مدلبندی داده های مقادیر فرین آب و هوای شهر تبریز

علی اکبر حیدری
گروه آمار، دانشگاه تبریز

چگونگی رفتار بیشترین یا کمترین مقادیر (یعنی مقادیر فرین) یک مجموعه داده، بویژه در پدیده‌هایی طبیعی مثل امواج دریا، دبی رودخانه، سرعت باد، درجه حرارت و میزان بارندگی در بسیاری از مسائل کاربردی مورد علاقه می‌باشد. بنابراین اگر بتوانیم رفتار اینگونه داده ها را بوسیله فرمولهای ریاضی بیان کنیم می توانیم چگونگی رفتار آنها را در آینده پیش بینی نماییم. مثلا می توانیم پیش بینی کنیم که به طور متوسط هر چند سال یکبار در یک منطقه خشکسالی رخ می‌دهد و متوسط طول مدت خشکسالی در هر رخداد چقدر است.
یکی از روشهای تحلیل مقادیر فرین بر اساس توزیع توأم مجانبی آماره‌های مرتب و توزیع مقادیر فرین تعمیم یافته می‌باشد. در روش فزونیها که توسط پیکندز (۱۹۷۵) ارائه شد، تمامی فزونیهای بالای یک سرحد بزرگ، با استفاده از توزیع پارتوی تعمیم یافته مدلبندی می‌شوند که در مقایسه با روش قبلی از دقت بیشتری برخوردار است.
در این مقاله ضمن معرفی روشهای اشاره شده، برآورد پارامترها و چندکهای توزیعهای مربوطه صورت گرفته سپس برای مقادیر فرین فاکتورهای مهم اقلیمی شهر تبریز (از جمله درجه حرارت، سرعت باد، میزان بارندگی و ...) مدل مناسب آماری ارئه شده، نیکویی برازش آنها بررسی شده و در نهایت برخی پیش بینی‌های لازم صورت گرفته است.

واژه‌های کلیدی: توزیع مقادیر فرین تعمیم یافته، توزع پارتوی تعمیم یافته، برآورد، نیکویی برازش، داده‌های اقلیمی.

۱ مقدمه

در مسائل کاربردی، گاهی مطالعه مقادیر فرین[1] یعنی مقادیر واقع در دم توزیع برخی از پدیده‌ها، مانند امواج، سرعت باد، درجهٔ حرارت، زمین لرزه، سیل، غرامتهای بیمه، غلظت آلاینده‌های جوی، استحکام مواد و ...مورد علاقه می‌باشند.

[1] Extreme Values

مواجه شدن با مسائلی مشابه موارد ذکر شده موجب توسعهٔ سریع مدلبندی مقادیر فرین در سالهای اخیر شده است. قبل از هر چیز، چند تعریف اساسی بیان می شوند:

تعریف ۱ : برای مشاهدات $x_1,\ldots,x_n$ و یک سرحد[2] از پیش تعیین شدهٔ u، هر x_iای که از u بیشترباشد، یک فزونی[3] نامیده می‌شود.

تعریف ۲: برای هر فزونی y_j، اندازهٔ $y_j - u$ مقدارفزونی[4] بالای u نامیده می‌شود.

تعریف ۳: فرض کنید سری داده‌ها در n دورهٔ متوالی با اندازهٔ m بصورت

$$x_{1,1}, x_{1,2}, \ldots, x_{1,m}, \ldots, x_{i,1}, \ldots, x_{i,m}, x_{n,1}, \ldots, x_{n,m}$$

مشاهده شده باشند. در این صورت $z_i = Max\{x_{i,1},\ldots,x_{i,m}\}$ ماکزیمای[5] دوره iام نامیده می‌شود.

معمولاً دو روش برای مدلبندی مقادیر فرین، بکار برده می‌شود. یکی از این روشها براساس توزیع توأم مجانبی آمارهٔ ترتیبی بالایی است و هنگامی به داده‌ها برازش داده میشود که داده‌ها مشتمل بر مجموعه‌ای از ماکزیماها باشند. مثلاً اگر ماکزیمم سرعت باد روزانهٔ یک نقطه، در طول ۵۰ سال در دسترس باشند، در این روش ماکزیمم سرعت باد هر سال، ماکزیمای آن سال نامیده می‌شود. بنابراین ۵۰ مقدار از مجموعهٔ داده‌ها در نظر گرفته می‌شود و با استفاده از قضایای حدی، یک توزیع پارامتری با عنوان توزیع مقادیر فرین تعمیم یافته[6] به داده‌ها برازش داده می‌شود. هنگامیکه مقادیر بزرگ متعددی در یک دوره وجود داشته باشد، استفاده از ماکزیمای هر دوره ممکن است موجب از دست دادن اطلاعات زیادی گردد در چنین مواردی از روش دیگری استفاده می‌شود که در آن تمام داده‌هایی که از یک سرحد معین بزرگتر هستند برای تجزیه و تحلیل و برازش توزیع، بکار گرفته می‌شوند. این روش به روش فزونیهای بیشتر از یک سرحد[7] معروف است و در آن تمامی داده‌های بزرگتر از یک سرحد (که معمولاً از پیش تعیین شده است) از داده‌ها استخراج شده و به آنها یک مدل پارامتری با عنوان توزیع پارتوی تعمیم یافته[8] برازش داده می‌شود. به این ترتیب بطور مثال در مثال ماکزیمم سرعت باد روزانه، ممکن است از یک سال چند داده وجود داشته باشد که از

---

[2]Threshold

[3]Exceedance

[4]Excesses

[5]Maxima

[6]Generalized Extreme Value Distribution

[7]*Peak Over Threshold*

[8]Generalized Pareto Distribution

سرحد داده شده بزرگتر باشند و در بعضی سالها ممکن است اصلاً چنین داده‌ای وجود نداشته باشد.

این روش مدلبندی که در این مقاله مورد مطالعه قرار خواهد گرفت، اولین بار توسط پیکندز(۱۹۷۵) معرفی شد . و هاسکینگ و والیس (۱۹۸۷) آنرا برای مدلبندی سیلهای سالیانهٔ رودخانهٔ Nidd انگلستان، و گریمشاو(۱۹۹۳) برای مدلبندی داده‌های مقاومت کششی یک نمونه تصادفی از رشته‌های قالی، بکار بردند. نمونه‌های دیگری از این مدلبندی را می‌توان در مقالات کاستیلو (۱۹۹۴)، دیویسن(۱۹۹۴) و روتزن و تجویدی (۱۹۹۷) مشاهده کرد.

۱-۱ توزیع مقادیر فرین تعمیم یافته:

روش کلاسیک تحلیل مقادیر فرین، مبتنی بر توزیعهای حدی ماکزیمای نمونه‌ای است. اگرچه بررسی کامل این روش در دستور کار این تحقیق نیست ولی بعلت اهمیت موضوع و خواص مشترکی که با روش فزونیهای بیشتر از یک سرحد دارد در اینجا بطور مختصر شرح داده میشود.

فرض کنید $X_1, X_2, \ldots, X_n$ دنباله‌ای از متغیرهای تصادفی دو به دو مستقل با تابع توزیع مشترک $F(x)$ باشند و برای $n \in N$, $M_n = Max(X_1, X_2, \ldots, X_n)$ فرض کنید دنباله‌هایی از اعداد حقیقی مانند a_n و b_n وجود دارند بگونه‌ای که

$$\lim_{n\to\infty} P(\frac{M_n - b_n}{a_n} \leq x) = \lim_{n\to\infty} F^n(a_n x + b_n) = G(x)$$

که در آن $G(x)$ یک توزیع ناتباهیده[9] است. در این صورت گوئیم $F(x)$ در دامنهٔ جذب[10] $G(x)$ قرار دارد و با نماد $F \in D(G)$ نمایش داده می‌شود. یک نتیجهٔ اصلی از قضیهٔ مقادیر فرین کلاسیک بیانگر آن است که $G(x)$ بایستی یک تابع توزیع مقادیر فرین تعمیم یافته باشد.

تمام توزیعهای حدی ممکن در رابطهٔ (۱)، یعنی توزیعهای مقادیر فرین را میتوان بصورت خانوادهٔ پارامتری

$$\begin{cases} G_{\gamma,\mu,\sigma} = exp(-(1 + \gamma(\frac{(x-\mu)}{\sigma}))^{-1/\gamma}) & 1 + \gamma x > 0 \quad , \gamma \neq 0 \\ G_{0,\mu,\sigma} = exp(-e^{\frac{-(x-\mu)}{\sigma}}) & \text{برای همهٔ مقادیر } x \end{cases}$$

[9] Nondegenerate

[10] Domain of Attraction

نوشت، که در آنها μ و σ به ترتیب پارامترهای مکانی و مقیاس هستند و γ پارامتر شکل [11] نامیده میشود. با توجه به اینکه

$$\lim_{\gamma\to\circ}(\text{۱}+\gamma x)^{\text{۱}/\gamma}=exp(x)$$

ملاحظه می‌شود $\lim_{\gamma\to\circ}G_\gamma(x)=G_\circ(x)$.

۱-۲ توزیع پارتوی تعمیم یافته

با توجه به اینکه ، توابع توزیع پارتوی تعمیم یافته (GPD) توابع توزیع پارامتری مناسبی برای فزونیها هستند، شناخت فرم و خواص آنها برای این تحقیق، بسیار مهم و ضروری است. فرم کلی توابع توزیع پارتوی تعمیم یافته بصورت

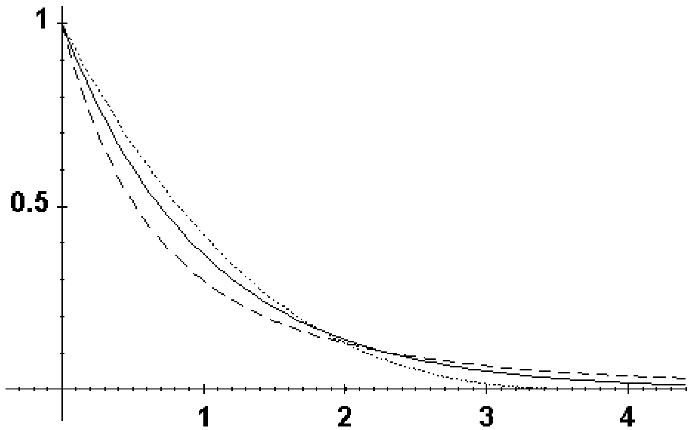
$$W_{k,\sigma}(x)=\begin{cases}\text{۱}-exp(-\frac{x}{\sigma}) & \text{برای}\ \ k=\circ,\sigma>\circ\\ \text{۱}-(\text{۱}-\frac{kx}{\sigma})^{\frac{1}{k}} & \text{برای}\ \ k\neq\circ,\sigma>\circ\end{cases}$$

می‌باشد، که در آن σ پارامتر مقیاس و k پارامتر شکل است. توجه شود که برای $k\leq\circ$، $x>\circ$ است و برای $k>\circ$، محدودهٔ x بصورت $\circ<x<\sigma/k$ می‌باشد. در اینجا نیز $\lim_{k\to\circ}W_{k,\sigma}(x)=W_{\circ,\sigma}(x)$ است. در توابع توزیع پارتوی تعمیم یافته:

- هنگامیکه $k=\circ$ است، تابع توزیع پارتوی تعمیم یافته به تابع توزیع $EXP(\sigma)$ تبدیل می‌شود.
- هنگامیکه $k=\text{۱}$ است، تابع توزیع پارتوی تعمیم یافته به تابع توزیع $UNIF(\circ,\sigma)$ تبدیل می‌شود.
- هنگامیکه $k<\circ$ است، تابع توزیع پارتوی تعمیم یافته به تابع توزیع $Pareto(k,\sigma)$ تبدیل می‌شود.

شکل ۱ نمودار توابع چگالی پارتوی تعمیم یافته را برای kهای مختلف نشان می‌دهد.

[11] Shape Parameter



شکلِ ۱: چگالی نمایی (خط ممتد) و چگالیهای پارتوی تعمیم یافته برای پارامترهای شکل$k = -۰/۵$(خط چین) و $k = ۰/۲۵$ (نقطه چین)

۲ برآورد پارامترهای توزیع پارتوی تعمیم یافته

در این بخش، چهار روش متفاوت برای برآورد پارامترهای توزیع پارتوی تعمیم یافته تشریح می‌شوند.

۲-۱ برآوردگرهای ماکزیمم درستنمایی

فرض کنید که $X = \{X_۱, \dots, X_n\}$ یک نمونهٔ تصادفی از توزیع $GP(k,\sigma)$ و $X_{(n)}$ بزرگترین آمارهٔ ترتیبی آن باشد. لگاریتم تابع درستنمایی بصورت

$$L(k,\sigma;x) = \begin{cases} -n\log\sigma + (۱/k - ۱)\sum_{i=۱}^{n}\log(۱ - \frac{kx_i}{\sigma}), & k \neq ۰ \\ n\log\sigma - ۱/\sigma\sum_{i=۱}^{n} x_i, & k = ۰ \end{cases}$$

است، که در آن محدودهٔ σ برای $k \leq ۰$ برابر $\sigma > ۰$ و برای $k > ۰$ برابر $\sigma > kx_{(n)}$ میباشد.
برای $k \neq ۰$،درایه‌های بردار مشتق لگاریتم تابع درستنمایی GPD بصورت

$$\begin{aligned} \frac{\partial L(k,\sigma;x)}{\partial k} &= n/k(۱/k - ۱) - \frac{۱}{k^۲}\sum_{i=۱}^{n}\ln(۱ - kx_i/\sigma) \\ &- ۱/k(۱/k - ۱)\sum_{i=۱}^{n}(۱ - kx_i/\sigma)^{-۱} \end{aligned}$$

و

$$\frac{\partial L(k,\sigma;x)}{\partial \sigma} = -\frac{n}{k\sigma} + ۱/\sigma(۱/k-۱)\sum_{i=۱}^{n}(۱-\frac{kx_i}{\sigma})^{-۱}.$$

هستند. اگر $k > ۱$ باشد، هیچگونه برآورد ماکزیمم درستنمایی وجود ندارد، زیرا برای هر $k > ۱$ داریم:

$$\lim_{\sigma/k \to x_{(n)}^{+}} L(k,\sigma;x) = \infty$$

بنابراین شرط $k \leq ۱$ برای بدست آوردن یک ماکزیمم متناهی برای لگاریتم درستنماییGPD ضروری است. برای $k = ۰$، در بر دار مشتق لگاریتم درستنمایی داریم

$$\lim_{k \to ۰} \frac{\partial L(k,\sigma;x)}{\partial k} = \sum_{i=۱}^{n} \frac{x_i^۲}{۲\sigma^۲} - \sum_{i=۱}^{n} \frac{x_i}{\sigma}$$

$$\lim_{k \to ۰} \frac{\partial L(k,\sigma;x)}{\partial \sigma} = ۱/\sigma(\sum_{i=۱}^{n} \frac{x_i}{\sigma} - n)$$

که برابر صفر هستند اگر و فقط اگر $۲\bar{x}^۲ = \frac{۱}{n}\sum_{i=۱}^{n} x_i^۲$ باشد. بنابراین اگر این شرط برقرار نباشد، حالت $k = ۰$ میتواند از فرضیات کاسته شود. بردار مشتق لگاریتم درستنمایی را در نظر بگیرید. جواب معادلات

$$\begin{cases} \frac{\partial L(k,\sigma;x)}{\partial k} = ۰ \\ \frac{\partial L(k,\sigma;x)}{\partial \sigma} = ۰ \end{cases}$$

را می‌توان بصورت

$$\begin{cases} n(\hat{k}-۱) = \sum_{i=۱}^{n} \ln(۱-\frac{\hat{k}x_i}{\hat{\sigma}}) + (\hat{k}-۱)\sum_{i=۱}^{n}(۱-\frac{\hat{k}x_i}{\hat{\sigma}})^{-۱} \\ n = (\hat{k}-۱)\sum_{i=۱}^{n}(۱-\frac{\hat{k}x_i}{\hat{\sigma}})^{-۱} \end{cases}$$

یا

$$\begin{cases} (۱+\frac{۱}{n}\sum_{i=۱}^{n} \ln(۱-\frac{\hat{k}x_i}{\hat{\sigma}}))(\frac{۱}{n}\sum_{i=۱}^{n}(۱+\frac{\hat{k}x_i}{\hat{\sigma}})^{-۱}) = ۱ \\ \hat{k} = -\frac{۱}{n}\sum_{i=۱}^{n} \ln(۱-\frac{\hat{k}x_i}{\hat{\sigma}}) \end{cases}$$

نوشت. برآوردهای ماکزیمم درستنمایی بایستی بصورت عددی محاسبه شوند؛ زیرا آماره‌های بسندهٔ مینی‌مال[12] برای GPD، آماره‌های ترتیبی[13] هستند و معادلات درستنمایی غیر خطی به هیچ وجه ساده‌تر نمیشوند. هاسکینگ و والیس (۱۹۷۸) برای پیدا کردن برآورد ماکزیمم درستنمایی حالت تعدیل شده‌ای[14] از آلگوریتم نیوتن-رافسون را بکار بستند. همچنین گریمشاو (۱۹۹۳) آلگوریتم کاملی را برای پیدا کردن برآورد ماکزیمم درستنمایی ارائه داده است.

۲-۲ برآوردگرهای گشتاوری

اگر $k > -1/r$ باشد، rامین گشتاور تابع توزیع پارتوی تعمیم یافته با پارامترهای σ و k، موجود است. لذا اگر متغیر تصادفی X دارای توزیع پارتوی تعمیم یافته با $1 + rk > 0$ باشد، آنگاه

$$E(1 + kX/\sigma)^r = 1/(1 + rk)$$

است. بنابراین میانگین و واریانس آن بشرط وجود بترتیب عبارتند از

$$E(X) = \sigma/(1 + k)$$
$$Var(X) = \sigma^2/(1 + k)^2(1 + 2k)$$

بنابراین برآوردهای گشتاوری σ و k بصورت

$$\hat{\sigma} = \frac{1}{2}\bar{x}(\bar{x}^2/s^2 + 1) \qquad \text{و} \qquad \hat{k} = \frac{1}{2}(\bar{x}^2/s^2 - 1)$$

خواهند بود، که در آنها $\bar{x}$ و s^2، میانگین و واریانس نمونه می‌باشند.

۳-۲ روش گشتاور وزنی-احتمالی

گشتاورهای وزنی-احتمالی متغیر تصادفی X با تابع توزیع F کمیتهای

$$M_{p,r,s} = E(X^p(F(X))^r(1 - F(X))^s)$$

[12] Minimal Sufficient Statistics

[13] Order Statistics

[14] Modified

هستند، که در آن s، p و r اعدادی حقیقی هستند.
برای توزیع GP استفاده از گشتاورهای

$$\alpha_s = M_{1,0,s} = E(X(1-F(X))^s) = \frac{\sigma}{(s+1)(s+1+k)}$$

که برای $k > -1$ وجود دارند و توسط هاسکینگ و والیس (۱۹۸۷) برای $-1/2 < k < 1/2$ ارائه شده‌اند، مناسب هستند. در این صورت پارامترهای توزیع GP بصورت

$$\sigma = \frac{2\alpha_0\alpha_1}{\alpha_0 - 2\alpha_1} \qquad \text{و} \qquad k = \frac{\alpha_0}{\alpha_0 - 2\alpha_1} - 2$$

داده می‌شوند. برآوردگر گشتاورهای وزنی-احتمالی $\hat{\sigma}$ و $\hat{k}$ با جایگزین کردن برآوردگرهای حاصل از یک نمونه مشاهده شده به اندازهٔ n، به جای α_0, α_1 در رابطهٔ (۱) بدست می‌آیند. یک برآورد برای α_r عبارتست از

$$a_s = n^{-1}\sum_{j=1}^{n}\frac{(n-j)(n-j-1)\ldots(n-j-s+1)}{(n-1)(n-2)\ldots(n-s)}x_{j:n}$$

که در آن $x_{1:n} \leq \ldots \leq x_{n:n}$ نمونه‌های مرتب شده هستند و $p_{j:n} = (j)/(n+1)$ است . یک مسئلهٔ جدی برای برآوردهای MOM و PWM این است که امکان دارد با مقادیر نمونه‌ای مشاهده شده سازگار نباشند؛ یعنی ممکن است بعضی از مقادیر نمونه‌ای خارج از محدودهٔ پیشنهاد شده بوسیلهٔ مقادیر پارامتر برآورد شده قرار گیرند. این وضعیت وقتی رخ می‌دهد که $\hat{\sigma}/\hat{k} < x_{n:n}$ باشد، که در آن $x_{n:n}$ بزرگترین آمارهٔ مرتب نمونه‌ای به اندازهٔ n است.

۲-۴ روش صدکهای اصلی

در این روش که توسط کاستیلو و هادی (۱۹۹۷) ارائه شده است، در مرحلهٔ اول برآوردهایی مقدماتی برای پارامترها محاسبه می‌شوند و با ترکیب آنها در مرحلهٔ دوم برآوردگرهایی برای همهٔ پارامترها بدست می‌آید. برای مشاهده این الگوریتمها و نحوه محاسبه برآوردگرهای این روش، به کاستیلو و هادی (۱۹۹۷) مراجعه فرمایید.

۳ مدلبندی مقادیر فرین فاکتورهای مهم اقلیمی شهر تبریز

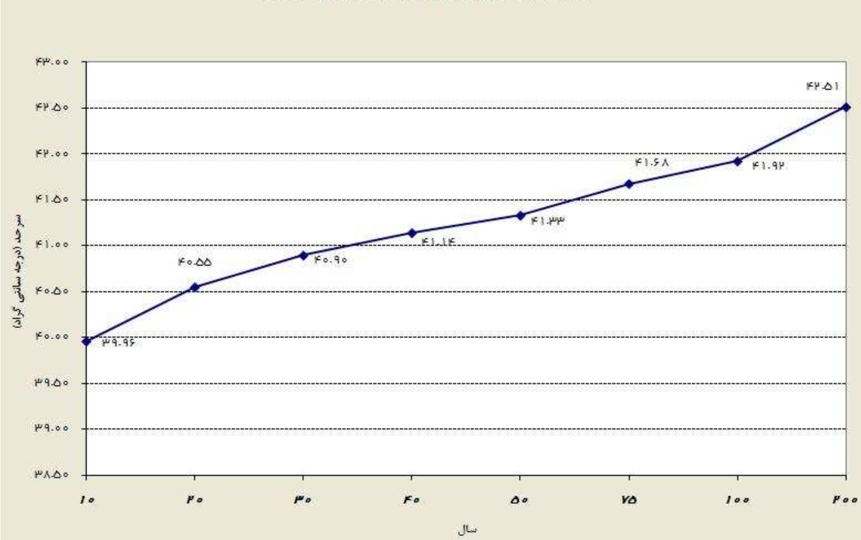
در این بخش داده های مقادیر فرین آب و هوای شهر تبریز شامل عوامل «ماکزیمم سرعت باد»، «ماکزیمم درجه حرارت » و «میزان بارندگی » مورد تجزیه و تحلیل قرار گرفته اند. به این ترتیب که ابتدا به هر کدام از عاملها مدل مناسب برازش داده شده، پارامترهای مدل برآورد شده نیکویی برازش مورد ارزیابی قرار گرفته و در نهایت برخی پیش بینیها انجام شده است. محاسبات توسط نرم افزار$Xtremes$ انجام شده است. این نرم افزار به روشهای مختلف، برآورد پارامترها را انجام می دهد که ما با توجه به نکات عنوان شده در قسمتهای قبلی بهترین آنها را انتخاب نموده و نتایج را ارائه داده ایم. سرانجام برخی پیش بینیها راجع به میانگین دوره بازگشت را برای هر عامل با محاسبه سرحد Tـسال و رسم نمودار آن برای ۱۰, ۲۰, ۳۰, ۴۰, ۵۰, ۷۵, ۱۰۰, ۲۰۰ سال به انجام رسانده ایم. سرحدی مانند u را که در آن میانگین اولین زمان فزونی از u برابرT است سرحد Tـسال نامیده می شود و برای توزیع پارتوی تعمیم یافته چندک $(1 - \frac{1}{T})$ ام تابع توزیع F است. در روش فزونیها یکی از نکات بسیار مهم انتخاب سرحد است. یکی از روشهای معمول این است که ابتدا به روش ماکزیما توزیع GEV را به داده ها برازش داده سپس پارامتر میانگین آنرا به عنوان سرحد برای روش فزونیها قرار می دهند. (اسمیت (۱۹۸۴)). بنابراین ما نیز در اتنخاب سرحدها برای تمامی فاکتورهاازاین روش استفاده کرده‌ایم.

۳-۱ ماکزیمم درجه حرارت مطلق شهر تبریز

برای این عامل، ماکزیمم درجه حرارت مطلق ماهیانه شهر تبریز بر حسب سانتیگراد از فروردین سال ۱۳۸۰ تا اسفند سال ۱۳۸۶ را در اختیار داشتیم. با در نظر گرفتن ماکزیمای سالیانه توزیع مقادیر فرین با $\gamma = 0.23, \mu = 37.99, \sigma = 1.26$ به داده ها برازش شد. بنابراین با انتخاب سرحد ۳۸ برای داده های ماکزیمم درجه حرارت ماهیانه شهر تبریز، تعداد ۵۵ داده که بیش از این سرحد بودند انتخاب شده و به آنها توزیع GP با $k = -0.22, \sigma = 0.68$ بدست آمد. با انجام آزمون $H_0 : k = 0$ (فرض پیروی از توزیع نمایی) $P - Value = 0.57$ بدست آمد. بنابراین با محاسبه مجدد پارامترها با این فرض، در نهایت توزیع نمایی با $\sigma = 0.85$ به مقدار فزونیها برازش شد که نیکویی برازش بسیار خوبی را نشان داد.

برای انجام پیش بینیها مقادیر سرحد T سال را با توجه به مدل پارتوی تعمیم یافته برازش

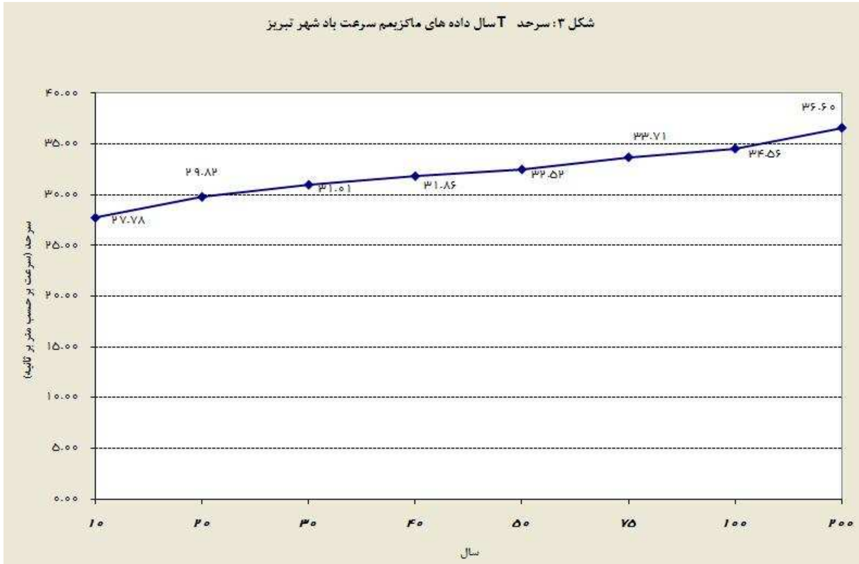
شکل ۲: سرحد T سال ماکزیمم درجه حرارت مطلق شهر تبریز



داده شده به داده ها آورده ایم. شکل ۲ نشان دهنده مقادیر سرحد T سال ماکزیمم درجه حرارت مطلق شهر تبریز است. با توجه به این شکل مثلا بطور متوسط هر ۳۰ سال یک بار دمای ماکزیمم از ۴۰.۹ درجه بیشتر می شود. همچنین بطور مثال برای اینکه ماکزیمم دما از ۴۲ درجه بیشتر شود بطور متوسط ۱۰۰ سال زمان لازم است.

۳-۲ ماکزیمم سرعت باد شهر تبریز

برای این عامل ماکزیمم سرعت باد ماهیانه شهر تبریز بر حسب متر بر ثانیه را از فروردین سال ۱۳۶۱ تا اسفند سال ۱۳۸۶ در اختیار داشتیم. با گرفتن ماکزیمای سالیانه توزیع GEV با $\gamma = ۰.۱, \mu = ۲۰.۴۸, \sigma = ۲.۱۶$ به داده ها برازش شد. با انتخاب سرحد ۲۱ متر بر ثانیه برای این داده ها تعداد ۱۹ ماه که در آنها ماکزیمم سرعت باد بیش از این سرحد بودند انتخاب شده و به آنها توزیع GP با $k = ۰.۲۸, \sigma = ۳.۸۳$ بدست آمد. آزمون تبعیت از توزیع نمایی منجر به $P-Value = ۰.۵۵$ شد و با قبول این فرض و برآورد مجدد پارامترها توزیع نمایی با $\sigma = ۲.۹۴$ را به عنوان مدل نهایی پذیرفتیم که برازش خوبی را نشان می دهد.
شکل ۳ نشان دهنده مقادیر سرحد T سال ماکزیمم سرعت باد شهر تبریز است. با توجه به این شکل مثلا بطور متوسط هر ۵۰ سال یکبار سرعت باد ماکزیمم از ۳۲.۵۲ متر بر ثانیه



شکل ۳: سرحد T سال داده های ماکزیمم سرعت باد شهر تبریز

بیشتر می شود. همچنین بطور مثال برای اینکه سرعت باد از ۲۷.۷۸ متر بر ثانیه بیشتر شود بطور متوسط ۱۰ سال زمان لازم است.

۳-۳ میزان بارندگی شهر تبریز

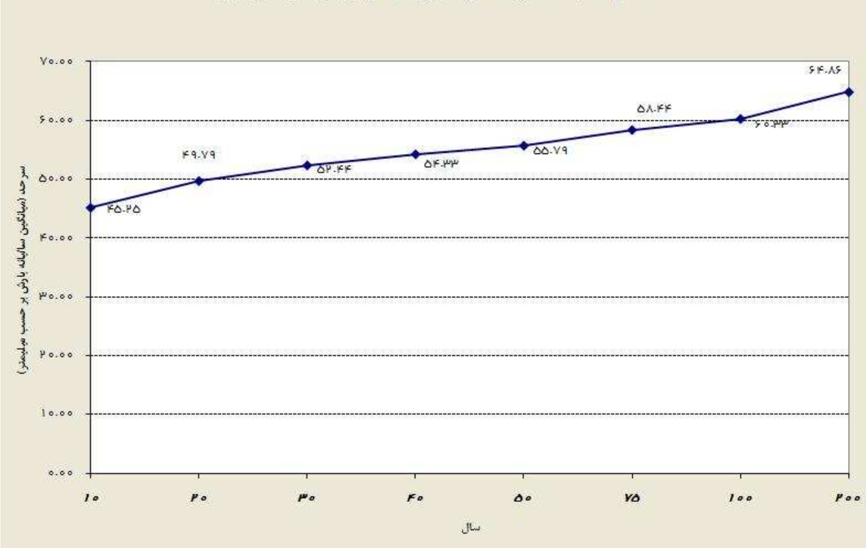
برای این عامل میانگین میزان بارندگی ماهیانه شهر تبریز بر حسب میلیمتر از فروردین سال ۱۳۳۰ تا اسفند ۱۳۸۶ را در اختیار داشتیم. ما برای این عامل دو مقدار سرحد را یکی برای مقادیر زیاد و دیگری برای مقادیر کم بارش انتخاب نمودیم.

با انتخاب سرحد ۳۰ میلیلمتر، تعداد ۱۰ سال که در آنها میانگین سالیانه میزان بارندگی بیش از این سرحد بودند انتخاب شده و به آنها توزیع GP با $k = ۰.۱۷, \sigma = ۵.۴۹$ برازش داده ایم. آزمون $H_۰ : k = ۰$ منجر به $P - Value = ۰.۵۱$ شد و با قبول فرض، توزیع نمایی با $\sigma = ۶.۵۴$ را به عنوان مدل نهایی پذیرفتیم.

شکل ۴ نشان دهنده مقادیر سرحد T سال مقادیر زیاد بارندگی شهر تبریز است که با توجه به آن مثلا بطور متوسط هر ۵۰ سال یکبار میانگین بارندگی سالیانه از ۵۵.۷۹ میلیمتر بیشتر می شود. همچنین بطور مثال برای اینکه میانگین بارندگی سالیانه از ۶۰.۳۳ بیشتر شود بطور متوسط ۱۰۰ سال زمان لازم است.

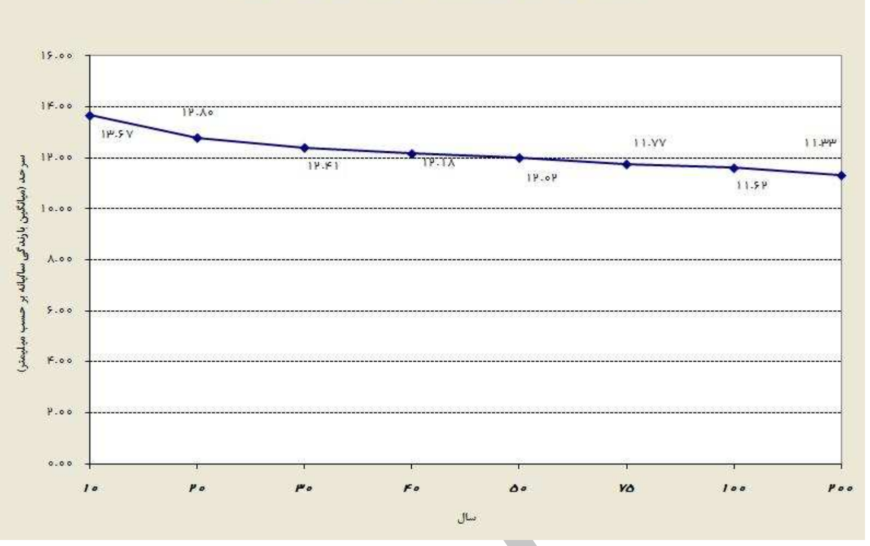
با انتخاب سرحد پایین ۲۰ میلیلمتر ، تعداد ۱۹ سال که در آنها میانگین سالیانه میزان

شکل ۴: سرحد T سال داده های مقادیر زیاد میانگین بارندگی سالیانه شهر تبریز



بارندگی کمتر از این سرحد بودند انتخاب شده و به آنها توزیع GP با $k = ۰.۴۸, \sigma = ۴.۵$ برازش داده ایم. آزمون $H_۰ : k = ۰$ منجر به $P - Value = ۰.۱۱$ شد که با توجه به کم بودن آن فرض پیروی از توزیع نمایی را نپذیرفته و همان مدل قبلی را لحاظ کردیم .

شکل ۵ نشان دهنده مقادیر سرحد T سال مقادیر کم بارندگی شهر تبریز است که با توجه به آن مثلا بطور متوسط هر ۳۰ سال یکبار میانگین بارندگی سالیانه از ۱۲.۴۱ میلیمتر کمتر می شود. همچنین بطور مثال برای اینکه میانگین بارندگی سالیانه از ۱۱.۷۷ کمتر شود بطور متوسط ۷۵ سال زمان لازم است.



شکل ۵: سرحد T سال داده های مقادیر کم میانگین بارندگی سالیانه شهر تبریز

مراجع

[1] Castillo, E. (1994), Extremes in Engineering Applications,. *in Proceeding of the Conference on Extreme Value Theory and Applications,. eds. J. Galambos et al., Geithburg, MD: Kluwer, pp. 15 - 42*

[2] Castillo, E. and Hadi, A. (1997), Fitting Generalized Pareto Distribution to Data,. *JASA, Vol. 92, No. 440, Theory and Method,* 1609 - 1620

[3] Davison, A. C. (1994), Modelling Excesses over High Thresholds, with an Application,. *in Statistical Extremes and Applications, ed. J. Tiago de Oliveira, NATO ASI Series, Dordresht: Reidel, pp.*461 - 482.

[4] Davison, A. C. and Smith, R. L. (1990), Models for Exceedances over High Thresholds,. *J. R. Statist. Soc. B 52, No. 3,* 393 - 442.

[5] Grimshaw, S. D. (1993), Computing Maximum Likelihood Estimation for the Generalized Pareto Distribution,. *Technometrics, Vol. 35, No. 2,* 185 - 191

[6] Hosking, J. R. M. and Wallis, J. R. (1987), Parameter and Quantile

Estimation for the Generalized Pareto Distribution,. *Technometrics, 29,* 339 - 349.

[7] Hosking, J. R. M., Wallis, J. R. and Wood, E. F. (1985), Estimation of The Generalized Extreme Value Distribution by the Method of Probability - Weighted Moments,. *Technometrics, 27,* 251 - 261

[8] Pickands, J. (1975), Statistical Inference Using Extreme Order Statistics,. *The Annals of Statistics, Vol. 3, No. 1*, 119 - 131.

[9] Reiss, R.- D. (1997),. *Statistical Analysis of Extreme Value, Birkhauser, Basel.*

[10] Rootzen, H. and Tajvidi, N. (1997), extreme Value Statistics and Wind Storm Losses: A Case Study,. *Scandinavian Acturial Journal, 1:* 70 - 94.

[11] Smith, R.L.(1984),Threshold Method for Sample Extremes. *In Statistical Extremes and Applications, ed T. Tiago de Oliveiria, Dordrecht: Reidal,* 261-638 .