مدیریت آب و آبیاری

دوره ۳ ■ شماره ۱ ■ بهار و تابستان ۱۳۹۲

صفحه‌های ۹۵-۸۳

توسعهٔ مدل ناپارامتری شبیه‌ساز داده‌های ماهانهٔ هیدرولوژیکی

سلمان شریف‌آذری*[۱]، شهاب عراقی‌نژاد[۲]

۱. مربی گروه مهندسی آب، دانشکدهٔ آب و خاک، دانشگاه زابل، زابل – ایران

۲. استادیار گروه مهندسی آبیاری و آبادانی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج – ایران

تاریخ وصول مقاله: ۱۳۹۰/۱۰/۷ تاریخ پذیرش مقاله: ۱۳۹۲/۲/۲۱

چکیده

یکی از ملزومات مدلسازی منابع آب در سطح حوضه‌ها، آگاهی کافی از سری‌های درازمدت پارامترهای هواشناسی و هیدرولوژی است. در تحقیق حاضر، روش خودگردان‌ساز نزدیک‌ترین همسایهٔ ارائه‌شده توسط لل و شارما توسعه داده شد. در مدل توسعه‌داده‌شده به‌جای استفاده از منحنی رگرسیون ناپارامتری محلی در روش پریرا، از مدل رگرسیون ناپارامتری نزدیک‌ترین همسایه برای پیش‌بینی سری زمانی استفاده شد. این کار مشکلات ناشی از تخمین درجهٔ چندجمله‌ای ناپارامتری در روش پریرا را رفع کرده و شبیه‌سازی سری‌های زمانی را به‌صورت چندمتغیره میسر می‌کند. کارایی مدل توسعه‌داده‌شده با استفاده از اطلاعات تاریخی ایستگاه‌های هیدرومتری کوتر و هواشناسی سینوپتیک واقع در حوضهٔ آبریز رودخانهٔ مهاباد ارزیابی شد. نتایج نشان داد که مدل علاوه بر رفع مشکل روش خودگردان نزدیک‌ترین همسایهٔ کلاسیک در تولید اعداد بی‌سابقهٔ تاریخی، قادر به حفظ مشخصات مهم آماری سری مشاهداتی است. همچنین مدل توسعه‌داده‌شده، شبیه‌سازی سری‌های زمانی هیدرولوژیکی به‌صورت چندمتغیره را امکان‌پذیر می‌کند.

کلیدواژه‌ها: چندمتغیره، خودگردان‌ساز نزدیک‌ترین همسایه، رودخانهٔ مهاباد، سری‌های زمانی، مدل ناپارامتری.

* نویسنده مسئول Email: s.sharifazari@ut.ac.ir

مقدمه

در اختیار نداشتن آمار آبدهی کامل و طولانی‌مدت در برخی از رودخانه‌ها و دست‌خورده بودن آنها، تجزیه‌وتحلیل داده‌ها را دشوار می‌کند. در نتیجه یکی از ملزومات مدلسازی منابع آب در سطح حوضه‌ها، آگاهی کافی از سری‌های درازمدت پارامترهای هواشناسی و هیدرولوژی به‌ویژه آبدهی طبیعی در محل سازه‌ها و نقاط خاص حوضه است. در این راستا تولید و تخمین سری‌های هیدرولوژیکی در آینده نقش مهمی در تحقیقات حوضه‌ای ایفا می‌کند. مدلسازی‌های تصادفی در گذشته با تحلیل ساختاری فرایندهای تصادفی زمانی و مکانی نظیر تحلیل خطاها، انواع روند، ترکیب و تناوب انجام می‌گرفت. هیدرولوژی استوکستیکی از فناوری‌های در دسترس برای سازمان‌های آبی در چند دهۀ گذشته بوده است. برای مثال، شرکت مهندسی ارتش آمریکا[1] مدل HEC-4 را توسعه داد که به‌طور وسیع برای تعمیم و ساخت داده‌های ماهانۀ هیدرولوژیکی استفاده شده است (۱۹). همچنین ادارۀ آبادانی آمریکا (BOR)[2] بستۀ نرم‌افزاری LAST را در اواخر سال ۱۹۷۰ با هدف مدلسازی و شبیه‌سازی جریان رودخانه برای سیستم شبکۀ جریان با چند سایت توسعه داد. علاوه بر دو مدل HEC-4 و LAST، مدل‌های SPIGOT و SAMS بسته‌های نرم‌افزاری تخصصی‌ای هستند که برای شبیه‌سازی هیدرولوژیکی توسعه یافته‌اند. نرم‌افزار SAMS با همکاری دانشگاه ایالتی کلرادو و ادارۀ آبادانی آمریکا توسعه یافته است (۱۷). در بیشتر این مدل‌ها از روش‌های پارامتری استفاده شده است. به‌طور کلی می‌توان گفت روش‌های پارامتری بسیار مفیدند، ولی به کاستی‌های مهمی نیز دچارند (۱۳). اولین و مهم‌ترین کاستی آنها این است که به اندازۀ کافی جنبه‌های مختلف وابستگی زمانی و مکانی متغیرها را بازتولید نمی‌کنند. در این مدل‌ها به‌منظور شروع مدلسازی، یک فرض اولیه نسبت به فرم توزیع احتمالاتی متغیرها باید در نظر گرفته شود که اغلب تابع قضاوت فردی است. بنابراین برای رفع بسیاری از این مشکلات، از مدل‌های ناپارامتری استفاده می‌شود. مروری جامع بر روش‌های ناپارامتری و کاربرد آنها در حوزه‌های مختلف هیدرولوژی در سال ۱۹۹۵ انجام گرفت (۵). روش‌های مبتنی بر برآورد توابع چگالی هسته‌ای برای تولید داده‌های آبدهی توسعه داده شدند (۱۵ ، ۱۸). این مدل‌ها با موفقیت برای مدلسازی بارندگی (٤ ، ۷)، تحلیل فراوانی سیلاب (٦ ، ۹)، پیش‌بینی جریان (١٦) و آب زیرزمینی به‌کار رفتند (۲). روش‌های مبتنی بر برآورد توابع چگالی هسته‌ای، برای مسائل حدی نیز کاربرد دارند (۵ ، ۸ ، ۱۲). این کار مستلزم استفاده از هسته‌های متفاوت است (١٤). چنین روش‌هایی در مسائلی با ابعاد بزرگ (تعداد زیاد متغیرها یا سری‌های زمانی) بسیار پیچیده خواهند بود. از دیگر روش‌های غیرپارامتری برای تولید داده‌های آب‌وهوایی، روش خودگردان‌ساز K-NN است که اخیراً برای تولید مصنوعی داده‌های آب‌وهوایی استفاده شده است. روش خودگردان‌ساز نزدیک‌ترین همسایه برای مدلسازی سری‌های زمانی توسعه داده شد و برای شبیه‌سازی جریان رودخانه به‌کار رفت (۹). کاربرد این روش برای شبیه‌سازی تک‌منطقه‌ای بارش و دمای روزانه برای چندین ایستگاه در حوضۀ راین توصیف شد (۳). در تحقیقی دیگر روش خودگردان‌ساز نزدیک‌ترین همسایه با مدل سری زمانی پارامتری مقایسه شد که برتری روش نزدیک‌ترین همسایه به اثبات رسید (۱۱). با استفاده از روش برآورد هسته‌ای توابع چگالی احتمال و برآورد چگالی به روش k نزدیک‌ترین همسایه و همچنین تلفیق روش‌های

1. US Army Corps of Engineers
2. US Bureau of Reclamation

خودگردان‌ساز نزدیک‌ترین همسایه با شبکه‌های عصبی مصنوعی، مدل‌هایی ناپارامتری برای تولید داده‌هایی ساختگی از دبی جریان توسعه داده شدند (۱). روش خودگردان‌ساز نزدیک‌ترین همسایه ارائه‌شده برای مدلسازی تک‌متغیرهٔ سری‌های زمانی برای تولید مقادیر بی‌سابقه در سری زمانی مشاهداتی، توسعه داده شد (۱۰). با توجه به تحقیقات صورت‌گرفته و نتایج آنها، می‌توان دریافت که روش‌های ناپارامتری بخش بزرگی از نارسایی‌های روش‌های پارامتری را در شبیه‌سازی سری‌های زمانی رفع کرده‌اند. استفاده از روش برآورد تابع چگالی هسته‌ای از سری روش‌های ناپارامتری با توجه به وجود مسائل حدی (وجود مقادیر حدی در سری زمانی و اهمیت شبیه‌سازی این مقادیر) و نیز ضرورت شبیه‌سازی سری‌های هیدرولوژیکی به‌صورت چندمتغیره پیچیدگی زیادی را سبب می‌شود و شبیه‌سازی را دشوار می‌کند. از طرفی با اینکه روش خودگردان‌ساز نزدیک‌ترین همسایه، شبیه‌سازی سری زمانی را همراه با امکان تولید مقادیر بی‌سابقه در سری زمانی مشاهداتی مقدور می‌کند، این وضعیت تنها در صورت شبیه‌سازی تک‌متغیرهٔ سری زمانی ممکن است. ازاین‌رو در این تحقیق سعی می‌شود مدل خودگردان‌ساز نزدیک‌ترین همسایه طوری توسعه داده شود که علاوه بر رفع مشکلات مدلسازی سری‌های زمانی با روش‌های پارامتری، شبیه‌سازی سری‌های زمانی را به‌صورت چندمتغیره و با امکان تولید مقادیر بی‌سابقه در سری‌های زمانی مشاهداتی میسر کند.

مواد و روش‌ها

روش خودگردان‌ساز نزدیک‌ترین همسایه در راستای بهبود مدلسازی مسائل حدی در روش‌های شبیه‌سازی که بر پایهٔ تابع کرنل هستند، توسعه یافت و از آن در شبیه‌سازی دبی جریان استفاده شد (۹). در این روش، ابتدا نزدیک‌ترین همسایه‌های متغیر در ماه قبلی بر اساس فاصلهٔ اقلیدسی تعیین شده و همسایه‌ها بر اساس یک تابع وزن که بیشترین وزن را به نزدیک‌ترین همسایه می‌دهد بازنمونه‌گیری می‌شوند. این روش در واقع تابع توزیع احتمال شرطی ماه فعلی را نسبت به ماه قبلی تخمین می‌زند و شبیه‌سازی را با استفاده از آن انجام می‌دهد. به‌علاوه این روش می‌تواند اطلاعات اضافی ویژگی‌های بزرگ‌مقیاس اقلیمی مانند پدیدهٔ انسو[1] را نیز به‌دست دهد. این روش همچنین در شبیه‌سازی چندمتغیرهٔ متغیرهای هواشناسی روزانه به‌کار رفت (۲۰). تنها عیب این روش، بازتولید نکردن مقادیر مشاهده‌نشده در سری تاریخی است. برای رفع این مشکل، روش خودگردان‌ساز نزدیک‌ترین همسایهٔ ارائه‌شده توسعه داده شد. در مدل توسعه‌یافته ابتدا یک رگرسیون ناپارامتری محلی بر نمودار پراکندگی هر ماه نسبت به ماه قبلی(برای مثال ماه می نسبت به آوریل) برازش داده می‌شود، به‌طوری‌که در هر نقطه به k از نزدیک‌ترین همسایه‌های آن یک چندجمله‌ای محلی برازش داده می‌شود (۱۰). تعداد همسایه‌ها و درجهٔ چندجمله‌ای در این مدل از روش اعتبارسنجی متقابل تعمیم‌یافته به‌دست می‌آید.

در تحقیق حاضر به‌منظور بهبود روش خودگردان‌ساز نزدیک‌ترین همسایه و بر مبنای الگوریتم ارائه‌شده در مدل توسعه‌یافتهٔ آن، روش مورد بحث توسعه داده شد. در مدل توسعه‌داده‌شده به‌جای برازش رگرسیون ناپارامتری بر نمودار پراکندگی هر ماه نسبت به ماه قبلی، از مدل رگرسیون ناپارامتری نزدیک‌ترین همسایه برای پیش‌بینی سری زمانی استفاده شد. به‌این‌ترتیب مشکلات ناشی از تخمین درجهٔ چندجمله‌ای ناپارامتری رفع شده و شبیه‌سازی سری‌های زمانی به‌صورت چندمتغیره نیز میسر می‌شود. شبیه‌سازی سری زمانی در این مدل شامل سه

1. El Nino Southern Oscillation (ENSO)

مرحلهٔ زیر است:

۱. برازش مدل رگرسیون ناپارامتری(KNN) برای پیش‌بینی سری‌های مشاهداتی به‌منظور به‌دست آوردن باقی‌مانده‌ها؛

در این مرحله پس از تعیین مقدار پیش‌بینی‌شده برای ماه اول با انتخاب تصادفی از بین ماه‌های مشابه، بردار ویژگی برای ماه فعلی (t ام) با تأخیر یک ماه تعیین می‌شود و فاصلهٔ اقلیدسی آن با بردار ویژگی ماه‌های مشابه (مشاهداتی) به‌دست می‌آید. در نهایت با رتبه‌دهی همسایه‌ها براساس فاصلهٔ اقلیدسی تعداد k نزدیک‌ترین همسایه انتخاب شده و با استفاده از تابع وزن‌دهی (تابع توزیع احتمالاتی تجمعی برازش‌داده‌شده بر k همسایه) مقدار پیش‌بینی‌شده برای ماه فعلی به‌صورت زیر به‌دست می‌آید:

$$y_{i,t}^{j} = \sum_{k=1}^{K} w_k x_k^j \qquad [1]$$

در این رابطه، $y_{i,t}^{j}$ مقدار پیش‌بینی‌شده برای ماه فعلی در ایستگاه j ام و x_k^j همسایهٔ k ام ماه فعلی و w_k وزن متناسب با همسایهٔ k ام به‌دست‌آمده از تابع وزن تعریف شده، است. با تکرار مراحل بالا، سری پیش‌بینی‌شدهٔ داده‌ها به‌دست می‌آید.

۲. محاسبهٔ باقی‌مانده‌ها (با کم کردن مقادیر پیش‌بینی‌شده از مقادیر مشاهداتی سری زمانی)؛

$$\varepsilon_{i,t}^{j} = (x_{i,t}^{j} - y_{i,t}^{j}) \qquad [2]$$

۳. ساخت اطلاعات با استفاده از روش خودگردان‌ساز نزدیک‌ترین همسایه.

در این مرحله، مقدار ساختگی برای هر کدام از ماه‌های سال اول به‌صورت تصادفی انتخاب شده و سپس با تعیین بردار ویژگی برای ماه فعلی با تأخیر یک ماه، و محاسبهٔ فواصل اقلیدسی و تعیین k نزدیک‌ترین همسایه، یکی از همسایه‌ها با استفاده از تابع توزیع احتمالاتی برازش داده شده برای محاسبهٔ وزن همسایه‌ها و تولید عدد تصادفی بین صفر و یک (r) به‌صورت زیر انتخاب می‌شود:

a. اگر $p_1 < r < p_k$ ، همسایهٔ j ام که احتمال تجمعی آن(p_j) به مقدار r نزدیک‌تر است به‌عنوان x_t انتخاب می‌شود؛

b. اگر $r \le p_1$ ، همسایهٔ مطابق با d_1 به عنوان x_t انتخاب می‌شود؛

c. اگر $r = p_k$ ، همسایهٔ مطابق با d_k به‌عنوان x_t انتخاب می‌شود.

همسایهٔ انتخاب‌شده برای تولید مقدار متغیر در ماه فعلی به‌کار می‌رود. مقدار باقی‌مانده نیز به‌همین صورت و با استفاده از تابع توزیع احتمالاتی تجمعی برازش‌داده‌شده و عدد تصادفی بین صفر و یک ($r' \subset (0,1)$) به‌دست می‌آید:

a. اگر $p_1 < r' < p_k$ ، باقی‌ماندهٔ همسایهٔ j ام که احتمال تجمعی آن(p_j) به مقدار r نزدیک‌تر است انتخاب می‌شود؛

b. اگر $r' \le p_1$ ، باقی‌ماندهٔ همسایهٔ مطابق با d_1 انتخاب می‌شود؛

c. اگر $r' = p_k$ ، باقی‌ماندهٔ همسایهٔ مطابق با d_k انتخاب می‌شود.

در نتیجه مقدار متغیر شبیه‌سازی‌شده برای ماه فعلی به‌صورت زیر به‌دست می‌آید:

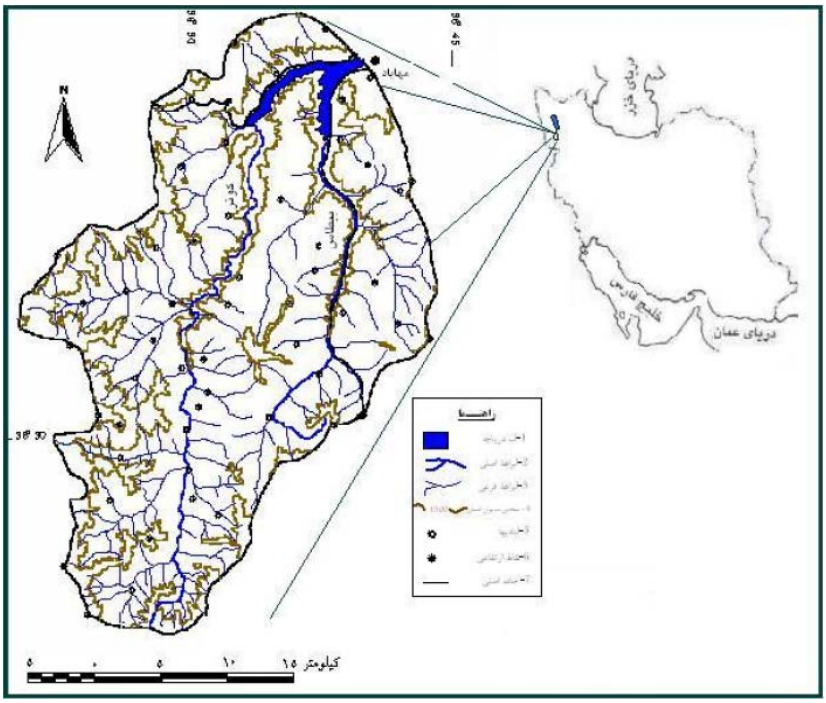
$$z_{i,t}^{j} = x_{i,t}^{j} + \varepsilon_t^j \qquad [3]$$

اختلاف محسوسی بین استفاده از روش تجربی و روش بهینه‌سازی اعتبارسنجی متقابل در تعیین تعداد همسایگان (k) وجود ندارد (۱۰). ازاین‌رو در این مدل نیز از روش تجربی برای تعیین تعداد همسایه‌ها استفاده شد، به‌طوری‌که: $k = \sqrt{N}$، که N برابر با تعداد سال‌های سری زمانی مشاهداتی است.

برای بررسی کارایی مدل توسعه‌داده‌شده در شبیه‌سازی سری زمانی داده‌های هیدرولوژیکی از اطلاعات ثبت‌شده

در ایستگاه‌های هیدرومتری (واقع در سرشاخهٔ کوتر، یکی از دو سرشاخهٔ اصلی رودخانهٔ مهاباد) و ایستگاه هواشناسی سینوپتیک استفاده شد. این اطلاعات شامل سری‌های زمانی ثبت‌شدهٔ آبدهی، دمای متوسط و مجموع بارش ماهانه در حوضهٔ آبریز مهاباد هستند. طول سری‌های زمانی ۱۸ سال است (از مهر ۱۳۶۹ تا شهریور ۱۳۸۷). حوضهٔ آبریز رود مهاباد در جنوب دریاچهٔ ارومیه، بین ۳۶ درجه و ۲۶ دقیقه تا ۳۶ درجه و ٤٦ دقیقه شمالی و ٤٥ درجه و ۲۵ دقیقه تا ٤٥ درجه و ٤٦ دقیقه طول شرقی واقع شده و مساحت آن تا محل سد مخزنی حدود ۸۲۹/۲۸ کیلومتر مربع است. این حوضه از شمال به جلگهٔ مهاباد، از شرق به حوضهٔ آبریز سیمینه‌رود، از غرب به حوضهٔ آبریز گادار و از جنوب به حوضهٔ آبریز زاب منتهی می‌شود. شکل ۱ نشان‌دهندهٔ موقعیت جغرافیایی حوضهٔ آبریز مهاباد است.

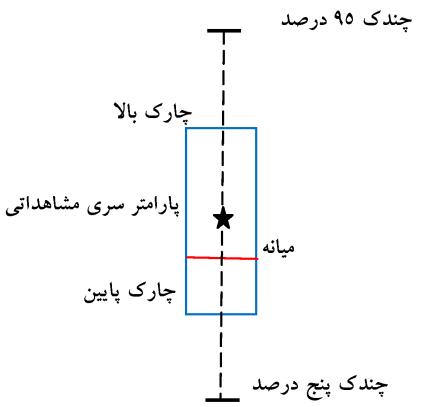
هدف از شبیه‌سازی سری‌های زمانی (تولید داده‌های بلندمدت)، بازتولید سری‌هایی با مشخصات آماری مشابه با سری مشاهداتی است، درحالی‌که در داده‌های آن آشفتگی ایجاد شده است. ازاین‌رو به‌منظور ارزیابی کارایی مدل توسعه‌داده‌شده در این تحقیق برای شبیه‌سازی سری‌های زمانی هیدرولوژیکی، از مقایسهٔ مشخصات آماری سری مشاهداتی و سری‌های شبیه‌سازی‌شده برای هر متغیر و نیز بررسی مقایسه‌ای توابع چگالی احتمال سری‌های شبیه‌سازی‌شده با سری مشاهداتی استفاده شد. مشخصات آماری مورد نظر برای مقایسه شامل میانگین، ضریب چولگی، ضریب خودهمبستگی با تأخیر یک ماه (R_1)، بیشینه و کمینهٔ متغیرهای مشاهده‌شده برای هر سری است. از آنجا که یک مدل شبیه‌ساز چندمتغیره باید قادر به بازتولید سری زمانی چند متغیر به‌صورت همزمان با حفظ همبستگی آنها در صورت معناداری باشد، دیگر ویژگی بررسی شده، نحوهٔ بازتولید ضرایب همبستگی متقاطع بین سری‌های زمانی متغیرهای به‌کاررفته در تحقیق بود.



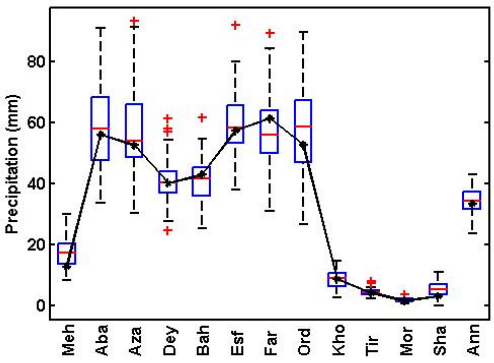
شکل ۱. موقعیت قرارگیری حوضهٔ آبریز رودخانهٔ مهاباد در ایران

بحث و نتیجه‌گیری

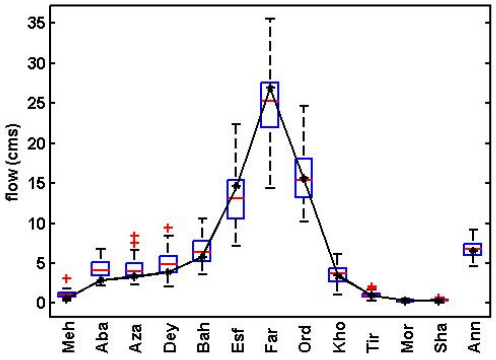
با استفاده از مدل توسعه‌داده‌شده در این تحقیق، به اندازهٔ دورهٔ زمانی مشاهداتی و برای هر کدام از متغیرهای ثبت‌شده، ۵۰ سری زمانی شبیه‌سازی شد. همچنین برای مقایسهٔ مدل توسعه‌داده‌شده با مدل خودگردان‌ساز نزدیک‌ترین همسایهٔ اولیه، شبیه‌سازی سری‌های زمانی با استفاده از مدل ذکرشده نیز انجام گرفت. به‌منظور ارزیابی کارایی مدل توسعه‌داده‌شده در این تحقیق، پارامترهای آماری میانگین، بیشینه، کمینه، ضریب چولگی و ضریب خودهمبستگی با تأخیر یک ماه در طول دورهٔ شبیه‌سازی، از هر یک از سری داده‌ها برای هر متغیر و همچنین از داده‌های مشاهده‌شده استخراج شد. برای بررسی نحوهٔ بازسازی پارامترهای ذکرشده توسط مدل توسعه‌داده‌شده در تحقیق حاضر و مقایسهٔ آن با پارامترهای متناظر در سری مشاهداتی، از نمودار جعبه‌ای استفاده شد. در شکل‌های ۳ تا ۱۵ پارامترهای آماری مربوط به سری‌های مشاهده‌شده و پارامترهای متناظر آنها که از سری‌های شبیه‌سازی‌شده به‌دست آمده‌اند، با یکدیگر مقایسه شده‌اند. در هر یک از این شکل‌ها، خطوط شکسته که نقاط توپر را به یکدیگر متصل می‌کنند، نمودار پارامتر آماری سری مشاهده‌شده در ماه‌های مختلفند. برای مقایسهٔ ساده‌تر ، مقادیر میانه، چارک بالا، چارک پایین و چندک‌های ۵ و ۹۵ درصد پارامتر مربوط به سری‌های شبیه‌سازی‌شده (۵۰ سری شبیه‌سازی‌شده) در هر ماه به‌صورت خطوط افقی در نمودارها نمایش داده شده‌اند . دامنهٔ واقع بین دو چارک بالا و پایین که دامنهٔ میان چارکی نامیده می‌شود، در برگیرندهٔ ۵۰ درصد میانی از مقادیر مربوط و دامنهٔ واقع بین چندک‌های ۵ و ۹۵ درصد دربرگیرندهٔ ۹۰ درصد داده‌هاست



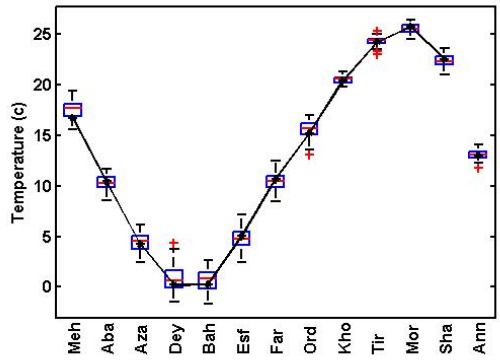
شکل ۲ . نحوهٔ نمایش داده‌ها (پارامترهای ۵۰ سری شبیه‌سازی‌شده در تحقیق حاضر) در نمودار جعبه‌ای



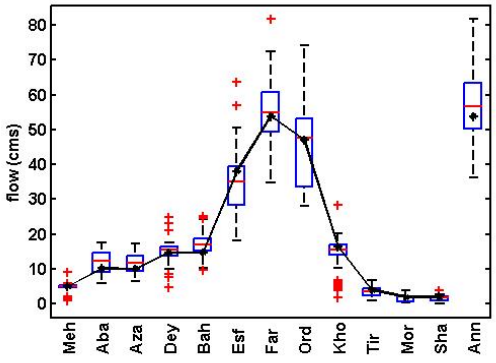
شکل ۳. نمودار جعبه‌ای میانگین ماهانهٔ آبدهی



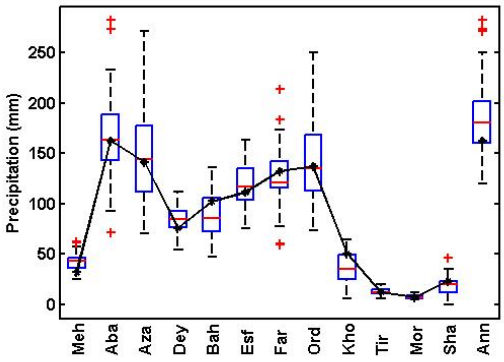
شکل ٤. نمودار جعبه‌ای میانگین مجموع بارش ماهانه



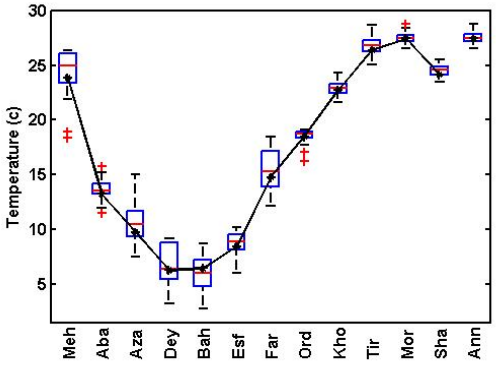
شکل ۵. نمودار جعبه‌ای میانگین متوسط دمای ماهانه



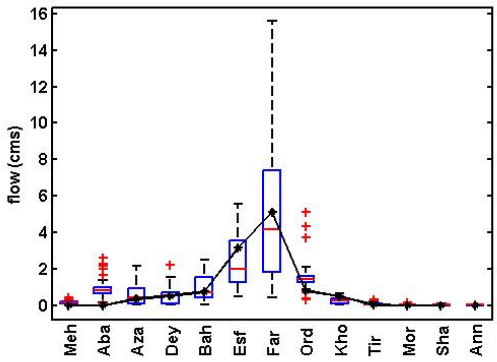
شکل ٦. نمودار جعبه‌ای حداکثر آبدهی ماهانه



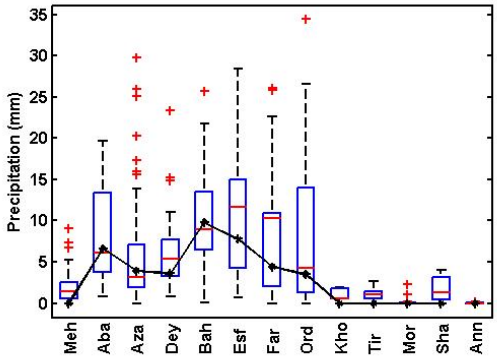
شکل ۷. نمودار جعبه‌ای حداکثر بارش مجموع ماهانه



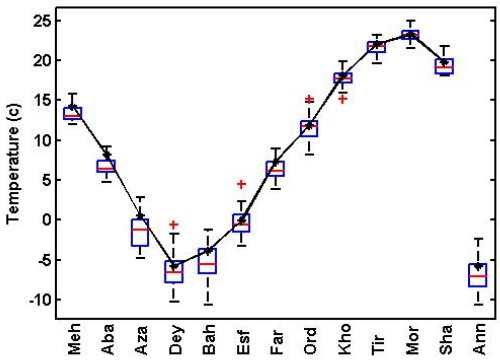
شکل ۸. نمودار جعبه‌ای حداکثر دمای ماهانه



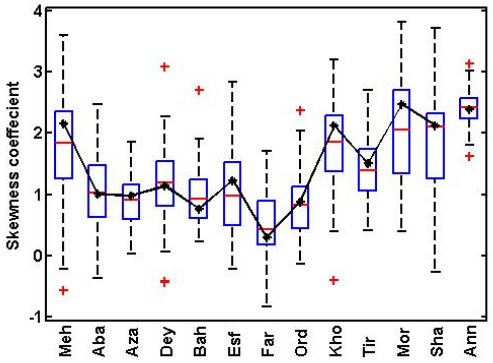
شکل ۹. نمودار جعبه‌ای حداقل آبدهی ماهانه



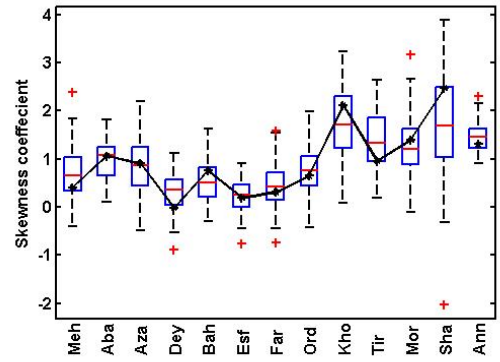
شکل ۱۰. نمودار جعبه‌ای حداقل مجموع بارش ماهانه



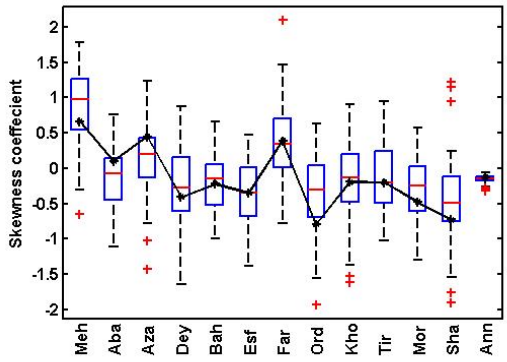
شکل ۱۱. نمودار جعبه‌ای حداقل آبدهی ماهانه



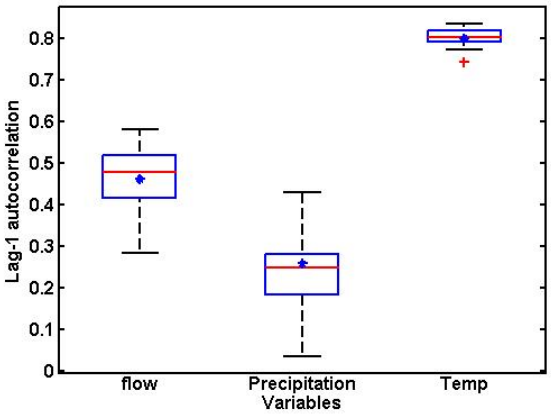
شکل ۱۲. نمودار جعبه‌ای ضریب چولگی آبدهی ماهانه



شکل ۱۳. نمودار جعبه‌ای ضریب چولگی مجموع بارش ماهانه

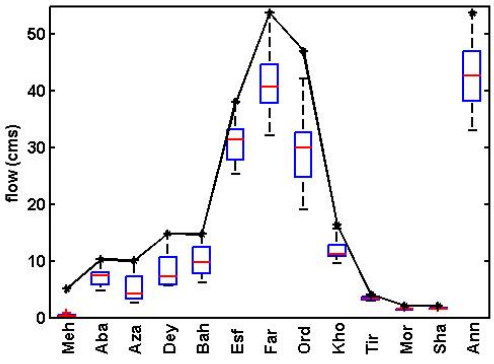


شکل ۱٤. نمودار جعبه‌ای ضریب چولگی دما متوسط ماهانه

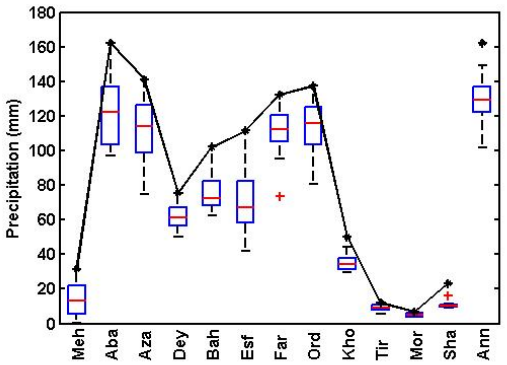


شکل ۱۵. نمودار جعبه‌ای ضریب خودهمبستگی با تأخیر یک ماه

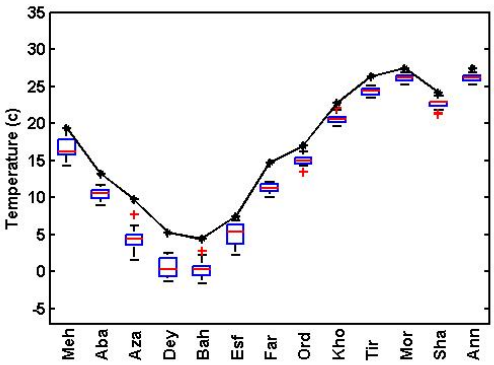
مقایسهٔ پارامترهای سری مشاهداتی هر متغیر (میانگین، بیشینه، کمینه، ضریب چولگی و ضریب خودهمبستگی) و مقادیر پارامترهای متناظر در سری‌های شبیه‌سازی‌شده (۵۰ سری تولیدشده) توسط مدل برای آن متغیر نشان می‌دهد که مدل کارایی بالایی در شبیه‌سازی داده‌های هیدرولوژیکی دارد. در اکثر ماه‌ها تفاوت چندانی بین مقادیر پارامترهای سری مشاهداتی و مقادیر متناظر در سری‌های شبیه‌سازی‌شده موجود نیست. در بیشتر ماه‌ها پارامتر سری مشاهداتی در محدودهٔ ۵۰ درصد پارامترهای سری‌های شبیه‌سازی‌شده قرار دارد. نمودارهای مربوط به پارامترهای بیشینه و کمینه نشان می‌دهند که مدل، علاوه بر تولید اعداد بی‌سابقه از متغیر در طول سری زمانی، قادر به تولید این اعداد در مقیاس ماهانه و برای هر ماه است. درحالی‌که مدل اولیهٔ خودگردان‌ساز نزدیک‌ترین همسایه توسط لل و شارما فاقد این قابلیت است و توانایی تولید مقادیر خارج از محدودهٔ ثبت‌شدهٔ متغیرها را ندارد. شکل‌های ۱۶ تا ۲۱ مقایسهٔ سری‌های شبیه‌سازی‌شده توسط مدل لل و مقادیر بیشینه و کمینهٔ سری‌های مشاهداتی برای هر متغیر را نشان می‌دهند.



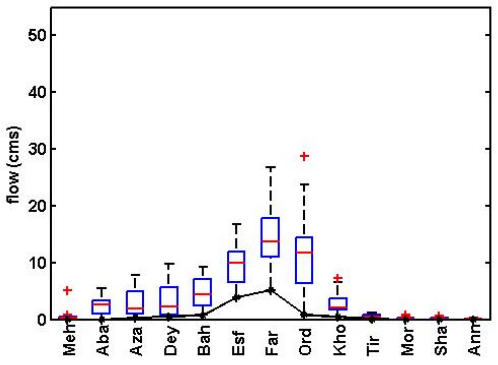
شکل ۱۶. نمودار جعبه‌ای حداکثر آبدهی ماهانه در مدل لل



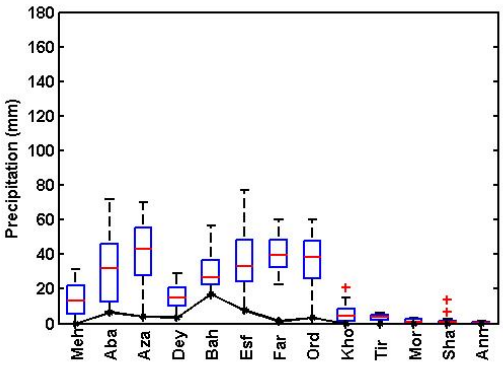
شکل ۱۷. نمودار جعبه‌ای حداکثر مجموع بارش ماهانه در مدل لل



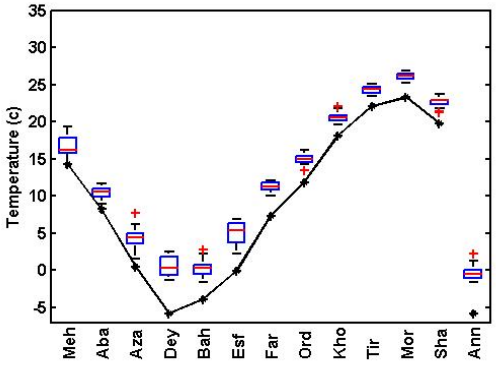
شکل ۱۸. نمودار جعبه‌ای حداکثر دما ماهانه در مدل لل



شکل ۱۹. نمودار جعبه‌ای حداقل آبدهی ماهانه در مدل لل

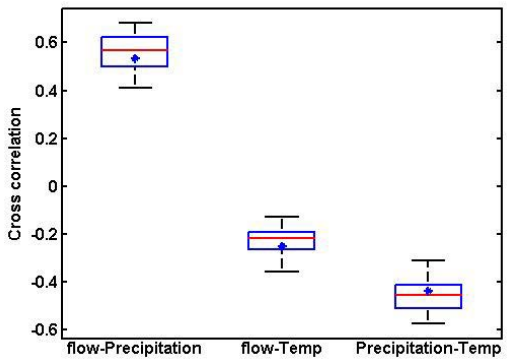


شکل ۲۰. نمودار جعبه‌ای حداقل مجموع بارش ماهانه در مدل لل



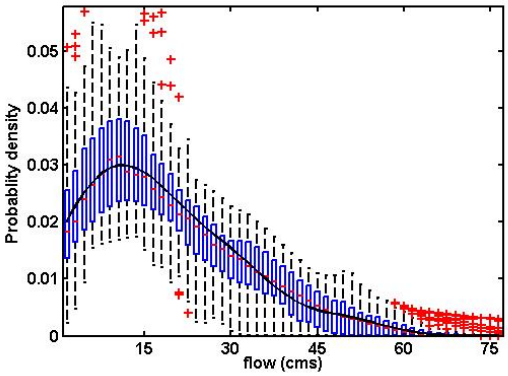
شکل ۲۱. نمودار جعبه‌ای حداقل دما ماهانه در مدل لل

یکی دیگر از قابلیت‌های مدل توسعه‌داده‌شده در این تحقیق شبیه‌سازی سری‌های زمانی به‌صورت چندمتغیره است. در واقع مدل‌های شبیه‌ساز چندمتغیره، قابلیت شبیه‌سازی همزمان سری‌های زمانی چند متغیر را با حفظ همبستگی بین آنها دارند. برای بررسی کارایی مدل توسعه‌داده‌شده در زمینهٔ شبیه‌سازی چندمتغیرهٔ سری‌های زمانی نحوهٔ بازتولید ضرایب همبستگی متقاطع بین متغیرها بررسی شد. در شکل ۲۲ ضرایب همبستگی متقاطع سری‌های مشاهده‌شده و مقادیر متناظر آنها که از سری‌های شبیه‌سازی‌شده به‌دست آمده، مقایسه شده‌اند.

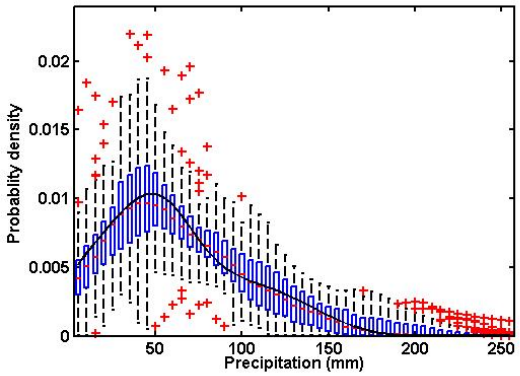


شکل ۲۲. نمودار جعبه‌ای ضرایب همبستگی تقاطع بین متغیرها

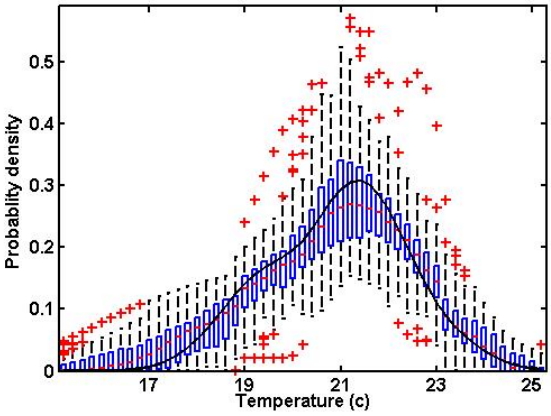
بررسی چشمی ضرایب همبستگی متقاطع بین متغیرها در سری مشاهداتی و سری‌های شبیه‌سازی‌شده نشان می‌دهد که تطابق شایان توجهی بین این مقادیر وجود دارد. به‌عبارتی ضرایب همبستگی متقاطع بین متغیرها در سری مشاهداتی در محدودهٔ ۵۰ درصد مقادیر متناظر در سری‌های شبیه‌سازی‌شده قرار دارند. ازاین‌رو مدل توسعه‌داده‌شده کارایی مناسبی در بازتولید این ضرایب در سری‌های شبیه‌سازی‌شده دارد و می‌تواند به‌عنوان نوعی مدل چندمتغیره برای شبیه‌سازی سری‌های زمانی به‌کار رود. همان طور که پیشتر ذکر شد، مدل‌های پارامتری برای شروع روند شبیه‌سازی فرض اولیه‌ای را نسبت به توزیع احتمالاتی متغیرها درنظر می‌گیرند که در بیشتر مواقع با واقعیت تطابق ندارد. بنابراین این مدل‌ها، اغلب، توانایی بازتولید مناسب توزیع احتمالاتی متغیرها را ندارند. روش‌های ناپارامتری بدون درنظر گرفتن فرض خاصی نسبت به توزیع احتمالاتی متغیرها، فرایند مدلسازی را شروع می‌کنند و با استفاده از بازنمونه‌گیری از سری‌های مشاهداتی، شبیه‌سازی سری‌های مشاهداتی را انجام می‌دهند. شکل‌های ۲۳ تا ۲۵ نحوهٔ بازتولید توزیع احتمالاتی متغیرهای به‌کاررفته در این تحقیق را در ماه اردیبهشت نشان می‌دهند. خط مشکی ممتد در شکل‌ها نشان‌دهندهٔ توزیع احتمالاتی سری مشاهداتی و نمودار جعبه‌ای مقادیر احتمال متناظر در سری‌های شبیه‌سازی‌شده هستند.



شکل ۲۳. نمودار جعبه‌ای توزیع احتمالاتی آبدهی اردیبهشت



شکل ۲٤. نمودار جعبه‌ای توزیع احتمالاتی مجموع بارش اردیبهشت



شکل ۲۵. نمودار جعبه‌ای توزیع احتمالاتی دما اردیبهشت

مقایسهٔ نمودارهای توزیع احتمالاتی متغیرهای به‌کاررفته در این تحقیق نشان می‌دهد که توزیع احتمالاتی سری‌های شبیه‌سازی‌شدهٔ متغیر در اردیبهشت‌ماه توسط مدل توسعه‌داده‌شده تطابق زیادی با توزیع احتمالاتی سری مشاهداتی متغیر دارد. نمودارهای توزیع احتمالاتی سایر ماه‌ها همانند اردیبهشت نشان‌دهندهٔ بازتولید مناسب توزیع احتمالاتی سری مشاهداتی در سری‌های شبیه‌سازی‌شده بود.

نیاز به آمار و اطلاعات پدیده‌های هیدرولوژیکی به‌منظور برنامه‌ریزی و پیش‌بینی مسائل مربوط به منابع آب، ضروری است. یکی از مشکلات اساسی در زمینهٔ هیدرولوژی خاصه در ایران، کوتاه بودن آمار سری‌های ثبت‌شدهٔ داده‌های هیدرولوژیکی به‌ویژه آبدهی است که مانع تجزیه‌وتحلیل دوره‌های خشک و تر می‌شود. با داشتن آمار و اطلاعات جریان رودخانه در ایستگاه‌ها با تعداد زیاد و اجرای آنالیز آماری می‌توان به پیش‌بینی روند وقوع ترسالی و خشکسالی و مدیریت مخازن در آینده پرداخت. در تحقیق حاضر، به‌منظور شبیه‌سازی سری‌های زمانی هیدرولوژیکی، مدل اولیهٔ خودگردان‌ساز نزدیک‌ترین همسایه توسعه داده شد. مدل توسعه‌داده‌شده با استفاده از

سری زمانی داده‌های ثبت‌شده در حوضهٔ آبریز رودخانهٔ مهاباد آزمون شد. نتایج نشان داد که مدل توسعه‌داده‌شده، در شبیه‌سازی سری زمانی داده‌ها کاراست و برخلاف مدل کلاسیک خودگردان‌ساز نزدیک‌ترین همسایه، قابلیت تولید اعداد خارج از محدودهٔ ثبت‌شده در سری زمانی را دارد. مزیت دیگر این مدل شبیه‌سازی چندمتغیره به‌صورت همزمان است.

تشکر و قدردانی

بدین وسیله از آقایان جلال خورشیدی، خالد احمدآلی، هادی همت‌یار و شیرکو جعفری تشکر و قدردانی می‌شود.

منابع

۱. اشرف‌زاده، ا؛ خلقی، م (۱۳۸٤). «توسعه و کاربرد یک مدل ناپارامتری برای شبیه‌سازی دبی جریان رودخانه». *مجلهٔ علوم کشاورزی ایران*، ۳٦ (٤): ۹۹۹-۹۹۱.

2. Adamowski, K. and Feluch, W. (1991) Application of nonparametric regression to groundwater level prediction. Canadian Journal of Civil Engineering 18(4), 600-606.

3. Brandsma, T. and Buishand, T.A. (1998) Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resam pling. Hydrology and Earth System Sciences Discussions 2(2/3), 195-209.

4. Harrold, T.I., Sharm a, A. and Sheather, S.J. (2003) A nonparametric model for stochastic generation of daily rainfall amounts. Water resources research 39(12), 1343.

5. Lall, U. (1995) Recent advances in nonparametric function estim ation: Hydrologic applications. Reviews of Geophy sics 33(S2), 1093-1102.

6. Lall, U., Moon, Y.-i. and B osworth, K. (1993) Kernel flood frequency estimators: Bandwidth selection and kernel c hoice. Water resources research 29(4), 1003-1015.

7. Lall, U., R ajagopalan, B. and Tarboton, D.G. (1996) A nonparametric wet/dry spell model for resampling daily precipitation. Water resources research 32(9), 2803-2823.

8. Lall, U. and Sharm a, A. (1996) A nearest neighbor bootstrap for resampling hydrologic time series. W ater resources research 32(3), 679-693.

9. Moon, Y.I. and Lall, U. (1994) Kernel quantite function estimator for flood frequency analysis. Water resources research 30(11), 3095-3103.

10. Prairie, J.R., Rajagopalan, B., Fulp, T.J. and Zagona, E.A. (2006) Modified K-NN model for stochastic streamflow simulation. Journal of Hydrologic Engineering 11(4), 371-378.

11. Rajagopalan, B. and Lall, U. (1999) A k-nearest-neighbor simulator for daily precipitation and other weather variables. Water resources research 35(10), 3089-3101.

12. Rajagopalan, B., Lall, U., Tarboton, D. and Bowles, D. (1997) Multivariate nonparametric resampling scheme for generation of daily weather variables. Stochastic Hydrology and Hydraulics 11(1), 65-93.

13. Sharif, M., Burn, D.H. and Wey, K.M. (2007) Daily and hourly weather data generation using a k-nearest neighbour approach. Canadian Hydrotechnical Conference.

14. Sharma, A. and O' Neill, R. (2002) A nonparametric approach for representing interannual dependence in m onthly streamflow sequences. Water resources research 38(7), 1100.

15. Sharma, A., Tarboton, D.G. and Lall, U. (1997) Streamflow simulation: A nonparam etric approach. Water resources research 33(2), 291-308.

16. Smith, J.A. (1991). LONG-R ANGE STREAMFLOW FORCASTING USING NONPARAMETRIC REGRESSION1. JAWRA Journal of the Am erican Water R esources Association 27(1), 39-46.

17. Sveinsson, O.G., Salas, J.D., Lane, W.L. and Frevert, D.K. (2007) Stochastic Analy sis, Modeling, and Sim ulation (SAMS) Version 2007, User's Manual. Computing Hydrology Laboratory, Department of C ivil and Environmental Engineering, C olorado State University, Fort Collins, Colorado.

18. Tarboton, D.G., Sharma, A. and Lall, U. (1998) Disaggregation procedures for stochastic hydrology based on nonparam etric density estimation. Water resources research 34(1), 107-119.

19. Us Army Corps of Engineering (1971) HEC-4 Monthly Stream flow Sim ulation. User's manual.

20. Yates, D., Gangopadhy ay, S., R ajagopalan, B. and Strzepek, K. (2003) A technique for generating regional clim ate scenarios using a nearst-neighbor algorithm water resources research 39 (7).