بررسی تأثیر استفاده از آمار تبدیل‌شده در طراحی شبکه‌ی پایش باران بابهره‌گیری از مفهوم ورگشت

محمود رضا شقاقیان[۱]

تاریخ دریافت: ۹۵/۲/۱۳ تاریخ پذیرش: ۹۵/۶/۶

چکیده

در میان پدیده‌های مختلف چرخه‌ی آبشناسی، بارندگی به عنوان یکی از مهمترین آن‌ها شناخته می‌شود. برداشت صحیح از آمار در این زمینه به وسیله‌ی مجموعه‌ی ایستگاه‌های پایش باران (شبکه‌ی پایش)، کمک مهمی به شبیه‌سازی این پدیده و برنامه‌ریزی‌های مربوط به حوضه‌ی آبخیز می‌کند. از این‌رو طراحی یک شبکه‌ی پایش مناسب بارندگی در منطقه، شامل تعداد ایستگاه‌های بهینه را می‌توان از اولویت‌های هر مطالعه‌ی آبشناسی دانست. شاخص‌های مختلفی در انعکاس دقت و حجم اطلاعات برداشت شده از شبکه‌ی پایش استفاده می‌گردد، که یکی از پرکاربردترین آنها ورگشت می‌باشد. در مطالعات گذشته با فرض بهنجار بودن آمار بارندگی، از معادلات منطبق با این توزیع آماری استفاده گردیده که نبودن این فرض، می‌تواند به عدم دقت لازم نتایج منجر گردد. با توجه به عدم دقت به این مسأله در مطالعات گذشته، در این تحقیق روش نگاشت داده‌های اولیه با استفاده از تبدیل باکس-کاکس با فراسنج‌های بهینه، به منظور اطمینان استفاده از فرض توزیع آماری بهنجار، پیشنهاد گردیده است. بدین منظور با طراحی یک نمایشنامه، داده‌های ماهانه مربوط به ۲۴ ایستگاه پایش بارندگی استان فارس که دارای توزیع آماری غیر بهنجار می‌باشند، بهنجار گردیده و بعد از آن شبیه‌سازی‌های مرسوم طراحی شبکه‌های پایش با استفاده از مفهوم ورگشت برای این داده‌ها استفاده قرار گرفته است. علاوه بر آن، طراحی شبکه‌ی بارانسنجی با استفاده از روشهای متداول نیز به منظور مقایسه با روش پیشنهادی، انجام پذیرفته است. نتایج به‌دست آمده نشان می‌دهند که نرخ رشد ورگشت با افزایش تعداد ایستگاه‌ها در حالت داده‌های بهنجار شده کاهش می‌یابد، در حالی‌که در روش متداول طراحی این نرخ ثابت می‌باشد. علاوه بر آن اولویت ایستگاه‌های انتخابی در روش‌های مطرح شده تغییر نموده است.

واژگان کلیدی: طراحی شبکه پایش باران، توزیع بهنجار، ورگشت، تبدیل باکس-کاکس

[۱] گروه عمران، دانشکده مهندسی، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران
ایمیل (نویسنده مسئول): shaghaghian_mr@iaushiraz.ac.ir

Archive of SID
www.SID.ir

مقدمه

از جمله اطلاعات مورد لزوم در برنامه‌ریزی، ساخت و بهره‌برداری از سازه‌های هیدرولیکی، داده‌های مربوط به بارندگی می‌باشد. این داده‌ها عمدتاً از شبکه‌های پایش باران حاصل گردیده و دقت در آنها به مؤثر بودن و اقتصاد پروژه‌هایی که بر مبنای این آمار به دست آمده است کمک می‌کند. از این‌رو شبکه‌های پایش بایستی با حداقل ابزار پایش و حداکثر دقت به اندازه‌گیری بارندگی در مقیاس زمانی و مکانی مورد نظر بپردازد.

واریانس مقادیر مشاهده، واریانس خطا، حداقل فاصله، بعد فراکتال و ورگشت از جمله فراسنجهایی می‌باشند که برای بیان دقت شبکه پایش مورد استفاده قرار می‌گیرد (چنگ و همکاران، ۲۰۰۸؛ مازرالا و ترافاگلیا، ۲۰۰۰؛ وای و همکاران، ۲۰۱۴ و شقاقیان و عابدینی، ۲۰۱۳). از بین این فراسنجها ورگشت به علت در نظر گرفتن اثرات مقیاس زمانی از مقبولیت خاصی برخوردار است. گاهی اوقات نیز با ترکیب این فراسنج با مفاهیم زمین‌آماری، مقیاس‌های مکانی نیز در نظر گفته می‌شود (چن و همکاران، ۲۰۰۸).

با توجه به اینکه در اغلب روش‌های پیشنهادی طراحی شبکه‌های پایش باران نیاز به محاسبه ورگشت توأمان می‌باشد، تعیین سازوکاری برای محاسبه احتمال توأمان دو پدیده الزامی‌می‌باشد (حسین، ۱۹۸۹). با توجه به مشکلات محاسباتی در تابع توزیع احتمال دلخواه پدیده معمولا در اغلب مطالعات، تابع توزیع احتمال در محاسبه ورگشت توأمان به صورت بهنجار در نظر گرفته می‌شود (کریستناویچ ۱، ۱۹۹۲؛ کریستناویچ ۲، ۱۹۹۲؛ یه و همکاران، ۲۰۱۱؛ حسین، ۱۹۸۷؛ الظهرانی و حسین، ۱۹۹۸ و معمارزاده و همکاران، ۲۰۱۲). این موضوع ممکن است در صورت عدم تطابق تابع توزیع بهنجار با توابع توزیع منطبق بر پدیده، ممکن است باعث عدم حصول جواب‌های دقیق می‌گردد که با بررسی‌های صورت گرفته به وسیله این محقق در مطالعات پیشین به آن توجهی نگردیده است.

به‌منظور افزایش دقت نتایج به دست آمده از طراحی شبکه‌های پایش با استفاده از مفهوم ورگشت و برطرف نمودن این خلا تحقیقاتی، دو رویکرد مجزا را می‌توان پیشنهاد نمود. در روش نخست، با در نظر گرفتن توزیع آماری واقعی به جای توزیع بهنجار، از روابط متناسب با آن توزیع استفاده نموده و روابط را بهبود می‌بخشد (به عنوان مثال، با توجه به توزیع نمایی بارندگی ماهیانه، تابع احتمال چندمتغیره نمایی به جای احتمال چندمتغیره بهنجار، برای بارندگی ماهانه در نظر گرفته شود (یه و همکاران، ۲۰۰۸) در روش پیشنهادی دوم، به جای بهبود روابط مورد استفاده که دارای فرایند پیچیده‌ای می‌باشد، با تبدیل داده‌های اولیه و انطباق توزیع بهنجار بر آن‌ها، شرایط برای استفاده از روابط متداول مهیا می‌گردد که به عنوان یک راهکار جدید در این تحقیق مورد استفاده قرار می‌گیرد.

در این مقاله ابتدا طراحی یک شبکه پایش بارش با داده‌های طبیعی ماهانه، با درنظر گرفتن شبه‌سازی‌های مرسوم مبتی بر ورگشت و با فرض توزیع بهنجار به‌دست می‌آید. پس از آن، به منظور ارزیابی روش پیشنهادی، با انتخاب یک تبدیل مناسب داده‌ها بهنجار می‌گردند. در ادامه با استفاده از داده‌های بهنجار شده مجددا طراحی شبکه پایش انجام می‌پذیرد. در پایان نتایج هر دو نمایشنامه با یکدیگر مقایسه شده و اثر عدم توجه به توزیع آماری اولیه درمورد داده‌ها مشخص می‌گردد.

مواد و روش‌ها

معرفی ورگشت

ورگشت اطلاعات که توسط شانون معرفی شده است بر پایه احتمالات می‌باشد (شانون، ۱۹۴۸). در حقیقت این مفهوم نشان‌دهنده عدم قطعیت می‌باشد و نشان‌دهنده حجم اطلاعات می‌باشد. از نظر ریاضی، در صورتیکه متغیر X یک متغیر تصادفی پیوسته با توزیع احتمال p(x) درنظر گرفته شود، ورگشت از رابطه زیر محاسبه می‌گردد:

$$H(X) = -\int p(x) \ln p(x) dx \qquad (۱)$$

Archive of SID

www.SID.ir

که در آن H(X) مقدار ورگشت p(x) احتمال وقوع x می‌باشد. در صورتی‌که متغیر تصادفی X، مقدار بارندگی در یک ایستگاه پایش باران باشد، مقدار H(X) مشخص کننده حجم اطلاعات مربوط به آن ایستگاه می‌باشد. در صورتی‌که حجم اطلاعات مربوط به دو یا چند ایستگاه مد نظر باشد، از ورگشت توأم استفاده می‌گردد. دراین‌صورت مقدار ورگشت به شکل زیر محاسبه می‌گردد:

$$(X^1,X^2,...,X^n)= \int p(x^1,x^2,...,x^n)\ln p(x^1,x^2,...,x^n)dx^1dx^2..dx^n \quad (۲)$$

که در آن $H(X^1,X^2,...,X^n)$ و $p(x^1,x^2,...,x^n)$ احتمال وقوع توأم x^1 و x^2 و ... می‌باشد. از ورگشت شرطی برای محاسبه مقدار ورگشت یک ایستگاه به شرط داشتن اطلاعات m ایستگاه دیگر استفاده می‌شود. در این‌صورت می‌بایستی از احتمالات شرطی به شکل زیر استفاده کرد:

$$p(X^{m+1}\mid X^1,X^2,...,X^m)=\frac{p(x^1,x^2,...,x^{m+1})}{p(x^1,x^2,...,x^m)} \quad (۳)$$

که در آن $p(x^1,x^2,...,x^m)$ احتمال وقوع توأم x^1 و x^2 و ... و x^m می‌باشد. در این حالت مشخص است که:

$$H(X^1, X^2, ..., X^{m+1})= H(X^1, X^2, ..., X^m) +H(X^{m+1}| X^1, X^2, ..., X^m) \quad (۴)$$

بنابراین با استفاده از رابطه ۴، ورگشت شرطی را می‌توان به صورت زیر محاسبه نمود:

$$H(X^{m+1}| X^1, X^2, ..., X^m)= H(X^1, X^2, ..., X^{m+1})- H(X^1, X^2, ..., X^m) \quad (۵)$$

محاسبه ورگشت

در اکثر مطالعات انجام شده، به منظور سهولت، تابع بهنجار به عنوان تابع توزیع احتمال بر پدیده در نظر گرفته شده است (هارمانچقلو و یویویچ، ۱۹۸۷ و آلفنسو و همکاران، ۲۰۱۰). در اینصورت ورگشت تنها با داشتن واریانس (□□)، به شکل زیر محاسبه می‌گردد:

$$H(X)=\frac{1}{2}\ln(2\pi e\sigma^2) \quad (۶)$$

همچنین برای محاسبه ورگشت توأم، با در نظر گرفتن واریانس ایستگاه‌های مختلف، از رابطه زیر استفاده می‌گردد:

$$H(X^1,X^2,...,X^n)= \frac{k}{2}[1+\ln(2\pi)]+\frac{1}{2}\ln|\Sigma| \quad (۷)$$

که k تعداد متغیرها و □ ماتریس واریانس-کواریانس می‌باشد. به عنوان مثال در صورتی که دو ایستگاه در نظر گرفته شود،حجم انتقال اطلاعات توسط دو ایستگاه (ورگشت توأم) از رابطه زیر محاسبه می‌گردد:

$$H(X^1,X^2)=[1+\ln(2\pi)] +\frac{1}{2}\ln[\mathrm{var}(X^1)\mathrm{var}(X^2) -\mathrm{cov}(X^1,X^2)] \quad (۸)$$

در این‌صورت ورگشت شرطی (ورگشت ایستگاه X2 به شرط داشتن اطلاعات مربوط به ایستگاه X1) با استفاده از روبط (۵) و (۷) به صورت زیر محاسبه می‌گردد:

$$H(X^1|X^2)= \frac{1}{2}[1+\ln(2\pi)]+\frac{1}{2}\ln\frac{|\Sigma^{1,2}|}{\mathrm{var}(X^1)} \quad (۹)$$

که در آن □□□ ماتریس واریانس-کواریانس بین متغیر X^1 و X^2 می‌باشد.

بهینه‌سازی شبکه پایش

روش مرسوم

اهمیت هر ایستگاه پایش باران در شبکه با توجه به ورگشت آن مشخص می‌گردد. به بیان دیگر، هر چه که ورگشت یک ایستگاه بیشتر باشد، عدم قطعیت مقادیر مشاهده شده در آن ایستگاه بیشتر بوده و در نتیجه آن ایستگاه در اولویت انتخاب قرار می‌گیرد. بعد از تثبیت موقعیت اولین ایستگاه که دارای بیشترین ورگشت می‌باشد، بقیه ایستگاه‌ها به صورت مرحله‌ای با توجه حداکثرسازی ورگشت به شرط داشتن اطلاعات ایستگاه‌های انتخاب شده می‌باشد (وای و همکاران، ۲۰۱۴؛ حسین، ۱۹۸۷ و کریستناویچ ۱، ۱۹۹۲).

در محاسبه ایستگاه (m+1) از n ایستگاه موجود که m ایستگاه قبلا انتخاب گردیده است (m<n)، می‌بایستی از

Archive of SID

www.SID.ir

ورگشت شرطی استفاده نمود. بدین شکل که ایستگاه m+1 به شکلی انتخاب گردد که مقدار زیر که در حقیقت ورگشت ایستگاه m+1 با در نظر گرفتن ایستگاه‌های 1, 2, …, m حداکثر گردد:

$$\mathrm{Max}\{H(X_{m+1}|\, X_1, X_2, \ldots, X_m)\} \quad (۱۰)$$

بعد از انتخاب و اولویت بندی مرحله به مرحله ایستگاه‌ها، ورگشت توأم ایستگاه‌های انتخاب شده بر حسب تعداد ایستگاه‌ها رسم می‌گردد. بعد از این مرحله، ممکن است ملاحظه گردد که با افزایش تعداد ایستگاه‌ها، تأثیر زیادی بر مقدار ورگشت توأم گذاشته نخواهد شد. در این حالت می‌توان تعداد بهینه ایستگاه‌های پایش با در نظر گرفتن اولویت آنها را مشخص نمود.

تصحیح در روش مرسوم:

روش مرسوم در استفاده از آمارهای تبدیل نیافته و بهنجار نشده می‌باشد. در این مقاله، با عنایت به رابطه‌های مربوط به محاسبه ورگشت، می‌بایستی داده‌ها قبل از استفاده در فرایند طراحی شبکه بهینه به استفاده از یک تبدیل مناسب بهنجار شده و شاخص بهنجار بودن بررسی گردد. تبدیل داده‌ها با استفاده از تبدیل باکس-کاکس که به صورت زیر تعریف می‌شود، انجام می‌پذیرد (آلفونسو و همکاران، ۲۰۱۰):

$$Y = \frac{(X+\lambda_2)^{\lambda_1}-1}{\lambda_1 g^{(\lambda_1-1)}} \quad \lambda_1 \neq 0$$
$$Y = g\ln(X+\lambda_2) \quad \lambda_1 = 0 \quad (۱۱)$$

که در آن λ_1 و λ_2 فراسنجهای تبدیل، g میانگین هندسی متغر تصادفی $X+\lambda_2$ و Y داده‌های تبدیل شده می‌باشند.

منطقه و داده‌های مورد مطالعه:

داده‌های مورد استفاده در این تحقیق مربوط به داده‌های ماهانه ۲۴ ایستگاه‌های پایش مورد مطالعه در این تحقیق در جنوب غربی ایران، در استان فارس با حداقل دوره آماری ۲۰ سال می‌باشند. این ایستگاه‌ها بین طول‌های جغرافیایی ′۲۲ °۲۷ و ′۳۱ °۳۲ و عرض‌های ′۳۲ °۵۱ و ′۲۵ °۵۴ در مساحتی حدود ۱۲۰،۰۰۰ کیلومتر مربع واقع گردیده‌اند (جدول شماره ۱). این منطقه در شمال داری اقلیم کوهستانی و در جنوب اغلب به صورت دشت می‌باشد و بازه ارتفاعی آنها بین حدود ۴۵۰ تا ۲۳۰۰ متر از تراز دریا می‌باشد. در شکل ۱ پراکنش جغرافیایی ایستگاه‌ها نشان داده شده است.

منطقه مورد مطالعه بیشتر تحت تأثیر جبهه مدیترانه‌ای از سمت شمال غرب واقع است. میانگین بارندگی آن بین ۶۷۹/۳ میلیمتر (ایستگاه سپیدان) و ۱۳۴/۴ میلیمتر (ایستگاه آباده) می‌باشد.۲همانطور که از شکل ۲ ملاحظه می‌گردد، تغییرات میانگین بارندگی سالانه دارای جهت تقریبی شمال غرب-جنوب شرق می‌باشد. این موضوع تأثیر جبهه غالب را نشان می‌دهد که فعالیت آن در فصل زمستان می‌باشد.

نتایج و بحث

در این بخش تحلیل‌ها و نتایج در دو قسمت ارائه می‌گردند. ابتدا نتایج با استفاده از داده‌های اصلی ارائه شده و در ادامه با استفاده از تبدیل مناسب، ابتدا داده‌هایی با توزیع بهنجار به‌دست آمده و سپس فرایند طراحی شبکه پایش انجام شده است.

طراحی شبکه با استفاده از داده‌های تبدیل نیافته

در این حالت ابتدا ایستگاه پایش با حداکثر مقدار ورگشت به عنوان اولویت اول انتخاب می‌گردد. همانطور که از رابطه (۶) مشخص است ورگشت و واریانس نسبت مستقیم داشته و در نتیجه ایستگاه با حداکثر واریانس ماهیانه دارای حدکثر ورگشت می‌باشد.

با استفاده از جدول (۲) ایستگاه ۲۱ به عنوان اولویت ۱ انتخاب می‌گردد. پس از آن با فرض انتخاب ایستگاه ۲۱ اولویت ۲ به گونه‌ای انتخاب می‌گردد که ورگشت شرطی بدست آمده از رابطه (۹) حداکثر گردد. این فرایند ادامه یافته تا بقیه ایستگاه‌های پایش نیز اولویت بندی گردند.

با استفاده از جدول (۲) ایستگاه ۲۱ به عنوان اولویت ۱ انتخاب می‌گردد. پس از آن با فرض انتخاب ایستگاه ۲۱ اولویت ۲ به گونه‌ای انتخاب می‌گردد که ورگشت شرطی بدست آمده از رابطه (۹) حداکثر گردد. این فرایند ادامه یافته تا بقیه ایستگاه‌های پایش نیز اولویت بندی گردند.

شکل (۳) تغییرات حجم اطلاعات (ورگشت توأم) بر حسب اولویت ایستگاه‌ها نشان می‌دهد. همانطور که در این

[۱] Box-Cox transformation

[۲] مقدار برف به حساب نیامده است.

Archive of SID

www.SID.ir

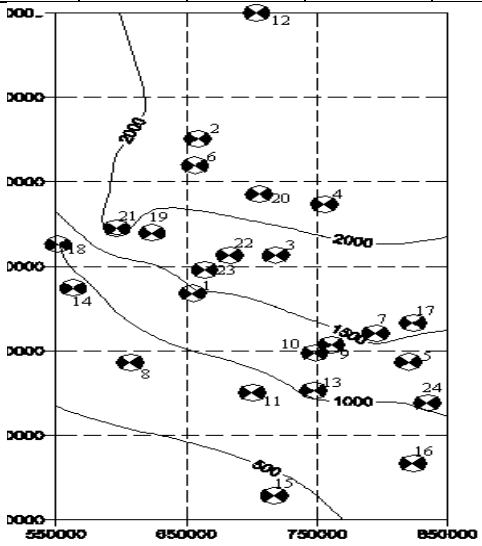
شکل ملاحظه می‌گردد، با توجه به ایستگاه‌های موجود، با افزایش تعداد ایستگاه‌ها، حجم ورگشت نیز به صورت تقریبا خطی افزایش می‌یابد. بنابراین تعداد ایستگاه‌های پایش بهینه مورد نیاز برای بهره جستن از حداکثر حجم اطلاعات منطقه بیشتر از ایستگاه‌های موجود بوده و با عنایت به روند رشد ورگشت تعداد ایستگاه‌های پایش بهینه میسر نمی‌باشد.

طراحی شبکه با استفاده از داده‌های بهنجار شده

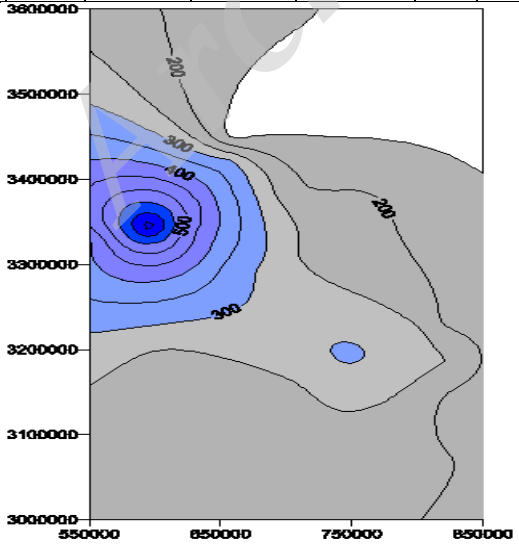
با عنایت به اینکه اغلب رابطه‌های محاسبه ورگشت بر اساس داده‌های با توزیع بهنجار می‌باشند، شایسته است قبل از استفاده از این روابط، بهنجار بودن داده‌ها بررسی گردد. در صورت عدم وجود تطابق توزیع بهنجار این داده‌ها، یا می‌بایستی از روابط مناسب برای توزیع منطبق بر داده‌ها استفاده کرد یا با استفاده از تبدیل مناسب داده‌ها به توزیع بهنجار نزدیک گردند. گزینه نخست دارای پیچیدگی‌های ریاضی بالایی می‌باشد و استفاده از این رویکرد ساده به‌نظر نمی‌رسد. بنابراین بهتر است برای بهره جستن از روابط مربوط به محاسبه ورگشت با توزیع بهنجار، از بهنجارسازی داده‌های اولیه بهره جست.

جدول شماره ۱: مشخصات ایستگاه‌های منطقه مورد مطالعه

شماره ایستگاه	نام ایستگاه	طول جغرافیایی (شرقی)	عرض جغرافیایی (شمالی)	ارتفاع از سطح دریا (m)	میانگین بارندگی سالیانه	شماره ایستگاه	نام ایستگاه	طول جغرافیایی (شرقی)	عرض جغرافیایی (شمالی)	ارتفاع از سطح دریا (m)	میانگین بارندگی سالیانه
۱	شیراز	۵۲°۳۶'	۲۹°۳۲'	۱۴۸۴	۳۲۳/۱	۱۳	جهرم	۵۳°۳۲'	۲۸°۲۹'	۱۰۸۲	۲۷۷/۹
۲	آباده	۵۲°۴۰'	۳۱°۱۱'	۲۰۳۰	۱۳۴/۴	۱۴	کازرون	۵۱°۳۹'	۲۹°۳۶'	۸۶۰	۳۸۰/۰
۳	ارسنجان	۵۳°۱۶'	۲۹°۵۶'	۱۷۳۰	۲۳۲/۲	۱۵	لامرد	۵۳°۱۲'	۲۷°۲۲'	۴۰۵	۲۱۶/۰
۴	بوانات	۵۳°۴۰'	۳۰°۲۸'	۲۲۳۱	۲۱۴/۳	۱۶	لار	۵۴°۱۷'	۲۷°۴۲'	۷۹۲	۲۰۳/۰
۵	داراب	۵۴°۱۷'	۲۸°۴۷'	۱۰۹۸	۲۵۲/۵	۱۷	نیریز	۵۴°۲۰'	۲۹°۱۲'	۱۶۳۲	۱۸۵/۰
۶	اقلید	۵۲°۳۸'	۳۰°۵۴'	۲۳۰۰	۳۲۰/۴	۱۸	نورآباد مسنی	۵۱°۳۲'	۳۰°۰۴'	۹۷۲	۳۹۵/۰
۷	استهبان	۵۴°۰۲'	۲۹°۰۵'	۱۶۹۰	۲۷۰/۸	۱۹	سد دردزن	۵۲°۱۷'	۳۰°۱۱'	۱۶۵۰	۴۷۸/۰
۸	فراشبند	۵۲°۰۶'	۲۸°۴۸'	۷۸۲	۲۲۷/۴	۲۰	صفاشهر	۵۳°۰۹'	۳۰°۳۵'	۲۲۵۱	۲۰۷/۸
۹	فسا	۵۳°۴۱'	۲۸°۵۸'	۱۲۸۸	۲۹۰/۲	۲۱	سپیدان	۵۲°۰۰'	۳۰°۱۴'	۲۲۱۰	۶۷۹/۳
۱۰	فیروزآباد	۵۲°۳۳'	۲۸°۵۳'	۱۳۶۲	۳۱۳/۳	۲۲	تخت جمشید	۵۲°۵۴'	۲۹°۵۶'	۱۶۰۵	۳۰۲/۶
۱۱	قیر و کارزین	۵۳°۰۳'	۲۸°۲۸'	۷۴۶	۲۳۲/۶	۲۳	زرقان	۵۲°۴۲'	۲۹°۴۷'	۱۵۹۶	۳۱۶/۲
۱۲	ایزدخواست	۵۲°۷۰'	۳۲°۳۱'	۲۱۸۸	۱۵۸/۳	۲۴	زرین دشت	۵۴°۲۵'	۲۸°۲۱'	۱۰۲۹	۱۶۹/۱



شکل ۱- موقعیت ایستگاه‌های مورد مطالعه



شکل ۲- پهنه‌بندی میانگین بارش سالیانه در منطقه مورد مطالعه

Archive of SID

www.SID.ir

جدول شماره ۲- مشخصات توزیع آماری بارندگی ماهیانه ایستگاه‌ها برای آمار تبدیل نیافته

شماره ایستگاه	واریانس (mm)	ورگشت	چولگی	κ^2 (نسبت به توزیع بهنجار)	شماره ایستگاه	واریانس (mm)	ورگشت	چولگی	κ^2 (نسبت به توزیع بهنجار)
۱	۴۱/۹	۵/۱۵	۲/۱۶	۶۰۸/۴۱	۱۳	۴۰/۹	۵/۱۳	۳/۱۲	۳۰۱/۲۶
۲	۱۹/۶	۴/۳۹	۳/۲۷	۴۵۸/۶۷	۱۴	۴۶/۳	۵/۲۵	۱/۹۲	۸۰/۷۳
۳	۲۸/۸	۴/۷۸	۱/۹۲	۱۰۳/۹۸	۱۵	۳۳/۳	۴/۹۳	۳/۰۸	۱۸۰/۶۲
۴	۲۷/۷	۴/۷۴	۲/۲۵	۱۴۵/۳۷	۱۶	۳۳/۱	۴/۹۲	۴/۱۲	۱۹۷/۸۵
۵	۳۸/۸	۵/۰۸	۲/۵۱	۳۵۱/۲۳	۱۷	۲۶/۳	۴/۶۹	۲/۷۷	۹۰/۹۴
۶	۴۱/۲	۵/۱۴	۲/۳۵	۲۴۷/۱۵	۱۸	۵۵/۰	۵/۴۳	۱/۷۹	۴۳/۱۴
۷	۳۱/۶	۴/۸۷	۲/۵۹	۱۴۴/۱۶	۱۹	۶۰/۳	۵/۵۲	۱/۹۵	۲۲۵/۰۶
۸	۳۷/۳	۵/۰۴	۲/۷۴	۷۵/۹۸	۲۰	۲۸/۴	۴/۷۷	۲/۶۸	۷۱/۲۳
۹	۴۳/۸	۵/۲۰	۳/۳۶	۶۲۶/۲۶	۲۱	۷۴/۳	۵/۷۳	۱/۸۲	۵۹/۹۴
۱۰	۴۲/۴	۵/۱۷	۱/۹۰	۷۵/۳۱	۲۲	۴۲/۸	۵/۱۸	۲/۸۲	۱۶۱/۳۲
۱۱	۳۸/۳	۵/۰۶	۳/۴۷	۵۷/۲۰	۲۳	۴۲/۳	۵/۱۶	۲/۴۷	۲۴۳/۷۳
۱۲	۲۱/۱	۴/۴۷	۲/۴۹	۲۲۴/۹۷	۲۴	۳۲/۵	۴/۹۰	۳/۷۶	۵۷/۹۶

شاخص‌های متعددی برای بررسی تطابق داده‌ها بر توزیع بهنجار وجود دارد. از جمله این شاخص‌ها می‌توان به چولگی ، که اختصاصا برای بررسی توزیع بهنجار مورد استفاده قرار می‌گیرد و κ^2 که برای هر توزیع دلخواه قابل استفاده است، اشاره نمود. این فراسنجها به شکل زیر محاسبه می‌گردند:

$$\text{چولگی} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^3}{\left[\left(\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2\right)\right]^{\frac{3}{2}}} \qquad (۱۱)$$

$$\kappa^2 = \sum_{i=1}^{NC}\frac{\left(f_i-\hat{f}_i\right)^2}{\hat{f}_i} \qquad (۱۲)$$

که در این روابط x_i داده‌ها، $\bar{x}$ میانگین داده‌ها، N تعداد داده‌ها، f_i تواتر داده‌ها برای یک بازه، $\hat{f}_i$ تواتر داده‌ها در برای یک توزیع آماری مشخص (در اینجا توزیع بهنجار) می‌باشد. مقدار نزدیک به صفر این دو شاخص نمایانگر تطابق بیشتر با توزیع می‌باشد.

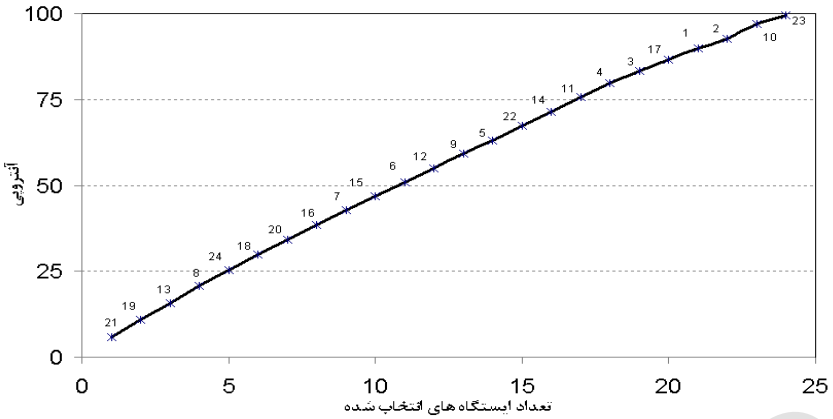
شکل ۴ پراکنش داده‌های مربوط به سه ایستگاه پایش نسبت به توزیع بهنجار و جدول ۲ شاخص‌های چولگی و κ^2 را برای داده‌های مربوط به کل ایستگاه‌های پایش نشان می‌دهند. همانطور که از جدول ۲ ملاحظه می‌گردد، داده‌های[۴] تبدیل نیافته دارای عدم تطابق کامل با توزیع بهنجار می‌باشند (ماکیوال و یوها، ۲۰۱۲). بنابراین این داده‌ها برای استفاده از روش‌های مرسوم طراحی شبکه با استفاده از مفهوم ورگشت نیاز به بهنجار شدن دارند.

معمولا، در صورت لزوم، برای تغییر توزیع آماری داده‌ها از شکل تبدیل یافته آنها استفاده می‌گردد. در این مقاله با استفاده از رابطه ۱۱ از تبدیل باکس-کاکس برای بهنجار کردن داده‌های اولیه استفاده گردیده است. با استفاده از کمینه‌سازی شاخص κ^2 نسبت به توزیع بهنجار ضرایب بهینه λ_1 و λ_2 برای کلیه ایستگاه‌ها محاسبه می‌گردد، شکل ۵ تغییرات شاخص κ^2 نسبت به توزیع بهنجار را بر حسب ضرایب λ_1 و λ_2 برای ایستگاه شماره ۱ (به عنوان نمونه) و جدول ۳، مقادیر λ_1 و λ_2 برای شاخص κ^2 بهینه (کمینه) نشان می‌دهد. شکل ۶ پراکنش داده‌های تبدیل یافته مربوط به سه ایستگاه پایش ذکر شده در شکل ۴ را نسبت به توزیع بهنجار نشان می‌دهد. کلیه داده‌ها پس از تبدیل و بهنجار شدن امکان استفاده بهتر از روش‌های مبتی بر ورگشت را دارا می‌باشند.

[۴] Skewness

Archive of SID

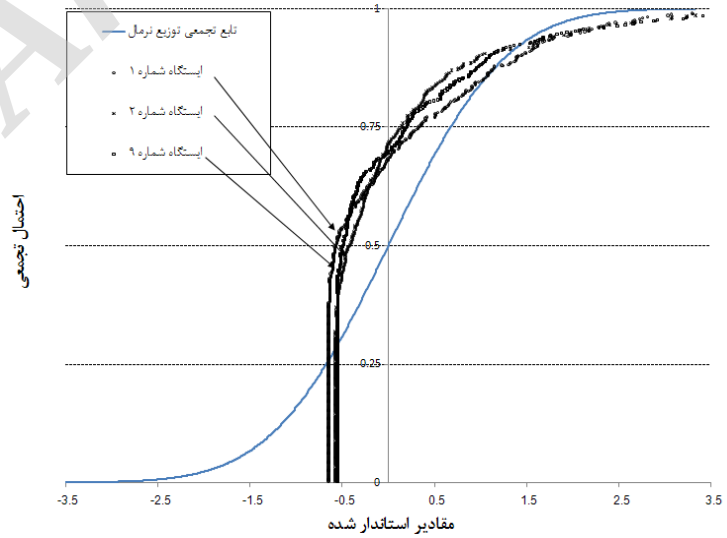
www.SID.ir



شکل ۳- تغییرات حجم ورگشت با تعداد ایستگاه‌های انتخاب شده اولویت بندی شده بدون بهنجارسازی (تبدیل) داده‌ها

جدول شماره ۳- فراسنجهای تبدیل بهینه برای بهنجار سازی داده‌ها

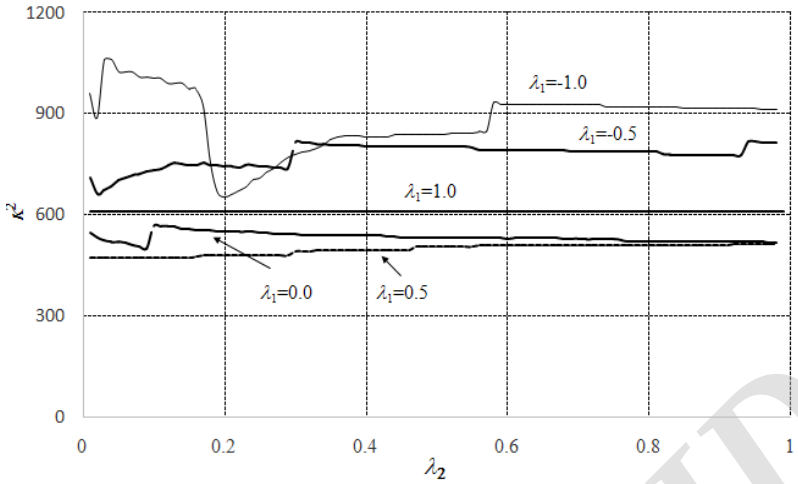
شماره ایستگاه	واریانس (mm)	$\square_1$	$\square_2$	$\square^2$ (نسبت به توزیع بهنجار)	شماره ایستگاه	واریانس (mm)	$\square_1$	$\square_2$	$\square^2$ (نسبت به توزیع بهنجار)
۱	۱۲/۰	۰/۴	۰/۲۵	۲۸۰/۰۳	۱۳	۲/۹	۰/۳	۰/۰۱	۳۰۱/۲۶
۲	۴/۴	۰/۵	۰/۰۱	۱۳۴/۹۲	۱۴	۱۴/۸	۰/۳	۰/۸۲	۸۰/۷۳
۳	۷/۱	۰/۶	۰/۰۱	۸۶/۶۸	۱۵	۴/۴	۰/۳	۰/۰۵	۱۸۰/۶۲
۴	۱۳/۰	۰/۵	۰/۸۵	۶۹/۴۵	۱۶	۵/۷	۰/۴	۰/۰۵	۱۹۷/۸۵
۵	۹/۳	۰/۳	۰/۳۴	۸۷/۴۵	۱۷	۳/۸	۰/۴	۰/۰۱	۹۰/۹۴
۶	۷/۷	۰/۴	۰/۰۱	۹۷/۵۰	۱۸	۱۱/۵	۰/۵	۰/۰۱	۴۳/۱۴
۷	۱۰/۲	۰/۰	۰/۴۴	۴۹/۸۶	۱۹	۱۸/۷	۰/۴	۰/۵۱	۲۲۵/۰۶
۸	۱۲/۰	۰/۴	۰/۵۱	۴۳/۹۵	۲۰	۸/۲	۰/۵	۰/۰۷	۷۱/۲۳
۹	۹/۰	۰/۳	۰/۲۵	۳۵۱/۲۱	۲۱	۱۹/۱	۰/۵	۰/۰۶	۵۹/۹۴
۱۰	۶/۰	۰/۴	۰/۰۱	۴۵/۷۳	۲۲	۵/۱	۰/۴	۰/۰۱	۱۶۱/۳۲
۱۱	۷/۳	۰/۴	۰/۰۸	۳۷/۴۲	۲۳	۱۱/۸	۰/۴	۰/۲۶	۲۴۳/۷۳
۱۲	۷/۹	۰/۴	۰/۴۶	۸۸/۳۲	۲۴	۴/۸	۰/۲	۰/۱۸	۵۷/۹۶



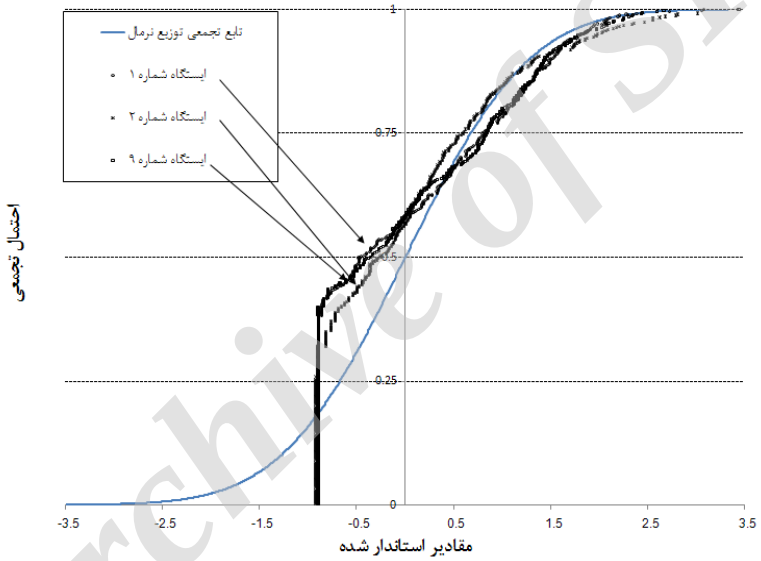
شکل ۴- پراکنش قرائتهای استاندارد شده مربوط به سه ایستگاه نمونه قبل از تبدیل داده‌ها

Archive of SID

www.SID.ir



شکل ۵- تغیرات شاخص با ضرایب تبدیل و برای ایستگاه شماره ۱



شکل ۶- پراکنش قرائتهای استاندارد شده مربوط به سه ایستگاه نمونه بعد قبل از تبدیل داده‌ها

شکل (۷) تغییرات حجم اطلاعات (ورگشت توأم) بر حسب اولویت ایستگاه‌ها با استفاده از داده‌های بهنجار شده را نشان می‌دهد. همانطور که در این شکل نیز ملاحظه می‌گردد، با توجه به ایستگاه‌های موجود، با افزایش تعداد ایستگاه‌ها، حجم ورگشت افزایش می‌یابد. اما مقدار اختلاف بین ورگشت‌های توام با افزایش تعداد ایستگاه‌ها کاهش می‌یابد. بنابراین بعد از انتخاب تعدادی ایستگاه شبکه اشباع شده و ورگشت توأم افزایش چشمگیری نخواهد داشت. همانطور که مشاهده می‌شود، با توجه به تعداد ایستگاه‌های پایش موجود، هنوز شبکه اشباع نشده است و برای بدست آمدن حداکثر اطلاعات از شبکه نیاز به ایستگاه‌های پایش بیشتر می‌باشد.

با مقایسه شکل‌های ۳ و ۷ که در حقیقت طراحی شبکه پایش با داده‌های خام و بهنجار شده می‌باشند، ملاحظه می گردد که نرخ افزایش اطلاعات و همچنین اولویت انتخاب ایستگاه‌های پایش در دو حالت متفاوت است. در حالت نخست اطلاعات کلی شبکه (ورگشت توأم) با انتخاب ایستگاه‌های متوالی کاهش چشمگیری نیافته در حالی که در حالت دوم (داده‌های تبدیل یافته)، با افزایش تعداد ایستگاه‌ها نرخ افزایش ورگشت کاهش میابد. از طرفی اولویت انتخاب ایستگاه‌های پایش در دو حالت متفاوت می‌باشد. اما همانطور که ملاحظه می‌گردد این تفاوت چشمگیر نمی‌باشد.

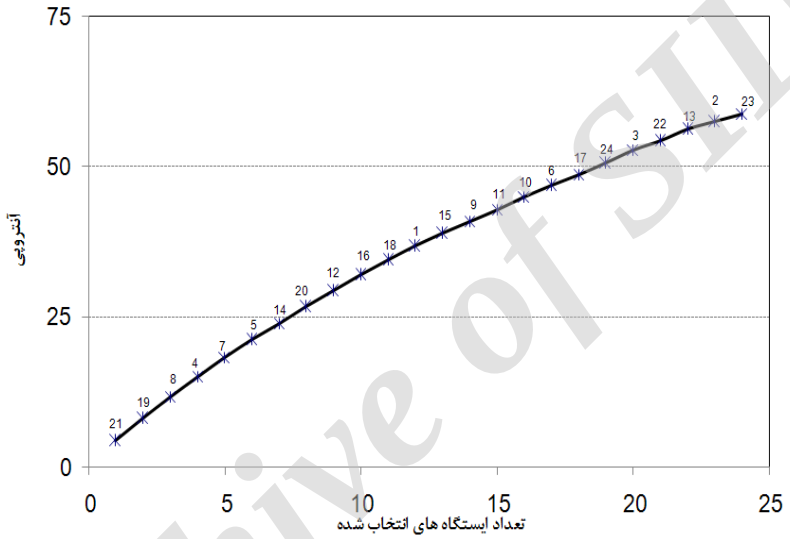
Archive of SID

www.SID.ir

نتیجه گیری

در اغلب مطالعات مبتنی بر ورگشت، برای محاسبه ورگشت توأم، از روابط مبتنی بر داده‌های بهنجار استفاده می‌شود. عدم تطابق داده‌ها با تابع توزیع بهنجار ممکن است باعث بوجود آمدن خطا در نتایج گردد. به منظور بررسی این فرض اولیه و غلبه بر خطای احتمالی، در این مقاله سعی گردید که علاوه بر طراحی شبکه پایش با استفاده از روش‌های متداول کنونی، روش پیشنهادی بهنجار کردن داده‌های ماهیانه با استفاده از تبدیل کاکس و باکس مورد استفاده قرار گیرد.

تأثیر داده‌های خام با وضعیت استفاده از داده‌های بهنجار شده بر ورگشت نشان داد که در حالت نخست، با افزایش تعداد ایستگاه‌ها افزایش نرخ ورگشت تغییر محسوسی نمی‌نماید. بر خلاف آن، هنگام استفاده از داده‌های بهنجار شده (منطبق با روابط ورگشت)، با افزایش تعداد ایستگاه‌ها افزایش نرخ ورگشت کاهش می‌یابد که این موضوع بیشتر منطبق بر واقعیت می‌باشد (چن و همکاران، ۲۰۰۸). همچنین اولویت انتخاب ایستگاه‌ها هم در دو حالت نسبتا متفاوت می‌باشد.



شکل ۷- تغییرات حجم ورگشت با تعداد ایستگاه‌های انتخاب شده اولویت بندی شده بعد از بهنجارسازی (تبدیل) داده‌ها

منابع

1) Alfonso L., Lobbrecht A., and Price R. 2010. Information theory–based approach for location of monitoring water level gauges in polders. Water Resour. Res. (46): W03528

2) Al-Zahrani M., and Husain T. 1998. An algorithm for designing a precipitation network in the south-western region of Saudi Arabia. J. Hydrol. 205: 205-216,

3) Box G.E.P., and Cox D.R. 1964. An analysis transformation. J. Roy. Stat. Soc. B26: 211-252

4) Chen Y.C., Wei C, and Yeh H.C. 2008. Rainfall network design using Kriging and entropy. Hydrol. Process. 22: 340-346

5) Cheng K.C., Lin Y.C., and Liou J.J. 2008. Rain-gauge network evaluation and augmentation using geostatistics. Hydrol. Process. 22(14): 2554-2564

6) Harmancioglu N., and Yevjevich V. 1987. Transfer of hydrologic information among river points. J. Hydrol. 91: 103-118

7) Husain T. 1987 .Hydrologic Network Design Formulation. Canadian Water Resources Journal. 12(1): 44-63

8) Husain T. 1989. Hydrologic uncertainty measure and network design. Water Resour. Bull., 25(3): 527-534

9) Krstonavic (1) P.F. 1992 Evaluation of rainfall network using entropy: I. theoretical development. Water Resour. Manag. 6: 279-293

10) Krstonavic (2) P.F. 1992. Evaluation of rainfall network using entropy: II. Application. Water Resour. Manag. 6: 295-314

11) Machiwal D., and Jha M.K. 1012. Hydrologic time series analysis. Capital Publishing Company: pp 272

12) Mazzarella A., and Tranfaglia G. 2000. Fractal characterization of geophysical

Archive of SID

www.SID.ir

measuring networks and its implication for an optimal location of additional stations: an application to a rain-gauge network, Theor. Appl. Climatol. 65: 157-163

13) Memarzadeh M., Mahjouri N., and Kerachian R. 2013. Evaluating sampling locations in river water quality monitoring networks: application of dynamic factor analysis and discrete entropy theory. Environ. Earth Sci. 70: 2577-2585

14) Shaghaghian M.R., and Abedini M.J. 2013. Rain gauge network design using coupled geostatistical and multivariate techniques. Scientia Iranica A. 20(2): 259-269

15) Shannon C.E. 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27: 623-656

16) Wei C., Yeh H.C., and Chen Y.C. 2014. Spatiotemporal scaling effect on rainfall network design using entropy. Entropy. 16(8): 4626-4647

17) Yeh H.C., Chen Y.C., Wei C., Chen R.H. 2011. Entropy and Kriging Approach to rainfall Network Design. Paddy Water Environ. 9: 343-355

Archive of SID

www.SID.ir