Archive of SID

نشریه حفاظت منابع آب و خاک، سال هشتم، شماره سوم، بهار ۱۳۹۸

ISSN 2251-7480

ارزیابی مدل ترکیبی موجک – حداقل مربعات ماشین بردار پشتیبان در ریزمقیاس کردن مکانی – زمانی سری‌های زمانی بارش

نیما فربودفام[۱]، وحید نورانی[۲*] و بابک امین نژاد[۳]

۱) دانشجوی دکتری گروه عمران – آب، دانشکده فنی و مهندسی، واحد رودهن، دانشگاه آزاد اسلامی، رودهن، ایران

۲*) استاد گروه عمران – آب، دانشکده عمران، دانشگاه تبریز و استاد گروه عمران – آب، دانشکده فنی و مهندسی، واحد رودهن، دانشگاه آزاد اسلامی، رودهن، ایران

* نویسنده مسئول مکاتبات: v.nourani2018@gmail.com

۳) استادیار گروه عمران – آب، دانشکده فنی و مهندسی، واحد رودهن، دانشگاه آزاد اسلامی، رودهن، ایران

تاریخ دریافت: ۱۳۹۶/۱۰/۲۲ تاریخ پذیرش: ۱۳۹۷/۱۲/۰۵

چکیده

با توجه به نیاز شبیه‌سازی سری‌های زمانی بارش در مقیاس‌های مختلف برای مقاصد مهندسی از یک طرف و عدم ثبت این پارامترها در مقیاس‌های ریز بدلیل مشکلات اجرایی و اقتصادی از طرف دیگر، ریزمقیاس کردن بارش به مقیاس مورد نظر، یک امر ضروری می‌باشد. در این مطالعه، برای ریزمقیاس کردن سری زمانی بارش ایستگاه‌های تبریز و سهند، با توجه به ویژگی‌های غیرخطی مقیاس‌های زمانی، مدل ترکیبی موجک – حداقل مربعات ماشین بردار پشتیبان (WLSSVM) پیشنهاد شده و داده‌های سری زمانی ماهانه شش ایستگاه و روزانه چهار ایستگاه بارش حوضه آبریز دریاچه ارومیه، برای ۱۰ سال بوسیله تبدیل موجک به زیرسری‌های زمانی تجزیه شده و سپس با استفاده از معیارهای اطلاعات متقابل و ضریب همبستگی، زیرسری‌ها رتبه‌بندی شده و برای ریزمقیاس کردن سری زمانی ماهانه ایستگاه‌های تبریز و سهند به روزانه، زیرسری‌های برتر به‌عنوان داده‌های ورودی به مدل حداقل مربعات ماشین بردار پشتیبان (LSSVM) وارد شد. نتایج حاصل از مدل WLSSVM، با نتایج کاربرد روش LSSVM و روش کلاسیک رگرسیون چند متغیره خطی، مقایسه شد. در مجموع نتایج مدل WLSSVM نسبت به مدل‌های LSSVM و رگرسیون چند متغیره خطی برای اعتبارسنجی در حالت بهینه ایستگاه تبریز به ترتیب ۱۰٪ و ۳۷٫۵٪ و در حالت بهینه ایستگاه سهند، به ترتیب ۲۴٫۵٪ و ۴۶٫۷٪ افزایش نشان داد. لذا ملاحظه گردید که روش WLSSVM نسبت به دو روش دیگر، دقت بالاتری داشته و به‌عنوان روشی مناسب جهت ریزمقیاس کردن سری‌های زمانی بارش پیشنهاد می‌گردد.

کلید واژه‌ها: تبدیل موجک؛ حداقل مربعات ماشین بردار پشتیبان؛ ریزمقیاس کردن؛ سری‌های زمانی بارش؛ مدل ترکیبی

مقدمه

رشد فزاینده جمعیت و به تبع آن افزایش نیازهای آبی از یک طرف و حجم محدود آب‌های شیرین در دسترس از طرف دیگر، برنامه‌ریزی و مدیریت صحیح منابع آب را بیش از پیش ضروری ساخته است. بارش متغیری تصادفی بوده و عدم قطعیت در تخمین آن، برنامه‌ریزی و مدیریت بهینه منابع آب را برای متولیان و تصمیم‌گیرندگان مشکل می‌سازد. کمبود اطلاعات در بسیاری از زمینه‌ها و تحلیل‌های هیدروکلیماتولوژیکی، منجر به ارایه روش‌هایی جهت تولید اطلاعات مصنوعی از روی اطلاعات

www.SID.ir

Archive of SID

مشاهداتی شده است و در اکثر اوقات، یک فرآیند تصادفی هیدروکلیماتولوژیکی باید در مقیاس‌های زمانی مختلف مورد مطالعه قرار گیرد. لیکن چگونگی تولید سری‌های زمانی سازگار در مقیاس‌های زمانی بزرگ و کوچک مشکل می‌باشد.

بارش از جمله فرآیندهای بسیار مهم در مدیریت منابع آب بوده و باران سنج‌های متعددی وجود دارد که اغلب برای چند دهه مورد بهره‌برداری بوده، در حالی‌که تعداد باران سنج‌هایی که داده‌ها با مقیاس کوچکتر را ثبت نمایند محدود می‌باشد. این وضعیت، نشان دهنده کمبود نسبی در تعداد باران سنج‌ها و سری‌های ثبت شده بارش برای مقیاس‌های کوچکتر، می‌باشد. استفاده از داده‌های تاریخی بزرگ مقیاس (سالانه، فصلی و یا ماهانه) که به‌عنوان مثال در مدل‌های گردش عمومی (*General Circulation Models*) مطرح می‌شوند، باعث افزایش عدم قطعیت گردیده و جهت رسیدن به حداقل عدم قطعیت و افزایش دقت مدل‌سازی‌ها و همچنین نیاز به داده‌های ریزمقیاس برای کاربردهای هیدروکلیماتولوژیکی، استفاده از تکنیک‌های مناسب برای ریزمقیاس کردن داده‌ها را ضروری می‌سازد.

سری‌های زمانی عموماً به دو روش آماری (*Statistical*) و دینامیکی (*Dynamical*)، ریز مقیاس (*disaggregation*) می‌شوند. مدل‌هایی که برای ریزمقیاس کردن دینامیکی استفاده می‌شوند بسیار شبیه مدل‌های گردش عمومی بوده ولی گام‌های زمانی و مکانی شبکه ریزتر و دقیق‌تر هستند. در ریزمقیاس کردن آماری یک ارتباط آماری با استفاده از رگرسیون ساده، چند متغیره، شبکه عصبی و ... بین رفتار واقعی سنجنده و خروجی مدل ایجاد می‌شود. مدل‌های ریزمقیاس کردن آماری از رفتار سیستم در دوره آماری استفاده کرده و برای مکان مورد نظر می‌توانند وضعیت سری‌های زمانی بارش را مدل نمایند. این مدل‌های آماری نمی‌توانند رفتاری خارج از رفتار دوره آماری و مشاهداتی را تولید نمایند.

روش‌های ریاضی و آماری کلاسیک متعددی در خصوص ریز مقیاس کردن سری‌های زمانی بارش توسط پژوهشگران مورد استفاده قرار گرفته است. Koutsoyiannis و Onof (۲۰۰۱) مدل ریز مقیاس کردن بارش با استفاده از روش دقیق‌سازی بر روی مدل خوشه پواسون را ارائه نمودند وKoutsoyiannis (۲۰۰۱) مدل ریز مقیاس کردن بارش در چند مکان را به‌منظور ریز مقیاس کردن همزمان زمانی و مکانی در مقیاس زمانی ارائه نمود. علاوه بر موارد ذکر شده، برای روش‌های کلاسیک، از مطالعات دیگر که عموماً بر روی بارش متمرکز می‌باشند، می‌توان به کارهای Hershenhorn و Woolhiser (۱۹۸۷)، Ormsbee (۱۹۸۹)، Glasbey و همکاران (۱۹۹۵)، Connolly و همکاران (۱۹۹۸)، Olsson و Berndtsson (۱۹۹۸)، Durrans و همکاران (۱۹۹۹)، Sivakumar و همکاران (۲۰۰۱)، Socolofsky و همکاران (۲۰۰۱)، Gyasi-Agyei (۲۰۰۵)، Zhang و همکاران (۲۰۰۸) و Knoesen و Smithers (۲۰۰۹) اشاره کرد. عموماً روش‌های کلاسیک ریزمقیاس کردن، دارای ذات خطی بوده و در مواجهه با فرآیندهای بارش که ماهیتی غیرخطی دارند دارای محدودیت می‌باشند. با عنایت به ماهیت غیر خطی سری‌های زمانی بارش، در سال‌های اخیر، مطالعات مختلفی برای مدل‌سازی بارش با مدل‌های غیرخطی هوش مصنوعی توسعه یافته و برای ریز مقیاس کردن سری‌های زمانی بارش نیز مورد استفاده قرار گرفته است. Burian و همکاران (۲۰۰۰) استفاده از شبکه‌های عصبی مصنوعی را برای ریزمقیاس کردن داده‌های بارش ساعتی به داده‌های کوچکتر از ساعتی مورد ارزیابی قرار دادند. Burian و همکاران (۲۰۰۱) تاثیر توانایی مشخصات مدل‌های مختلف شبکه های عصبی مصنوعی و آموزش آن‌ها شامل استاندارد نمودن داده‌ها، موقعیت مکانی داده‌های آموزش، تعداد داده‌های آموزش، تعداد تکرار آموزش و تعداد نرون‌های مخفی در مدل‌های ریزمقیاس کردن بارش را مورد ارزیابی قرار دادند.

www.SID.ir

Archive of SID

Burian و Durrans (۲۰۰۲) نتایجی را از توانایی مدل شبکه‌های عصبی مصنوعی خصوصاً آزمون اینکه چه خطاهایی در هایتوگراف ریزمقیاس شده بارش همسان با خطاهای پیش‌بینی هایتوگراف بارش می‌باشد ارائه نمودند. Tripathi و همکاران (۲۰۰۶) مدل های متعدد ریزمقیاس کردن بر پایه مدل SVM (*Support Vector Machine*) و ANN (*Artificial Neural Network*) برای ریزمقیاس کردن بارش ماهانه در نقاط مختلف هند را توسعه دادند. Raje و Mujumdar (۲۰۱۱) سه روش ریزمقیاس کردن CRF (*Conditional random field*)، KNN (*K-nearest neighbor*) و SVM را برای ریزمقیاس کردن بارش روزانه در ناحیه پنجاب در هند برای رژیم فصلی را توسعه دادند و نتایج را با هم مقایسه کردند. Kalra و همکاران (۲۰۱۳) مدل‌های ریزمقیاس کردن آماری با استفاده از مدل SVM و ANN رابرای پیش‌بینی حجم سیلاب سالانه توسعه دادند. Kim و Singh (۲۰۱۵) از مدل های ANN شامل MLP (*Multiple perceptron*) و KSOFM (*Kohonen self-organizing feature map*) برای ریز مقیاس کردن مکانی بارش منطقه‌ای در حوضه وای استریم استفاده نمودند که نتایج نشان از دقت قابل قبول هر دو روش در ریزمقیاس کردن بود. با وجود قابلیت مناسب مدل‌های هوش مصنوعی در مدل‌سازی سری‌های زمای هیدروکلیماتولوژیکی، هنگامیکه نوسانات سری زمانی تا حد زیادی ناایستا باشند و فرآیند مربوطه در دامنه بزرگی از مقیاس‌های زمانی از چند دقیقه تا چند روز عمل کند، ممکن است مدل‌های هوش مصنوعی به نتایج مطلوبی منجر نشده و در صورتیکه پیش پردازش مناسبی روی داده‌های مدل‌های هوش مصنوعی انجام نشود، هوش مصنوعی در مدلسازی داده‌های ناایستا توانایی مطلوب نخواهد داشت (Cannas *et al.*, 2006). لذا روش‌های هوش مصنوعی ریزمقیاس کردن با وجود ماهیت غیر خطی، در مواجه با داده‌های ناایستا دارای ضعف بوده و با عنایت به اینکه اکثر سری‌های زمانی بارش دارای ناایستایی می‌باشند، استفاده از روش‌هایی که کارایی بیشتری در لحاظ کردن ناایستایی‌ها در مدلسازی داشته باشد احساس می‌شود.

در حالت کلی برای تشخیص اطلاعاتی که در سیگنال مستتر می‌باشد باید از تبدیلات ریاضی استفاده نمود که تبدیل فوریه یکی از این تبدیلات می‌باشد (Polikar, 1996). تبدیل موجک (*Wavelet Transform*) نسخه بهبود یافته تبدیل فوریه می‌باشد. تبدیل فوریه در آنالیز سیگنال‌های ناایستا ضعیف بوده در حالیکه تبدیل موجک اجازه می‌دهد که سیگنال‌های نا ایستا نیز به خوبی آنالیز شوند (Sifuzzaman *et al.*, 2009). با عنایت به ناایستایی سری‌های زمانی بارش، در سال‌های اخیر پژوهش‌های هرچند محدود در خصوص ریزمقیاس کردن بارش با استفاده از تبدیل موجک صورت گرفته است (Rashid *et al.*, 2015; Kim *et al.*, 2017). ترکیب تئوری موجک و حداقل مربعات ماشین بردار پشتیبان منجر به توسعه و کاربرد مدل‌های حداقل مربعات ماشین بردار پشتیبان در مطالعات هیدروکلیماتولوژیکی شده است که در این مدل‌های موجک – حداقل مربعات ماشین بردار پشتیبان (*Wavelet – Least Square Support Vector Machine*) علاوه بر پارامترهای مدل حداقل مربعات ماشین بردار پشتیبان، موقعیت و مقیاس موجک‌ها بهینه‌سازی می‌شوند (Nourani *et al.*, 2014; Himanshu *et al.*, 2017; Sharma and Goyal, 2016; Farajzadeh and Alizadeh, 2017; Baydaroglo *et al.*, 2017; Kim *et al.*, 2017).

در تحقیق حاضر از قابلیت‌های مدل هوش مصنوعی حداقل مربعات ماشین بردار پشتیبان به‌عنوان مدل جعبه سیاه و نیز تبدیل موجک که یک ابزار مناسب برای پیش پردازش داده‌های ناایستا است، بهره گرفته شده و با استفاده از مدل ترکیبی موجک – حداقل مربعات ماشین بردار پشتیبان اقدام به ریزمقیاس کردن مکانی و زمانی سری‌های زمانی بارش شده و در نهایت نتایج با مدل حداقل مربعات ماشین بردار پشتیبان و یک روش کلاسیک رگرسیونی مقایسه گردیده است که با عنایت به پیشینه

www.SID.ir

Archive of SID

تحقیق‌های صورت گرفته، تاکنون از چنین مدل ترکیبی برای ریزمقیاس کردن سری‌های زمانی بارش استفاده نشده است که نشان از نوآوری تحقیق دارد.

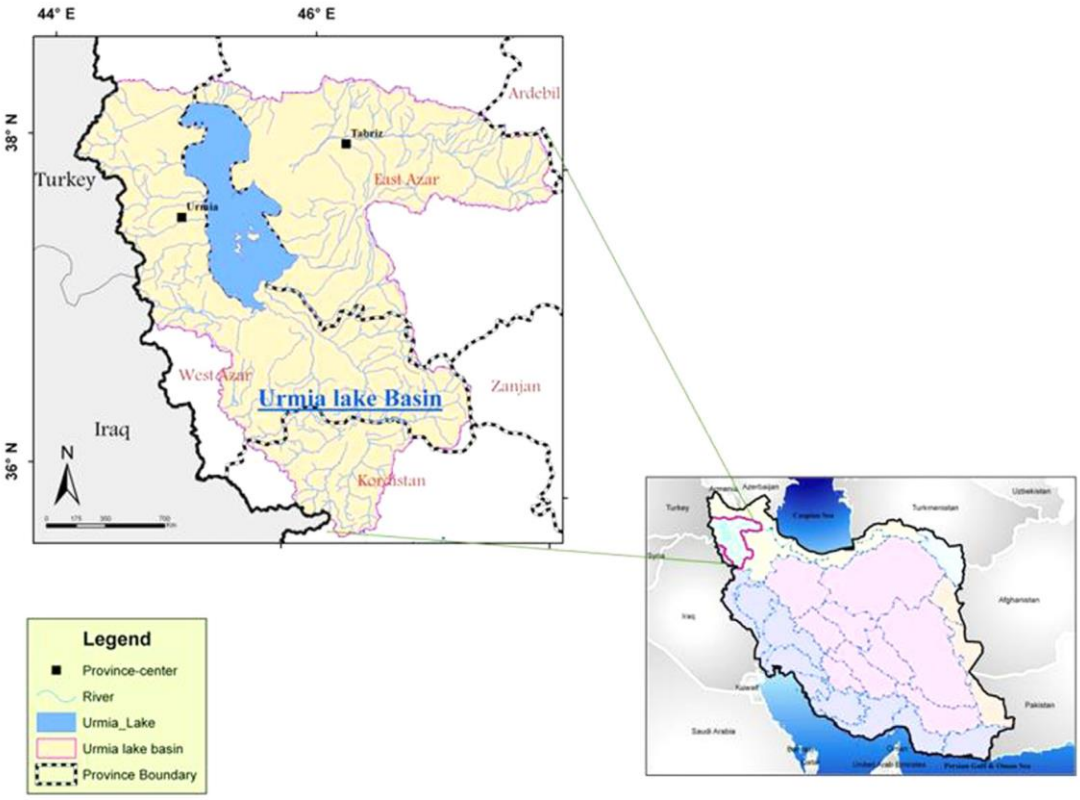
مواد و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه در این تحقیق مربوط به حوضه آبریز دریاچه ارومیه و علی‌الخصوص محدوده واقع در استان آذربایجان شرقی می‌باشد. حوضه آبریز ارومیه در شمال‌غرب ایران دارای مساحتی برابر با ۵۱۸۷۶ کیلومترمربع می‌باشد. محدوده غربی حوضه، مماس به ارتفاعات مرز بین ایران و ترکیه است. بارش سالانه حوضه از ۲۰۰ الی ۶۹۰ میلیمتر متغیر بوده و به مقدار میانگین ۳۴۵ میلی‌متر می‌باشد.

حوضه دارای ایستگاه‌های باران سنج متعدد می‌باشد که با عنایت به همبستگی بیشتر بین ایستگاه‌ها، داده‌های شش ایستگاه تبریز، سراب، سهند، مراغه، بناب و اهر انتخاب و در این مطالعه مورد استفاده قرار گرفته است. برای هر شش ایستگاه، سری زمانی داده‌های روزانه برای ده سال در دوره ۲۰۰۰ الی ۲۰۰۹ به صورت پیوسته موجود بوده و جهت مدل‌سازی استفاده گردید. ۷۵ درصد از این داده‌ها برای آموزش و ۲۵ درصد آن برای صحت سنجی در نظر گرفته شد. شکل (۱) و جدول (۱) به ترتیب منطقه مورد مطالعه و مشخصات آماری ایستگاه‌های مورد استفاده در این تحقیق، را نشان می‌دهد.

ریزمقیاس کردن با استفاده از داده‌های روزانه چهار ایستگاه (سراب، مراغه، بناب و اهر) و داده‌های ماهانه هر شش ایستگاه، برای یافتن داده‌های روزانه ایستگاه‌های تبریز و سهند انجام شد.



شکل ۱. حوضه آبریز دریاچه ارومیه با شبکه رودخانه‌ها و استان‌های درگیر

www.SID.ir

Archive of SID

جد.ول ۱. مشخصات آماری سری‌های زمانی روزانه ایستگاه‌های مورد مطالعه

	واحد	ایستگاه ها					
		تبریز	سهند	سراب	مراغه	بناب	اهر
تعداد داده ها	*	۳۶۵۳	۳۶۵۳	۳۶۵۳	۳۶۵۳	۳۶۵۳	۳۶۵۳
میانگین	میلی متر	۰.۶۵۴	۰.۵۴۳	۰.۶۵۵	۰.۷۱۰	۰.۷۳۳	۰.۷۰۶
مقدار ماکزیمم	میلی متر	۳۷	۴۲	۳۱	۴۲	۴۲	۲۹.۴
انحراف معیار استاندارد	میلی متر	۲.۴۳۸	۲.۱۴۲	۲.۱۶۰	۲.۸۰۹	۲.۹۲۵	۲.۴۳۹
چولگی	*	۵.۹۲۸	۷.۱۴۲	۵.۳۰۲	۶.۸۰۹	۶.۸۹۴	۵.۴۱۱

روش پیشنهادی

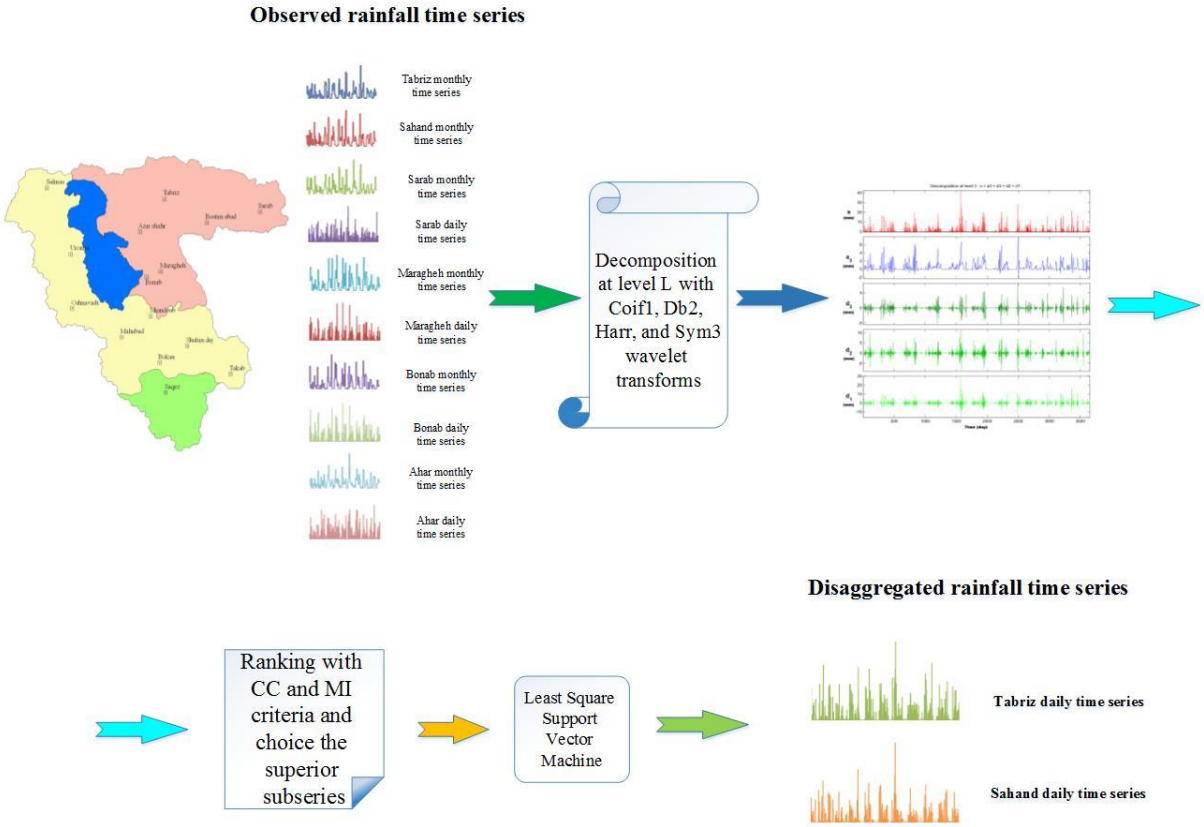
در این تحقیق، از مدل حداقل مربعات ماشین بردار پشتیبان به‌عنوان یکی از ابزارهای هوش مصنوعی و محاسباتی نرم (*soft computing*) و تبدیل موجک برای ریزمقیاس کردن مکانی (چند ایستگاهه) و زمانی بارش استفاده شده است. برای این منظور در ابتدا، داده‌های ورودی در مقیاس‌های بزرگ زمانی برای چند ایستگاه تهیه و نرمالسازی و پیش پردازش داده‌ها صورت گرفت. سپس در دوحالت با تبدیل موجک و بدون تبدیل موجک اقدام به ریزمقیاس کردن شد. در حالت با تبدیل موجک، ابتدا سری‌های زمانی با استفاده از تبدیل موجک و در مرتبه L به زیر سری‌های زمانی در مقیاس‌های مختلف تجزیه شده و سپس زیر سری‌ها با معیارهای اطلاعات متقابل (*Mutual Information*) و ضریب همبستگی (*Correlation Coefficient*) رتبه‌بندی شده و زیرسری‌های برتر انتخاب گردیده و این زیر سری‌ها به‌عنوان ورودی به مدل حداقل مربعات ماشین بردار پشتیبان اعمال شد. در حالت بدون تبدیل موجک نیز، داده‌های خام اصلی وارد مدل حداقل مربعات ماشین بردار پشتیبان شده و خروجی حاصل شد. در مرحله بعد نتایج حاصل از مدل‌های ریزمقیاس کردن مختلف این تحقیق، صحت سنجی گردید و نتایج حاصله با مدل رگرسیون خطی چند متغیره، مقایسه و مدل دقیق‌تر انتخاب گردید.

در این تحقیق برای ریزمقیاس کردن ایستگاه‌ها از سری‌های زمانی ایستگاه‌های مختلف (مکان‌های مختلف) استفاده شده است که بیانگر ماهیت مکانی ریزمقیاس کردن بوده و با توجه به ریز مقیاس کردن سری‌های زمانی بارش ماهانه به روزانه، بیانگر ماهیت زمانی ریزمقیاس کردن می‌باشد و در حقیقت برای مکان‌های مختلف به صورت مدل‌های مجزا از سری‌های زمانی مکان‌های دیگر استفاده شده است .

در شکل (۲)، شمای کلی از روند ریزمقیاس کردن توسط مدل موجک – حداقل مربعات ماشین بردار پشتیبان آورده شده است. در ادامه جزئیات مدل‌های مورد استفاده به اختصار آورده می‌شود.

www.SID.ir

Archive of SID



شکل۲. روند ریز مقیاس کردن بارش با استفاده از مدل ترکیبی موجک – حداقل مربعات ماشین بردار پشتیبان

تبدیل موجک

تبدیل موجک یک روش جدید برای تحلیل زمان – فرکانس سیگنال‌های غیرایستا می‌باشد. این تبدیل انعطاف‌پذیری بین مقیاس زمان و فرکانس را که در تبدیل فوریه وجود نداشت، فراهم می‌کند. موجک‌ها امکان تحلیل سیگنال‌ها در سطوح مختلف از مقیاس زمان را فراهم می‌آورند. موجک یک تابع زمانی است که طبیعت پرموج یا نوسانی دارد و انرژی آن به یک مدت متناهی از زمان محدود شده است. اگر $\psi(x)$ جهت مشخص کردن موجک مادر استفاده شود، آن گاه تبدیل موجک پیوسته توسط رابطه (۱) مشخص می‌شود (Addison *et al.*, 2001).

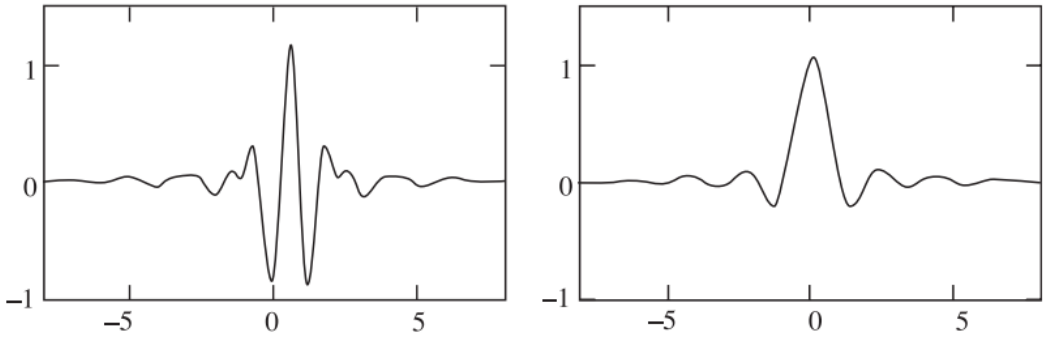
$$w(a,b) = \int f(t) \times (1/\sqrt{a}) \times \psi\,(t-b/a)dt \quad (۱)$$

پارامتر a به‌عنوان مقیاس در نظر گرفته می‌شود و میزان کشش یا مدت زمان موجک را بیان می‌کند. مقادیر بالای مقیاس، اتساع موجک و در نتیجه کاهش قابلیت تفکیک زمانی و افزایش مقیاس فرکانس را باعث می‌گردد. بالعکس، مقدار پایین مقیاس، متناظر با مقیاس زمانی بالا و مقیاس فرکانسی کم می‌باشد. پارامتر b پارامتر انتقال بوده و موقعیت موجک را بر روی محور زمان بیان می‌کند. این پارامتر تمرکز زمانی مورد نیاز را فراهم می‌کند.

شکل (۳) یک نمونه موجک است که به نام موجک میر (*Meyer*) معروف است و در سمت راست آن تابع مقیاس آن موجک مشخص می‌باشد.

www.SID.ir

Archive of SID



شکل ۳. **موجک میر و تابع مقیاس مربوطه (سمت راست)** (Addison, 2002)

از آنجایی که در کارهای هیدروکلیماتولوژیکی اغلب سری‌های زمانی گسسته مورد نظر می‌باشند، برای تجزیه سری زمانی می‌توان از فرم گسسته تبدیل موجک (*Discrete wavelet transform*) استفاده کرد. ضرایب موجک در هرنقطه از سیگنال (b) و برای هر مقدار از مقیاس (a) با رابطه (۲) قابل محاسبه است (Addison *et al.*, 2001).

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left[\frac{x-b}{a}\right] \qquad (۲)$$

در فرم گسسته تبدیل موجک، پارامترهای انتقال و مقیاس به صورت غیر پیوسته و با بهره‌گیری از روند دودویی انتخاب می‌شوند، به نحوی که:

$$a = 2^m, b = 2^m n \qquad (۳)$$

که m و n اعداد صحیح هستند. در نتیجه با جایگذاری a و b در رابطه (۲):

$$\psi_{m,n}(x) = 2^{-m/2}\psi[2^{-m}x - n] \qquad (۴)$$

برای سری زمانی گسسته x_i، تبدیل دودویی موجک به صورت زیر بیان می‌شود:

$$T_{m,n} = 2^{-m/2}\sum_{i=0}^{N-1}\psi(2^{-m}i - n)x_i \qquad (۵)$$

که در این رابطه، N طول سری زمانی و $T_{m,n}$ ضریب موجک برای a و b منتخب است و معکوس آن به شکل زیر محاسبه می‌شود:

$$x_i = \overline{T}(t) + \sum_{m=1}^{M}\sum_{n=0}^{M-m_{-1}} T_{m,n}2^{-m/2}g(2^{-m}i - n) \qquad (۶)$$

که در آن M طوری انتخاب شده که $N = 2^M$. رابطه فوق به صورت زیر خلاصه می‌شود:

$$x_i = \overline{T}(t) + \sum_{n=1}^{M} W_m(t) \qquad (۷)$$

که در رابطه اخیر، $\overline{T}$ زیر سری تخمین (*Approximation*) (که دوره تناوب فصلی غالب سری زمانی اصلی را دارد) از مرتبه M و W_m ها زیر سری زمانی جزئی (*Detailed*) از مراتب m=1,2,3,…,M هستند. به بیان دیگر هر یک از زیر سری‌ها، یک دوره تناوب از سری اصلی را شامل می‌شوند که خواص همان دوره را بیان می‌کنند. از این رو در انتخاب تعداد زیرسری‌ها و یا به عبارتی درجه تجزیه یک سری بایستی تعداد الگوهای تناوب موجود و یا محتمل در سری زمانی ملاک انتخاب قرار گیرد.

در این مطالعه در مرحله اول ابتدا با استفاده از رابطه تجربی (۸) مرتبه تجزیه موجک تعیین و سری‌های زمانی اصلی با استفاده از نرم‌افزار متلب و با به‌کارگیری چهار موجک مادر متداول db2 (*Daubechies-2*), Harr , Coif1 (*Coiflet-1*) و sym3 (*Symmlet-3*) و توسط تبدیل موجک گسسته به زیر سری‌های معین تجزیه و سپس با استفاده از معیار اطلاعات متقابل و ضریب همبستگی طبق روابط (۹) و (۱۰) ، زیرسری‌ها رتبه‌بندی شده و زیر سری‌های برتر هر دو معیار به‌عنوان ورودی مدل حداقل مربعات ماشین بردار پشتیبان استفاده گردید (Nourani *et*

www.SID.ir

Archive of SID

al., 2014; Yang *et al.*, 2000; Swinscow and Campbell, 1997). در اغلب مطالعات قبلی، جهت یافتن ورودی‌های مهم برای حداقل مربعات ماشین بردار پشتیبان از معیار ضریب همبستگی بین ورودی‌های مورد نظر و خروجی استفاده شده است ولی با عنایت به خطی بودن این معیار، نمی‌تواند وزن ارتباط‌های غیر خطی را به‌طور مناسب تعیین کند، لذا علاوه بر معیار ضریب همبستگی از معیار اطلاعات متقابل که مبتنی بر مفهوم آنتروپی شانون (*Shannon entropy*) می‌باشد نیز استفاده شده است. سری زمانی داده‌های مشاهداتی نظیر بارش، ماهیتی گسسته دارند. بنابراین از تبدیل موجک گسسته برای تجزیه و بازسازی سری‌های زمانی استفاده شده است.

$$L = int[log(N)] \qquad (۸)$$

$$MI(X,Y) = -\int\int p(x,y) log \frac{p(x,y)}{p(x)p(y)} dxdy \qquad (۹)$$

$$CC = \frac{\sum_{i=1}^{N}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\overline{x})^2 \sum_{i=1}^{N}(y_i-\overline{y})^2}} \qquad (۱۰)$$

در رابطه‌های اخیر، L مرتبه تجزیه، N تعداد داده های سری زمانی، int[.] قسمت صحیح تابع، X سری مشاهداتی، Y زیرسری محاسباتی، p(x,y) احتمال وقوع x و y، p(x) احتمال وقوع x، p(y) احتمال وقوع y، x_i مقدار مشاهده شده در گام زمانی i ام، y_i مقدار مشاهده شده در همان زمان، $\overline{x}$ میانگین مقادیر مشاهداتی و $\overline{y}$ نیز میانگین مقادیر محاسباتی می‌باشد. با عنایت به اینکه طول سری‌های زمانی ۳۶۵۳ روز می باشد با استفاده از رابطه (۸)، مرتبه تجزیه موجک، برابر با سه تعیین گردید. در تجزیه سیگنال توسط موجک، همه اطلاعات پنهان در سیگنال استخراج می‌شود، روشهای ابتدایی پردازش داده‌ها، نظیر نرمالسازی داده‌های خام، تأثیری در تجزیه ندارد و نتایج تجزیه حاصل از هر دو یکسان است بنابراین بدون نرمال‌سازی اقدام به تجزیه سری‌های زمانی شد.

حداقل مربعات ماشین بردار پشتیبان (LSSVM)

ماشین بردار پشتیبان نوعی سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که توسط Vapnik و Cortes (۱۹۹۵) ارائه شد. در این مدل از اصل استقرای کمینه‌سازی خطای ساختاری استفاده و به یک جواب بهینه کلی منجر می‌شود. در این مدل برای حل معادلات از برنامه‌ریزی دوگانه استفاده می‌شود. بنابراین حل مسائل در ابعاد بزرگ با استفاده از مدل یادشده، هزینه محاسباتی را به شدت افزایش می‌دهد. برای رفع این مشکل، Suykens و همکاران (۲۰۰۲) روش حداقل مربعات ماشین بردار پشتیبان را ارایه نمودند. در این مدل از معادلات خطی ساده برای حل مسائل استفاده می‌شود که به طور مؤثری پیچیدگی الگوریتم را کاهش می‌دهد.

تابع رگرسیونی استفاده شده در مدل LSSVM به صورت رابطه (۱۱) است.

$$y(x_i) = w^T.\emptyset(x_i) + b \qquad (۱۱)$$

که در آن $\emptyset(x_i)$ ترسیم غیرخطی ورودی‌ها در فضای ویژگی با ابعاد بزرگ است. پارامترهای w و b به ترتیب مقادیر وزن‌ها و بایاس تابع رگرسیون هستند که از طریق حداقل‌سازی تابع هدف در رابطه (۱۵) تعیین می‌گردند.

$$min_{w,e,b} j(w,e) = \frac{1}{2}w^T w + \frac{\gamma}{2}\sum_{i=1}^{N} e_i^2 \qquad (۱۲)$$

با محدودیت:

$$y_i = w^T\emptyset(x_i) + b + e_i \quad i = 1,2,\dots,N \qquad (۱۳)$$

که در آن γ پارامتر تنظیم کننده بخش خطا است و e_i خطای داده‌های آموزشی می باشد. در نهایت، تابع تخمین مدل LSSVM به صورت رابطه (۱۴) تعریف می‌شود.

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x_i, x_j) + b \qquad (۱۴)$$

که در آن $K(x_i, x_j)$ تابع کرنل نامیده می‌شود که مطابق رابطه (۱۵)، به‌عنوان تابعی با ایجاد ضرب داخلی در فضای ویژگی معرفی می‌شود.

$$K(x_i, x_j) = \emptyset(x_i).\emptyset(x_j) \quad i,j = 1,2,\dots,N \qquad (۱۵)$$

برای تحلیل حساسیت مدل LSSVM می‌توان با استفاده از روش آزمون و خطا، تابع کرنل را انتخاب نمود. با توجه به اینکه در مسائل دارای رگرسیون عموماً با مسائل غیرخطی موجه هستیم، بایستی روش حل را برای حالت غیرخطی نیز تعمیم دهیم. به این منظور از تابع

www.SID.ir

Archive of SID

کرنل پایه شعاعی (*Radial Basis Function*) مطابق رابطه (۱۶) که در مسائل مدیریت منابع آب بیشترین سرعت و کمترین خطا را داشته، در این تحقیق استفاده شده است. (Suykens *et al.*, 2002)

$$K(x_i, x_j) = exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \qquad (۱۶)$$

که در آن σ پارامتر پهنه کرنل می باشد.

در مدلسازی حداقل مربعات ماشین بردار پشتیبان، برای سری زمانی لازم است که داده‌های سری به دو دسته داده‌های آموزشی (واسنجی) و داده‌های آزمایشی (صحت سنجی) تقسیم گردند که این تقسیم‌بندی می‌تواند به نسبت ۷۵ به ۲۵ درصد به ترتیب برای داده‌های واسنجی و صحت سنجی تقسیم گردند. همچنین نرمال کردن داده‌ها در بهبود نتایج و افزایش دقت پیش‌بینی سری زمانی تأثیر بسزایی دارد (Cannas *et al.*, 2006; Nourani *et al.*, 2009).

با وجود عدم نرمالسازی داده‌ها در مدل ترکیبی موجک – حداقل مربعات ماشین بردار پشتیبان، در مدل حداقل مربعات ماشین بردار پشتیبان، برای نرمال کردن ورودی‌ها و خروجی حداقل مربعات ماشین بردار پشتیبان از فرمول (۱۷) استفاده شده است.

$$O_i = \frac{(X_i - X_{min})}{(X_{max} - X_{min})} \qquad (۱۷)$$

که X_i سری زمانی بارش و X_{min} و X_{max} به ترتیب مینیمم و ماکزیمم بارش می باشند.

انتخاب پارامترهای بهینه پهنه کرنل (σ) و تنظیم کننده بخش خطا (γ) موضوع مهم در مدل‌سازی توسط LSSVM می‌باشند که باید مورد توجه قرار گیرد. به‌طوریکه انتخاب مناسب این دو پارامتر سبب پیشرفت مدلسازی در مرحله آموزش و تست می‌شود. در این تحقیق با انتخاب پارامتر پهنه کرنل ۱ الی ۱۵۰ و پارامتر تنظیم کننده بخش خطا ۱ الی ۵۰ اقدام به یافتن مدل بهینه با روش جستجوی شبکه (*Grid Search*) گردید و ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد برای صحت‌سنجی در نظر گرفته شد (Nourani and Andalib 2015).

رگرسیون چند متغیره خطی

در این تحقیق، برای بررسی کارایی مدل پیشنهادی، نتایج با نتایج یک مدل کلاسیک رگرسیون چند متغیره خطی مقایسه شده است.

آنالیز رگرسیونی یک ابزار آماری برای بررسی روابط بین متغیرها است. در آنالیز رگرسیونی روابطی را که تابع رگرسیون نامیده می‌شود بین یک متغیر وابسته y و متغیرهای متعدد مستقل x_i مورد مطالعه قرار داده می‌شود. همچنین تابع رگرسیون از پارامترهای نامشخص b_i نیز تشکیل یافته است. اگر یک تابع رگرسیون در پارامترها خطی باشد (نه الزاماً در متغیرهای وابسته)، مدل خطی رگرسیون نامیده می‌شود. در غیراینصورت مدل غیر خطی نامیده می‌شود. مدل‌های رگرسیون خطی با بیش از یک متغیر مستقل، مدل خطی چندمتغیره نامیده می‌شود. رگرسیون چندمتغیره، روشی برای ساختن معادله مدل از سری داده‌های گذشته می‌باشد (Shimizukawa *et al.*, 2009). این روش یکی از قویترین روش‌های کلاسیک برای تبیین ارتباط بین متغیرهای مدل شده و مشاهداتی به شمار می‌رود. شکل کلی معادله همبستگی چند متغیره، به صورت رابطه (۱۸) می‌باشد (Afifi and Azen, 1979; Gunst and Mason, 1980).

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots b_p x_p \qquad (۱۸)$$

که در آن y متغیر وابسته (که با یک مدل رگرسیون تخمین زده می‌شود) ، x_i $(i = 1,2, \ldots p)$ متغیر مستقل i ام از مجموعه‌ای از p متغیر، b_i $(i = 1,2, \ldots p)$ ضریب متناظر به x ، b_0 یک ثابت و p تعداد متغیرهای مستقل می‌باشد.

هدف اصلی آنالیز رگرسیونی چندمتغیره، تخمین بهترین پارامترهای b_i می‌باشد به نحوی که مدل مقادیر تجربی متغیر وابسته را با بهرین دقت تخمین بزند (یعنی

www.SID.ir

Archive of SID

مقادیر محاسباتی y_j بایستی به مقادیر تجربی y_j^* نزدیک باشند).

معیار ارزیابی مدل

برای ارزیابی مقدار کارایی و دقت مدل‌ها از ریشه مربعات خطاها (*Root Mean Square Errors*) و ضریب تبیین (*Determination Coefficient*) با فرمول‌های (۱۹) و (۲۰) استفاده شده است (Nourani *et al.*, 2009). با عنایت به اینکه در سری‌های زمانی بارش در مقیاس روزانه، خیلی از روزها به دلیل عدم بارش، می‌تواند صفر باشد، در صورت استفاده از ضریب تبیین به جز صفر (*Except zero Determination Coefficient*) با فرمول (۲۱)، می‌توان بهتر به مقایسه و ارزیابی نتایج پرداخت. لذا در این تحقیق از این معیار نیز استفاده گردید

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(O_{obs_i}-O_{com_i}\right)^2}{N}} \quad (۱۹)$$

$$DC = 1 - \frac{\sum_{i=1}^{N}\left(O_{obs_i}-O_{com_i}\right)^2}{\sum_{i=1}^{N}\left(O_{obs_i}-\bar{O}_{obs}\right)^2} \quad (۲۰)$$

$$DC_{Except\ zero} = 1 - \frac{\sum_{i=1}^{N}\left(O_{obs_{Except\ zero}}-O_{com_{Except\ zero}}\right)^2}{\sum_{i=1}^{N}\left(O_{obs_{Except\ zero}}-\bar{O}_{obs_{Except\ zero}}\right)^2} \quad (۲۱)$$

در روابط فوق، N تعداد داده‌های مشاهداتی، O_{obs_i} داده‌های مشاهداتی، $O_{com\,i}$ مقادیر محاسبه شده و $\bar{O}_{obs}$ میانگین داده‌های مشاهداتی، $O_{obs_{Except\ zero}}$ داده‌های مشاهداتی به جز صفرها، $O_{com\ Except\ zero}$ مقادیر محاسبه شده به جز صفرها و $\bar{O}_{obs_{Except\ zero}}$ میانگین داده‌های مشاهداتی به جز صفر هستند.

نتایج و بحث

در این تحقیق، مدل ریزمقیاس کردن با استفاده از داده‌های روزانه چهار ایستگاه (سراب، مراغه، بناب و اهر) و داده‌های ماهانه هر شش ایستگاه (تبریز، سهند، سراب، مراغه، بناب و اهر)، برای یافتن داده‌های روزانه ایستگاه‌های تبریز و سهند صورت گرفت. ابتدا بدون هیچ گونه پیش پردازشی و صرفاً با نرمالسازی داده‌ها بین صفر و یک، مدل حداقل مربعات ماشین بردار پشتیبان برای ریزمقیاس کردن بارش ایستگاه‌های تبریز و سهند مدنظر گرفته شد. در مدل حداقل مربعات ماشین بردار پشتیبان، از دوازده الگوی ورودی مختلف به شرح جداول شماره (۲) و (۳) برای ریز مقیاس کردن بارش ایستگاه‌های تبریز و سهند استفاده شد.

جدول ۲. الگو های ورودی مدل حداقل مربعات ماشین بردار پشتیبان برای ایستگاه بارش تبریز

شماره الگوی ورودی	الگوی ورودی
۱	$DR1(t) = f\{MR2(t), \mathrm{MR1(t)}\}$
۲	$DR1(t) = f\{MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۳	$DR1(t) = f\{MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۴	$DR1(t) = f\{MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۵	$DR1(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۶	$DR1(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t), \mathrm{MR1(t)}\}$
۷	$DR1(t) = f\{DR6(t), DR5(t), DR4(t), DR3(t), \mathrm{MR1(t)}\}$
۸	$DR1(t) = f\{MR6(t), DR6(t-1), MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۹	$DR1(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t-1), MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۱۰	$DR1(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t), MR4(t), DR4(t-1), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۱۱	$DR1(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t-1), MR2(t), \mathrm{MR1(t)}\}$
۱۲	$DR1(t) = f\{MR6(t), DR6(t-1), MR5(t), DR5(t-1), MR4(t), DR4(t-1), MR3(t), DR3(t-1), MR2(t), \mathrm{MR1(t)}\}$

توجه : MR مخفف بارش ماهانه و DR مخفف بارش روزانه و شماره ایستگاه‌های بارش تبریز، سهند، سراب، مراغه، بناب و اهر به ترتیب یک الی شش می‌باشد.

www.SID.ir

Archive of SID

جدول ۳. الگوهای ورودی مدل حداقل مربعات ماشین بردار پشتیبان برای ایستگاه بارش سهند

شماره الگوی ورودی	الگوی ورودی
۱	$DR2(t) = f\{MR2(t), \mathrm{MR1(t)}\}$
۲	$DR2(t) = f\{MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۳	$DR2(t) = f\{MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۴	$DR2(t) = f\{MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۵	$DR2(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۶	$DR2(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t), \mathrm{MR2(t)}\}$
۷	$DR2(t) = f\{DR6(t), DR5(t), DR4(t), DR3(t), \mathrm{MR2(t)}\}$
۸	$DR2(t) = f\{MR6(t), DR6(t-1), MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۹	$DR2(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t-1), MR4(t), DR4(t), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۱۰	$DR2(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t), MR4(t), DR4(t-1), MR3(t), DR3(t), MR2(t), \mathrm{MR1(t)}\}$
۱۱	$DR2(t) = f\{MR6(t), DR6(t), MR5(t), DR5(t), MR4(t), DR4(t), MR3(t), DR3(t-1), MR2(t), \mathrm{MR1(t)}\}$
۱۲	$DR2(t) = f\{MR6(t), DR6(t-1), MR5(t), DR5(t-1), MR4(t), DR4(t-1), MR3(t), DR3(t-1), MR2(t), \mathrm{MR1(t)}\}$

جداول شماره (۴) و (۵) نتایج ریزمقیاس کردن را با استفاده از مدل LSSVM برای ایستگاه‌های تبریز و سهند نشان می‌دهند که در ایستگاه تبریز، الگوی شماره نه با ورودی سری‌های زمانی بارش روزانه و ماهانه ایستگاه‌های اهر، سراب و مراغه و سری‌های زمانی بارش ماهانه ایستگاه‌های تبریز، سهند و بناب و سری زمانی بارش روزانه یک روز قبل ایستگاه بناب، دارای بهترین ساختار در بین الگوهای ورودی با پارامتر بهینه پهنه کنترل (σ) ۵۰ و پارمتر بهینه تنظیم کننده بخش خطا (ϒ) ۵۰ بوده و ضریب تبیین ۰/۷۴ برای آموزش و ۰/۶۰ برای اعتبارسنجی می‌باشد. در خصوص ایستگاه سهند نیز، الگوی شماره ده با ورودی سری‌های زمانی بارش روزانه و ماهانه ایستگاه‌های اهر، سراب و بناب و سری‌های زمانی بارش ماهانه ایستگاه‌های تبریز، سهند و بناب و سری زمانی بارش روزانه یک روز قبل ایستگاه مراغه، دارای بهترین ساختار در بین الگوهای ورودی با پارامتر بهینه پهنه کنترل (σ) ۵۰ و پارمتر بهینه تنظیم کننده بخش خطا (ϒ) ۵۰ بوده و ضریب تبیین ۰/۷۷ برای آموزش و ۰/۵۳ برای اعتبارسنجی می باشد.

نتایج حاکی از تأثیر سری‌های زمانی بارش تمامی ایستگاه‌ها در مدل ریز مقیاس کردن داشته و الگوهای شماره نه و ده در هر دو ایستگاه، نتایج مناسب تری را ارائه داده که در هر دو الگوی شماره نه و ده، با توجه به دور بودن ایستگاه‌های بناب و مراغه به دو ایستگاه تبریز و سهند، تأثیر بارش روزانه یک روز قبل ایستگاه‌های بناب و مراغه بیشتر بوده و نتایج نشان‌دهنده اجرای نسبتاً مناسب مدل حداقل مربعات ماشین بردار پشتیبان در مدل‌سازی ریزمقیاس کردن می‌باشد.

جدول ۴. نتایج ضریب تبیین و خطاهای حاصل از ریزمقیاس کردن مدل حداقل مربعات ماشین بردار پشتیبان در ایستگاه تبریز

شماره الگوی	تابع کرنل	بهترین پارامتر ϒ	بهترین پارامتر σ	ضریب تبیین (DC) آموزش	ضریب تبیین (DC) اعتبارسنجی	جذر میانگین مربعات خطا (RMSE) (میلی متر) آموزش	جذر میانگین مربعات خطا (RMSE) (میلی متر) اعتبارسنجی	رتبه بندی
۱		۵۰	۲۰	۰.۰۹	۰.۰۵	۲.۴۳۵	۱.۹۹۳	۱۲
۲		۱۰۰	۵۰	۰.۵۴	۰.۴۹	۱.۷۳۱	۱.۴۶۷	۱۰
۳		۵۰	۵۰	۰.۶۳	۰.۵۵	۱.۵۵۰	۱.۳۷۰	۵
۴		۵۰	۵۰	۰.۶۶	۰.۵۳	۱.۴۸۲	۱.۴۱۱	۷
۵		۵۰	۵۰	۰.۷۴	۰.۵۸	۱.۳۰۱	۱.۳۲۱	۲
۶	تابع کرنل پایه	۵۰	۵۰	۰.۷۳	۰.۵۸	۱.۳۱۶	۱.۳۳۷	۳
۷	شعاعی	۲	۸	۰.۷۰	۰.۵۲	۱.۳۹۹	۱.۴۲۳	۸
۸		۵۰	۵۰	۰.۷۰	۰.۵۰	۱.۳۹۱	۱.۴۴۷	۹
۹		۵۰	۵۰	۰.۷۴	۰.۶۰	۱.۳۱۲	۱.۳۰۱	۱
۱۰		۱۰۰	۵۰	۰.۷۴	۰.۵۷	۱.۳۰۹	۱.۳۴۵	۴
۱۱		۲	۸	۰.۷۲	۰.۵۴	۱.۳۴۶	۱.۳۹۴	۶
۱۲		۱	۵۰	۰.۱۶	۰.۰۹	۲.۳۴۱	۱.۹۵۸	۱۱

www.SID.ir

Archive of SID

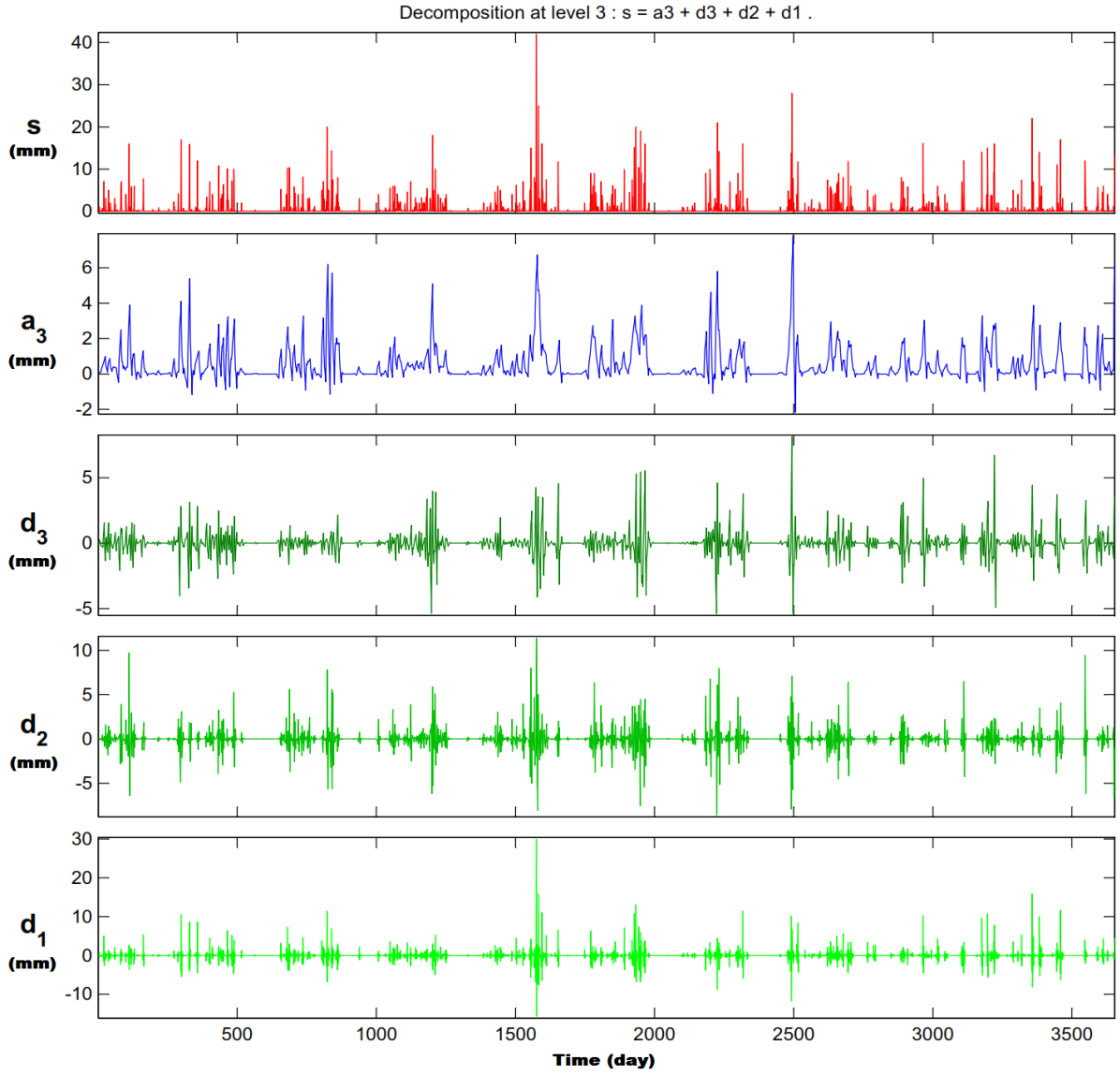
جدول ۵. نتایج ضریب تبیین و خطاهای حاصل از ریزمقیاس کردن مدل حداقل مربعات ماشین بردار پشتیبان در ایستگاه سهند

شماره الگوی	تابع کرنل	بهترین پارامتر		ضریب تبیین (DC)		جذر میانگین مربعات خطا (RMSE) (میلی متر)		رتبه بندی
		γ	σ	آموزش	اعتبارسنجی	آموزش	اعتبارسنجی	
۱		۵۰	۵۰	۰.۰۹	۰.۰۶	۲.۱۰۴	۱.۸۹۲	۱۲
۲		۵۰	۵۰	۰.۵۴	۰.۴۲	۱.۴۸۹	۱.۴۷۷	۱۰
۳		۵۰	۵۰	۰.۶۸	۰.۴۴	۱.۲۴۶	۱.۴۵۹	۷
۴		۲	۴۰	۰.۶۵	۰.۴۳	۱.۳۰۰	۱.۴۶۴	۸
۵		۲	۲۰	۰.۷۳	۰.۴۹	۱.۱۴۶	۱.۳۸۹	۳
۶	تابع کرنل پایه	۲	۲۰	۰.۷۲	۰.۴۸	۱.۱۶۵	۱.۳۹۸	۴
۷	شعاعی	۵۰	۲۰	۰.۷۷	۰.۵۰	۱.۰۴۹	۱.۳۸۲	۲
۸		۲	۵۰	۰.۶۵	۰.۴۳	۱.۳۰۱	۱.۴۶۶	۹
۹		۲	۴۰	۰.۶۷	۰.۴۷	۱.۲۶۱	۱.۴۱۶	۵
۱۰		۵۰	۵۰	۰.۷۷	۰.۵۳	۱.۰۶۵	۱.۳۳۷	۱
۱۱		۲	۸	۰.۷۹	۰.۴۷	۱.۰۱۹	۱.۴۲۰	۶
۱۲		۱	۵۰	۰.۱۶	۰.۱۰	۲.۰۱۹	۱.۸۴۳	۱۱

با عنایت به اینکه در مدل حداقل مربعات ماشین بردار پشتیبان و هنگام آموزش، به مقیاسی که بیش‌ترین اثر را داشته باشد، وزن بیش‌تری اختصاص داده می‌شود، در مرحله بعد به‌منظور افزایش دقت مدل، داده‌های پیش پردازش شده توسط تبدیل موجک گسسته، به‌عنوان ورودی مدل حداقل مربعات ماشین بردار پشتیبان در نظر گرفته شده و به این ترتیب تمام ویژگی‌های اصلی در مقیاس ریز و درشت سری زمانی نمایان می‌شوند و مدل حداقل مربعات ماشین بردار پشتیبان به هر کدام از زیرسری‌های تجزیه‌ای وزن خاصی اختصاص می‌دهد. به‌طوریکه وزن‌های بالا به سطوح با ارزش سیگنال اختصاص می یابند که این خود باعث افزایش دقت مدل می‌گردد. در این تحقیق از چهار موجک مادر متداول db2, Harr, Coif1 و sym3 برای تجزیه به زیر سری‌های معین استفاده شد. برای نمونه در شکل شماره (۴) زیر سری‌های جزئی و تقریبی d_1 تا d_3 و a_3 برای سری زمانی بارش روزانه ایستگاه سهند، در محدوده زمانی این تحقیق با مرتبه تجزیه سه، حاصل از تجزیه موجک با موجک مادر db2 نشان داده شده است.

www.SID.ir

Archive of SID



شکل ۴. زیر سری‌های جزئی و تقریبی حاصل از تجزیه موجک با موجک مادر db2 سری زمانی بارش روزانه ایستگاه سهند (مرتبه سه)

خلاصه نتایج مدلسازی با این روش برای ایستگاه‌های تبریز و سهند به ترتیب در جداول شماره (۶) و (۷) مشاهده می‌شود.

www.SID.ir

Archive of SID

جدول ۶. نتایج ضریب تبیین و خطاهای حاصل از ریزمقیاس کردن مدل موجک – حداقل مربعات ماشین بردار پشتیبان در ایستگاه تبریز

شماره الگوی ورودی	تابع کرنل	موجک مادر	تراز تجزیه	بهترین پارمتر		ضریب تبیین (DC)		جذر میانگین مربعات خطا (RMSE) (میلی متر)		رتبه بندی
				γ	σ	آموزش	اعتبارسنجی	آموزش	اعتبارسنجی	
۵	تابع کرنل پایه شعاعی	coif1	۳	۲	۵۰	۰.۸۷	۰.۵۹	۰.۹۰۷	۱.۳۱۸	۱۲
		db2	۳	۲	۵۰	۰.۸۷	۰.۶۱	۰.۹۱۴	۱.۲۸۱	۱۱
		Harr	۳	۲	۵۰	۰.۸۶	۰.۶۲	۰.۹۵۴	۱.۲۶۶	۷
		sym3	۳	۲	۵۰	۰.۸۷	۰.۶۲	۰.۹۰۶	۱.۲۷۲	۹
با رتبه بندی بوسیله معیار ضریب همبستگی (CC)	تابع کرنل پایه شعاعی	coif1	۳	۱	۲۰	۰.۷۴	۰.۶۶	۱.۳۰۶	۱.۲۰۴	۲
		db2	۳	۲	۲۰	۰.۷۸	۰.۶۲	۱.۱۸۷	۱.۲۵۹	۴
		Harr	۳	۵۰	۵۰	۰.۸۳	۰.۶۲	۱.۰۴۲	۱.۲۶۵	۵
		sym3	۳	۵۰	۵۰	۰.۸۰	۰.۶۱	۱.۱۴۹	۱.۲۷۹	۱۰
با رتبه بندی بوسیله معیار اطلاعات متقابل (MI)	تابع کرنل پایه شعاعی	coif1	۳	۵۰	۵۰	۰.۸۳	۰.۶۶	۱.۰۶۶	۱.۲۰۲	۱
		db2	۳	۵۰	۵۰	۰.۸۱	۰.۶۲	۱.۱۲۵	۱.۲۷۱	۸
		Harr	۳	۲	۲۰	۰.۷۹	۰.۶۲	۱.۱۷۷	۱.۲۶۵	۵
		sym3	۳	۲	۸	۰.۸۶	۰.۶۵	۰.۹۶۹	۱.۲۰۵	۳

جدول ۷. نتایج ضریب تبیین و خطاهای حاصل از ریزمقیاس کردن مدل موجک – حداقل مربعات ماشین بردار پشتیبان در ایستگاه سهند

شماره الگوی ورودی	تابع کرنل	موجک مادر	تراز تجزیه	بهترین پارمتر		ضریب تبیین (DC)		جذر میانگین مربعات خطا (RMSE) (میلی متر)		رتبه بندی
				γ	σ	آموزش	اعتبارسنجی	آموزش	اعتبارسنجی	
۵	تابع کرنل پایه شعاعی	coif1	۳	۲	۵۰	۰.۸۷	۰.۵۹	۰.۷۸۳	۱.۲۵۲	۱۲
		db2	۳	۲	۵۰	۰.۸۷	۰.۶۱	۰.۷۸۹	۱.۲۱۷	۱۰
		Harr	۳	۲	۵۰	۰.۸۶	۰.۶۲	۰.۸۲۴	۱.۲۰۲	۷
		sym3	۳	۲	۵۰	۰.۸۷	۰.۶۲	۰.۷۸۲	۱.۲۰۹	۹
با رتبه بندی بوسیله معیار ضریب همبستگی (CC)	تابع کرنل پایه شعاعی	coif1	۳	۱	۲۰	۰.۷۴	۰.۶۶	۱.۱۲۷	۱.۱۴۴	۳
		db2	۳	۲	۲۰	۰.۷۸	۰.۶۲	۱.۰۲۶	۱.۱۹۶	۵
		Harr	۳	۲	۲۰	۰.۷۷	۰.۶۰	۱.۰۵۰	۱.۲۳۶	۱۱
		sym3	۳	۲	۲۰	۰.۷۹	۰.۶۳	۱.۰۱۱	۱.۱۷۹	۴
با رتبه بندی بوسیله معیار اطلاعات متقابل (MI)	تابع کرنل پایه شعاعی	coif1	۳	۵۰	۵۰	۰.۸۲	۰.۶۶	۰.۹۲۷	۱.۱۴۱	۱
		db2	۳	۲	۲۰	۰.۷۶	۰.۶۲	۱.۰۸۵	۱.۲۰۳	۸
		Harr	۳	۲	۲۰	۰.۷۹	۰.۶۲	۱.۰۱۶	۱.۱۹۶	۶
		sym3	۳	۲	۸	۰.۸۵	۰.۶۶	۰.۸۳۹	۱.۱۴۲	۲

در مدل موجک – حداقل مربعات ماشین بردار پشتیبان، در حالت بدون رتبه‌بندی، کل زیرسری‌های الگوی ورودی شماره پنج در هر دو ایستگاه وارد مدل حداقل مربعات ماشین بردار پشتیبان گردید که با عنایت به تعداد بالای ورودی‌ها، می‌تواند نوفه (*Noise*) را وارد فرآیند کند و با وجود ارتقای نتایج آموزش، در اعتبارسنجی نتایج مناسب نخواهد بود لذا در سه حالت بدون رتبه‌بندی و رتبه‌بندی با دو معیار ضریب همبستکی و اطلاعات متقابل و انتخاب ده زیرسری برتر، اقدام به مدل‌سازی گردید که در هر دو ایستگاه حالت رتبه‌بندی با معیار اطلاعات متقابل، دارای نتایج بهتری بوده و ده زیر سری برتر انتخاب شده در ایستگاه‌های سهند و تبریز، مربوط به زیر

www.SID.ir

Archive of SID

سری‌های ایستگاه‌های نزدیک‌تر بود، بدین نحو که ایستگاه‌های نزدیک دارای تخمین و تقریب بزرگ‌تر و ایستگاه‌های دورتر دارای تخمین و تقریب کوچک‌تری بود. در هر دو ایستگاه، بهترین ساختار در بین الگوهای ورودی با پارامتر بهینه پهنه کنترل (σ) ۵۰ و پارمتر بهینه تنظیم کننده بخش خطا (γ) ۵۰ با موجک مادر coif1 حاصل شده است. برای ایستگاه تبریز، ضریب تبیین ۰/۸۳ برای آموزش و ۰/۶۶ برای اعتبار سنجی و برای ایستگاه سهند، ضریب تبیین ۰/۸۲ برای آموزش و ۰/۶۶ برای اعتبار سنجی مدل‌سازی ریزمقیاس کردن، حاکی از ارتقای دقت مدل نسبت به مدل حداقل مربعات ماشین بردار پشتیبان داشته و نشان دهنده اجرای مناسب مدل موجک – حداقل مربعات ماشین بردار پشتیبان در مدل‌سازی ریزمقیاس کردن می باشد.

برای اعتبارسنجی روش‌ها، از روش کلاسیک رگرسیون خطی چند متغیره، نیز برای ریزمقیاس کردن داده‌های سری‌های زمانی مدنظر، استفاده شده که بهترین رابطه بین متغیر وابسته و متغیرهای مستقل برای این روش کلاسیک برای ایستگاه‌های تبریز و سهند، به ترتیب به صورت روابط (۲۲) و (۲۳) می‌باشد:

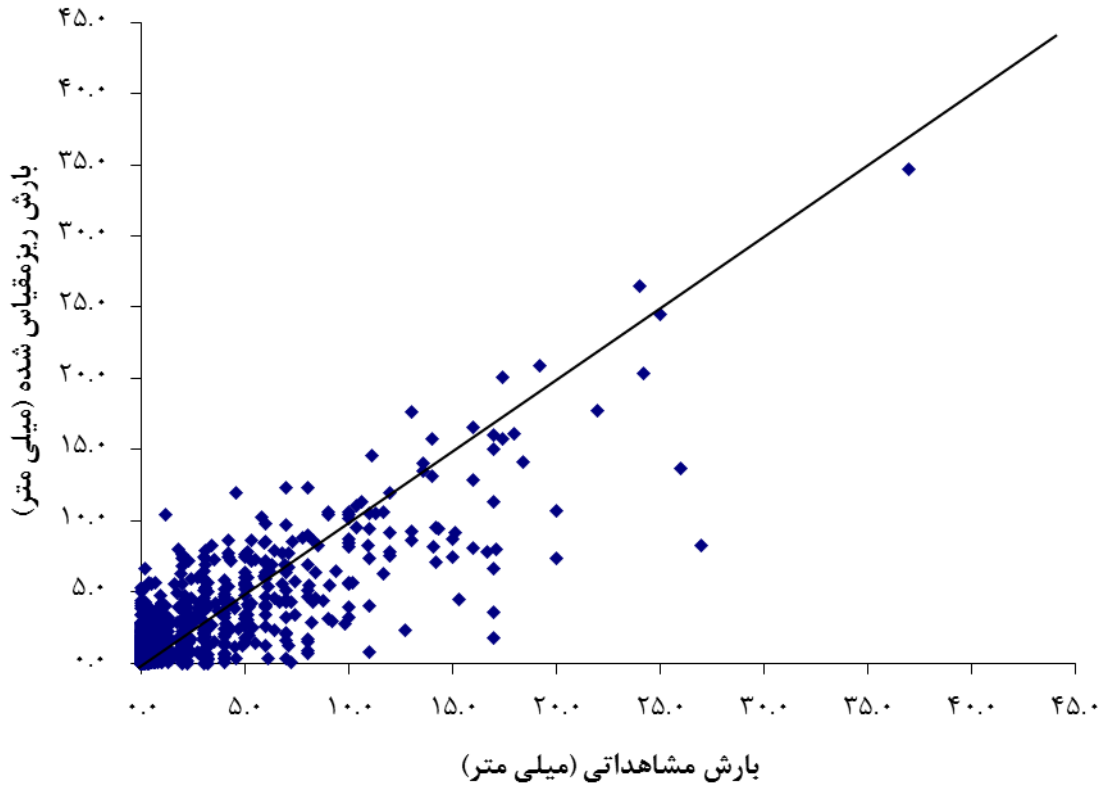
$$DR1(t) = 0.0005 + 0.0328\,MR1(t) + 0.0002\,MR2(t) - 0.0088\,MR3(t) + 0.2670\,DR3(t) - 0.0082\,MR4(t) + 0.2485\,DR4(t) - 0.0030\,MR5(t) + 0.0902\,DR5(t) - 0.0078\,MR6(t) + 0.2350\,DR6(t) \quad (۲۲)$$

$$DR2(t) = 0.0006 + 0.00004\,MR1(t) + 0.0330\,MR2(t) - 0.0073\,MR3(t) + 0.2202\,DR3(t) - 0.0078\,MR4(t) + 0.2349\,DR4(t) - 0.0039\,MR5(t) + 0.1180\,DR5(t) - 0.0055\,MR6(t) + 0.1637\,DR6(t) \quad (۲۳)$$

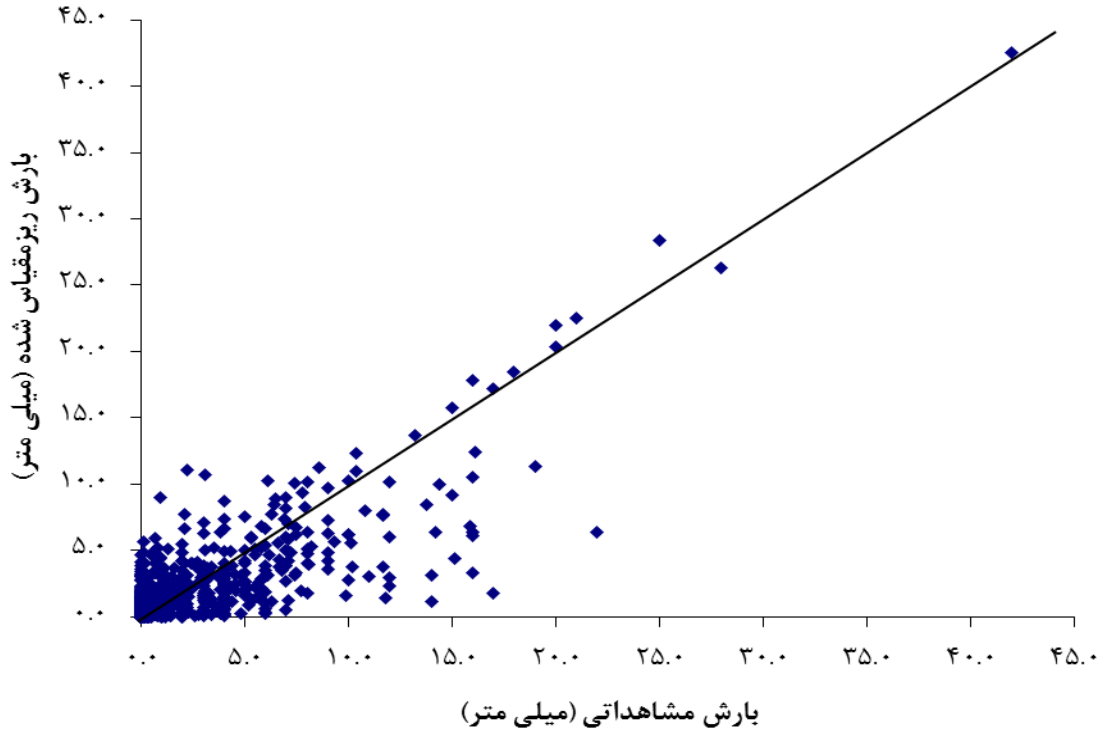
براساس روابط (۲۲) و (۲۳)، ورودی‌هایی که دارای ضریب بیش‌تری می‌باشند دارای اثر بیش‌تری بوده و برای هر دو ایستگاه تبریز و سهند، داده‌های روزانه به استثنای داده‌های روزانه ایستگاه بناب دارای اثر بیش‌تری می‌باشد که با نتایج حاصل از مدل حداقل مربعات ماشین بردار پشتیبان و الگوی شماره نه، دارای مطابقت بوده و حاکی از دقت مدلسازی می‌باشد. در شکل‌های (۵) و (۶) نیز به ترتیب پراکندگی داده‌های مشاهداتی و محاسباتی بارش حول نیمساز در بهینه‌ترین حالت حاصل از مدل ریزمقیاس کردن موجک – حداقل مربعات ماشین بردار پشتیبان در مرحله اعتبارسنجی برای ایستگاه‌های تبریز و سهند قابل مشاهده می‌باشد. همانطور که شکل نشان می‌دهد این مدل قادر به یافتن الگویی مناسب برای ریزمقیاس کردن بوده و توانایی شبیه‌سازی پیک‌ها را دارا می‌باشد.

www.SID.ir

Archive of SID



شکل ۵. پراکندگی داده‌های مشاهداتی و محاسباتی بارش (میلی‌متر) حول نیمساز در بهینه‌ترین حالت حاصل از مدل ریزمقیاس کردن WLSSVM برای ایستگاه تبریز



شکل ۶. پراکندگی داده های مشاهداتی و محاسباتی بارش (میلی‌متر) حول نیمساز در بهینه‌ترین حالت حاصل از مدل ریزمقیاس کردن WLSSVM برای ایستگاه سهند

www.SID.ir

Archive of SID

در شکل‌های (۵) و (۶)، تجمع داده ها بر روی نیمساز مؤید دقت رضایت بخش روش موجک – حداقل مربعات ماشین بردار پشتیبان، در ریزمقیاس کردن هر دو ایستگاه تبریز و سهند می‌باشد.

در جدول شماره (۸)، نتایج مقایسه‌ای سه روش ریزمقیاس کردن در دو ایستگاه تبریز و سهند آورده شده است. در این تحقیق، علاوه بر معیارهای ارزیابی ضریب تبیین و جذر میانگین مربعات خطا، با عنایت به اینکه در سری‌های زمانی بارش روزانه، بخش زیادی از داده‌ها، صفر می‌باشد، از ضریب تبیین به جز صفر نیز استفاده شده و دقت مدل‌ها در ریزمقیاس کردن سری‌های زمانی به غیر از صفر نیز بررسی شد. نتایج حاکی از این است که با توجه به این معیار نیز، مدل ریزمقیاس کردن موجک – حداقل مربعات ماشین بردار پشتیبان ، نسبت به دو مدل دیگر نتایج بهتری را نشان داده است. دلیل نتایج بهتر مدل موجک – حداقل مربعات ماشین بردار پشتیبان این است که ایستگاه‌ها از لحاظ مکانی و زمانی می‌توانند با تأخیرهای مختلف تأثیرگذار باشند که در این مدل ترکیبی، علاوه بر پیش پردازش و تقسیم داده‌ها به زیر سری‌ها، به هر زیر سری وزن متناسب با آن زیر سری اعمال و سپس مدلسازی شده که با توجه به ارتقای نتایج نسبت به مدل حداقل مربعات ماشین بردار پشتیبان، حاکی از تأثیر مثبت تبدیل موجک در مدل‌سازی و ریزمقیاس کردن می‌باشد.

جدول ۸. نتایج مقایسه ای ضریب تبیین و خطاهای حاصل از ریزمقیاس کردن بارش با سه مدل

مدل	نوع مدل	محل مورد مطالعه	ضریب تبیین (DC)		جذر میانگین مربعات خطا (RMSE) (میلی متر)		$DC_{ExceptZero}$
			آموزش	اعتبارسنجی	آموزش	اعتبارسنجی	اعتبارسنجی
رگرسیون چند متغیره خطی	خطی	تبریز	۰.۵۵	۰.۴۸	۱.۷۱۷	۱.۴۸۱	۰.۳۸
حداقل مربعات ماشین بردار پشتیبان	غیر خطی		۰.۷۴	۰.۶۰	۱.۳۱۲	۱.۳۰۱	۰.۵۴
موجک – حداقل مربعات ماشین بردار پشتیبان	ترکیبی		۰.۸۳	۰.۶۶	۱.۰۶۶	۱.۲۰۲	۰.۶۱
رگرسیون چند متغیره خطی	خطی	سهند	۰.۵۸	۰.۴۵	۱.۴۳۰	۱.۴۴۲	۰.۳۰
حداقل مربعات ماشین بردار پشتیبان	غیر خطی		۰.۷۷	۰.۵۳	۱.۰۶۵	۱.۳۳۷	۰.۴۹
موجک – حداقل مربعات ماشین بردار پشتیبان	ترکیبی		۰.۸۲	۰.۶۶	۰.۹۲۷	۱.۱۴۱	۰.۶۰

نتیجه‌گیری

در این تحقیق از سری زمانی داده‌های روزانه شش ایستگاه تبریز، سراب، سهند، مراغه، بناب و اهر مربوط به حوضه آبریز دریاچه ارومیه برای ده سال در پریود ۲۰۰۰ الی ۲۰۰۹ استفاده شد و ریزمقیاس کردن با استفاده از داده‌های روزانه چهار ایستگاه (سراب، مراغه، بناب و اهر) و داده های ماهانه هر شش ایستگاه، برای یافتن داده‌های روزانه ایستگاه‌های تبریز و سهند صورت گرفت. در این مطالعه نتایج مدل ترکیبی موجک – حداقل مربعات ماشین بردار پشتیبان با مدل حداقل مربعات ماشین بردار پشتیبان و مدل کلاسیک رگرسیون خطی چند متغیره، برای ریز مقیاس کردن بارش از مقیاس ماهانه به روزانه در حوضه آبریز دریاچه ارومیه مقایسه گردید و نتیجه‌ای که از این مقایسه گرفته شد حاکی از برتری مدل ریز مقیاس کردن موجک – حداقل مربعات ماشین بردار پشتیبان نسبت به دو مدل دیگر بود. با اینکه مدل LSSVM بخاطر در اختیار داشتن هسته غیرخطی توانایی مدلسازی رفتارهای غیرخطی را داراست، ولی با توجه به ویژگی‌های ناایستایی و غیرخطی بودن سری زمانی بارش، به‌منظور ریزمقیاس کردن و شبیه‌سازی، استفاده از مدل موجک – حداقل مربعات ماشین بردار پشتیبان به نتایج بهتری منجر گردید.

www.SID.ir

Archive of SID

چرا که تبدیل موجک با تجزیه سری زمانی به فرکانس‌های بالا و پایین ویژگی‌های چند مقیاسی سری زمانی را در اختیار گذاشته و دقت مدل را تا حد قابل توجهی بالا می‌برد.

در این مطالعه از چهار موجک مادر متداول db2, Harr, Coif1 و sym3 برای تجزیه به زیر سری‌های معین استفاده شد، که در هر دو ایستگاه، بهترین ساختار در بین الگوهای ورودی با پارامتر بهینه پهنه کنترل (σ) ۵۰ و پارمتر بهینه تنظیم کننده بخش خطا (γ) ۵۰ با موجک مادر coif1 حاصل شده است. در حالت بهینه ایستگاه تبریز، در مجموع نتایج مدل موجک – حداقل مربعات ماشین بردار پشتیبان نسبت به مدل‌های حداقل مربعات ماشین بردار پشتیبان و رگرسیون چند متغیره خطی برای اعتبارسنجی به ترتیب ۱۰٪ و ۳۷/۵٪ و در حالت بهینه ایستگاه سهند، در مجموع نتایج مدل موجک – حداقل مربعات ماشین بردار پشتیبان نسبت به مدل‌های حداقل مربعات ماشین بردار پشتیبان و رگرسیون چند متغیره خطی برای اعتبارسنجی به ترتیب ۲۴/۵٪ و ۴۶/۷٪ افزایش نشان داد. ملاحظه گردید که در هر دو ایستگاه، روش موجک – حداقل مربعات ماشین بردار پشتیبان نسبت به دو روش دیگر، دقت بالاتری داشته و به عنوان روشی مناسب جهت ریزمقیاس کردن پارامترهای هیدروکلیماتولوژیکی پیشنهاد گردید. همچنین در مدل ترکیبی، ضریب تبیین به جز صفر نیز بالا بوده و تطابق زیادی در بین سری‌های زمانی به جز صفر مشاهداتی و محاسباتی بارش وجود دارد.

ریزمقیاس کردن بارش با مدل‌های ترکیبی سایر مدل‌های هوش مصنوعی با تبدیل موجک و مقایسه نتایج با این تحقیق و ریزمقیاس کردن سایر پارامترهای هیدروکلیماتولوژیکی با مدل‌های این تحقیق می‌تواند از تحقیقات پیش روی در آینده باشد. همچنین ریزمقیاس کردن به مقیاس زمانی ریزتر مثل مقیاس ساعتی با مدل ترکیبی پیشنهادی و همینطور ریزمقیاس کردن سری‌های بارش مربوط به داده‌های مدل‌های گردش عمومی (GCMs)، می‌تواند از تحقیقات پیش روی در آینده باشد. ریزمقیاس کردن سری‌های زمانی دو یا چند ایستگاه به‌صورت همزمان با مدل ترکیبی پیشنهادی و همچنین مقایسه نتایج مدل ترکیبی پیشنهادی با سایر مدل‌های کلاسیک ریزمقیاس کردن من جمله مدل‌های Hyetos یا Mudrain و یا سایر مدل‌های هوش مصنوعی، می‌تواند برای ادامه این تحقیق، پیشنهاد گردد.

منابع مورد استفاده

Addison, P. 2002. The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance. Institute of Physics Publishing Bristol and Philadelphia, london, 368pp.

Addison, P., Murray, K. and Watson, J. 2001. Wavelet transform analysis of open channel wake flows. Journal of Engineering Mechanics, 127(1): 58-70.

Afifi, A. and Azen, S. 1979. Statistical analysis: A computer oriented approach. New York, Academic press, 442p

Baydaroğlu, Ö., Koçak, K. and Duran, K. 2017. River flow prediction using hybrid models of support vector regression with the wavelet transform, singular spectrum analysis and chaotic approach. Meteorology and Atmospheric Physics, 1(1): 1–11.

Burian, S. and Durrans, S. 2002. Evaluation of an artificial neural network rainfall disaggregation model. Water Science and Technology, 45(2): 99-104.

Burian, S., Durrans, S., Nix, S. and Pitt, R. 2001. Training artificial neural networks to perform rainfall disaggregation. Journal of Hydrologic Engineering, 6(1): 43-51.

Burian, S., Durrans, S., Tomić, S., Pimmel, R. and Wai, C. 2000. Rainfall disaggregation using artificial neural networks. Journal of Hydrologic Engineering, 5(3): 299-307.

www.SID.ir

Archive of SID

Cannas, B., Fanni, A., See, L. and Sias, G. 2006. Data preprocessing for river flow forecasting using neural networks: Wavelet transforms and data partitioning. Physics and Chemistry of the Earth, Parts A/B/C, 31(18): 1164-1171.

Connolly, R., Schirmer, J. and Dunn, P. 1998. A daily rainfall disaggregation model. Agricultural and Forest Meteorology, 92(2): 105-117.

Cortes, C. and Vapnik, V. 1995. Support vector machine. Machine learning Journal, 20(3): 273–297.

Durrans, S., Burian, S., Nix, S., Hajji, A., Pitt, R., Fan, C.Y. and Field, R. 1999. Polynomial-based disaggregation of hourly rainfall for continous hydrologic simulation. Journal of The American Water Resources Association, 35(5): 1213-1221.

Farajzadeh, J. and Alizadeh, F. 2017. A hybrid linear–nonlinear approach to predict the monthly rainfall over the Urmia Lake watershed using wavelet-SARIMAX-LSSVM conjugated model. Journal of Hydroinformatics, 19(6): 1-17.

Glasbey, C., Cooper, G. and McGechan, M. 1995. Disaggregation of daily rainfall by conditional simulation from a point-process model. Journal of Hydrology, 165(1): 1-9.

Gunst, R. and Mason, R. 1980. Regression analysis and its application: A data-oriented approach. Marcel Dekker Inc, NewYork, 424pp.

Gyasi-Agyei, Y. 2005. Stochastic disaggregation of daily rainfall into one-hour time scale. Journal of Hydrology, 309(1): 178-190.

Hershenhorn, J. and Woolhiser, D. 1987. Disaggregation of daily rainfall. Journal of Hydrology, 95(3): 299-322.

Himanshu, S. K., Pandey, A. and Yadav, B. 2017. Assessing the applicability of TMPA-3B42V7 precipitation dataset in wavelet-support vector machine approach for suspended sediment load prediction. Journal of Hydrology, 550(1): 103-117.

Kalra, A., Miller, W., Lamb, K., Ahmad, S. and Piechota, T. 2013. Using large-scale climatic patterns for improving long lead time streamflow forecasts for gunnison and san juan river basins. Hydrological Processes, 27(11): 1543-1559.

Kim, S., Kisi, O., Seo, Y., Singh, V. P. and Lee, C. J. 2017. Assessment of rainfall aggregation and disaggregation using data-driven models and wavelet decomposition. Hydrology Research, 48 (1): 99-116.

Kim, S. and Singh, V. 2015. Spatial disaggregation of areal rainfall using two different artificial neural networks models. Water Journal, 7(6): 2707-2727.

Knoesen, D. and Smithers, J. 2009. The development and assessment of a daily rainfall disaggregation model for South Africa. Hydrological Sciences Journal, 54(2): 217-233.

Koutsoyiannis, D. 2001. Coupling stochastic models of different timescales. Water Resources Research, 37(2): 379-391.

Koutsoyiannis, D. and Onof, C. 2001. Rainfall disaggregation using adjusting procedures on a poisson cluster model. Journal of Hydrology, 246(1): 109-122.

Nourani, V. and Andalib, G. 2015. Wavelet Based Artificial Intelligence Approaches for Prediction of Hydrological Time Series. Artificial Life and Computational Intelligence, 1(1): 422-435.

Nourani, V., Baghanam, A.H., Adamowski, J. and Kisi, O. 2014. Applications of hybrid wavelet–artificial intelligence models in hydrology: A review. Journal of Hydrology, 514(1): 358-377.

Nourani, V., Komasi, M. and Mano, A. 2009. A multivariate ann-wavelet approach for rainfall–runoff modeling. Water Resources Management, 23(14): 2877-2894.

Olsson, J. and Berndtsson, R. 1998. Temporal rainfall disaggregation based on scaling properties. Water Science and Technology, 37(11): 73-79.

Ormsbee, L. 1989. Rainfall disaggregation model for continuous hydrologic modeling. Journal of Hydraulic Engineering, 115(4): 507-525.

Polikar, R. 1996. Fundamental concept & an overview of the wavelet theory. The Wavelet Tutorial Part I, Rowan University, College of Engineering Web Servers, 15(1): 1-15.

Raje, D. and Mujumdar, P. 2011. A comparison of three methods for downscaling daily precipitation in the punjab region. Hydrological Processes, 25(23): 3575-3589.

Rashid, M., Beecham, S. and Chowdhury, R. 2015. Statistical downscaling of rainfall: A non-stationary and multi-resolution approach. Theoretical and Applied Climatology, 124(3): 919-933.

Sharma, A. and Goyal, M. 2016. A comparison of three soft computing techniques, bayesian regression, support vector regression, and wavelet regression, for monthly rainfall forecast. Journal of Intelligent Systems, 26(4): 641-655.

Shimizukawa, J., Chen, C.Y., Iba, K., Hida, Y., Yokoyama, R., Tanaka, K. and Yabe, K. 2009. Multi-regression model for peak load forecast in demand side like university campus. In: The International Conference on Electrical Engineering, China, 1-6.

Sifuzzaman, M., Islam, M.R. and Ali, M.Z. 2009. Application of wavelet transform and its advantages compared to fourier transform. Journal of Physical Sciences, 13(1): 121-134.

www.SID.ir

Archive of SID

Sivakumar, B., Sorooshian, S., Gupta, H.V. and Gao, X. 2001. A chaotic approach to rainfall disaggregation. Water Resources Research, 37(1): 61-72.

Socolofsky, S., Adams, E. and Entekhabi, D. 2001. Disaggregation of daily rainfall for continuous watershed modeling. Journal of Hydrologic Engineering, 6(4): 300-309.

Suykens, J. A. K., Van Gestel, T., De Brabanter, J., De Moor, B. and Vandewalle, J. 2002. Least squares support vector machines. World Scientific, Singapore, 310pp.

Swinscow, T.D.V. and Campbell, M.J. 1997. Statistics at square one, nineth edition. BMJ Publishing Group, University of Southampton, 140pp.

Tripathi, S., Srinivas, V. and Nanjundiah, R. 2006. Downscaling of precipitation for climate change scenarios: A support vector machine approach. Journal of Hydrology, 330(3): 621-640.

Yang, H.H., Vuuren, S.V., Sharma, S. and Hermansky, H. 2000. Relevance of time-frequency features for phonetic and speaker-channel classification. Speech Communication, 31(1): 35-50.

Zhang, J., Murch, R., Ross, M., Ganguly, A. and Nachabe, M. 2008. Evaluation of statistical rainfall disaggregation methods using rain-gauge information for west-central florida. Journal of Hydrologic Engineering, 13(12): 1158-1169.

www.SID.ir

Archive of SID

ISSN 2251-7480

Evaluation of wavelet – least square support vector machine hybrid model to rainfall time series spatiotemporal disaggregation

Nima Farboudfam[1], Vahid Nourani[2]* and Banak Aminnejad[3]

1) Ph.D. Student, Department of Water Resources Engineering, Engineering Faculty, Roudehen Branch, Islamic Azad University, Roudehen, Iran.

2*) Professor, Department of Water Resources Engineering, Civil Engineering Faculty, Tabriz University and Professor, Department of Water Resources Engineering, Engineering Faculty, Roudehen Branch, Islamic Azad University, Roudehen, Iran..

3) Assistant Professor, Department of Water Resources Engineering, Engineering Faculty, Roudehen Branch, Islamic Azad University, Roudehen, Iran

* Corresponding Author Email: v.nourani2018@gmail.com

Received: 12-01-2018 Accepted: 24-02-2019

Abstract

The need to simulate rainfall time series at different scales for engineering purposes on the one hand and lack of recording such parameters in small scales because of administrative and economic problems, on the other hand, disaggregation of rainfall time series to the desired scale is an essential topic. In this study, for disaggregating the Tabriz and Sahand rain gauges time series, according to nonlinear characteristics of time scales, wavelet- Least Square Support Vector Machine (WLSSVM) hybrid model is proposed and daily data of four rain gauges and monthly data of six rain gauges from Urmia Lake Basin for ten years were decomposed with wavelet transform and then by using mutual information and correlation coefficient criteria, the subseries were ranked and superior subseries were used as input data of Least Square Support Vector Machine (LSSVM) model for disaggregating the Tabriz and Sahand rain gauges monthly rainfall time series to the daily time series. Results obtained from the WLSSVM disaggregation model were compared with the results of LSSVM and traditional multiple linear regression models. The results of WLSSVM model to LSSVM and multiple linear regression models at validation stage in the optimized case for Tabriz rain gauge were increased 10% and 37.5% and in the optimized case for Sahand rain gauge were increased 24.5% and 46.7% respectively. It was concluded that hybrid WLSSVM model has a higher accuracy than two other methods and can be considered as an accurate disaggregation model to disaggregate the rainfall time series.

Keywords: Disaggregation؛ Hybrid Model؛ Least Square Support Vector Machine؛ Rainfall Time series؛ Wavelet Transform

www.SID.ir