

مقاله علمی-پژوهشی

بررسی پارامترهای تاثیرگذار در مدل‌سازی بارش ماهانه با استفاده از مدل‌های تلفیقی هوشمند

مبتنی بر تجزیه سری زمانی

کریم امینی نیا^{۱*}، سید مهدی ثاقبیان^۲

تاریخ دریافت: ۱۴۰۱/۰۲/۲۵ تاریخ پذیرش: ۱۴۰۱/۰۳/۲۵

چکیده

پیش‌بینی بارش در بسیاری از جنبه‌های مختلف مدیریت حوضه‌ها نظیر سیستم‌های هشدار سیل و خشک‌سالی اهمیت دارد. تغییرات زمانی و مکانی بارش موجب دشوار شدن پیش‌بینی بارش می‌شود. در تحقیق حاضر، پیش‌بینی بارش ماهانه ایستگاه‌های ارومیه و ماکو با استفاده از روش‌های هوشمند مبتنی بر کرنل و بر پایه روش تجزیه‌ی یکپارچه مد تجزیه‌ی (EEMD) و تبدیل موجک (DWT) مورد بررسی قرار گرفته است. بدین منظور، مدل‌های متفاوتی بر اساس شاخص‌های پیوند از دور و عناصر اقلیمی شامل بارش، رطوبت و دمای ماه‌های پیشین تعریف شد و تأثیر این پارامترها در دقت مدل‌سازی مورد بررسی قرار گرفت. نتایج به‌دست آمده کارایی بالای روش‌های ترکیبی به‌کار رفته را در مدل‌سازی بارش ماهانه نشان داد. ملاحظه گردید که در پیش‌بینی بارش ماهانه، استفاده هم‌زمان از شاخص‌های پیوند از دور و عناصر اقلیمی مربوط به ماه‌های گذشته موجب بهبود دقت مدل‌ها تا حدود ۳۵ درصد می‌شود. بررسی روش‌های تجزیه یکپارچه مد تجزیه‌ی و تبدیل موجک گسسته نشان داد که تجزیه سری زمانی بر اساس تبدیل موجک منجر به نتایج دقیق‌تری می‌گردد. نتایج آنالیز حساسیت نشان داد که پارامتر بارش با تأخیر سه ماهه تأثیرگذارترین پارامتر در مدل‌سازی بارش ماهانه می‌باشد.

واژه‌های کلیدی: الگوهای پیوند از دور، تجزیه مد تجزیه‌ی، عناصر اقلیمی، مدل‌سازی بارش

مقدمه

مستقل از مدل‌های دینامیکی سیستم هستند در جستجوی راه‌هایی برای پیشرفت و پیش‌بینی پارامترهای مهم هواشناسی هستند. در این بین سری زمانی به مجموعه‌ای از دیده‌بانی‌ها با مقادیر ثبت شده از یک متغیر گفته می‌شود که برحسب زمان مرتب شده است. هدف از سری زمانی، تعیین قانونمندی و شناسایی رفتار آن جهت پیش‌بینی آینده است. پیش‌بینی بارش به‌ویژه در نواحی مختلف سرزمین کم‌آبی نظیر ایران که منابع آب آن متکی به بارش و با تقاضای روزافزون جمعیت روبه رشد همراه است، از اهمیت ویژه‌ای برخوردار است. در این راستا، برنامه‌ریزی‌های صحیح و کسب اطلاعات از طریق انجام مطالعات علمی و پژوهش‌های دقیق ضروری است. فلاحی و همکاران (۱۳۹۰) از مدل رگرسیون درختی (M5) به‌منزله روشی کارآمد برای پیش‌بینی بارندگی به‌منظور کنترل سیل در سه ایستگاه اراک، انزلی و قم استفاده کردند که نتایج بیانگر توانایی و دقت بالای این مدل بود. بختیاری و همکاران (۱۳۹۲) احتمالات بارش روزانه با استفاده از مدل زنجیره مارکف را در اقلیم‌های مختلف ایران بررسی کردند. ایشان نشان دادند که احتمال عدم بارش در ایستگاه‌های مطالعاتی با اقلیم معتدل و مرطوب کمتر از اقلیم‌های خشک و نیمه‌خشک است. ستاری و همکاران (۱۳۹۳) بارش ماهانه را با

پیش‌بینی بارش، به‌عنوان یکی از مهم‌ترین متغیرهای اقلیمی در حوزه مدیریت منابع آب از اهمیت ویژه‌ای برخوردار است. از طرفی، وجود رابطه‌های غیرخطی پیچیده در معادلات حاکم، مدل‌سازی بارش را امری مشکل‌نموده است. در صورتی که برآوردهای قابل‌اعتمادی از وضعیت اقلیمی چند ماه یا چند سال آینده در دسترس باشد، برنامه‌ریزان، مدیران و اقشار مختلف اجتماعی می‌توانند ضمن آمادگی جهت مقابله با حوادث نامطلوب، نسبت به بهره‌برداری بهینه از منابع آب و خاک و نیروی انسانی اقدام نمایند. در این راستا پیش‌بینی بارش برای کشورهای در حال توسعه که اقتصاد آن‌ها مبتنی بر کشاورزی است، خیلی مهم است. به‌طور مشخص، اقلیم و بارش پدیده‌های غیرخطی در طبیعت هستند. امروزه محققین، با ابداع و پیشرفت علمی چون روش‌های هوشمند که ابزاری توانمند، انعطاف‌پذیر و

۱- استادیار گروه جغرافیا، واحد اهر، دانشگاه آزاد اسلامی، اهر - ایران

۲- استادیار گروه عمران، واحد اهر، دانشگاه آزاد اسلامی، اهر - ایران

*-نویسنده مسئول: (Email: k-amininia@iau-ahar.ac.ir)

SVM برای پیش‌بینی بارش- رواناب حوضه تریجوآلد^۱ واقع در کشور هلند استفاده کردند (Siviapragasam and Liang, 2001). روشنگر و قاسم‌پور با استفاده از روش هوشمند رگرسیون بردار پشتیبان دبی بار بستر را در لوله‌های انتقال آب و فاضلاب را بررسی کردند (Roushangar and Ghasempour, 2017). کیشی و سیمین با استفاده از روش ماشین بردار پشتیبان به پیش‌بینی بارش روزانه پرداختند (Kisi and Cimen, 2012). مارزانو و همکاران به بررسی کارایی روش شبکه عصبی در مقایسه با روش‌های رگرسیون در پیش‌بینی شدت بارش پرداختند (Marzano et al., 2006). کومار و همکاران در مقایسه روش‌های سری زمانی و شبکه عصبی مصنوعی، روش شبکه عصبی مصنوعی را با توجه به سری زمانی پیچیده بارش-های هند مناسب‌تر دانستند (Kumar et al., 2013). نایاک و همکاران با مروری بر روش‌های پیش‌بینی بارش در هند، شبکه عصبی مصنوعی را در مقایسه با روش‌های پیش‌بینی عددی مناسب‌تر معرفی کردند (Nayak et al., 2013).

بررسی منابع صورت گرفته، نشان از کاربرد بالای روش‌های هوشمند مانند شبکه عصبی مصنوعی در زمینه پیش‌بینی بارش است. این مدل‌ها به‌طور معناداری، خطاهای کوچک‌تری را در پیش‌بینی نسبت به مدل‌های آماری نشان می‌دهند و دارای سرعت پیش‌بینی بالاتری هستند. با این حال، مطالعات چندانی در زمینه کاربرد روش‌های مبتنی بر کرنل مانند رگرسیون فرآیند گاوسی و ماشین بردار پشتیبان در تخمین میزان بارش ماهانه انجام نشده است. از طرفی، در سال‌های اخیر استفاده از روش‌های تجزیه سیگنال به عنوان روشی نوین و بسیار موثر در زمینه تحلیل سیگنال‌ها و سری‌های زمانی مورد توجه قرار گرفته است. در این روش سیگنال اصلی به چندین زیرسری تجزیه شده و داده‌های حاصل از تجزیه با معادلات مناسب برازش می‌یابد و مدل بدست‌آمده برای پیش‌بینی استفاده می‌شود. از جمله این روش‌ها روش جدید تجزیه مد تجربی یکپارچه است که در روند تجزیه‌سازی آن، حجم محدودی از نویز سفید وارد سیگنال اصلی شده و با استفاده از جنبه‌های مثبت آماری نویز سفید که توزیع متوازن در حوزه فرکانس دارد، اثر نویز متناوب از روند تجزیه حذف می‌شود. بنابراین، با توجه به اهمیت دقت پیش‌بینی بارش در هیدرولوژی و پیش‌بینی سیلاب، در تحقیق کنونی کارایی روش‌های ماشین یادگیری سریع مبتنی بر هسته غیرخطی (KELM) و ماشین بردار پشتیبان (SVM) به عنوان روش‌های نوین داده‌کاوی در تخمین بارش ماهانه ایستگاه سنوپتیک ارومیه مورد بررسی قرار گرفت. مدل-های مختلفی بر اساس داده‌های مربوط به عناصر اقلیمی مانند بارش، میانگین دما و رطوبت نسبی ماهانه با تاخیر زمانی یک تا چهار ماهه و

استفاده از ANN و مدل M5 در ایستگاه اهر پیش‌بینی کردند. نتایج نشان داد که هر دو روش نتایج نسبتاً دقیقی را برای پیش‌بینی ارائه می‌کنند. اما از آنجاکه مدل M5 روابط خطی ساده‌ای را در اختیار کاربر می‌گذارد، این روش کاربردی‌تر است. سلطانی و همکاران بهترین مدل سری زمانی در پیش‌بینی بارندگی سالانه ایستگاه‌های منتخب استان آذربایجان غربی را بررسی کردند و به این نتیجه رسیدند که مدل $ARIMA(1,0,0)$ ، $ARIMA(0,1,1)$ و مدل $ARIMA(0,1,1)$ به ترتیب در ایستگاه‌های ارومیه، ماکو و مهاباد به‌عنوان مدلی مناسب جهت پیش‌بینی بارندگی سالانه هستند (Soltani et al., 2017). سامانتارای و همکاران از روش‌های مختلف شبکه عصبی مصنوعی برای مطالعه پیش‌بینی بارندگی در ناحیه بولانگیر، هند با استفاده از داده‌های ۴۸ سال گذشته استفاده کردند. ترکیب‌های مختلفی از مقادیر بارندگی گذشته به‌عنوان ورودی‌های پیش‌بینی برای ارزیابی کارایی تقریب مدل هوشمند تشکیل شد. تکنیک‌های SVM، RNN و ANFIS برای تجزیه و تحلیل بارندگی استفاده شد. نتایج نشان داد که SVM با دقت بالاتر عملکرد بهتری را ارائه می‌دهد (Samantaray et al., 2020). چانگ و همکاران جهت پیش‌بینی بارش بر روی حوضه رودخانه لنگات از ادغام تبدیل موجک (WT) و شبکه عصبی کانولوشنال (CNN) استفاده کردند. روش پیشنهادی شامل استفاده از CNN برای استخراج ویژگی برای یادگیری کارآمد از مجموعه داده خام بارش بود. پیش‌بینی بارندگی ماهانه با CNN و پیش‌بینی بارش روزانه با CNN نشان داد که مدل پیشنهادی می‌تواند به‌عنوان یک الگوریتم کارآمد و قوی برای پیش‌بینی سری‌های زمانی بارش تصادفی عمل کند (Chong et al., 2020).

با وجود اینکه مدل‌های هیدرولوژیکی مختلفی جهت مدل‌سازی بارندگی پیشنهاد شده‌اند، اما با توجه به ماهیت غیرخطی و ویژگی‌های متغیر زمانی و مکانی در سیستم گردش آبی، هیچ‌یک از مدل‌های آماری و مفهومی پیشنهاد شده به‌منظور الگوسازی دقیق بارش نتوانسته‌اند به‌عنوان یک مدل برتر و توانا شناخته شوند و عدم قطعیت در پیش‌بینی‌ها از اعتبار برآوردها از طریق مدل‌های مختلف می‌کاهد. امروزه شبکه‌های غیرخطی به‌عنوان یکی از سیستم‌های هوشمند در پیش‌بینی چنین پدیده‌های پیچیده بسیار مورد استفاده قرار می‌گیرند. این روش‌ها که اکثراً الهام گرفته از پدیده‌های موجود در طبیعت هستند و می‌توانند برای پیش‌بینی هرگونه متغیری که در آن: رابطه بین متغیرهای مربوطه به خوبی درک نمی‌شود، پیدا کردن اندازه و شکل راه حل نهایی دشوار است، و روش‌های تحلیلی ریاضی رایج قادر به حل آن‌ها نبوده و یا بسیار زمان‌بر هستند، استفاده شوند. به‌عنوان ابزاری توانمند در حل مسائل پیچیده به شمار می‌آیند و معمولاً دارای جواب بهتری نسبت به روش‌های کلاسیک هستند (ASCE, 2000). به‌عنوان نمونه سیواپراگاسام و لیئونگ از مدل

1- Trygvevalde

2- Kernel-based Extreme Learning Machine

3- Support Vector Machine

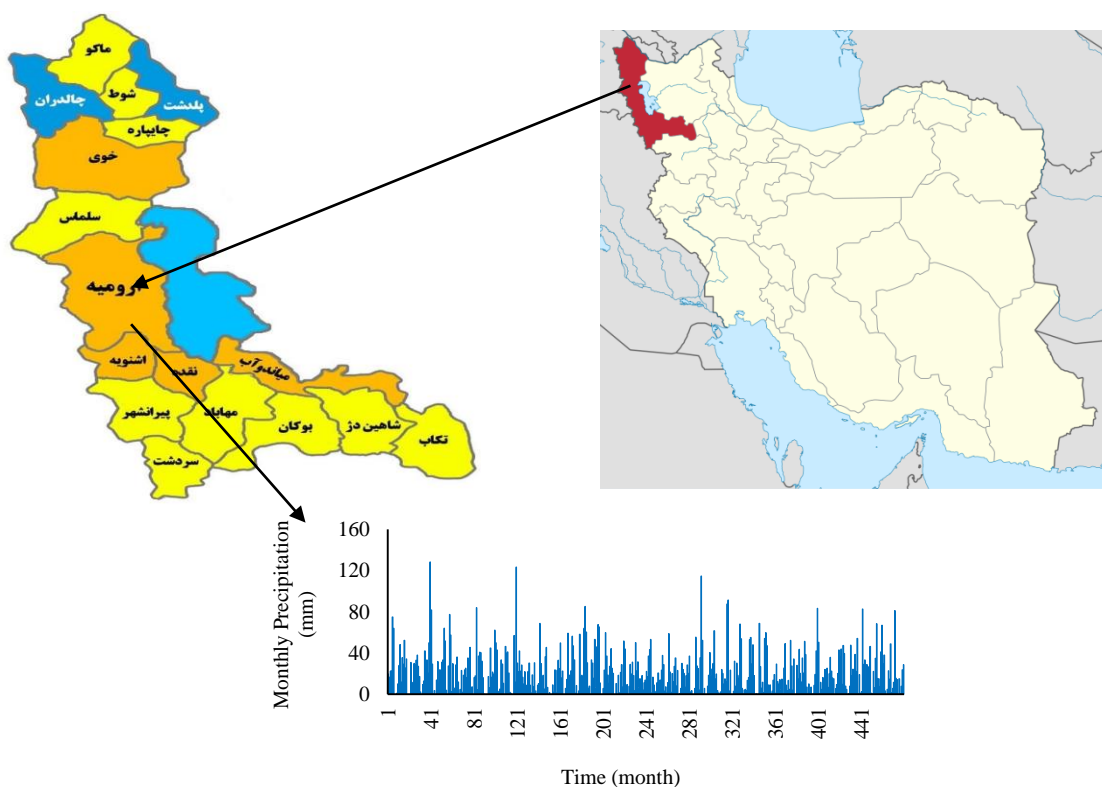
جغرافیایی در ۲۰ کیلومتری دریاچه ارومیه قرار دارد. ارتفاع این شهر ۱۳۳۲ متر از آب های آزاد است. ارومیه بین دریاچه ارومیه و دیواره کوه های غرب استان واقع شده است. آب و هوای ارومیه با متوسط درجه حرارت ۹/۸ درجه سانتی گراد در تابستان تقریباً گرم و در زمستان سرد است. دوره بارندگی از اواخر مهر و ابتدای آبان شروع شده و تا خرداد ادامه دارد. میانگین دوره بارش ۹۳ روزه در این شهر ۳۶۰ میلی متر و میانگین بارش بلندمدت آن ۲/۲۳۸ میلی متر است. در شکل ۱ منطقه مطالعاتی و اطلاعات آماری سری زمانی بارندگی ماهانه ایستگاه سینوپتیک ارومیه در بازه زمانی ۱۹۸۰-۲۰۱۸ نمایش داده شده است.

الگوهای پیوند از دور در بازه زمانی (۱۹۸۰-۲۰۱۸) تعریف گردیده و دو حالت استفاده از داده های اصلی (بدون تجزیه) و استفاده از داده های تجزیه شده با استفاده از تبدیل موجک گسسته و روش جدید تجزیه مد تجزیه کامل ارزیابی شد. در نهایت با استفاده از آنالیز حساسیت موثرین پارامتر در مدلسازی بارش ماهانه تعیین شد.

مواد و روش ها

موقعیت و ویژگی های منطقه مورد مطالعه

منطقه مورد مطالعه شهر ارومیه است که با مختصات جغرافیایی ۳۷ درجه و ۳۲ دقیقه طول شرقی و ۴۵ درجه و ۲ دقیقه عرض شمالی از نصف النهار گرینویچ واقع شده است. ارومیه با داشتن موقعیت ممتاز



شکل ۱- موقعیت منطقه مطالعاتی و متوسط بارندگی ماهانه در ایستگاه سینوپتیک ارومیه (۱۹۸۰-۲۰۱۸)

کوتاه مدت بودن را دارد. $\psi(x)$ تابعی موجک است اگر و فقط اگر تبدیل فوریه آن $\Psi(\omega)$ شرط زیر را ارضا کند (Lau and Weng, 1995).

$$\int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|}{|\omega|^2} d\omega < +\infty \quad (1)$$

این شرط با عنوان شرط پذیرفتگی برای موجک $\psi(x)$ شناخته می شود $\psi(x)$ تابع موجک مادر است (رابطه ۲) که ضرایب a و b در

تبدیل موجک

موجک (Wavelet) دسته ای از توابع ریاضی است که برای تجزیه سیگنال پیوسته به مؤلفه های فرکانسی آن بکار می رود که سطح تفکیک هر مؤلفه برابر با مقیاس آن است. تبدیل موجک تجزیه یک تابع بر مبنای توابع موجک است. موجک ها (که به عنوان موجک های مادر شناخته می شوند) نمونه های انتقال یافته و مقیاس شده یک تابع (موجک مادر) با طول متناهی و نوسانی شدیداً میرا هستند. تابع موجک، تابعی است که دو ویژگی مهم نوسانی بودن و

آن نقش انتقال و مقیاس سیگنال مورد نظر را دارند.

$$\Psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (2)$$

توابع موجک دارای انواع بسیاری هستند که مهم‌ترین و پرکاربردترین آن‌ها شامل تابع موجک Mexican hat، Meyer، Haar، Symlet، Morlet، Coiflet، Daubechies است. تبدیل موجک پیوسته به عنوان روش جایگزین تبدیل فوریه زمان-کوتاه ارائه گردید و هدف آن، فائق آمدن بر مشکلات مربوط به تفکیک پذیری در تبدیل فوریه زمان-کوتاه است. در عمل اکثر سری‌های زمانی هیدرولوژیکی به صورت گسسته در زمان‌هایی مشخص اندازه‌گیری می‌شوند. بنابراین کاربرد موجک گسسته می‌تواند تناسب بهتری با این نوع سری زمانی داشته باشد. تبدیل موجک گسسته در عمل همان تبدیل موجک پیوسته است (نسخه گسسته شده آن) که در آن یک سری موجک از تبدیل موجک پیوسته نمونه گرفته است و سیستم‌های کامپیوتری قابلیت محاسبه آن را با هزینه زمانی کمتر دارا است. اصول تبدیل موجک گسسته به روشی تحت عنوان کدینگ زیرباند بر می‌گردد که در سال ۱۹۷۶ سنگ بنای اولیه آن گذارده شد و در آن نوعی توصیف زمان-مقیاس از سیگنال گسسته با استفاده از فیلترهای دیجیتال ارائه می‌گردد. تبدیل موجک، حاصل شباهت سنجی (کورولیشن) بین محتوای فرکانسی (مقیاسی) سیگنال و تابع موجک در مقیاس‌های مختلف است (Adamowski et al., 2009). تجزیه کردن سری‌های زمانی با استفاده از فیلترهای مشخص (موجک و توابع مقیاس‌گذاری شده) دو نوع زیرسری را به وجود می‌آورند؛ بردارهای تقریبی یا باقی مانده و بردارهای جزییات (Chou, 2011). این ضرایب حاصل از در هم پیچیدن سری زمانی اصلی با فیلتر پایین-گذر و فیلتر بالا-گذر می‌باشد. فیلتر پایین-گذر همان تابع مقیاس‌گذاری است و فیلتر بالاگذر همان تابع موجک است. تابع تبدیل موجک قابلیت تجزیه سری زمانی به چندین زیرسری زمانی با مقیاس‌های مختلف را دارد و با بررسی زیرسری‌های زمانی منتجه از سری زمانی کلی، رفتار کوچک مقیاس و بزرگ مقیاس یک فرآیند هیدرولوژیکی را آنالیز می‌کند. روابطی که بیشتر برای تخمین سطح تجزیه مورد استفاده قرار می‌گیرد به صورت زیر است که در آن L سطح تجزیه، N تعداد نقاط داده در سری زمانی و p برابر با زمان محو شونده برای موجک مادر انتخاب شده است.

$$L = \text{int}[\log(N)] \quad , \quad L = \frac{\log\left(\frac{N}{2^{p-1}}\right)}{\log 2} \quad (3)$$

۳-۲- روش تجزیه مد تجربی EMD^۱

روش تجزیه مد تجربی روشی برای تجزیه سیگنال‌های گوناگون است که این عمل در فرآیندی تحت عنوان غربال کردن صورت می‌گیرد. طی این فرآیند سیگنال اصلی به تعدادی مولفه با محتوای بسامدی متفاوت تجزیه می‌شود. روش EMD طبق رابطه (۴) سیگنال اصلی $x(n)$ را به تعدادی مد ذاتی (IMF^T) تجزیه می‌کند (Amirat et al., 2018).

$$x(n) = \sum_{i=1}^n c_i(x) + r_n(x) \quad (4)$$

که $r_n(x)$ همان مولفه باقی مانده، بعد از n تعداد IMFها و $c_i(x)$ است. یک داده ممکن است در یک زمان دارای چندین مد ذاتی باشد. این مدهای نوسانی، تابع‌های مد ذاتی (IMF) نام دارند و دارای دو شرایط هستند: ۱) در کل داده‌ها، تعداد نقاط فرین (اکسترمم) و نقاط صفر با هم برابر و یا حداکثر دارای یک واحد اختلاف هستند، ۲) هر نقطه میانگین پوش برازش داده شده بر نقاط بیشینه محلی و پوش برازش داده شده بر نقاط کمینه محلی باید صفر باشد. به علت وجود تناوب و نويز در سیگنال‌ها، در برخی موارد به علت اختلاط مدها؛ توزیع حوزه زمان-فرکانس دچار انقطاع می‌شود و عملکرد EMD دچار نقض می‌گردد (چون میانگین IMFها حالتی نامشخص به خود می‌گیرد). جهت رفع این ایراد، وو و هوانگ (Wu and Huang, 2004) روشی متفاوت با عنوان روش تجزیه مد تجربی یکپارچه ($EEMD^2$) را پیشنهاد کردند. در روند تجزیه‌سازی آن، حجم محدودی از نويز سفید وارد سیگنال اصلی می‌شود. با استفاده از جنبه‌های مثبت آماری نويز سفید که توزیع متوازن در حوزه فرکانس دارد، اثر نويز متناوب از روند تجزیه حذف می‌شود. در روش تجزیه مد تجربی یکپارچه کامل ($CEEMD^3$) نوفه سفید بصورت دوتایی (یکی مثبت و دیگری منفی) به داده اصلی اضافه می‌شود تا دو سری IMF مجموع ساخته شود. بنابراین ترکیبی متشکل از داده اصلی و نوفه‌ی اضافی داریم که مجموع IMFها برابر سیگنال اصلی می‌شود و نوفه‌ی اضافی دامنه‌ی موجی از مرتبه‌ی pm خواهد بود.

روش‌های هوشمند مبتنی بر کرنل

یکی از روش‌های متداول برای حل مسائل غیرخطی استفاده از توابع کرنل است؛ این توابع بر اساس ضرب داخلی داده‌های مفروض تعریف می‌شود و به‌عنوان ابزاری توانمند در حل مسائل پیچیده به شمار می‌آیند و معمولاً دارای جواب بهتری نسبت به روش‌های کلاسیک هستند. طراحی روش‌های رگرسیون مبتنی بر کرنل مانند ماشین بردار پشتیبان و فرآیند گاوسی نیز شامل استفاده از مفهوم تابع

2- Intrinsic Mode Functions

3- Ensemble Empirical Mode Decomposition

4- Complementary Ensemble Empirical Mode Decomposition

1- Empirical Mode Decomposition

تیین (DC) که همبستگی خطی بین مقادیر پیش بینی شده و واقعی را نشان می دهد و یکی از معیارهای مهم ارزیابی قابلیت یک مدل است و ریشه میانگین مربعات خطاها (RMSE) که بیانگر میزان انحراف بین مقادیر پیش بینی شده است. هر چه مقدار R و DC برای یک مدل به یک نزدیک تر و مقدار RMSE کوچکتر باشد به معنی مطلوب بودن آن مدل است. روابط این پارامترهای آماری به صورت زیر است:

$$DC = 1 - \frac{\sum_{i=1}^N (I_o - I_p)^2}{\sum_{i=1}^N (I_o - \bar{I}_p)^2}, \quad R = \frac{\sum_{i=1}^N (I_o - \bar{I}_o) \times (I_p - \bar{I}_p)}{\sqrt{\sum_{i=1}^N (I_o - \bar{I}_o)^2 \times \sum_{i=1}^N (I_p - \bar{I}_p)^2}} \quad (5)$$

در روابط بالا I_o : مقدار اندازه گیری شده، \bar{I}_o : متوسط مقدار اندازه گیری شده، I_p : مقدار پیش بینی شده، \bar{I}_p : متوسط مقدار پیش بینی شده و N تعداد داده ها است. از آنجا که وارد کردن داده ها به صورت خام باعث کاهش سرعت و دقت شبکه می شود لذا نرمالیزه کردن داده ها کمک شایانی به آموزش بهتر و سریع تر مدل می کند. در تحقیق کنونی داده ها مطابق رابطه (۶) نرمالیزه شدند که در آن x_{max} ، x_{min} به ترتیب حداقل و حداکثر مقدار داده های مشاهده شده و x_n داده نرمال شده است.

$$x_n = 0.05 + 0.95 \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right) \quad (6)$$

نتایج و بحث

تعریف مدل های ورودی

با توجه به اینکه انتخاب مدل های ورودی در سامانه های هوشمند می تواند بر روی دقت جواب های حاصل از تحلیل تأثیر گذارد، لذا سعی گردید در مدل سازی پارامترهای مناسب و مؤثر در تعیین میزان بارش ماهانه انتخاب گردد. بررسی منابع متعدد نشان می دهد که ارتباط بین الگوهای پیوند از دور با بارش و دما در مناطق وسیعی از کره زمین به اثبات رسیده است، الگوهای پیوند از دور منعکس کننده ی تغییرات در مقیاس وسیع در امواج اتمسفری و موقعیت رودبادهای بوده و بر درجه حرارت، بارش، مسیر توفان ها و موقعیت و شدت رودباد بر روی مناطق وسیع تأثیر می گذارند. اداره ملی اقیانوس و اتمسفر ایالات متحده امریکا الگوهای ارتباط از دور اساسی را جهت بررسی سهم و درک نقش آن ها در سیستم اقلیم سیاره ای شناسایی نموده و در قالب شاخص هایی ارائه داده است (NOAA, 2009). از انواع این شاخص ها می توان شاخص مد جنوبی اقیانوس اطلس (AMM)، آتلانتیک شرقی (EA)، نوسانات اطلس شمالی (NAO)، شاخص چند متغیره انسو (MEI)، شاخص های دمای سطح اقیانوس آرام (Nino) و شاخص بارش موسمی ناحیه جنوب غربی (SW

کرنل است. مبنای کاری SVM دسته بندی خطی داده ها است و در تقسیم خطی داده ها سعی می کند خطی را انتخاب کند که حاشیه اطمینان بیشتری داشته باشد. حل معادله پیدا کردن خط بهینه برای داده ها به وسیله روش های QP، که روش های شناخته شده ای در حل مسائل محدودیت دار هستند، صورت می گیرد. قبل از تقسیم خطی برای اینکه ماشین بتواند داده های با پیچیدگی بالا را دسته بندی کند داده ها را به وسیله تابع کرنل به فضای با ابعاد خیلی بالاتر انتقال داده می شود. در واقع، با یک تبدیل غیرخطی از فضای ورودی به فضای خصیصه با ابعاد بیشتر (حتی نامتناهی) می توان مسائل را به صورت خطی تفکیک پذیر ساخت. با تبدیل نمونه ها از فضای ورودی به فضای ویژگی، تفکیک کننده غیرخطی به حالت خطی تبدیل خواهد شد. از مهمترین توابع کرنل می توان به کرنل چند جمله ای ساده، چند جمله ای نرمال شده، تابع شعاع محور و تابع کرنل سیگموئید اشاره کرد. این روش ها می توانند برای پیش بینی هرگونه متغیری که در آن رابطه بین متغیرهای مربوطه به خوبی درک نمی شود، پیدا کردن اندازه و شکل راه حل نهایی دشوار است، و روش های تحلیلی ریاضی رایج قادر به حل آن ها نبوده و یا بسیار زمان بر هستند استفاده شوند.

ماشین یادگیری سریع (Extreme learning machine) یک شبکه عصبی پیشخور است که برای طبقه بندی آماری، تحلیل رگرسیون، خوشه بندی، اسپارس تقریبی، مقایسه و آموزش ویژگی های یک تک لایه یا چند لایه در گره مخفی می باشد به صورتی که نیازی به پارامترهای گره های پنهان نباشد. این گره های پنهان می توانند به صورت تصادفی تخصیص داده شوند و هرگز بروزسانی وزن ها برایش صورت نگیرد. در اغلب موارد، وزن خروجی گره های پنهان معمولاً در یک مرحله گنجانده می شود که اساساً به یادگیری یک مدل به صورت خطی می پردازد. روش ماشین یادگیری سریع مبتنی بر هسته غیر خطی (KELM) یکی از روش های قدرتمند ارائه شده در حوزه یادگیری ماشین است. اساس مرحله یادگیری در این روش مبتنی بر ساخت ماتریس هسته نمونه های برچسب دار و محاسبه معکوس آن می باشد. از این رو، بکارگیری این روش در محیط های مقیاس بزرگ امکان پذیر است. کاربرد مدل ماشین یادگیری افراطی آسان است و هیچ پارامتری به جز معماری شبکه از پیش تعریف شده نیاز ندارد. بنابراین این مدل از عوارضی همچون نرخ یادگیری، دوره یادگیری و کمینه محلی که الگوریتم های بر اساس شیب با آنها مواجه هستند، به دور هستند (Deo et al., 2016).

معیارهای ارزیابی

به منظور ارزیابی کارایی روش استفاده شده در تخمین میزان بارش ماهانه، از سه پارامتر آماری استفاده گردید که عبارتند از: ضریب همبستگی بین مقادیر مشاهداتی و محاسباتی (R)، ضریب

تبدیل موجک و تجزیه یکپارچه مد تجزیه به زیر سری‌هایی تجزیه گردید و سپس این زیر سری‌ها به عنوان ورودی مدل‌های مبتنی بر کرنل استفاده شد. جدول ۱ بیان‌گر مدل‌های تعریف شده است. لازم به توضیح است که داده‌های سری زمانی بارش از سال ۱۹۸۰-۲۰۰۹ به عنوان داده‌های آموزش و بقیه داده‌ها به عنوان داده‌های آزمون مورد استفاده قرار گرفتند.

(monsoon) را نام برد. شاخص‌های پیوند از دور و عناصر اقلیمی مانند بارش (P)، دما (T) و رطوبت نسبی (R) در فرآیند پیش‌بینی می‌توانند به عنوان متغیرهای مستقل به کار روند؛ بنابراین به عنوان پیش‌بینی در مدل ترکیبی مورد استفاده قرار گرفتند. در تحقیق کنونی دو نوع مدل-سازی در نظر گرفته شده است: مدل‌سازی بر اساس فقط داده‌های بارش و مدل‌سازی بر اساس عناصر اقلیمی و عناصر پیوند از دور. بنابراین، پس از نرمالیزه کردن داده‌ها ابتدا سری‌های زمانی توسط

جدول ۱- مدل‌های تعریف شده در تحقیق

مدلسازی بر اساس بارش		مدلسازی بر اساس عناصر اقلیمی و عناصر پیوند از دور	
مدل	پارامترهای ورودی	مدل	پارامترهای ورودی
(I)	(P_{t-1})	(I)	$(P_{t-1}, Nino3_{t-1}, T_{t-1})$
(II)	(P_{t-1}, P_{t-2})	(II)	$(P_{t-1}, P_{t-2}, P_{t-3}, T_{t-1})$
(III)	$(P_{t-1}, P_{t-2}, P_{t-3})$	(III)	$(P_{t-1}, P_{t-2}, P_{t-3}, T_{t-1}, R_{t-1})$
(IV)	$(P_{t-1}, P_{t-2}, P_{t-3}, P_{t-4})$	(IV)	$(P_{t-1}, P_{t-2}, P_{t-3}, NAO_{t-1})$
		(V)	$(P_{t-1}, P_{t-2}, P_{t-3}, MEI_{t-1})$
		(VI)	$(P_{t-1}, P_{t-2}, P_{t-3}, Nino3_{t-1})$
		(VII)	$(P_{t-1}, P_{t-2}, P_{t-3}, Nino3_t, T_{t-1}, R_{t-1})$
		(VIII)	$(NOA_{t-1}, Nino3_{t-1}, T_{t-1}, MEI_{t-1})$

بدون تجزیه آن به چندین زیرسری، چهار مدل بر اساس داده‌های ماهانه بارش مربوط به ماه‌های قبل تعریف گردید و کارایی روش‌های مبتنی بر کرنل مورد بررسی قرار گرفت. بنابراین، این مدل‌ها با روش KELM و SVM آنالیز گردید و نتایج حاصله از تحلیل مدل‌ها در جدول ۲ و شکل ۳ نشان داده شد. با توجه به نتایج به دست آمده از سه معیار ارزیابی مدل برای داده‌های آموزش و آزمون، مشاهده می‌شود که مدل (III) با پارامترهای ورودی $P_{t-1}, P_{t-2}, P_{t-3}$ که مربوط به بارش ماهانه یک، دو و سه ماه قبل است، بیشترین کارایی را داراست. مطابق با نتایج ملاحظه می‌گردد که تنها استفاده از مقادیر بارش ماه قبل نمی‌تواند منجر به نتایج دقیقی گردد و همچنین مقادیر بارش چهار ماه قبل نیز در مدل‌سازی تأثیرگذار بوده و موجب بهبود نتایج شده است.

انتخاب تابع کرنل مناسب

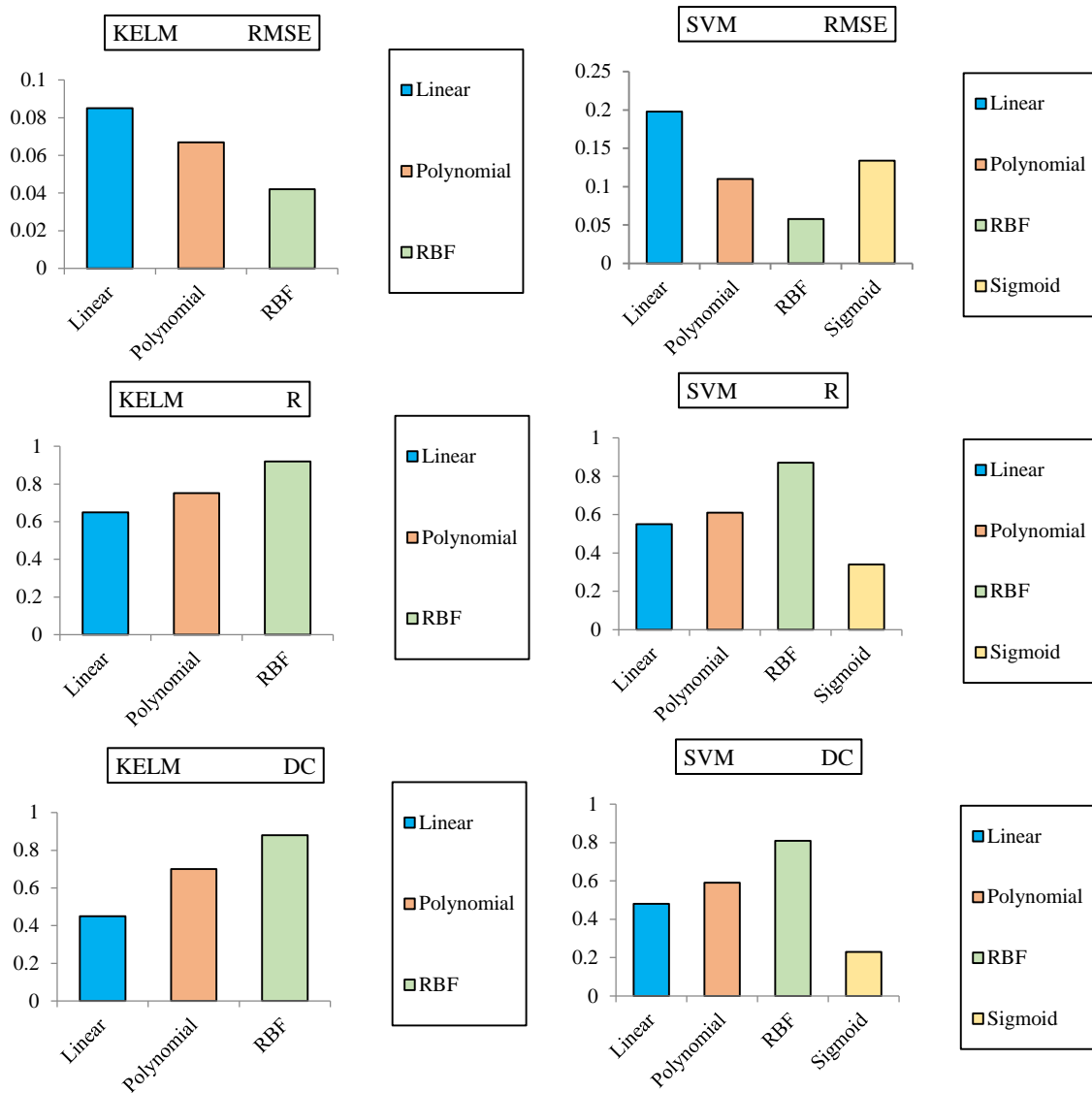
جهت انتخاب تابع کرنل مناسب در روش‌های ماشین یادگیری سریع و ماشین بردار پشتیبان، مدل (III) پس از تجزیه داده‌ها توسط تبدیل موجک با کرنل‌های مختلف مورد ارزیابی قرار گرفت. از سه معیار ارزیابی $RMSE$ ، R و DC جهت تعیین کارایی هر یک از کرنل‌ها استفاده گردید. مطابق شکل ۲، نتایج نشان داد که تابع کرنل شعاع مداری در تخمین میزان بارش ماهانه نسبت به توابع کرنل دیگر دارای ضریب همبستگی و تبیین بیشتر و خطای کمتری است. بنابراین در تجزیه و تحلیل تمامی مدل‌ها از کرنل شعاع مداری استفاده شد.

نتایج به دست آمده برای مدل‌های تعریف شده بدون تجزیه داده‌ها

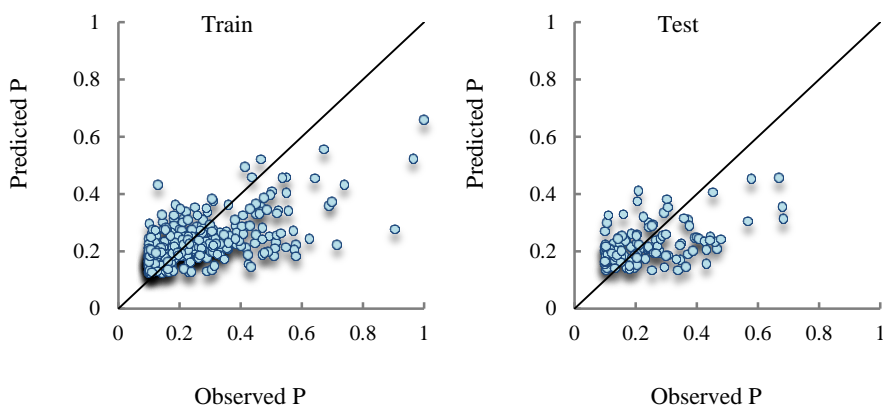
جهت تخمین میزان بارش ماهانه بر اساس سری زمانی اصلی و

جدول ۲- نتایج ارزیابی مدل‌های KELM و SVM بدون تجزیه داده‌ها

مدل		آزمون			تست		
		R	DC	RMSE	R	DC	RMSE
(I)	KELM	۰/۴۲۶	۰/۳۲۰	۰/۲۳۲	۰/۳۷۹	۰/۲۱۸	۰/۲۳۷
	SVM	۰/۴۰۴	۰/۳۱۱	۰/۲۴۵	۰/۳۶۲	۰/۲۱۲	۰/۲۵۰
(II)	KELM	۰/۵۸۲	۰/۳۷۳	۰/۲۱۸	۰/۵۳۵	۰/۳۳۷	۰/۲۲۱
	SVM	۰/۵۵۹	۰/۳۶۱	۰/۲۳۰	۰/۵۱۳	۰/۳۲۲	۰/۲۳۷
(III)	KELM	۰/۶۱۸	۰/۵۲۰	۰/۱۷۷	۰/۶۰۵	۰/۵۰۹	۰/۱۹۳
	SVM	۰/۵۸۶	۰/۵۰۴	۰/۱۸۷	۰/۵۸۰	۰/۴۹۱	۰/۲۰۶
(IV)	KELM	۰/۶۰۸	۰/۴۴۳	۰/۱۸۳	۰/۵۸۷	۰/۴۲۴	۰/۱۹۵
	SVM	۰/۵۸۳	۰/۴۲۹	۰/۱۹۳	۰/۵۶۲	۰/۴۰۶	۰/۲۰۸



شکل ۲- پارامترهای آماری مدل KELM و SVM با توابع کرنل متفاوت برای سری داده های آزمون

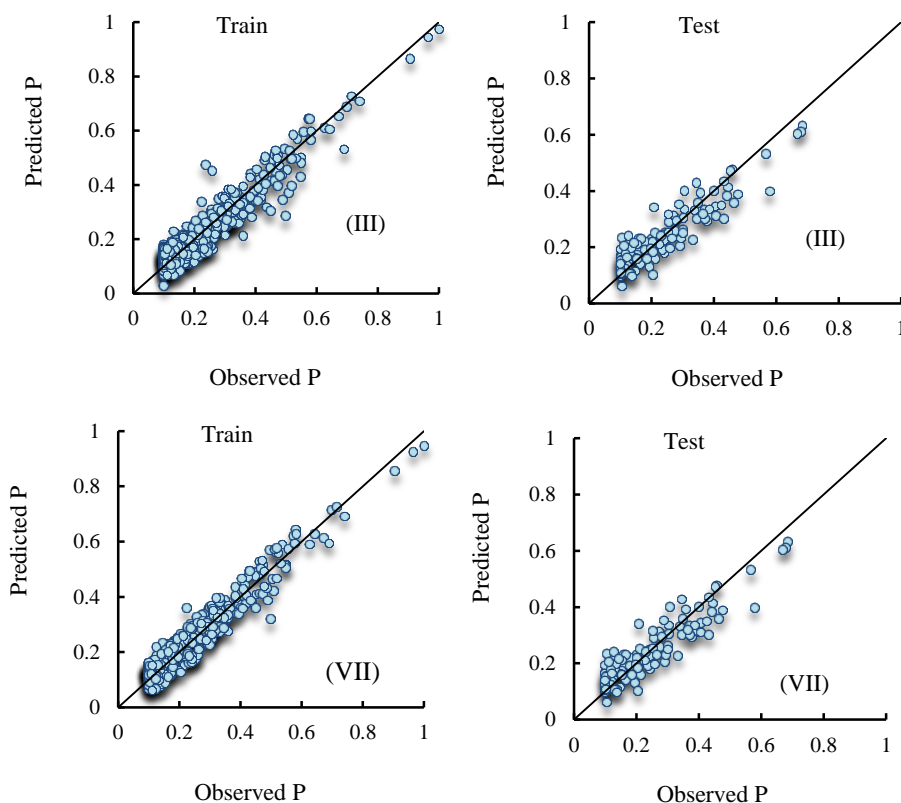


شکل ۳- نتایج مرحله آزمون و تست مدل برتر بدون تجزیه داده ها

نتایج به دست آمده برای مدل KELM بر پایه تجزیه ی یکپارچه مد تبدیل کامل و تجزیه ی یکپارچه مد

در این قسمت با استفاده از دو روش تجزیه ی یکپارچه مد تجربی و تبدیل موجک گسسته ابتدا سری های زمانی تجزیه گردید و سپس زیر سری های به دست آمده به عنوان ورودی به مدل KELM داده شد. مدل سازی نیز در دو حالت بر اساس تنها داده های بارش و بر اساس عناصر اقلیمی شامل میانگین بارش، دما و رطوبت نسبی ماهانه و الگوهای پیوند از دور شامل NOA, Nino3, MEI انجام شد. اولین قدم برای تجزیه سیگنال با تبدیل موجک، انتخاب موجک مادر است. در این تحقیق از موجک مادر db4 با دو سطح تجزیه استفاده شد. اساس روش EEMD نیز تجزیه سیگنال به IMF های مختلف و یک سیگنال باقیمانده است که در نهایت مجموعه ی این ها، همان سیگنال اصلی خواهد بود. روش تشکیل IMF ها بر اساس تفریق تابع پایه از سیگنال اصلی است. این پروسه تا وقتی ادامه می یابد که تقریباً سیگنال باقیمانده ثابت شود. در نهایت زیر سری های به دست آمده از هر دو روش، به عنوان ورودی به مدل KELM داده شد. نتایج حاصل از مدل سازی های در نظر گرفته شده در جدول ۳ و شکل ۴ ارائه شده است. با مقایسه نتایج جدول ۲ و ۳ ملاحظه می شود که تجزیه سری-

های زمانی تا حدود زیادی باعث بهبود نتایج گردیده و تاثیر تجزیه با استفاده از تبدیل موجک گسسته بیشتر از روش تجزیه ی یکپارچه مد تجربی است. مطابق با جدول ۳، در حالت مدل سازی بر اساس تنها داده های بارش، مدل (III) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} داشتن بیشتر R و DC و کمترین مقادیر خطای RMSE مدل برتر است. برای حالتی که از عناصر اقلیمی و الگوهای پیوند از دور استفاده شده است نیز ملاحظه می شود که مدل (VII) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} , $Nino3_{t-1}$, T_{t-1} , R_{t-1} را ارائه داده است. با مقایسه مدل های تعریف شده مشاهده می گردد که استفاده از پارامترهای میانگین دما و رطوبت نسبی ماهانه باعث افزایش دقت مدل ها شده اند. همچنین تاثیر الگوی پیوند از دور Nino3 از NOA و MEI بیشتر است. در حالت کلی، هر دو عناصر اقلیمی و الگوهای پیوند از دور در بهبود دقت نتایج تاثیر گذار هستند. باین حال، داده های اقلیمی در پیش بینی بارش موفق تر بوده و استفاده از داده های مربوط به الگوهای پیوند از دور در کنار داده های بارش ماه های گذشته دقت مدل را اندکی افزایش داده اند و در حالت استفاده تنها از الگوهای پیوند از دور مربوط به بازه زمانی قبل نتایج دقیقی به دست نیامده است (نتایج مدل (VIII)).



شکل ۴- نتایج مرحله آزمون و تست مدل برتر روش KELM بر اساس تجزیه داده ها با استفاده از روش DWT

جدول ۳- نتایج ارزیابی مدل ها بر اساس تجزیه سری های زمانی برای روش KELM

مدل	آزمون			تست		
	R	DC	RMSE	R	DC	RMSE
EEMD- KELM						
(I)	۰/۷۳۴	۰/۵۳۱	۰/۱۰۸	۰/۶۲۳	۰/۴۸۴	۰/۱۱۸
(II)	۰/۸۷۷	۰/۷۶۲	۰/۰۷۳	۰/۷۶۶	۰/۵۴۸	۰/۰۹۱
(III)	۰/۸۷۵	۰/۸۰۲	۰/۰۶۶	۰/۸۲۲	۰/۷۲۱	۰/۰۸۱
(IV)	۰/۹۰۱	۰/۸۰۱	۰/۰۶۶	۰/۸۰۵	۰/۶۹۹	۰/۰۸۶
DWT- KELM						
(I)	۰/۷۶۳	۰/۵۸۲	۰/۱۰۱	۰/۷۲۶	۰/۵۲۶	۰/۰۹۹
(II)	۰/۹۴۳	۰/۸۸۲	۰/۰۵۰	۰/۹۰۷	۰/۸۱۶	۰/۰۵۶
(III)	۰/۹۵۸	۰/۹۱۷	۰/۰۴۴	۰/۹۱۵	۰/۸۳۵	۰/۰۵۳
(IV)	۰/۹۳۳	۰/۸۶۲	۰/۰۵۷	۰/۹۱۲	۰/۸۲۸	۰/۰۶۱
نتایج ارزیابی مدل ها بر اساس عناصر اقلیمی و الگوهای پیوند از دور						
EEMD- KELM						
(I)	۰/۶۵۱	۰/۵۳۵	۰/۱۰۱	۰/۶۳۸	۰/۵۰۸	۰/۱۰۶
(II)	۰/۸۹۵	۰/۸۲۲	۰/۰۵۶	۰/۸۳۴	۰/۷۳۵	۰/۰۶۹
(III)	۰/۸۹۶	۰/۸۲۳	۰/۰۵۶	۰/۸۴۶	۰/۷۴۳	۰/۰۶۵
(IV)	۰/۸۸۴	۰/۸۰۶	۰/۰۵۸	۰/۸۲۲	۰/۷۲۶	۰/۰۷۶
(V)	۰/۸۸۰	۰/۷۹۵	۰/۰۶۵	۰/۸۱۸	۰/۷۲۴	۰/۰۸۰
(VI)	۰/۸۸۳	۰/۸۱۲	۰/۰۵۸	۰/۸۲۷	۰/۷۳۲	۰/۰۷۲
(VII)	۰/۹۰۳	۰/۸۳۴	۰/۰۴۸	۰/۸۶۶	۰/۷۶۸	۰/۰۵۴
(VIII)	۰/۵۷۰	۰/۳۷۴	۰/۲۱۴	۰/۴۷۱	۰/۳۵۱	۰/۲۵۰
DWT- KELM						
(I)	۰/۷۱۱	۰/۶۰۱	۰/۰۸۹	۰/۷۰۸	۰/۵۸۸	۰/۰۹۱
(II)	۰/۹۷۴	۰/۹۳۸	۰/۰۳۷	۰/۹۲۷	۰/۸۵۰	۰/۰۴۶
(III)	۰/۹۷۶	۰/۹۴۱	۰/۰۳۷	۰/۹۳۷	۰/۸۵۹	۰/۰۴۲
(IV)	۰/۹۶۲	۰/۹۳۱	۰/۰۴۰	۰/۹۱۲	۰/۸۴۲	۰/۰۵۱
(V)	۰/۹۵۹	۰/۹۰۹	۰/۰۴۳	۰/۹۰۸	۰/۸۳۹	۰/۰۵۳
(VI)	۰/۹۶۳	۰/۹۲۹	۰/۰۳۹	۰/۹۱۸	۰/۸۴۷	۰/۰۴۸
(VII)	۰/۹۸۴	۰/۹۵۴	۰/۰۳۲	۰/۹۶۱	۰/۸۸۸	۰/۰۳۶
(VIII)	۰/۶۲۱	۰/۴۲۸	۰/۱۴۲	۰/۵۲۵	۰/۴۰۹	۰/۱۶۶

سری داده های آموزش و آزمون در شکل ۴ نشان داده شده است.

نتایج به دست آمده برای مدل SVM بر پایه تجزیه ی یکپارچه مد تجزیه کامل و تبدیل موجک

در این قسمت با استفاده از دو روش تجزیه ی یکپارچه مد تجزیه و تبدیل موجک گسسته ابتدا سری های زمانی تجزیه گردید و سپس زیر سری های به دست آمده به عنوان ورودی به مدل SVM داده شد. نتایج حاصل از مدل سازی های در نظر گرفته شده در جدول ۴ و شکل ۵ ارائه شده است. با مقایسه نتایج جداول ۲ و ۴ ملاحظه می شود که تجزیه سری های زمانی در این روش نیز تقریباً ۳۵ درصد باعث بهبود نتایج گردیده و تأثیر تجزیه با استفاده از تبدیل موجک گسسته بیشتر از روش تجزیه ی یکپارچه مد تجزیه است. در حالت مدل سازی بر

لازم به توضیح است که در مورد مدل انتخاب شده علاوه بر اینکه این مدل از نظر معیارهای ارزیابی جهت تخمین پارامتر وابسته باید دارای دقت مطلوبی باشد، تعداد پارامترهای به کار رفته در مدل و ساده بودن آن نیز مهم است. استفاده از مدل با تعداد پارامترهای کمتر هم راحت تر بوده و هم زمان و حجم محاسبات را کاهش می دهد. با در نظر گرفتن این نکته و با توجه به نتایج جدول ۳ و مدل (II) مشاهده می گردد که در صورت استفاده از پارامترهای بارش و دما، پس از تجزیه سری داده ها روش استفاده شده در تحقیق کارایی خوبی را در فرایند پیش بینی نشان داده است. همچنین مدل (III) نیز با استفاده از تنها پارامتر بارش نتایج قابل قبولی را ارائه داده است و می توان در صورت عدم دسترسی به سایر داده های اقلیمی یا پیوند از دور از این مدل جهت تخمین بارش ماهانه است کرد. نمودار مدل های برتر برای

روش KELM انجام گرفت. برای این منظور با حذف تک-تک پارامترهای مدل برتر و اجرای دوباره مدل و تعیین معیار ارزیابی RMSE، میزان تأثیر پارامتر حذف شده در کاهش دقت مدل مورد بررسی قرار گرفت. نتایج حاصل از آنالیز حساسیت به صورت شکل ۶ ارائه گردیده است. همان طور که از شکل مشخص است با حذف پارامتر P_{t-3} دقت مدل تا حدود زیادی کاهش یافته است. بنابراین، می توان نتیجه گرفت که پارامتر بارش سه ماه قبل بیشترین تأثیر را در تخمین میزان بارش ماهانه داراست.

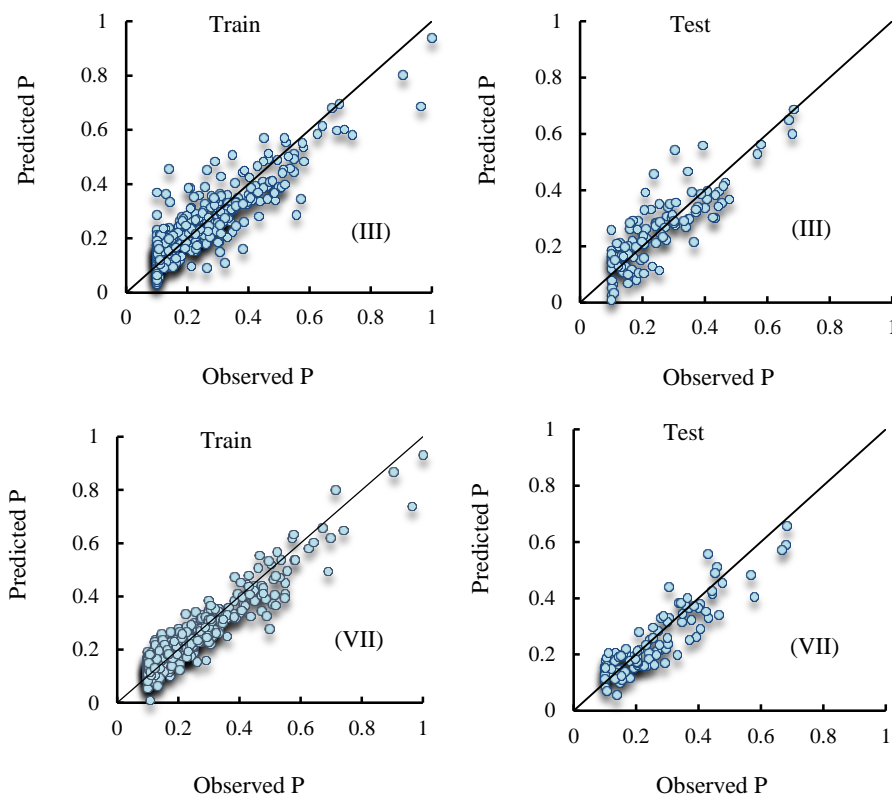
صحت سنجی نتایج حاصل برای روش پیشنهادی ارائه در تحقیق

در این بخش، جهت صحت سنجی نتایج حاصل از تحقیق از داده های ایستگاه ماکو استفاده شد و روش اعمالی بر روی این ایستگاه نیز بررسی شد. نتایج به صورت جدول ۵ و ۶ نشان داده شده است. با مقایسه نتایج این جداول ملاحظه می شود که تجزیه سری های زمانی در ایستگاه ماکو نیز باعث افزایش دقت مدل سازی تا حدود ۳۲ درصد شده است.

اساس تنها داده های بارش، مدل (III) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} با داشتن بیشتر R و DC و کمترین مقادیر خطای RMSE مدل برتر است. برای حالتی که از عناصر اقلیمی و الگوهای پیوند از دور استفاده شده است نیز مدل (VII) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} , Nino3_{t-1}, T_{t-1}, R_{t-1} مقایسه مدل های تعریف شده مشاهده می گردد که تأثیر الگوی پیوند از دور Nino3 از NOA و MEI بیشتر است. در حالت کلی مطابق با جدول ۴، هر دو عناصر اقلیمی و الگوهای پیوند از دور در بهبود دقت نتایج تأثیرگذار هستند. ولی داده های اقلیمی در پیش بینی بارش موفق تر بوده و استفاده از داده های مربوط به الگوهای پیوند از دور در کنار داده های بارش ماه های گذشته دقت مدل را اندکی افزایش داده اند و در حالت استفاده تنها از الگوهای پیوند از دور مربوط به بازه زمانی قبل نتایج دقیقی به دست نیامده است. با مقایسه نتایج دو روش مبتنی بر کرنل مشخص است که روش KELM نسبت به SVM منجر به نتایج دقیق تری شده است.

نتایج آنالیز حساسیت

جهت بررسی تأثیر پارامترهای به کار رفته در مدل برتر (مدل (VII)) در حالت تجزیه با تبدیل موجک گسسته) آنالیز حساسیت با



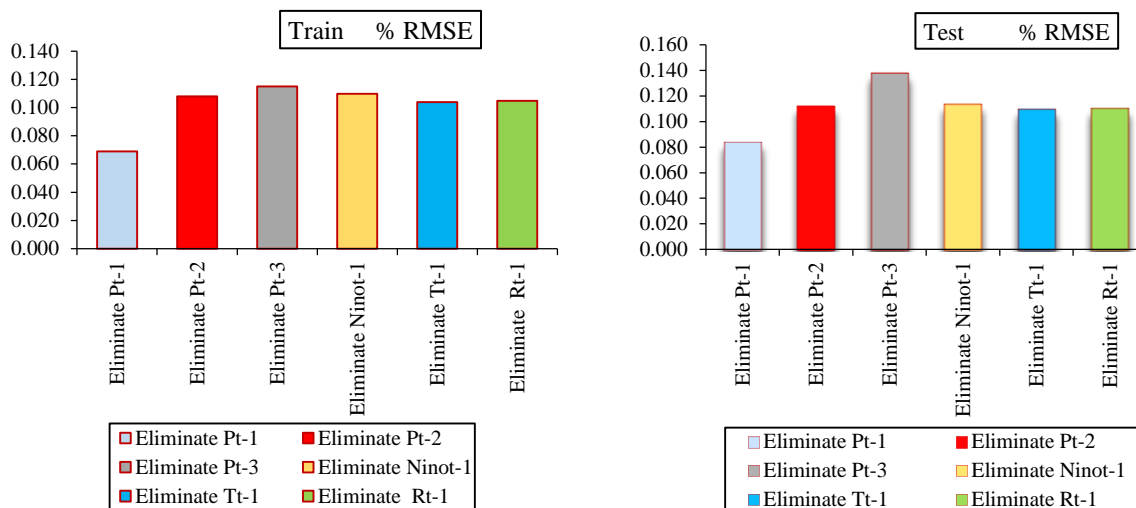
شکل ۵- نتایج مرحله آزمون و تست مدل برتر روش SVM بر اساس تجزیه داده ها با استفاده از روش DWT

جدول ۴- نتایج ارزیابی مدل ها بر اساس تجزیه سری های زمانی برای روش SVM

نتایج ارزیابی مدل ها بر اساس داده های بارش						
مدل	آزمون			تست		
	R	DC	RMSE	R	DC	RMSE
EEMD- SVM						
(I)	۰/۶۹۲	۰/۴۹۴	۰/۱۱۴	۰/۵۸۷	۰/۴۵۰	۰/۱۲۶
(II)	۰/۸۲۷	۰/۷۰۹	۰/۰۷۷	۰/۷۲۲	۰/۵۱۰	۰/۰۹۵
(III)	۰/۸۲۵	۰/۷۴۶	۰/۰۷۰	۰/۷۷۵	۰/۶۷۱	۰/۰۸۵
(IV)	۰/۸۵۰	۰/۷۴۵	۰/۰۷۰	۰/۷۵۹	۰/۶۵۰	۰/۰۹۱
DWT- SVM						
(I)	۰/۷۲۰	۰/۵۴۱	۰/۱۰۷	۰/۶۸۴	۰/۴۸۸	۰/۱۰۴
(II)	۰/۸۸۹	۰/۸۲۰	۰/۰۵۳	۰/۸۵۴	۰/۷۵۹	۰/۰۶۰
(III)	۰/۹۰۳	۰/۸۵۳	۰/۰۴۶	۰/۸۶۲	۰/۷۷۷	۰/۰۵۷
(IV)	۰/۸۸۰	۰/۸۰۲	۰/۰۶۰	۰/۸۶۱	۰/۷۷۰	۰/۰۶۵
نتایج ارزیابی مدل ها بر اساس عناصر اقلیمی و الگوهای پیوند از دور						
EEMD- SVM						
(I)	۰/۶۱۵	۰/۴۹۸	۰/۱۰۷	۰/۶۰۲	۰/۴۷۲	۰/۱۱۲
(II)	۰/۸۴۳	۰/۷۶۳	۰/۰۵۹	۰/۷۸۶	۰/۶۸۴	۰/۰۷۳
(III)	۰/۸۴۴	۰/۷۶۵	۰/۰۵۸	۰/۷۹۷	۰/۶۹۰	۰/۰۶۸
(IV)	۰/۸۳۳	۰/۷۴۹	۰/۰۶۲	۰/۷۷۵	۰/۶۷۵	۰/۰۷۹
(V)	۰/۸۳۰	۰/۷۳۹	۰/۰۶۹	۰/۷۷۱	۰/۶۷۳	۰/۰۸۴
(VI)	۰/۸۳۳	۰/۷۵۵	۰/۰۶۱	۰/۷۸۰	۰/۶۸۱	۰/۰۷۶
(VII)	۰/۸۵۲	۰/۷۷۶	۰/۰۵۱	۰/۸۱۷	۰/۷۱۴	۰/۰۵۷
(VIII)	۰/۵۳۸	۰/۳۴۸	۰/۲۲۶	۰/۴۴۴	۰/۳۲۶	۰/۲۶۴
DWT- SVM						
(I)	۰/۶۷۰	۰/۵۵۹	۰/۰۹۴	۰/۶۶۸	۰/۵۴۷	۰/۰۹۶
(II)	۰/۹۱۸	۰/۸۷۲	۰/۰۳۹	۰/۸۷۳	۰/۷۹۱	۰/۰۴۹
(III)	۰/۹۲۰	۰/۸۷۵	۰/۰۳۹	۰/۸۸۵	۰/۷۹۹	۰/۰۴۴
(IV)	۰/۹۰۷	۰/۸۵۷	۰/۰۴۲	۰/۸۶۰	۰/۷۸۲	۰/۰۵۳
(V)	۰/۹۰۴	۰/۸۴۵	۰/۰۴۶	۰/۸۵۶	۰/۷۷۹	۰/۰۵۶
(VI)	۰/۹۰۸	۰/۸۶۴	۰/۰۴۱	۰/۸۶۶	۰/۷۸۹	۰/۰۵۱
(VII)	۰/۹۲۸	۰/۸۸۷	۰/۰۳۴	۰/۹۰۶	۰/۸۲۷	۰/۰۳۸
(VIII)	۰/۵۸۶	۰/۳۹۸	۰/۱۵۱	۰/۴۹۳	۰/۳۷۹	۰/۱۷۵

جدول ۵- نتایج ارزیابی مدل های SVM و KELM بدون تجزیه داده ها برای ایستگاه ماکو

مدل		آزمون			تست		
		R	DC	RMSE	R	DC	RMSE
(I)	KELM	۰/۳۸۸	۰/۲۸۲	۰/۲۷۴	۰/۳۴۵	۰/۱۹۲	۰/۲۸۰
	SVM	۰/۳۶۸	۰/۲۷۴	۰/۲۸۹	۰/۳۲۹	۰/۱۸۷	۰/۲۹۵
(II)	KELM	۰/۵۳۰	۰/۳۲۸	۰/۲۵۷	۰/۴۸۷	۰/۲۹۷	۰/۲۶۱
	SVM	۰/۵۰۹	۰/۳۱۸	۰/۲۷۱	۰/۴۶۷	۰/۲۸۳	۰/۲۸۰
(III)	KELM	۰/۵۶۲	۰/۴۵۸	۰/۲۰۹	۰/۵۵۱	۰/۴۴۸	۰/۲۲۸
	SVM	۰/۵۳۳	۰/۴۴۴	۰/۲۲۱	۰/۵۲۸	۰/۴۳۹	۰/۲۴۳
(IV)	KELM	۰/۵۵۳	۰/۳۹۰	۰/۲۱۶	۰/۵۳۴	۰/۳۷۳	۰/۲۳۰
	SVM	۰/۵۳۱	۰/۳۷۸	۰/۲۲۸	۰/۵۱۱	۰/۳۵۷	۰/۲۴۵



شکل ۶- تأثیر نسبی هر یک از پارامترهای ورودی مدل برتر

جدول ۶- نتایج ارزیابی مدل‌ها بر اساس تجزیه سری‌های زمانی برای روش KELM برای ایستگاه ماکو

نتایج ارزیابی مدل‌ها بر اساس داده‌های بارش						
مدل	آزمون			تست		
	R	DC	RMSE	R	DC	RMSE
EEMD- KELM						
(I)	۰/۶۶۸	۰/۴۶۷	۰/۱۲۷	۰/۵۶۷	۰/۴۲۶	۰/۱۳۹
(II)	۰/۷۹۸	۰/۶۷۱	۰/۰۸۶	۰/۶۹۷	۰/۴۸۲	۰/۱۰۷
(III)	۰/۷۹۶	۰/۷۰۶	۰/۰۷۸	۰/۷۴۸	۰/۶۳۴	۰/۰۹۶
(IV)	۰/۸۲۰	۰/۷۰۵	۰/۰۷۸	۰/۷۳۳	۰/۶۱۵	۰/۱۰۱
DWT- KELM						
(I)	۰/۶۹۴	۰/۵۱۲	۰/۱۱۹	۰/۶۶۱	۰/۴۶۳	۰/۱۱۷
(II)	۰/۸۵۸	۰/۷۷۶	۰/۰۵۹	۰/۸۲۵	۰/۷۱۸	۰/۰۶۶
(III)	۰/۸۷۲	۰/۸۰۷	۰/۰۵۲	۰/۸۳۳	۰/۷۳۵	۰/۰۶۳
(IV)	۰/۸۴۹	۰/۷۵۹	۰/۰۶۷	۰/۸۳۰	۰/۷۲۹	۰/۰۷۲
نتایج ارزیابی مدل‌ها بر اساس عناصر اقلیمی و الگوهای پیوند از دور						
EEMD- KELM						
(I)	۰/۵۹۲	۰/۴۷۱	۰/۱۱۹	۰/۵۸۱	۰/۴۴۷	۰/۱۲۵
(II)	۰/۸۱۴	۰/۷۲۳	۰/۰۶۶	۰/۷۵۹	۰/۶۴۷	۰/۰۸۱
(III)	۰/۸۱۵	۰/۷۲۴	۰/۰۶۶	۰/۷۷۰	۰/۶۵۴	۰/۰۷۷
(IV)	۰/۸۰۴	۰/۷۰۹	۰/۰۶۸	۰/۷۴۸	۰/۶۳۹	۰/۰۹۰
(V)	۰/۸۰۱	۰/۷۰۰	۰/۰۷۷	۰/۷۴۴	۰/۶۳۷	۰/۰۹۴
(VI)	۰/۸۰۴	۰/۷۱۵	۰/۰۶۸	۰/۷۵۳	۰/۶۴۴	۰/۰۸۵
(VII)	۰/۸۲۲	۰/۷۳۴	۰/۰۵۷	۰/۷۸۸	۰/۷۱۲	۰/۰۶۴
(VIII)	۰/۵۱۹	۰/۳۲۹	۰/۲۵۳	۰/۴۲۹	۰/۳۰۹	۰/۲۹۵
DWT- KELM						
(I)	۰/۶۴۷	۰/۵۲۹	۰/۱۰۵	۰/۶۴۴	۰/۵۱۷	۰/۱۰۷
(II)	۰/۸۸۶	۰/۸۲۵	۰/۰۴۴	۰/۸۴۴	۰/۷۴۸	۰/۰۵۴
(III)	۰/۸۸۸	۰/۸۲۸	۰/۰۴۴	۰/۸۵۳	۰/۷۵۶	۰/۰۵۰
(IV)	۰/۸۷۵	۰/۸۱۰	۰/۰۴۷	۰/۸۳۰	۰/۷۴۱	۰/۰۶۰
(V)	۰/۸۷۳	۰/۸۰۰	۰/۰۵۱	۰/۸۲۶	۰/۷۳۸	۰/۰۶۳
(VI)	۰/۸۷۶	۰/۸۱۸	۰/۰۴۶	۰/۸۳۵	۰/۷۴۵	۰/۰۵۷
(VII)	۰/۸۹۵	۰/۸۴۰	۰/۰۳۸	۰/۸۷۵	۰/۷۸۱	۰/۰۴۲
(VIII)	۰/۵۶۵	۰/۳۷۷	۰/۱۶۸	۰/۴۷۸	۰/۳۶۰	۰/۱۹۶

می گردند. نتایج حاصل از آنالیز حساسیت نشان داد که پارامتر بارش سه ماه قبل بیشترین تأثیر را در تخمین میزان بارش ماهانه داراست. همچنین صحت سنجی دقت روش اعمالی بیانگر کارایی مطلوب روش اعمالی در تحقیق بود.

منابع

بختیاری، ب.، شهرکی، ن. و احمدی، م. م. ۱۳۹۲. برآورد احتمالات بارش روزانه با استفاده از مدل زنجیره مارکف در اقلیم های مختلف ایران. تحقیقات منابع آب ایران. ۲ (۱۰): ۴۴-۵۵.

ستاری، م.ت.، رضازاده جودی، ع. و نهرین، فرناز. ۱۳۹۳. پیش بینی مقادیر بارش ماهانه با استفاده از شبکه های عصبی مصنوعی و مدل درختی M5 مطالعه موردی: ایستگاه اهر. پژوهش های جغرافیای طبیعی. ۴۶ (۲): ۲۶۰-۲۴۷.

فلاحی، م. ر.، وروانی، ه. و گلیمان، س. ۱۳۹۰. پیش بینی بارش با استفاده از مدل رگرسیون درختی به منظور کنترل سیل. پنجمین کنفرانس سراسری آبخیزداری و مدیریت منابع آب و خاک کشور.

Adamowski, K., Prokoph, A. and Adamowski, J. 2009. Development of a new method of wavelet aided trend detection and estimation. *Hydrology Process.* 23(18): 2686-2696.

Amirat, Y., Benbouzid, M., Wang, T., Bacha, K. and Feld, G. 2018. EEMD-based notch filter for induction machine bearing faults detection. *Applied Acoustics.* 133: 202-209.

ASCE. 2000. Task Committee on Application of Artificial Neural Networks in Hydrology, *Artificial Neural Networks in hydrology. I: Preliminary concepts.* *Hydrological Engineering, ASCE.* 5(2): 115-123.

Chong, K. L., Lai, S. H., Yao, Y., Ahmed, A. N., Jaafar, W. Z. and El-Shafie, A. 2020. Performance enhancement model for rainfall forecasting utilizing integrated wavelet-convolutional neural network. *Water Resources Management.* 34(8): 2371-2387.

Chou, C. M. 2011. Complexity analysis of rainfall and runoff time series based on sample entropy in different temporal scales. *Stochastic Environmental Research and Risk Assessment.* 6: 1401-1408.

Deo, R. C., Samui, P. and Kim, D. 2016. Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models. *Stochastic Environmental Research and Risk Assessment.* 30: 1769-1786.

Kisi, O., Cimen, M. 2012. Precipitation forecasting by using wavelet-support vector machine conjunction

در این ایستگاه نیز تأثیر تجزیه با استفاده از تبدیل موجک گسسته بیشتر از روش تجزیه یکپارچه مد تجزیه است. در حالت مدل سازی بر اساس تنها داده های بارش، مدل (III) و برای حالتی که از عناصر اقلیمی و الگوهای پیوند از دور استفاده شده است مدل (VII) نتایج دقیقی را ارائه داده اند. مشاهده می گردد که استفاده از پارامترهای میانگین دما و رطوبت نسبی ماهانه باعث افزایش دقت مدل ها شده اند. همچنین، تأثیر الگوی پیوند از دور Nino3 در افزایش دقت مدل سازی بیشتر است. با این حال، هر دو عناصر اقلیمی و الگوهای پیوند از دور در بهبود دقت نتایج تأثیرگذار هستند. بر اساس نتایج حاصله می توان بیان کرد که روش ترکیبی پیشنهادی در این تحقیق از دقت مطلوبی جهت پیش بینی بارش ماهانه برخوردار است و با موفقیت می توان از آن برای پیش بینی بارش مناطق دیگر نیز استفاده شود.

نتیجه گیری

پیش بینی بارش به دلیل ماهیت تصادفی آن در مکان و زمان همواره با مشکلات بسیاری مواجه بوده است و این عدم قطعیت از اعتبار بسیاری از مدل های پیش بینی می کاهد. در این تحقیق به پیش بینی میزان بارش ماهانه با استفاده از روش هوشمند KELM و SVM بر پایه تجزیه ی یکپارچه مد تجزیه کامل و تبدیل موجک پرداخته شد. مدل های ورودی متفاوتی با استفاده از داده های مربوط به عناصر اقلیمی مانند بارش، میانگین دما و رطوبت نسبی ماهانه ایستگاه ارومیه در بازه زمانی (۱۹۸۰-۲۰۱۸) تعریف گردید. دو نوع مدل سازی با استفاده از داده های سری های زمانی اصلی و همچنین با استفاده از داده های سری های زمانی تجزیه شده با دو روش تجزیه ی یکپارچه مد تجزیه کامل و تبدیل موجک گسسته انجام شد. نتایج حاصله نشان داد که تجزیه سری های زمانی با هر دو روش EEMD و DWT تا حدود زیادی باعث بهبود نتایج می گردد و تجزیه داده ها با استفاده از تبدیل موجک گسسته منجر به نتایج دقیق تری می شود. مطابق با نتایج حاصله، برای مدل های تعریف شده بر اساس تنها داده های بارش، مدل (III) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} با داشتن بیشتر R و DC و کمترین مقادیر خطا مدل برتر گردید. در حالت مدل سازی بر اساس عناصر اقلیمی و الگوهای پیوند از دور نیز مدل (VII) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} , $Nino3_{t-1}$, T_{t-1} , R_{t-1} منجر به نتایج دقیق تری گردید. مشاهده شد که استفاده از پارامترهای میانگین دما و رطوبت نسبی ماهانه باعث افزایش دقت مدل ها می گردد و همچنین در بین الگوهای پیوند از دور نیز Nino3 نسبت به NOA و MEI تأثیرگذارتر است. در حالت کلی مشخص گردید هر دو عناصر اقلیمی و الگوهای پیوند از دور سبب بهبود دقت مدل ها

- boundary conditions using an evolutionary algorithm. *International Journal of Sediment Research*. 32(4): 564-574.
- Samantaray, S., Tripathy, O., Sahoo, A. and Ghose, D. K. 2020. Rainfall forecasting through ANN and SVM in Bolangir Watershed, India. In *smart intelligent computing and applications* (pp. 767-774). Springer, Singapore.
- Siviapragasam, C. and Liong, S. 2001. Rainfall and runoff forecasting with SSA-SVM approach. *Hydroinformation*. 3: 141-152.
- Soltani, A. S., Saberi, A. and Gheisouri, M. 2017. Determination of the best time series model for forecasting annual rainfall of selected stations of Western Azerbaijan province. *Researches in Geographical Sciences*. 17(44): 87-105.
- Wu, Z. and Huang, N. F. 2004. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc RS Lond* 460A: 1597-1611.
- model. *Engineering Application Artificial Intelligence*. 25: 783-792.
- Kumar, S., Tripathy, D., Nayak, S., Mohaparta, S. 2013. Prediction of rainfall in India using artificial neural network models. *International Journal of intelligent system and applications*. 12: 1-22.
- Lau, K. M. and Weng, H. Y. 1995. Climate signal detection using wavelet transform, How to make time series sing. *Bull Am Meteorol Soc*. 76: 2391-2402.
- Marzano, F. S., Fionda, E. and Ciotti, P. 2006. Neural-network approach to ground-based passive microwave estimation of precipitation intensity and extinction. *Hydrology*. 328: 121-131.
- Nayak, D., Mahapatra, A. and Mishra, P. 2013. A survey on rainfall prediction using artificial neural network. *International journal of computer applications*. 72(16): 32-40.
- NOAA. 2009. Earth System Research laboratory, <https://www.esrl.noaa.gov/psd/data/climateindices/list/>.
- Roushangar, K. and Ghasempour, R. 2017. Estimation of bedload discharge in sewer pipes with different

Investigation of Effective Parameters in Modeling Monthly Precipitation using Intelligent Integrated Models Based on Time Series Decomposition

K. Amininia^{1*}, S.M. Saghebian²

Received: May.15, 2022

Accepted: Jun.15, 2022

Abstract

Rainfall forecasting is important in many different aspects of watershed management, such as flood and drought warning systems. Spatiotemporal variations of rainfall cause its prediction to be difficult. In this study, the monthly rainfall of Urmia and Mako stations were assessed using the intelligent kernel-based methods using Ensemble Empirical Mode Decomposition (EEMD) and Wavelet Transform (DWT). For this aim, different models were developed based on teleconnection patterns and climatic elements including rainfall, humidity, and temperature of previous months, and the impact of these parameters on accuracy of the modeling process was investigated. The obtained results showed the high efficiency of the integrated methods used in modeling process. It was observed that in the monthly precipitation modeling, the simultaneous use of teleconnection patterns and climatic elements related to previous months improves the accuracy of the models by up to 35%. Comparison of the wavelet transform and ensemble empirical mode decomposition showed that time series decomposition based on wavelet transform led to more accurate outcomes. The results of sensitivity analysis showed that the precipitation parameter with three months lag was the most effective parameter in monthly precipitation modeling.

Keywords: Climatic elements, Empirical mode decomposition, Precipitation modeling, Teleconnection patterns

1- Assistant Professor, Department of Geography, Ahar Branch, Islamic Azad University - Ahar-Iran

2- Assistant Professor, Department of Civil Engineering, Ahar Branch, Islamic Azad University - Ahar-Iran

(*-Corresponding Author Email: k-amininia@iau-ahar.ac.ir)