions than the parametric distributions. In this study, Exponential, Log-Normal, Gamma and Gumbel as parametric distributions, Ortho-Normal Series method as a non-parametric method, and Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Mean Square Error (MSE) as precision criteria were applied to investigate the precision of mentioned methods and their sensitivity to reduction of annual maximum peak flow series. The precision and change in precision of fitness of parametric and non-parametric distributions to sample size reduction were investigated based on five hydrometry station (with 38 to 48 years recorded datasets) in Golestan province. The results show that the precision of non-parametric ortho-normal series method is considerably higher than parametric distributions. In addition, ortho-normal series method is less sensitive to sample size reduction than parametric distribution that makes it as suitable option where the recorded data sets are belonging to relatively short time period.

**Key words:** Ortho-Normal Series Method, Parametric Distributions, Sensitivity Analysis, Sample Size.

Archive of SID

1- Assistant Professor, Department of Water Engineering, Gorgan University of Agricultural Sciences and Natural Resources.

Email: meysam.salarijazi@gau.ac.ir