



بهبود دقت تخمین جریان ماهانه با کاهش نویز آشوبی (رودخانه‌های مورد مطالعه: نهندچای، اهرچای و ليقوان چای)

لیلا ملکانی^۱

تاریخ ارسال: ۱۳۹۸/۰۲/۲۰

تاریخ پذیرش: ۱۳۹۸/۰۵/۰۶

مقاله پژوهشی

چکیده

وجود ارتباط غیرخطی و پیچیده بین اجزای یک سیستم در فرایندهای هیدرولوژیکی و رفتار دینامیکی مابین آنها، استفاده از مدل‌های هوشمند را جهت مدل‌سازی ضروری می‌نماید. معمولاً در تحقیقات مختلف برای افزایش دقت نتایج مدل‌سازی، از مدل‌های جدیدتر با قابلیت محاسباتی بیشتر استفاده می‌شود. علاوه بر توانایی‌های محاسباتی مدل‌ها، بکارگیری اطلاعات ورودی صحیح به آنها نیز دارای اهمیت است و لازمه رسیدن به دقت مناسب در انواع روش‌های مدل‌سازی است. با توجه به اینکه خطا در داده‌های هیدرولوژیکی معمولاً وجود دارد، هدف از این تحقیق بررسی تاثیر حذف خطاهای ممکن در سیستم‌های هیدرولوژیکی در میزان افزایش دقت مدل‌ها است. در این تحقیق مقادیر جریان ماهانه حوضه‌های شاخص استان آذربایجان شرقی (رودخانه‌های نهند چای، ليقوان چای و اهرچای) در دو حالت با خطا و بدون خطا مورد بررسی قرار گرفته و سپس برای هر دو حالت، فرایند پیش‌بینی با مدل شبکه‌های عصبی مصنوعی مطالعه شده است. برای جداسازی خطای داده‌های ورودی از نظریه آشوب استفاده شده است. بر اساس معیار ارزیابی نش-ساتکلیف، دقت نتایج مدل‌سازی جریان ماهانه رودخانه‌های نهند چای، ليقوان چای و اهرچای برای داده‌های بدون نویز نسبت به داده‌های خام، بترتیب به میزان (۴۳/۲، ۲۷/۹ و ۵/۹) افزایش یافته و مقدار خطا، بر اساس معیار مجذور میانگین مربعات خطا به مقدار (۶۵/۲، ۶۳/۵ و ۲/۷) کاهش داشته است.

واژه‌های کلیدی: پیش‌بینی رواناب، نظریه آشوب، شبکه‌های عصبی مصنوعی، کاهش خطا

^۱ استادیار گروه عمران دانشکده فنی و مهندسی مرند، دانشگاه تبریز، ایران، تلفن تماس: ۰۴۱۴۲۶۰۰۰۳ آدرس پست الکترونیکی: Imalekani@tabrizu.ac.ir (نویسنده مسئول)



مقدمه

افزایش جمعیت و بدنبال آن نیاز به مطالعات متعدد در زمینه مهندسی آب و رودخانه، اهمیت تخمین جریان رودخانه‌ها را در مقیاس‌های مختلف زمانی روشن می‌نماید. اهمیت انجام پیش‌بینی جریان در تصمیماتی است که بر مبنای این پیش‌بینی‌ها به منظور مدیریت بهینه منابع آب صورت می‌گیرد.

پیش‌بینی‌ها به دو دسته کوتاه مدت و بلند مدت تقسیم می‌گردند. پیش‌بینی بلند مدت شامل پیش‌بینی در مدت زمان یک ماهه تا یکساله می‌باشد که در این تحقیق پیش‌بینی در افق زمانی ماهانه صورت گرفته است. تأثیرگذاری فاکتورها و عوامل مختلف و وجود رابطه غیرخطی بین آنها پیش‌بینی رواناب را امری پیچیده می‌کند. امروزه سیستم‌های هوشمند به منظور پیش‌بینی یک چنین پدیده‌های غیرخطی بسیار مورد استفاده قرار می‌گیرد که روش شبکه‌های عصبی مصنوعی، سیستم استنتاج فازی، سیستم استنتاج فازی عصبی تطبیقی و برنامه‌ریزی ژنتیک از جمله این روش‌ها است. در سال‌های اخیر علاوه بر مدل‌های هوشمند از نظریه آشوب و دیگر روش‌های فراکاوشی استفاده شده است.

چندوانی و همکاران (۲۰۱۵)، کیشی و همکاران (۲۰۱۳) و ماچادو و همکاران (۲۰۱۱) از ANN و سیستم فراکاوشی دیگر نظیر برنامه‌ریزی ژنتیک، فازی و سیستم فازی-عصبی تطبیقی برای پیش‌بینی جریان رودخانه بهره گرفتند. دهقانی و همکاران (۱۳۹۵) از ANN و ANFIS برای پیش‌بینی جریان در حوضه کیسیلیان در ایران استفاده نمودند. کلتی (۲۰۰۸) و سلیمانی (۲۰۰۹) برای پیش‌بینی رواناب از شبکه عصبی استفاده کردند و ضمن مقایسه نتایج مدل شبکه عصبی با داده‌های مشاهداتی به این نتیجه رسیدند که روش شبکه عصبی تخمین مناسبی از رواناب را ارائه می‌دهد.

پارامترهای هیدرولوژی همچون بارش و رواناب بکار رفته در مدل‌سازی اغلب مفروض به خطا می‌باشند و

مدل‌ها عموماً قادر به شناسایی و تفکیک این خطاها نیستند. قربانی و همکاران (۲۰۱۸) نیز در مقاله خود عنوان کردند که مدل ژنتیک در کاهش خطا خوب عمل نمی‌نماید. پرتویان و همکاران (۱۳۹۷) عنوان کردند که قابلیت روش‌های هوش مصنوعی برای داده‌هایی که دارای خطا هستند محدود می‌باشد. این کاستی به دلیل وابستگی زیاد مدل‌های داده محور به کیفیت داده‌های مورد استفاده می‌باشد. بنابراین توجه به بررسی صحت داده‌ها و کاهش میزان خطای پارامترهای ورودی به مدل‌ها به اندازه انتخاب مدل مناسب اهمیت دارد. اما عموماً اکثر محققان تلاش در راستای ایجاد و بکارگیری مدل‌های پیچیده‌تر داشته‌اند و کمتر به کاهش خطای داده‌ها توجه شده است. از جمله تحقیقاتی که در زمینه کاهش خطاها صورت گرفته است می‌توان به موارد زیر اشاره کرد:

الشورباگی و همکاران (۲۰۰۲) در مقاله‌ای تحقیقی را در زمینه کاربرد کاهش خطاها در سری‌های زمانی هیدرولوژیکی، بر روی داده‌های آبدی رودخانه انترابو در کشور کانادا انجام داده‌اند. در این تحقیق مسائل مربوط به کاهش خطا و تأثیرپذیری آن در افزایش دقت مدل‌های سری زمانی مورد بحث قرار گرفته است. در ابتدا مفهوم خطا و اثر کاهش خطا در داده‌های هیدرولوژیکی بحث شده و سپس عنوان گردیده که اغلب الگوریتم‌های رایج کاهش خطا ممکن است منجر به حذف قسمت‌های زیادی از داده‌ی اصلی گردند که صحیح نمی‌باشد و توصیه شده که الگوریتم کاهش خطاها با احتیاط مورد استفاده قرار گیرد. بر اساس نتایج این تحقیق استفاده از نظریه‌ی آشوب برای کاهش خطا توصیه شده است. سیواکومار (۱۹۹۹) در تحقیقی انواع روش‌های غیرخطی کاهش خطا را روی داده‌های بارش در سنگاپور به کار برده است و با توجه به معیارهای دقت پیش‌بینی به عنوان ابزار اصلی تشخیص در تعیین میزان کاهش خطای داده‌ها نشان داد که بهبود قابل توجهی در نتایج پیش‌بینی بارش حاصل شده است. سیواکومار و جایاواردنا (۲۰۰۲) و الشورباگی و همکاران

پیش‌بینی سری زمانی جریان دو رودخانه در Ciliwung در اندونزی استفاده کردند. آنها نشان دادند که استفاده از فرایند کاهش خطا بسته به مقدار خطا، دقت نتایج را تا ۱۵ الی ۴۰ درصد بهبود می‌بخشد.

سری زمانی داده‌های طبیعی و هیدرولوژیکی عموماً دارای خطا می‌باشند و با وجود اینکه کاهش خطای داده‌های ورودی به مدل‌ها، منجر به افزایش دقت نتایج مدلسازی می‌شود (رضایی و جباری، ۱۳۹۶، پرتویان و همکاران، ۱۳۹۷ و سیواکومار و جایاواردنا، ۲۰۰۲) همانطور که کاروالسینگا و لیونگ (۲۰۱۸) نیز مطرح کردند تحقیقات چندانی در این زمینه صورت نگرفته است. در این تحقیق سعی شده است تاثیر کاهش خطا با استفاده از نظریه آشوب بر روی نتایج مدلسازی جریان ماهانه‌ی سه رودخانه شاخص و مهم استان آذربایجانشرقی مورد بررسی قرار گیرد.

مواد و روش‌ها

منطقه و حوضه‌های مورد مطالعه

در این تحقیق سه حوضه آبریز مهم اهرچای، نهندچای و ليقوان چای را که به ترتیب حوضه بزرگ، متوسط و کوچک بحساب می‌آیند (امین علیزاده، ۱۳۹۴)، انتخاب گردید. با بررسی نتایج می‌توان تاثیر مساحت حوضه در مقدار نویز را تا حدودی مشخص نمود. شکل (۱) موقعیت حوضه‌های مورد مطالعه را در استان آذربایجانشرقی نشان می‌دهد.

حوضه آبریز اهرچای یکی از زیرحوضه‌های رودخانه ارس بوده و با وسعتی معادل ۲۲۳۲ کیلومتر مربع بخش قابل توجهی از حوضه آبریز ارس در ایران را پوشش می‌دهد. این حوضه بین مختصات $30^{\circ} 46'$ تا $40^{\circ} 47'$ طول شرقی و $20^{\circ} 38'$ تا $45^{\circ} 38'$ عرض شمالی از غرب به شرق گسترده شده است. متوسط ارتفاع منطقه ۱۸۸۰ متر از سطح دریا و متوسط شیب حوضه ۲۲ درصد می‌باشد. در این مطالعه از اطلاعات ایستگاه هیدرومتری ستارخان که در ورودی سد ستارخان بر روی رودخانه اهرچای قرار دارد استفاده شده است. حوضه آبریز رودخانه نهندچای از زیرحوضه‌های

(۲۰۰۲) نیز برای افزایش دقت نتایج از روش کاهش خطا استفاده کرده‌اند. این محققان در تحقیقات خود نشان داده‌اند آشوب می‌تواند در زمینه‌های مختلف هیدرولوژیکی که شامل حداقل کردن دامنه داده‌ها (Minimum Data Size)، حداقل کردن خطای داده‌ها، ارائه صفرها (Presence of zeros) و انتخاب پارامترهای بهینه، داده‌سازی و بازسازی داده‌های چند متغیره به کار برده شوند.

پرتویان و همکاران (۱۳۹۷) از روش ANN و ANFIS برای شبیه‌سازی جریان ایستگاه پل آنیان که بر روی رودخانه‌ی زرینه‌رود قرار دارد استفاده کردند. آنها خطای داده‌ها را با استفاده از روش موجک کاهش دادند و نتیجه‌گیری کردند که استفاده از داده‌هایی که خطای آنها کاهش یافته، مقدار ضریب همبستگی را در مدل‌های ANN و ANFF بترتیب ۲۳ و ۱۵ درصد افزایش می‌دهد.

رضایی و جباری (۱۳۹۶) با استفاده از آمار دبی‌های روزانه رودخانه نازلوچای تحلیل آشوبی شامل بررسی وجود آشوب با استفاده از روش بعد همبستگی و نیز شبیه‌سازی جریان رودخانه با استفاده از مدل تقریب موضعی انجام دادند. سپس به منظور بررسی تأثیر خطا در فرآیند تحلیل‌ها، کاهش خطای سری زمانی به روش غیرخطی مبتنی بر بازسازی فضای حالت را انجام دادند و نتیجه گرفتند که کاهش خطا منجر به کاهش $6/07$ درصدی بعد همبستگی و افزایش دقت مدل تقریب موضعی برای سری نویززدایی شده نسبت به سری اصلی داده‌ها (افزایش $1/09$ درصدی R^2 و کاهش ۴۸ درصدی RMSE) می‌گردد.

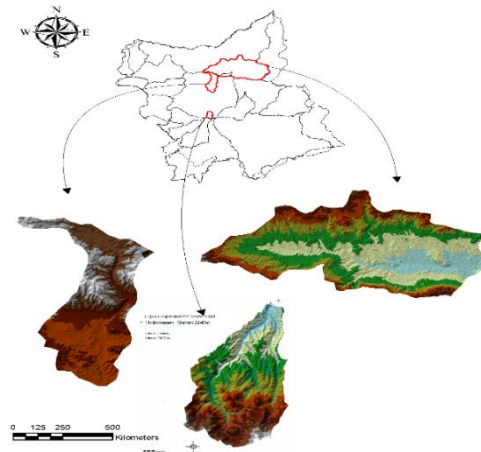
کاروناسینگا و لیونگ (۲۰۱۸) بیان کردند که پس از الشورباگی تحقیقات چندانی در زمینه کاهش خطا صورت نگرفته است و تاکید کردند زمانی که خطای داده‌های ورودی کاهش داده شود، دقت پیش‌بینی افزایش چشمگیری می‌یابد. بنابراین، تاکید کردند که نیاز به ارائه و توسعه روش‌های مختلف کاهش خطا می‌باشد. آنها از روش کاهش خطا فیلتر کالمن برای

در دامنه شمالی سهند طول شرقی $۴۶^{\circ} ۲۰'$ تا $۴۶^{\circ} ۲۷'$ و عرض شمالی $۳۷^{\circ} ۴۲'$ تا $۳۷^{\circ} ۴۹'$ واقع شده است. رودخانه ليقوان چای به عنوان زهکش حوضه مذکور می باشد و ارتفاع و شیب متوسط آن بترتیب ۲۶۷۵ متر و ۱۱ درصد می باشد.

مشخصات آماری داده های بکار رفته در این تحقیق برای حوضه های مربوطه در جدول ۱ نشان داده شده است.

دریاچه ارومیه با مساحتی برابر ۳۲۰ کیلومتر مربع در شمال شرق تبریز و بین عرض جغرافیایی $۳۸^{\circ} ۱۳'$ تا $۳۸^{\circ} ۲۹'$ و طول شرقی $۴۶^{\circ} ۲۰'$ تا $۴۶^{\circ} ۳۳'$ واقع گردیده است. ارتفاع متوسط حوضه ۲۳۶۸ متر و شیب متوسط ۲۸/۵ درصد می باشد. مساحت حوضه آبریز این رودخانه در بالادست سد نهند و ایستگاه هیدرومتری نهند ۲۱۹ کیلومتر مربع است.

حوضه معرف ليقوان چای نیز یکی از زیرحوضه های دریاچه ارومیه بوده و با وسعتی معادل ۷۶ کیلومتر مربع



شکل ۱. موقعیت حوضه های آبریز مطالعاتی

جدول (۱): مشخصات آماری داده‌های جریان رودخانه‌های مورد

مطالعه

مشخصه‌های آماری	دبی ماهانه (متر مکعب بر ثانیه)		
	نهندچای	اهرچای	لیقوانچای
دوره آماری	۴۹ تا ۱۳۸۸	۶۳-۱۳۸۳	۳۲-۵۴ و ۶۱-۸۹
تعداد داده‌ها	۴۴۴	۲۶۴	۶۱۲
میانگین	۱/۳	۲/۳۴	۰/۷۸
حداقل مقدار	۰	۰	۰
حداکثر مقدار	۱۲/۸۰۷	۱۵/۸۴	۵/۲۸
انحراف معیار	۳/۶۲۷	۲/۷۱	۰/۸
ضریب چولگی	۲/۶۸۲	۲/۳۸	۲/۴۸

شبکه عصبی مصنوعی (ANNs)

شبکه‌های عصبی مصنوعی یکی از انواع سیستم‌های دینامیکی هوشمند می‌باشد. شبکه عصبی یک مدل غیرخطی است بطوریکه با ساختار غیرخطی قادر به تقریب روابط غیرخطی بین داده‌هاست که برای پیش‌بینی و توسعه‌پذیری مدل‌های سری زمانی به کار برده می‌شود. ANNs با ساختار محاسباتی انعطاف‌پذیر در مسائل پیچیده غیرخطی می‌تواند با دقت بالایی مدل‌سازی کند. در این تحقیق از شبکه عصبی پرسپترون چندلایه (MLP) با الگوریتم آموزشی پس انتشار خطا (BP) استفاده شده است. هر ورودی در وزن متناظرش که بیانگر قدرت اتصالی است ضرب می‌شود و سپس همه ورودی‌های وزن‌دار با یکدیگر جمع می‌گردند تا سطح تحریک سلول عصبی را معین نمایند. شکل ۲ مدلی را نشان می‌دهد که این نظریه را به انجام می‌رساند در اینجا دسته‌ای از ورودی‌ها که به صورت (x_1, x_2, \dots, x_n) نشان داده شده‌اند، به سلول عصبی اعمال می‌شوند. این ورودی‌ها که جمعاً به عنوان یک بردار در نظر گرفته می‌شوند مشابه علایمی هستند که به سیناپس‌های سلول عصبی بیولوژیکی فرستاده می‌شوند هر سیگنال قبل از اینکه به واحد جمع که با

علامت \sum نشان داده شده است اعمال شود در یک وزن مربوط به خود (w_1, w_2, \dots, w_n) ضرب می‌گردد که هر وزن مشابه با قدرت یک اتصال سیناپس بیولوژیکی منفرد است (دسته وزن‌ها را جمعاً به صورت یک بردار W می‌توان در نظر گرفت) واحد جمع که شباهت خیلی کمی به جسم سلول بیولوژیکی دارد همه ورودی‌های وزن‌دار را به صورت جبری جمع و خروجی را تولید می‌کند که در اینجا با NET نشان داده می‌شود این روند ممکن است به صورت خلاصه با نماد برداری به صورت زیر بیان شود:



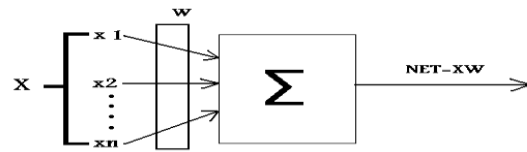
می‌شود در حالیکه بعد همبستگی به عنوان ابزار مکمل معرفی شده است.

کاهش نویز آشوبی

فرض می‌شود که داده‌های سری زمانی (η_n) از دو بخش متفاوت مقدار اصلی (X_n) و تصادفی (خطا، η_n) تشکیل شده‌اند $(S_n = X_n + \eta_n)$. منظور از کاهش خطا، جداسازی داده‌های سری زمانی به این دو جزء است به نحوی که یک جزء شامل سیگنال اصلی بوده و بخش دیگر حاوی جزء تصادفی باشد (رضایی و جباری ۱۳۹۶). کاهش خطا، یک گام اساسی در هر فرایند اندازه‌گیری است. ابزار مرسوم و کلاسیک آماری برای تمایز این دو جزء تبدیل فوریه^۱ و تبدیل موجک^۲ می‌باشد که بعنوان تبدیل‌های خطی شناخته شده‌اند (دوایی و نظرآهاری، ۱۳۹۴).

در روش‌های قدیمی، کاهش خطا با استفاده از تکنیک‌های خطی از قبیل تبدیل فوریه که بوسیله جدا کردن سیگنال از خطا در فضای فرکانس یا فضای خطی دیگر صورت می‌گیرد. عیب این روش‌ها این است که سیگنال‌های بسیاری ناحیه طیفی یکسان را به عنوان فضای خطا پوشش می‌دهند که باعث حذف بخشی از سیگنال اصلی می‌شود و لازم است تکنیک‌های فیلتر توسعه و بهبود داده شود. (شورایبر و کانتز، ۱۹۹۸). استفاده از روش تبدیل فوریه و طیف توانی^۳ حاصل از آن برای سیگنال‌هایی که از سیستم قطعی آشوبی به دست می‌آیند غیرقابل استفاده می‌باشد. زیرا احتمال حذف بخشی از سیگنال اصلی زیاد می‌باشد. چون اساس استفاده از این فیلترها به منظور جداسازی جزء تصادفی این است که خطا می‌تواند به عنوان مجموعه‌ای از اجزاء با فرکانس بالا مدل شده و از طیف توانی داده‌های ورودی کسر گردد. حال آنکه خروجی سیستم‌های دینامیکی قطعی آشوبی به خودی خود دارای طیف پهن باند بوده که می‌توان خواص طیف حاصل از آنها را به

$$NET = XW \quad (1)$$



شکل (۲): سلول عصبی مصنوعی

نظریه آشوب

در چند دهه اخیر، تحولی عظیم در شیوه درک و بیان پدیده‌های طبیعی توسط محققین صورت گرفته، که در سالین گذشته، تبیین‌های خود را در قالب‌های منظم و مشخص ارائه می‌دادند. نظریه آشوب به مطالعه سیستم‌هایی می‌پردازد که در نگاه اول به نظر می‌رسد رفتار تصادفی داشته باشند اما در واقع همین سیستم تحت حاکمیت قوانین مشخصی است و یا به عبارتی در هر بی‌نظمی، نظم نهفته است. چنین سیستمی به شرایط اولیه بسیار حساس است بگونه‌ای که ورودی‌های ظاهراً ناچیز و دلخواه قادرند تاثیرات شگرفی بر روی آن داشته باشند. به چنین سیستم‌هایی، سیستم‌های آشوبناک گفته می‌شود. پس ناپایداری، رفتار غیرپریودیک، سیستم‌های قطعی، غیر خطی بودن، در کنار همدیگر یک سیستم آشوبناک تعریف می‌کنند. نظریه آشوب اولین بار در سال ۱۹۶۵ توسط دانشمندی بنام ادوارد لورنز در هواشناسی به کار برده شد و سپس در حیطه تمام علوم و مباحث تجربی، ریاضی، رفتاری، مدیریتی و اجتماعی وارد شده و اساس تغییرات بنیادی در علوم بویژه هواشناسی، نجوم، مکانیک، فیزیک، ریاضی، زیست‌شناسی، اقتصاد و مدیریت را فراهم آورده است (کوجاک و همکاران، ۲۰۰۰). در تحقیق فعلی یک روش برای کاهش خطا در سری‌های زمانی ارائه داده می‌شود. این روش برای تخمین میزان خطاها روش مناسبی است و دقت کاهش خطاها عمدتاً بوسیله دقت پیش‌بینی بعنوان ابزار اصلی تشخیص صحت‌سنجی

³ Power Spectral

¹ Fourier Transform

² Wavelet Transform

برای انتخاب روش مناسب برای کاهش خطا لازم است آشوبناکی سری زمانی داده‌های بکار رفته بررسی گردد. گام نخست در انجام تحلیل‌های آشوبی، بازسازی فضای حالت سپس تعیین بعد همبستگی می‌باشد.

بازسازی فضای حالت

مفهوم فضای حالت، ابزاری سودمند برای مطالعه سیستم‌های دینامیک می‌باشد. مطابق این مفهوم، یک سیستم دینامیک می‌تواند توسط یک نمودار فضای حالت توصیف شود. این سیستم شامل یک مختصات با متغیرهایی هستند که در فرمول ریاضی آن وجود دارد و متغیرها نشان دهنده وضعیت سیستم در هر لحظه خاص می‌باشند. تکنیک متداولی که توسط تاکن ارائه شده از روش تاخیر برای نگاهستن یک سری زمانی واحد در یک فضای de بعدی استفاده می‌کند. روش تاکن به بیان ساده به این صورت است که یک شبکه‌ای به طول m از روی سری زمانی عبور داده می‌شود و به این ترتیب یک ماتریس با تعداد سطرهای برابر m تشکیل می‌یابد، برای سری زمانی:

$$x(t) = (x_0, x_1, x_2, x_3, \dots, x_i, \dots) \quad (2)$$

$$X = \begin{pmatrix} x_0 & x_1 & x_2 & \dots & x_{m-1} \\ x_1 & x_2 & x_3 & \dots & x_m \\ x_2 & x_3 & x_4 & \dots & x_{m+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix} \quad (3)$$

ماتریس برای بازسازی خط سیر اینگونه است:

تاکن نشان داده است که چنین ماتریسی کلیه ویژگی‌های هندسی سیستم دینامیکی اولیه را بدون اینکه با خودش برخوردی داشته باشد بیان می‌کند. انتخاب هوشیارانه تاخیرها (تغییر پریود نمونه برداری مقادیر متوالی x_i) با تاثیر بر وسعت جاذب نتیجه را بهتر می‌کند.

برای سری‌های زمانی اسکالر که $t = 1, 2, 3, \dots$ ، فضای حالت می‌تواند با استفاده از تاخیرها ایجاد شود. ایده اساسی درباره نحوه انتخاب زمان تاخیر آنست که ارزیابی هر متغیر مجزای سیستم توسط متغیرهای دیگر سیستم تعیین می‌شود که دارای اثر متقابل می‌باشند.

نویز تصادفی نسبت داد (رضایی و جباری، ۱۳۹۶ و سیواکمار و همکاران، ۱۹۹۹). لذا در سیستم‌های دینامیکی قطعی آشوبی از روش‌های غیرخطی برای کاهش خطا استفاده می‌گردد و اگر یک سیگنال مربوط به دیمانسیون در فضای آشوبناکی کم بعد باشد روش‌های خطی می‌تواند به صورت موفقیت‌آمیز استفاده شود.

پرپراتو و ریدولفی (۱۹۹۷) برای نخستین بار مفهوم خطا را وارد تحلیل‌های آشوبی سیستم‌های هیدرولوژیکی کردند. آنها با بکارگیری روش پیشنهادی شورایبر و گراسبرگر به کاهش نویز در سری‌های زمانی پرداختند. نتایج نشان از بهبود عملکرد پیش‌بینی در صورت کاهش خطا داشت. الگوریتم به کار رفته، شامل طرح‌ریزی برای جاذب‌های آشوبناکی در ابعاد کم است که در فضاهای بالاتر جایگزین شده است. این روش غیرخطی به خوبی برای سیستم‌های قطعی ثابت نیز به کار می‌رود.

هدف از کاهش خطا در سیستم‌های دینامیک قطعی آشوبی جایگزینی اندازه‌گیری‌های خطا دار با بهترین مقداری است که خطای کمتری دارد. ایده‌ی اصلی روش یاد شده به این صورت است که سیر تکامل سیستم توسط نگاهشت $X_n = f(X_{n-m}, \dots, X_{n-1})$ قطعی در نظر گرفته می‌شود و در حالیکه f برای ما ناشناخته بوده و تنها اطلاعات در دسترس از سیستم، سری زمانی اندازه‌گیری‌های خطا دار می‌باشد. در رابطه $S_n = X_n + \eta_n$ ، η_n خطای تصادفی با داشتن تابع همبستگی است که سرعت کاهش می‌یابد و بدون بعد همبستگی با سیگنال X_n است (بوستانی و همکاران، ۱۳۹۸).

Urbanowicz و Kantz (۲۰۰۷) اعلام کردند که عملکرد کلی روش غیرخطی برتر بوده و نشان دادند که الگوریتمی که ریشه در تئوری دینامیک آشوبناک و غیرخطی دارد، دارای پتانسیل بزرگی در کاهش خطا است.



فرایندهای تصادفی، D_2 بدون رسیدن به یک مقدار اشباع با افزایش m تغییر می‌کند در حالیکه برای فرایندهای قطعی مقدار D_2 بعد از یک m معین اشباع می‌شود. مقدار اشباع، بعد فرکتالی (بعد همبستگی) جاذب یا سری زمانی است (کوجاک و همکاران، ۲۰۰۰). به منظور بررسی آشوب‌پذیری و تخمین بعد همبستگی مناسب می‌توان از نمودار $\log C(r)/\log(r)$ در مقابل تغییرات $\log(r)$ استفاده نمود.

معیارهای ارزیابی مدل‌ها

زمانی عمل شبیه‌سازی به درستی انجام شده است که تمامی پارامترهای موثر در ساختار مدل طوری تعیین گردند که مقدار خطای مدل به حداقل مقدار خود برسد و همچنین مقادیر کارایی بیشترین مقدار را داشته باشند. در این تحقیق جهت ارزیابی مدل‌ها از معیارهای کارایی ضریب همبستگی (R^2) و نش-ساتکلیف و معیار نمایش خطای مجذور میانگین مربعات خطا (RMSE) استفاده شده است. روابط مربوط به این معیارها از فرمول‌های زیر محاسبه می‌شوند.

$$R^2 = \frac{\sum_{i=1}^N (Q_{mi} - \bar{Q}_m)(Q_{pi} - \bar{Q}_p)}{\left(\sum_{i=1}^N (Q_{mi} - \bar{Q}_m)^2 \right)^{0.5} \left(\sum_{i=1}^N (Q_{pi} - \bar{Q}_p)^2 \right)^{0.5}} \quad (8)$$

$$E = 1.0 - \frac{\sum_{i=1}^N (Q_{mi} - Q_{pi})^2}{\sum_{i=1}^N (Q_{mi} - \bar{Q}_m)^2} \quad (9)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_{mi} - Q_{pi})^2} \quad (10)$$

در این روابط Q_{mi} ، رواناب مشاهداتی و Q_{pi} ، رواناب محاسبه شده است. همچنین \bar{Q}_m و \bar{Q}_p بترتیب متوسط رواناب مشاهداتی و محاسباتی است و N برابر با تعداد داده‌های آموزش یا صحت‌سنجی می‌باشد هرچه مقدار E به عدد یک نزدیک باشد یعنی دقت مدل بالا بوده و برعکس هرچه مقدار RMSE به صفر نزدیکتر باشد بیانگر دقت بالای مدل است از طرفی مقدار R^2 بالا و نزدیک به یک بیانگر ارتباط قوی خطی بین مقادیر

بنابراین اطلاعات هر متغیر وابسته در تاریخچه هر متغیر مستقل دیگر سیستم وجود خواهد داشت. بر مبنای چنین معادل سازی، فضای حالت می‌تواند با استفاده از المان x_t سری‌های زمانی ایجاد شود و تاخیر آن مطابق با سری‌های زمانی جدید خواهد بود.

$$Y_t = \{x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(m-1)\tau}\} \quad (4)$$

τ ، نشان دهنده زمان تاخیر بوده و در سری‌های زمانی عددی، حاصلضرب زمان نمونه‌برداری می‌باشد و m ، اندازه محاط شده را نشان می‌دهد (کانل و همکاران، ۱۹۹۲).

بعد همبستگی

بعد همبستگی یکی از روش‌های مرسوم تعیین آشوبناکی سیستم و همچنین بعد آشوبی می‌باشد. برای فضای حالت m بعدی، تابع همبستگی $C(r)$ بصورت زیر تعریف می‌شود:

$$C(r) = \lim_{N \rightarrow \infty} \frac{2}{N(N-1)} \sum_{1 \leq i < j \leq N} H(r - |Y_i - Y_j|) \quad (5)$$

که در آن H ، یک تابع هویساید پله‌ای با $H(u) = 0$ بر $u \leq 0$ و $H(u) = 1$ بر $u > 0$ است. N تعداد نقاط در فضای مزبور، $r = |Y_i - Y_j|$ شعاع کره ساخته شده به مرکز Y_i یا Y_j می‌باشد. برای مقادیر مثبت r ، تابع همبستگی $C(r)$ با رابطه زیر به r مرتبط می‌شود:

$$C(r) \underset{\substack{r \rightarrow 0 \\ N \rightarrow \infty}}{\approx} ar^{D_2} \quad (6)$$

که در این رابطه a ، یک ضریب ثابت بوده و D_2 توان همبستگی می‌باشد که از رابطه زیر بدست می‌آید:

$$D_2 = \lim_{\substack{r \rightarrow 0 \\ N \rightarrow \infty}} \frac{\log C(r)}{\log(r)} \quad (7)$$

از آنجائیکه مجموعه داده‌ها پیوسته نخواهد بود، نمی‌تواند مقادیر نزدیک به صفر داشته باشد، در نتیجه $\log C(r)$ بر $\log(r)$ تقسیم شده و از آن حد گرفته می‌شود و سپس قسمت خطی نمودار حاصل از آن انتخاب می‌شود. در نتیجه این کار، مقدار D_2 حاصل می‌شود. با استفاده از رسم D_2 در مقابل m برای

مشاهداتی و محاسباتی است (اعلمی و همکاران، ۲۰۱۳).

نتایج و بحث

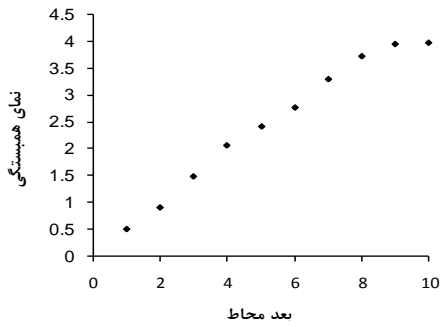
همانطور که کارناسینگها و لیونگ (۲۰۱۸) اشاره کردند با وجود اهمیت و تاثیری که کاهش خطا در میزان افزایش دقت نتایج پیش‌بینی‌ها دارد تحقیقات زیادی صورت نگرفته است. در این تحقیق نیز تاثیر کاهش خطا در پیش‌بینی ماهانه رواناب سه حوضه بزرگ، متوسط و کوچک با استفاده از شبکه‌ی عصبی مصنوعی مورد بررسی قرار گرفته است. در ابتدا برای تعیین روش کاهش خطا، آشوب‌پذیری داده‌های سری زمانی بررسی شده است و سپس از روش GHKSS که بوسیله گراسبرگر توسعه داده شده است و یک روش غیرخطی می‌باشد برای کاهش خطا استفاده شده است. همانطور که گفته شده بعد همبستگی شناخته شده‌ترین روش تعیین آشوبناکی می‌باشد. به منظور بررسی آشوب‌پذیری و تخمین بعد همبستگی مناسب می‌توان از نمودار $\log C(r)/\log(r)$ در مقابل تغییرات $\log(r)$ استفاده نمود. شکل (۳)، نمودار تغییرات نسبت $\log C(r)/\log(r)$ در مقابل $\log(r)$ را نشان می‌دهد. به دلیل وجود خطا در داده‌ها به ازای مقادیر کوچک $\log(r)$ نوساناتی در نمودار مشاهده می‌شود.

زمان تاخیر برای جریان ماهانه رودخانه‌های نهند چای، ليقوان چای و اهرچای بترتیب برابر ۳، ۲ و ۳ ماه محاسبه شده است. مطابق با نتایج شکل (۳) می‌توان نتیجه گرفت که هر سه جریان ماهانه هر سه رودخانه آشوبناک می‌باشد و مقدار بعد همبستگی نیز همانگونه که در شکل مشاهده می‌گردد به ترتیب برابر ۴، ۲/۵ و ۳/۵ می‌باشد.

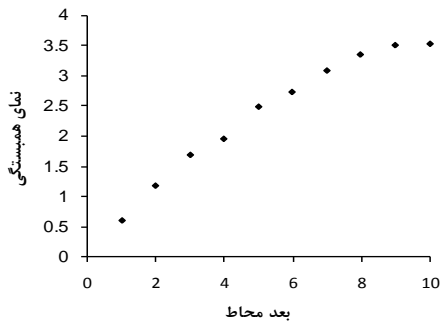
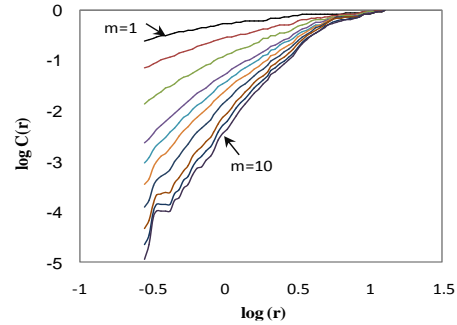
برنامه‌های کامپیوتری آشوب، به صورت اجرایی در بسته‌های نرم افزاری TISEAN وجود دارد که از هر برنامه برای محاسبه پارامترهایی مانند زمان تاخیر، بعد همبستگی، کاهش خطا و غیره استفاده می‌شود. با استفاده از برنامه GHKSS در این بسته و بهره‌گیری از نمونه دستور زیر، سری زمانی که خطای آن کاهش یافته تولید می‌گردد. این برنامه بر اساس نظریه گراسبرگر و همکاران (۱۹۹۳) نوشته شده است.

`-d3 -m1,5 -nahand.txt -GHKSS.exe -NRnahand.txt`

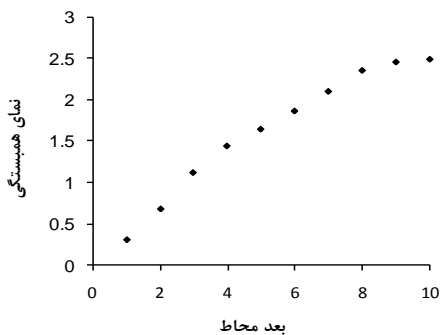
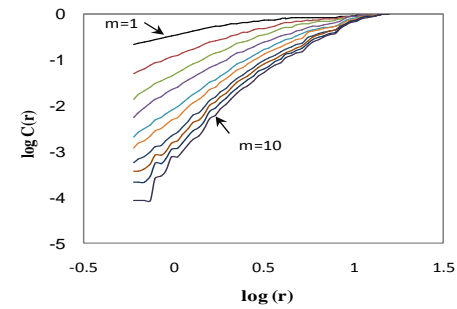
در این دستور nahand نام فایل ورودی و NRnahand نام فایل خروجی است. عدد ۳ مربوط به زمان تاخیر محاسبه شده است. ۱ و ۵ بترتیب مقدار حداقل و حداکثر بعد همبستگی می‌باشد و براساس مقدار ضریب همبستگی بدست آمده قابل تغییر می‌باشد.



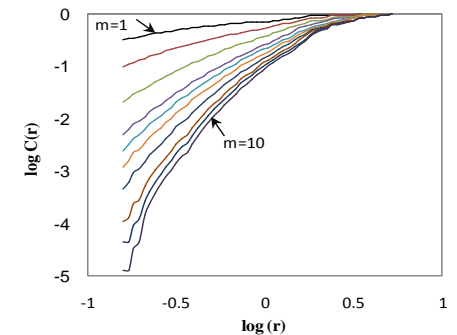
الف) حوضه نهندچای



ب) حوضه اهرچای



ج) حوضه ليقوانچای



شکل (۳): نمودار تغییرات $\log C(r)/\log(r)$ در مقابل $\log(r)$ و نمودار تغییرات بعد همبستگی با افزایش بعد محاط

برای مشخص شدن تاثیر هر کدام از موارد مذکور مدل-های ایجاد شده در گام اول شبکه با تک متغیر ورودی ماه قبل Q_{t-1} ، در گام‌های بعدی شبکه با دو متغیر ورودی Q_{t-1} و Q_{t-2} ، شبکه با سه متغیر ورودی Q_{t-1} و Q_{t-2} و Q_{t-3} و شبکه چهار متغیر ورودی Q_{t-1} و Q_{t-2} و Q_{t-3} و Q_{t-4} می‌باشد. نتایج بررسی‌ها نشان می‌دهد که مدل ANN با یک متغیر ورودی (رواناب ماه قبل) در حوضه‌های مورد مطالعه، از دقت کافی برخوردار نبوده و حالت سه متغیر ورودی Q_{t-1} و Q_{t-2}

پس از اینکه سری زمانی کاهش یافته با استفاده از برنامه GHKSS تولید گردید نوبت به اجرای برنامه شبکه عصبی مصنوعی می‌رسد. برای داشتن نتایج دقیق در هر دو سری داده خام و کاهش خطا یافته، تنها با در نظر گرفتن یک متغیر موثر بر رواناب (رواناب ماه قبل) بهترین ساختار شبکه عصبی مشخص گردیده و سپس تاثیر تعداد متغیرهای ورودی، تاثیر تعداد لایه‌های میانی و تعداد تکرار مورد بررسی قرار گرفته و دقیق‌ترین مدل بر اساس معیارهای کارایی و خطا تعیین می‌گردد.

مطالعات مشابه دیگری است که انجام شده است. کاروناسینگا و لیونگ (۲۰۱۸) نشان دادند که استفاده از روش کاهش خطا دقت نتایج پیش‌بینی جریان رودخانه‌های مطالعه شده را بسته به میزان خطا در سری‌های زمانی، ۱۵ الی ۴۰ درصد افزایش می‌دهد. در تحقیق دیگری که رضایی و جباری (۱۳۹۶) به منظور بررسی اثر کاهش خطا بر روی مدل‌سازی داده‌های دبی روزانه نازل‌وچای انجام دادند مشاهده گردید که نتایج مدل پیش‌بینی با استفاده از سری نویز‌زادایی شده دارای دقت بیشتری نسبت به مدل با استفاده از سری اصلی بوده و مقدار ضریب همبستگی به مقدار ۱/۰۹ افزایش و مقدار RMSE به میزان ۴۸ درصد کاهش یافته است. همچنین سیواکومار و همکاران (۱۹۹۹) نیز نشان دادند که با کاهش نویز دقت مدل شبیه‌سازی افزایش یافته است.

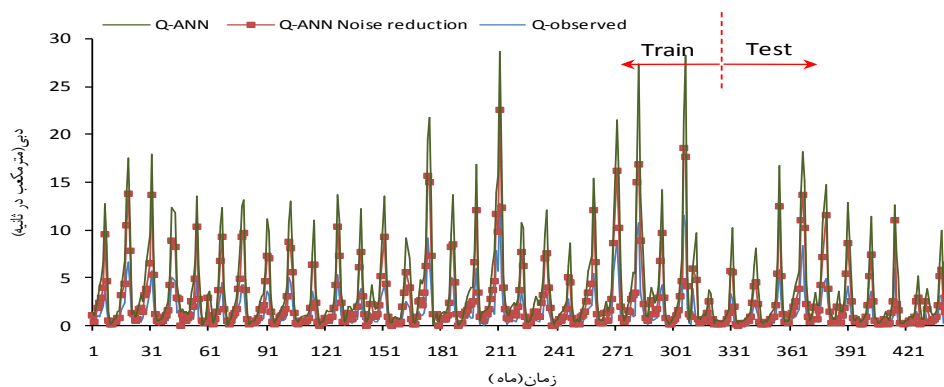
داده‌های مشاهداتی و خروجی مدل شبکه عصبی با داده‌های خام و داده‌های با کاهش خطا برای هر سه حوضه در شکل شماره (۴) نمایش داده شده است.

و Q_{1-3} بهترین جواب را برای هر سه حوضه می‌دهد اما نتایج فوق‌خصوص برای حوضه‌های نهندچای و اهرچای (حوضه‌های با مساحت متوسط و بزرگ) چندان رضایت بخش نمی‌باشد. این نتایج در جدول (۲) ارائه شده است.

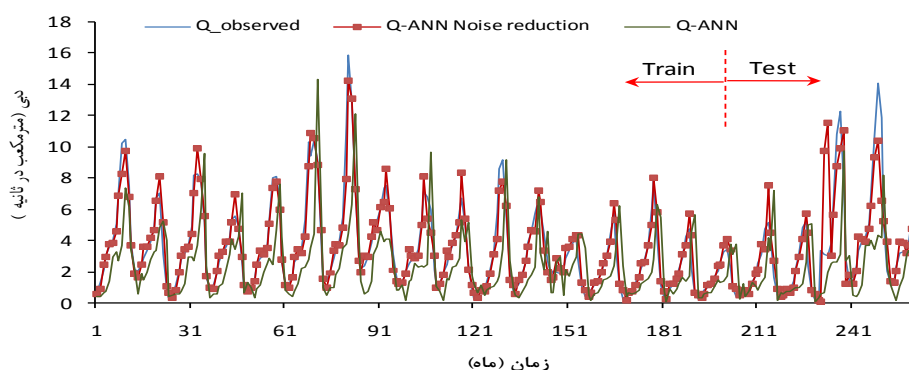
برای بالا بردن دقت مدل که هدف اصلی در بسیاری از کارهای تحقیقاتی و اجرایی است در این تحقیق روش کاهش خطاها روی داده‌های رواناب ماهانه معرفی شده در سری زمانی داده‌های مشاهداتی مربوط به حوضه‌های متوسط و بزرگ بیشتر از حوضه‌های کوچک بوده و روش کاهش خطا میزان دقت مدل را بصورت چشمگیری در دو حوضه آبریز اهرچای و نهندچای افزایش داده است به گونه‌ای که میزان خطا در حوضه‌های نهندچای و اهرچای برای داده‌های صحت‌سنجی بترتیب ۶۵ و ۶۳ درصد کاهش داشته است و همچنین دقت نتایج پیش‌بینی در این دو ایستگاه برای داده‌های صحت‌سنجی بر اساس معیار نش-ساتکلیف ۴۳ و ۲۸ درصد افزایش را نشان می‌دهد. این نتایج مطابق با

جدول (۲): مقایسه نتایج مدل شبکه عصبی با داده‌های خام و داده‌های با Noise-reduction

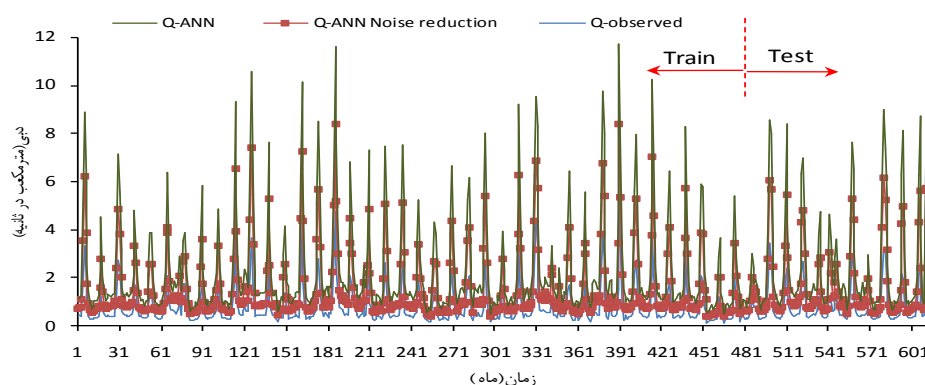
ایستگاه	ساختار شبکه عصبی		R^2		E		RMSE	
	آموزش	صحت‌سنجی	آموزش	صحت‌سنجی	آموزش	صحت‌سنجی	آموزش	صحت‌سنجی
نهند	۳-۴-۱	۳-۴-۱	۰/۸۹۷	۰/۸۱۵	۰/۱۸۰	۰/۱۸۶۹	۰/۱۶۸۲	۰/۱۶۸۲
	۳-۴-۱	۳-۴-۱	۰/۶۸۶	۰/۵۸۲	۰/۴۵۴	۱/۳۰۵	۱/۱۲۷	۱/۱۲۷
			۲۳/۵	۲۸/۶	۴۳/۲	۵۰/۲	۶۵/۲	۶۵/۲
اهر	۳-۸-۱	۳-۸-۱	۰/۸۰۸	۰/۹۳۶	۰/۶۹۷	۰/۶۶۸	۱/۰۹۲	۱/۰۹۲
	۳-۵-۱	۳-۵-۱	۰/۶۳۴	۰/۶۷۵	۰/۵۶۱	۱/۳۰۵	۱/۸۹۴	۱/۸۹۴
			۱۹/۵	۱۷/۰۴	۲۷/۹	۳۶/۱	۶۳/۵	۶۳/۵
لیقوان	۳-۶-۱	۳-۶-۱	۰/۸۵۰	۰/۸۸۱	۰/۷۳۳	۰/۲۶۵	۰/۴۵۰	۰/۴۵۰
	۳-۸-۱	۳-۸-۱	۰/۸۳۲	۰/۸۰۵	۰/۶۹۰	۰/۲۸۰	۰/۴۶۲	۰/۴۶۲
			۴/۵	۸/۶	۵/۹	۵/۷	۲/۷	۲/۷



الف) حوضه نهندچای



ب) حوضه اهرچای



ج) حوضه لیقوانچای

شکل (۴): سری زمانی خروجی مدل شبکه عصبی با داده‌های خام و بدون خطا و داده‌های مشاهداتی

منابع

- بوستانی، م، فرزین س، موسوی س ف، کرمی حجت. (۱۳۹۸). تأثیر نویز زدایی سری زمانی بر تحلیل آن با استفاده از نظریه نظم در بی‌نظمی (مطالعه موردی: رودخانه زاینده‌رود)، اکوهیدرولوژی، ۶(۱): ۱۵-۲۷
- پرتویان، ا، نورانی، و، اعلمی م ت. (۱۳۹۷). بهبود عملکرد نرم افزارهای هوش مصنوعی در شبیه سازی بارش-رواناب با استفاده از روش حذف-تزریق نوفه. فصلنامه علمی-پژوهشی مهندسی منابع آب، ۱۱(۳۶): ۸۱-۹۴.



دهقانی ن، وفاخواه م، بهره‌مند ع. (۱۳۹۵). مدل‌سازی بارش- رواناب با استفاده از شبکه عصبی مصنوعی و شبکه فازی- عصبی تطبیقی در حوزه آبخیز کسلیان. پژوهشنامه مدیریت حوزه آبخیز. ۷ (۱۳): ۱۲۸-۱۳۷

دوایی مرکزی اح، نظرآهاری م. ۱۳۹۴. کاربرد تبدیل فوریه گسسته در زمان در شناسایی سیگنال‌های صوتی کشتی‌ها با استفاده از روش‌های کاهش بُعد و یادگیری توده‌ای. مهندسی مکانیک مدرس. ۱۵ (۸): ۸۴-۷۵

رضایی ح، جبباری قره‌باغ، ث. (۱۳۹۶). تأثیر کاهش نویز در تحلیل آشوبی جریان رودخانه نازلوچای. دانش آب و خاک، ۲۷ (۳): ۲۵۰-۲۳۹

علیزاده، امین (۱۳۹۴)، اصول هیدرولوژی کاربردی، چاپ چهل، مشهد، انتشارات آستان قدس رضوی.

Alami, M.T., Ghorbani, M.A. and Malekani, L., 2013. Chaotic Analysis and Prediction of River Flows. *Journal of Hydraulic Structures*, 1(1), pp.54-64.

Chandwani, V., Vyas, S.K., Agrawal, V. and Sharma, G. 2015. Soft computing approach for rainfall-runoff modelling: A review. *Aquatic Procedia*, 4, pp.1054-1061.

Elshorbagy A. Simonovich S. P. and Panu U. S. 2002. Noise reduction in chaotic hydrologic time series: facts and doubts. *Journal of Hydrology*. V.256: 147-165.

Elshorbagy A. Simonovich S. P. and Panu U. S. 2002a. Estimation of missing streamflow data using principles of chaos theory. *Journal of Hydrology*. 255:123-133.

Ghorbani, M.A., Khatibi, R., Mehr, A.D. and Asadi, H. 2018. Chaos-based multigene genetic programming: A new hybrid strategy for river flow forecasting. *Journal of hydrology*, 562: 455-467.

Grassberger, P., Hegger, R., Kantz, H., Schaffrath, C. and Schreiber, T. 1993. On noise reduction methods for chaotic data. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 3(2), pp.127-141.

James E. J. Gopakumar R. 2002. Flood forecasting of Achencoil river basin using Artificial Neural Network. fifth International conference of Hydroinformation. 1245-1251.

Kalteh A. M. 2008. Rainfall-runoff modelling using artificial neural networks (ANNs): modelling and understanding. *Caspian Journal of Environmental Sciences*. 6(1): 53-58.

Karunasingha, D.S.K. and Liang, S.Y. 2018. Enhancement of chaotic hydrological time series prediction with real-time noise reduction using Extended Kalman Filter. *Journal of hydrology*, 565, pp.737-746.

Kennel M. Brown R and Abarbanel HDI. 1992. Determining Embedding Dimension for Phase-Space Reconstruction Using a Geometrical Construction. *Physical Review A*. 45(6): 3403-3411.

Kisi, O., Shiri, J. and Tombul, M. 2013. Modeling rainfall-runoff process using soft computing techniques. *Computers & Geosciences*, 51, pp.108-117.

Kocak K. Saylan L. and Sen O. 2000. Nonlinear Time Series Prediction of O₃ Concentration in Istanbul. *Atmosphere Environment*. 34: 1267-1271.

Machado, F., Mine, M., Kaviski, E. and Fill, H. 2011. Monthly rainfall-runoff modelling using artificial neural networks. *Hydrological Sciences Journal-Journal des Sciences Hydrologiques*, 56(3), pp.349-361.

Porporato A and Ridolfi L, 1997. Nonlinear analysis of river flow time sequences. *Water Resources Research* 33(6): 1353-1367.

Schreiber, T., Kantz. H. 1998. Nonlinear projective filtering II: Application to real time series", arXiv:chao-dyn/9805025v1.

Schreiber T and Grassberger P, 1991. A simple noise reduction method for real data *Physics Letters A* 160: 411-418.

Sivakumar, B., Phoon, K.-K., Liang, S. Y. and Liaw, C. Y. 1999. A systematic approach to noise reduction in chaotic hydrological time series", *Journal of Hydrology* 219:103-135.



Sivakumar, B. and Jayawardena, A.W. 2002. An investigation of the presence of low-dimensional chaotic behavior in the sediment transport phenomenon. *Journal of Hydrology Science* 47(3):405–416.

Solaimani, K. 2009. Rainfall-runoff Prediction Based on Artificial Neural Network (A Case Study: Jarahi Watershed). *American-Eurasian J. Agric. & Environ. Sci.*, 5 (6): 856-865.

Urbanowicz, K. and Kantz, H. 2007. Improvement of speech recognition by nonlinear noise reduction. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(2), p.023121.



Reduced chaotic noise to improve the accuracy of estimates of monthly flow (case study: Nahandchai, Aharchai and Lighvanchai Rivers)

Leila Malekani ¹

Abstract

The nonlinear and complex relationship between the components of a system in hydrological processes and the dynamic behavior between them makes it necessary to use intelligent models for modeling. Usually, in a variety of studies, to increase the accuracy of modeling results, newer models with more computational capabilities are used. In addition to the computational abilities of the models, the use of correct input information is also important to them, and it is necessary to achieve the appropriate accuracy in a variety of modeling methods. Because the error is usually in the hydrologic data, the purpose of this study is to investigate the effect of eliminating possible errors in hydrologic systems in increasing the accuracy of models.

In this paper, the monthly flows of Nahandchai, Lighvanchai, and Aharchai rivers have been investigated in two cases the original series and noise reduced time series. Then, for both cases, the runoff predictions with the ANN model have been done. Chaos theory used for the separation of input noise data. The results base on the evaluation criteria shows that this method, providing a possible higher accuracy in data without noise. The amount of Nashatcliff coefficient increased (43.2%, 27.9%, and 5.9%) respectively at the stations of Nahandchai, Aharchai, and Lighvanchai, and RMSE decreased to 65.2%, 65.5% and 7.7% at these stations

Keywords: Runoff Prediction, Chaos theory, ANN, Noise reduction

¹ Assistant Professor, Department of civil Engineering, Marand Faculty of Technical and Engineering, University of Tabriz, Iran. (lmalekani@tabrizu.ac.ir)