


## Validation of Combined Empirical Mode Decomposition with Normal Transformation in Climate Elements Forecasting

parviz, L.<sup>1</sup> 

1. Associate Professor, Faculty of Agriculture, Azarbaijan Shahid Madani University.

**Corresponding Author E-mail:** [laleh\\_parviz@yahoo.com](mailto:laleh_parviz@yahoo.com)

*(Received: 17 Feb 2025, Revise: 17 Mar 2025, Accepted: 19 May 2025, Published online: 23 Sep 2025)*

### Abstract

The non-stationary nature and skewness of time series due to various factors, including human activities, have created a new challenge in forecasting climate elements. To increase the accuracy of the forecasting process and overcome the skewness of time series, an algorithm was proposed and evaluated in three climates: arid, semi-arid, and very humid with maximum temperature data and total sunshine hours. The algorithm is based on time series decomposition using the empirical mode decomposition (EMD) approach and then normalizing the decomposed series. Series decomposition and then normalization of subseries were able to significantly improve modeling performance, so that the average reduction in root mean square error at five stations from the maximum temperature data without decomposition to series decomposition and normalization of decomposed subseries was 11.22 and 22.94 percent, respectively. The proposed algorithm had minimal error in modeling average temperature and total sunshine hours in dry and very humid climates, respectively. In the case of maximum temperature, except for the Birjand station, all other stations have underestimates, and in the case of sunshine hour data, only stations in arid climates have underestimates. Therefore, the type of climate can affect the modeling performance. The type of normalized sub-series can also affect the evaluation statistics. Combining empirical mode decomposition with normal transformation has satisfactory performance, especially in the case of non-stationary series, which can play an effective role as an efficient tool in complex modeling.

**Keywords:** Algorithm, Normalization, EMD, Combining

**Cite this article:** Parviz,L. (2025). Validation of Combined Empirical Mode Decomposition with Normal Transformation in Climate Elements Forecasting. *Nivar*, 49(130-131), 160-176. doi: 10.30467/nivar.2025.507159.1324

**E-mail:**



Publisher: Iran Meteorological Organization  
DOI: <https://doi.org/10.30467/nivar.2025.507159.1324>

Print ISSN: 1735-0565  
Online ISSN: 2645-3347

## صحت‌سنجی تجزیه مد تجربی تلفیقی با تبدیل نرمال در پیش‌بینی عناصر اقلیمی

لاله پرویز<sup>۱</sup>

۱. دانشیار دانشکده کشاورزی دانشگاه شهید مدنی آذربایجان.

رایانامه نویسنده مسئول: [lahle\\_parviz@yahoo.com](mailto:lahle_parviz@yahoo.com)

دریافت: ۱۴۰۳/۱۱/۲۹، بازنگری: ۱۴۰۳/۱۲/۲۷، پذیرش: ۱۴۰۴/۰۲/۲۹، انتشار آنلاین: ۱۴۰۴/۰۷/۰۱

### چکیده

ماهیت نایستا و چولگی سری‌های زمانی ناشی از عوامل مختلف از جمله فعالیت‌های انسانی چالش جدیدی را در پیش‌بینی عناصر اقلیمی ایجاد کرده است. جهت افزایش دقت در روند پیش‌بینی و غلبه بر چولگی سری‌های زمانی، الگوریتمی پیشنهاد و عملکرد آن در سه اقلیم خشک، نیمه‌خشک و بسیار مرطوب با داده‌های دمای بیشینه و مجموع ساعات آفتابی ارزیابی شد. اساس الگوریتم بر تجزیه سری زمانی با رویکرد تجزیه مد تجربی (EMD) و سپس نرمال‌سازی سری‌های تجزیه شده است. تجزیه سری و سپس نرمال‌سازی زیر سری‌ها توانست عملکرد مدل‌سازی را به شدت ارتقاء دهد، به طوری که میزان متوسط کاهش خطای جذر میانگین مربعات در پنج ایستگاه با داده‌های بیشینه از حالت بدون تجزیه به تجزیه سری و نرمال‌سازی زیر سری‌های به ترتیب برابر با ۱۱/۲۲ و ۲۲/۹۴ درصد شد. الگوریتم پیشنهادی در مدل‌سازی دمای متوسط و مجموع ساعات آفتابی به ترتیب در اقلیم خشک و بسیار مرطوب دارای کمینه خطا بود. در مورد دمای بیشینه بجز ایستگاه بیرجند بقیه ایستگاه‌ها دارای کم برآورد و در مورد داده‌های مجموع ساعت آفتابی فقط ایستگاه‌ها در اقلیم خشک دارای تخمین کم برآورد هستند. بنابراین نوع اقلیم می‌تواند عملکرد مدل‌سازی را تحت تاثیر قرار دهد. نوع زیر سری نرمال شده در روند مدل‌سازی هم می‌تواند بر دقت پیش‌بینی تاثیر داشته باشد. تلفیق تجزیه مد تجربی با تبدیل نرمال بخصوص در مورد سری‌های نایستا عملکرد رضایت بخشی دارد که می‌تواند به عنوان ابزار کارآمد در مدل‌سازی پیچیده نقش موثری داشته باشد.

کلید واژه‌ها: الگوریتم، نرمال‌سازی، EMD، تلفیق

### ۱. مقدمه

گروه‌های آسیب‌پذیر، آمادگی اضطراری طولانی‌مدت و پایدار نقش مهمی دارد (لین<sup>۲</sup> و همکاران، ۲۰۲۱). دمای بالای غیرعادی در تابستان باعث می‌شود افراد از انواع مختلفی از بلاایای حرارتی مانند اثرات بهداشتی به عنوان مثال گرم‌زدگی، خسارت محصولات کشاورزی، آتش‌سوزی‌ها و قطع برق رنج ببرند. تعیین دمای هوای بیشینه قابل اطمینان و دقیق در تابستان برای آماده شدن و پاسخگویی به بلاایای حرارتی ضروری است (چو<sup>۳</sup> و همکاران، ۲۰۲۲). افزایش سالانه دمای بیشینه می‌تواند مشکلات زندگی شهری را

بررسی نوسانات دمای هوا برای تجزیه و تحلیل اثرات تغییرات آب و هوا در کشاورزی، اکولوژی، محیط‌زیست و صنعت بسیار مهم است. پیش‌بینی دقیق دما برای حفاظت از جان و اموال حیاتی است و به طور قابل توجهی در تلاش برای برنامه‌ریزی استراتژیک کمک کننده است (لایسر<sup>۱</sup> و همکاران، ۲۰۲۵). سیستم پیش‌بینی دمای بیشینه در پایش رویدادهای با دمای بالا جهت کاهش آتش‌سوزی‌ها بخصوص در جنگل‌ها، هشدار زود هنگام در مناطق بحرانی و

<sup>3</sup> Cho<sup>1</sup> Licer<sup>2</sup> Lin

تشدید کند. گزارش شده است که افزایش دمای سالانه باعث افزایش مصرف آب آشامیدنی و کاهش پایداری سیستم‌های خنک کننده می‌شود. در ۸۳ درصد از مناطق شهر ناپایداری در بیشینه دمای سالانه تشخیص داده می‌شود (سالاریجازی<sup>۱</sup> و همکاران، ۲۰۲۳).

در طی چند دهه گذشته محققان تلاش زیادی برای توسعه مدل‌های پیش‌بینی داده‌های هواشناسی انجام داده‌اند که در دو دسته شامل مدل‌های فیزیکی و مبتنی بر داده تقسیم شده‌اند. مدل‌های فیزیکی شامل فرآیندهای محاسباتی پیچیده و نیاز به تعداد بیشتر داده‌های هواشناسی هستند. بر این اساس مدل‌های مبتنی بر داده به طور گسترده مورد استفاده قرار گرفته‌اند، چرا که به داده‌های کمتری نیاز دارند و دارای ساختار ساده‌ای می‌باشند. از الگوریتم‌های یادگیری ماشین مانند شبکه عصبی مصنوعی، رگرسیون خطی چندگانه، ماشین بردار پشتیبان و جنگل تصادفی در بازه زمانی ۱۹۸۱-۲۰۲۲ جهت برآورد دمای بیشینه در هند استفاده شد. نتایج نشان دادند که شبکه عصبی مصنوعی با خطای جذر میانگین مربعات برابر با ۱/۴۱ و ضریب همسانی (d) برابر با ۰/۸۹ و ضریب همبستگی ۰/۸۱ عملکرد بهتری نسبت به سایر روش‌ها داشت (ولیولی<sup>۲</sup> و همکاران، ۲۰۲۴). استفاده از این مدل‌های یادگیری آنچنان توسعه یافته است که در یک تحقیق در مراکش الگوریتم ترکیبی جدید که برای تعیین دمای بیشینه با ادغام چندین مدل یادگیری عمیق (با ترکیب مدل‌های پیشرفته‌ای مانند شبکه‌های کانولوشن موقت (TCN) و شبکه‌های حافظه کوتاه‌مدت بلند مدت (LSTM)) توسعه یافته است. ارزیابی با استفاده از معیارهایی خطا عملکرد برتر این الگوریتم را نسبت به تکنیک‌های سنتی نشان داد و پتانسیل مدل‌های یادگیری عمیق ترکیبی را در پیش‌بینی آب و هوای منطقه برجسته می‌کند. این تحقیق ابزار ارزشمندی برای برنامه‌ریزی و تصمیم‌گیری آگاهانه در بخش‌های حساس به تغییرات آب و هوایی است (لایسر و همکاران، ۲۰۲۵).

علی‌رغم برتری و پایداری مدل‌های داده محور، غیرایستایی داده‌ها منجر به ایجاد خطا در داده‌ها شده است. بسیاری از تحقیقات از ترکیب مدل‌های مبتنی بر داده و تکنیک‌های تجزیه برای حل این مشکل از جمله تبدیل موجک (WT)، تجزیه مد تجربی (EMD) و تجزیه حالت متغیر (VMD) تمرکز کرده‌اند. بنابراین در بسیاری از مطالعات از ایده تجزیه سری زمانی استفاده شده است تا با تجزیه سری زمانی عدم ایستایی سری کاهش یابد. در تعیین دمای متوسط هوا در شهرستان ساری از شبکه عصبی موجک استفاده شد. داده‌های ورودی مدل شامل رطوبت نسبی، دمای بیشینه، کمینه، سرعت باد و تخریر طی بازه زمانی ۱۳۸۲-۱۳۹۲ انتخاب شدند. در ترکیب ورودی از ۵ نوع ترکیب استفاده شد. ترکیب اول شامل رطوبت نسبی، ترکیب دوم شامل رطوبت نسبی و دمای کمینه، ترکیب سوم شامل رطوبت نسبی، دمای کمینه و بیشینه، ترکیب چهارم شامل رطوبت نسبی، دمای کمینه و بیشینه و سرعت باد در نظر گرفته شدند. براساس آماره‌های ارزیابی شامل ضریب همبستگی، ضریب نش ساتکلیف مدل شبکه عصبی موجک نسبت به شبکه عصبی مصنوعی عملکرد بهتری داشت (شاهی نژاد و دهقانی، ۱۳۹۵). با استفاده از فیلترهای بالا و پایین گذر از موجک symlet16 جهت حذف نویز داده‌های دما استفاده شد. پیش‌بینی داده‌های دما براساس دسته‌بندی ضرایب موجک با بکارگیری ماشین بردار پشتیبان و بازسازی سری‌های زمانی با عکس تبدیل موجک بود. بکارگیری تجزیه موجک توانست میزان خطا (۱/۰۸٪) را کاهش دهد (پزشکی و همکاران، ۱۳۹۷). دقت مدل حافظه طولانی کوتاه مدت با تلفیق دو روش پیش‌پردازنده تبدیل موجک گسسته (DWT) و تجزیه مد تجربی کامل (CEEMD) در مدل‌سازی دما در ارومیه و آمل با روش نوین LSTM ارزیابی شد. داده‌های ورودی مدل شامل دمای بیشینه، کمینه، بارش و تابش خورشید بود، به طوری که بیشینه دمای یک روز تابع داده‌های هواشناسی در همان روز، یک روز ما قبل و دو روز ما قبل بود. نتایج بیانگر آن بود که در منطقه نیمه‌خشک، داده‌های دمای بیشینه

<sup>2</sup> Velivelli

<sup>1</sup> Salarijazi

هوای سطحی مورد مطالعه قرار گرفت. هموارسازی با روش - های مختلف مزیت روش تجزیه داده‌های موجک را نشان داد که امکان فیلتر کردن مولفه‌های پرت و فرکانس بالا را فراهم کرد. در این تحقیق مدل تجزیه موجک ماهیت نوساناتی تغییرات دمای سطحی را بخوبی پیش‌بینی کرد (ولواج<sup>۲</sup> و همکاران، ۲۰۲۴). تحقیقات بیان شده با تاکید بیشتر بر دما است، به طور کلی افزایش دقت مدل‌سازی با تجزیه سری داده‌های ورودی در مورد بسیاری از متغیرها از جمله پیش‌بینی سطح آب زیرزمینی با ترکیب سری زمانی و موجک (صالحی و همکاران، ۱۳۹۷) به اثبات رسیده است، ولی مساله‌ای که در این حالت قابل توجه است، چولگی داده‌های ورودی است. برای غلبه بر ماهیت نایستایی و چولگی سری رواناب، از یک مدل ترکیبی چند متغیره براساس تجزیه- نرمال‌سازی و رگرسیون Lasso جهت پیش‌بینی رواناب استفاده شد. مراحل انجام تحقیق به این صورت بود که در ابتدا سری زمانی رواناب با استفاده از تبدیل موجک به سری‌های زمانی فرعی (زیرسری) با فرکانس‌های مختلف تجزیه شدند. سپس از تبدیل باکس- کاکس یا معکوس W-H جهت نرمال‌سازی سری‌های زمانی فرعی استفاده شد. در ادامه رگرسیون Lasso برای رتبه‌بندی اهمیت پیش‌بینی کننده‌های تبدیل شده بکار برده شده است و در نهایت مجموعه بهینه پیش‌بینی کننده‌ها انتخاب می شوند. از PSO برای بهینه‌سازی رگرسیون بردار پشتیبان استفاده شد و مرحله نهایی مربوط به پیش‌بینی رواناب بود. نتایج نشان دادند که حالت تجزیه نرمال‌سازی توانست عملکرد مدل‌ها را بهبود ببخشد، به طوری که ضریب نش بزرگتر از ۰/۹ شد. این مطالعه نشان داد که مدل ترکیبی (تجزیه- نرمال‌سازی) دارای عملکرد رضایت بخشی است که می‌تواند رواناب پیچیده را به طور دقیق پیش‌بینی کند (کانگ<sup>۳</sup> و همکاران، ۲۰۲۳).

تحقیقات انجام گرفته بیانگر افزایش دقت پیش‌بینی با رویکرد تجزیه سری زمانی است که کاربرد این رویکرد در مورد

و کمینه دو روز قبل و دمای بیشینه و کمینه یک روز قبل و دمای کمینه و تابش خورشیدی همان روز و در منطقه مرطوب داده‌های دمای بیشینه دو روز قبل و دمای بیشینه و کمینه یک روز قبل و دمای کمینه و تابش خورشیدی همان روز به عنوان داده‌های ورودی موثر بدست آمدند. استفاده از دو روش پیش‌پردازش منجر به بهبودی مدل شد و تبدیل موجک بیشترین تاثیر را در افزایش دقت مدل LSTM داشت. نوع داده‌های ورودی موثر در دو اقلیم با هم متفاوت است (روشگر و عدل زاده، ۱۴۰۲). تعیین مقدار تابش پراکنده خورشیدی با روش‌های دقیق و قابل اعتماد ضروری است، به همین جهت یک مدل جفت شده برای تخمین تابش افقی پراکنده خورشیدی با ادغام ماشین بردار پشتیبان و تبدیل موجک در کرمان بکار گرفته شد. تجزیه سری زمانی در روند مدل‌سازی ماشین بردار پشتیبان منجر به افزایش دقت در تخمین مقدار تابش پراکنده بخصوص نسبت به مدل تجربی درجه سه شد (شمشیربند<sup>۱</sup> و همکاران، ۲۰۱۶). از تلفیق روش مد تجربی جمعی مکمل چند بعدی (MCEEMD) با مدل RBFNN در تعیین دمای بیشینه در تایوان (شهر Taipei) در بازه زمانی ۱۹۶۰-۲۰۱۷ استفاده شد. مدل-MCEEMD RBFNN از مدل RBFNN در پیش‌بینی دمای بیشینه روزانه ۷ روز بعد در دوره آزمایشی عملکرد بهتری داشت. در نتیجه، این نتایج می‌تواند توسط مقامات محلی برای جلوگیری از آسیب‌های ناشی از گرما و حوادث ناشی از گرمای شدید استفاده شود. در پیش‌بینی کوتاه مدت، مدل گاهی اوقات دارای برخی اختلافات برای گرفتن مقادیر اوج، فراتر از یک پیش‌بینی ۳ روزه است. از نتایج دیگر تحقیق می‌توان به تنظیم پارامترهای RBFNN اشاره کرد. برای بدست آوردن دقت بالاتر در پیش‌بینی طولانی‌تر از ۳ روز، می‌توان مقادیر مختلفی از تعداد نورون‌های لایه پنهان را در مدل RBFNN بررسی کرد (لین و همکاران، ۲۰۲۱). در تحقیقی ترکیب طیفی و رابطه خطی میانگین دمای سالانه

<sup>3</sup> Kang

<sup>1</sup> Shamsheirband

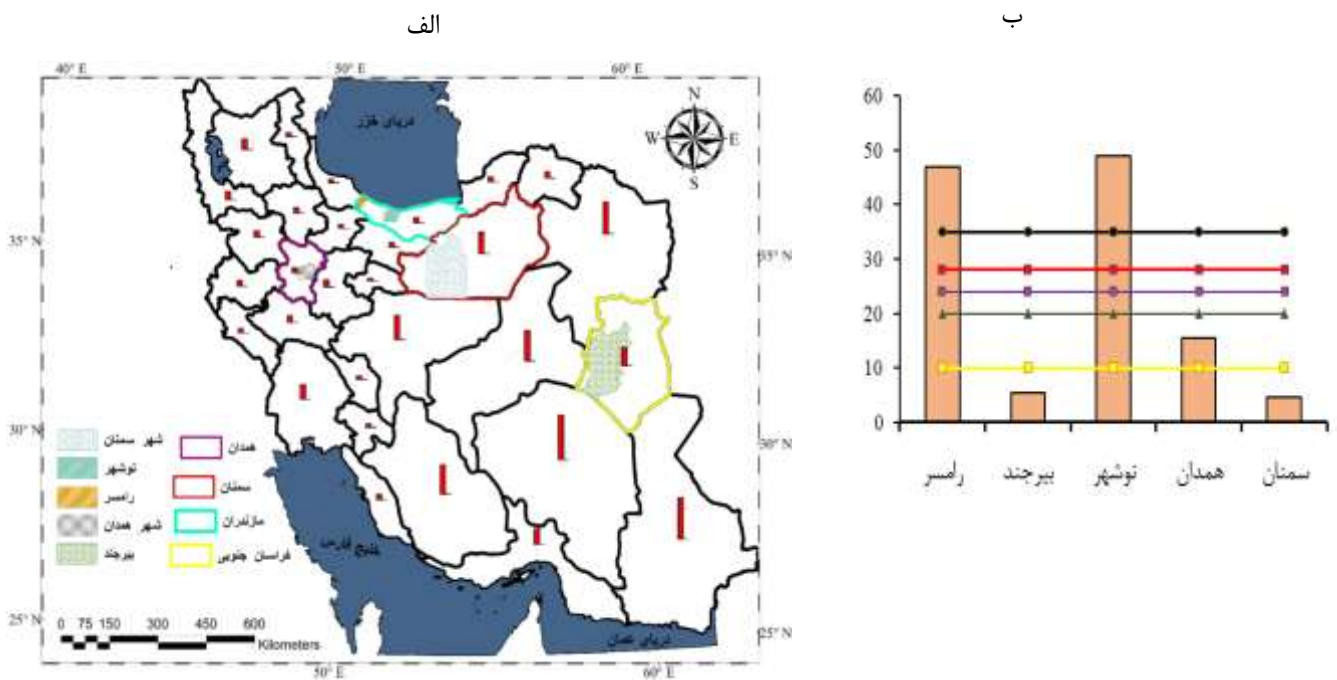
<sup>2</sup> Volvach

داده‌های مطالعاتی مربوط به ایستگاه‌های رامسر، نوشهر، همدان، سمنان و بیرجند می‌باشند که موقعیت جغرافیایی مربوطه در شکل ۱-الف نشان داده شده است. براساس روش دومارتن در شکل ۱-ب اقلیم ایستگاه‌های رامسر و نوشهر بسیار مرطوب، بیرجند و سمنان خشک و همدان نیمه‌خشک تعیین شد. در سال‌های اخیر تغییرات اقلیمی منجر به تاثیر در داده‌های هواشناسی بخصوص داده‌های دما شده است و نیاز به پایش این نوع داده‌ها کاملاً احساس می‌شود.

داده‌های مجموع ساعات آفتابی کم است. از سوی دیگر با توجه به وجود ناپایداری در سری‌های تجزیه شده سعی در رفع آن می‌باشد. هدف اصلی این مقاله بررسی پتانسیل تلفیق تجزیه سری زمانی با تبدیل نرمال است. در واقع سعی در پاسخگویی به این سوال است که آیا تبدیل نرمال برای پیش-بینی سری زمانی همراه با تجزیه سری زمانی مناسب است یا خیر؟

## ۲. مواد و روش‌ها

### ۲-۱. منطقه مورد مطالعه



شکل ۱. موقعیت مکانی مناطق بررسی شده در تحقیق (الف)، تعیین اقلیم ایستگاه‌ها با روش دومارتن (ب)

برداری از داده‌های ورودی و لا خروجی مربوطه است که فرض می‌شود از یک تابع مشخص تولید می‌شود. SVR به دنبال تقریب این تابع به صورت معادله ۱ است.

(۱)

$$y_i = f(x_i) = w^T x_i + b$$

$w$  و  $b$  به ترتیب نشان دهنده وزن و اریبی می‌باشند. تابع  $f$  در معادله ۱ معمولاً تابع تصمیم نامیده می‌شود. برای این مورد  $f$  یک رگرسیون خطی ساده را تعریف می‌کند.

الگوریتم پیشنهادی تحقیق شامل مدلسازی ساده با رگرسیون بردار پشتیبان، تجزیه سری زمانی با EMD و نرمال‌سازی است که روند نمای آن در شکل ۲ آورده شده است.

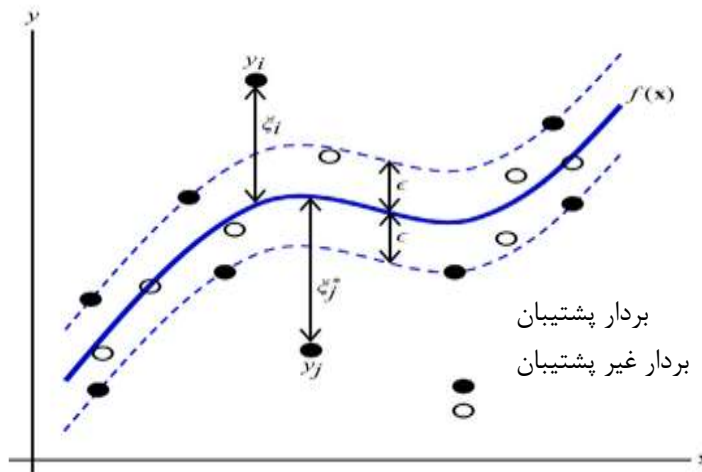
### ۲-۲. رگرسیون بردار پشتیبان (SVR)

رگرسیون بردار پشتیبان یک تکنیک محوری در زمینه پیش‌بینی است که به‌ویژه برای مدیریت داده‌های با ابعاد بالا و غیرخطی مناسب می‌باشد. فرض می‌شود که  $\{(x_i, y_i)\}_{i=1}^{N_1}$  مجموعه‌ای از داده‌های آموزشی هستند. در این حالت  $x_i$

(۲)

$$\ell_{\varepsilon}(y_i, f(x_i)) = \begin{cases} |y_i - f(x_i)| - \varepsilon, & |y_i - f(x_i)| > \varepsilon \\ 0 & |y_i - f(x_i)| \leq \varepsilon \end{cases}$$

$\ell_{\varepsilon}$ : تابع زیان برای تخمین بردار وزن و ثابت  $b$  به صورت کمینه سازی تابع براساس رابطه ۳ انجام می گیرد.



شکل ۳. نمایشی از بردار پشتیبان (موجتادی-المسیاه<sup>۱</sup> و همکاران، ۲۰۲۴).

همکاران (۱۹۹۸) پیشنهاد شده است. EMD یک روش تطبیقی فرکانس-زمان است که در تجزیه و تحلیل سری های زمانی غیرخطی و غیر ایستا استفاده می شود. این روش شامل تجزیه تطبیقی یک سیگنال به تعداد محدودی از توابع مد ذاتی (IMFs) است. IMF های به دست آمده یک دسته از توابع بر اساس ویژگی های محلی سیگنال اصلی هستند که آنها را کامل و تقریباً متعامد می کند. نشان داده شده است که روش EMD برای تجزیه و تحلیل سیگنال های غیرخطی و نایستا نسبت به روش هایی مانند تبدیل موجک یا فوریه مناسب تر است (استالون<sup>۳</sup> و همکاران، ۲۰۲۰). (et al., 2020). IMF ها باید دو شرط را رعایت کنند. اولین مورد

SVR بصورت خطی هر مشاهده را با یک خطای پیش بینی شده مطلق  $|y_i - f(x_i)| > \varepsilon$  جریمه می کند.  $\varepsilon$  یک پارامتر غیر منفی مشخص شده است. تفاوت بین خطا و  $\varepsilon$  با استفاده از تابع زیان  $\varepsilon$ - غیر حساس اندازه گیری می شود که در معادله ۲ و نمودار آن در شکل ۳ نشان داده شده است.

(۳)

$$\min R(C) = \frac{1}{2} w^T w + \frac{C}{N_1} \sum_{i=1}^{N_1} \ell_{\varepsilon}(y_i, f(x_i))$$

$\frac{1}{2} w^T w$ : همواری تابع،  $C$ : ثابتی که به عنوان تبادل بین خطای تقریبی و همواری عمل می کند (موجتادی-المسیاه و همکاران، ۲۰۲۴).

### ۳-۲. رویکرد تجزیه سری زمانی

در بین روش های تجزیه سری زمانی، روش نسبتاً جدیدی تحت عنوان تجزیه حالت تجربی (EMD) توسط هوانگ<sup>۲</sup> و

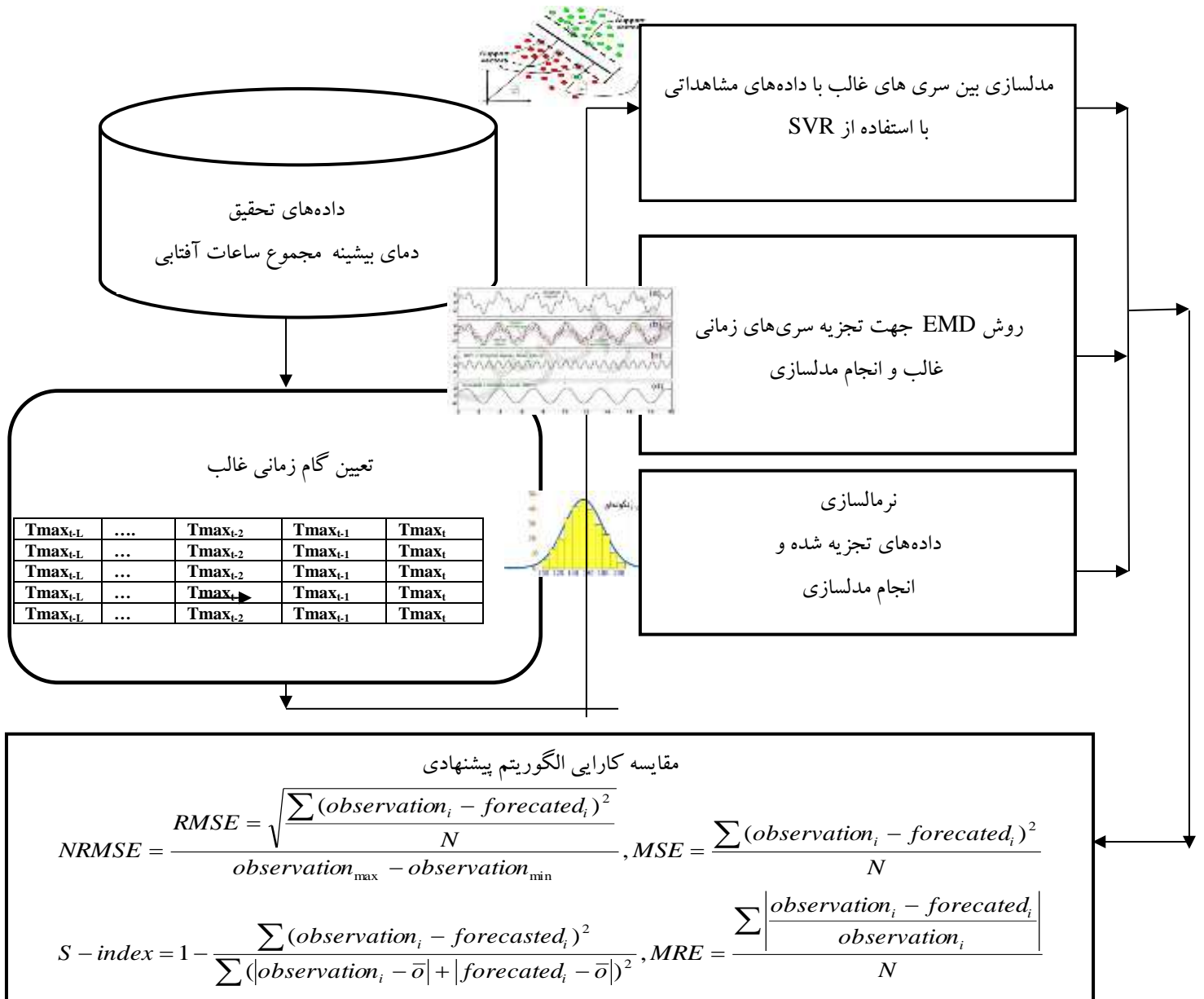
<sup>1</sup> Muchtadi-Alamsyah

<sup>2</sup> Huang

<sup>3</sup> Stallone

است که سیگنال حداقل دو اکستریم داشته باشد، یکی بیشینه و یکی کمینه، و مقیاس زمانی مشخصه با زمان بین اکستریم‌ها تعریف شود.

این است که تعداد گذرگاه‌های اکستریم و صفر باید برابر باشد یا بیش از یک برای کل مجموعه تغییر نکند. شرط دیگر این است که مقدار میانگین پوش از حداکثر محلی و پوش از حداقل محلی در هر لحظه صفر باشد. پیش شرط تجزیه این



شکل ۲. روندنمای انجام تحقیق

۱- تعیین اکستریم در داده‌ها (بیشینه و کمینه محلی)

مراحل روش تجزیه با EMD برای سری زمانی  $x(t)$  با طول  $t=1,2,3,\dots,k$

طول زمان تغییر می‌کند. با این حال، این روش‌ها فرض می‌کنند که تمام نقاط زمانی در یک نمونه دارای ویژگی‌های آماری یکسانی هستند، که بسیار ایده‌آل است و ممکن است منجر به بهبودهای نسبی کمتر از حد مطلوب شود. برای کاهش چنین مشکلی، تلاش‌های متعددی با کاهش عدم ایستایی با عملیات نرمال‌سازی انجام شده است. برای این منظور، یک نرمال‌سازی تطبیقی جدید پیشنهاد می‌شود (لیو<sup>۳</sup> و همکاران، ۲۰۲۴). نرمال‌سازی داده‌ها قبل از اجرای مدل تأثیر قابل توجهی بر عملکرد هر مدل دارد، زیرا هدف نرمال‌سازی داده‌ها، تضمین کیفیت داده‌ها قبل از ورود به مدل است. اثر بخشی هر الگوریتم یادگیری به شدت به روش نرمال‌سازی بستگی دارد. هدف اصلی روش نرمال‌سازی داده‌ها تولید داده‌هایی با کیفیت بالا است که می‌تواند به هر الگوریتم یادگیری وارد شود. داده‌های سری زمانی می‌توانند طیف وسیعی از مقادیر را داشته باشند، بنابراین باید به همان محدوده‌ای از مقادیر مقیاس‌بندی شوند. نرمال‌سازی یک مرحله پیش‌پردازش در سری زمانی است که شباهت مقایسات ثابت نسبت به انحرافات را تضمین می‌کند.

#### ۴-۲. الگوریتم پیشنهادی تحقیق

- ۱- تعیین دوره صحت‌سنجی و واسنجی سری زمانی
- ۲- تعیین گام زمانی غالب با تعیین ضریب همبستگی داده‌های مشاهداتی در زمان  $t$  با داده‌های مشاهداتی در تاخیرهای مختلف  $t-L, t-2, \dots, t-1$  (سری‌های غالب)
- ۳- مدل‌سازی بین سری‌های غالب با داده‌های مشاهداتی با استفاده از SVR (گروه اول)
- ۴- استفاده از روش EMD جهت تجزیه سری‌های زمانی غالب و انجام مدل‌سازی بین سری‌های تجزیه شده با داده‌های مشاهداتی با استفاده از SVR (گروه دوم)
- ۵- نرمال‌سازی داده‌های تجزیه شده و انجام مدل‌سازی بین سری‌های تجزیه شده نرمال شده با داده‌های مشاهداتی با استفاده از SVR (گروه سوم)

۲- تعیین پوش بالایی و پایینی با اتصال تمام ماکزیم‌های محلی و به طور جداگانه همه کمینه‌های محلی با یک خط اسپلاین مکعبی.

۳- تعیین میانگین پوش  $(m_1)$  و جزء اول  $(c_1)$  به عنوان

$$C_1 = x(t) - m_1$$

اگر شرایط تابع مد ذاتی (IMF) را برآورده کند، آنگاه  $c_1$  برابر با IMF1 است. در غیر این صورت، این روش تا زمانی که شرایط IMF برآورده شود،  $n$  بار تکرار می‌شود. در طول تکرارها،  $c_{1n}$  به این صورت در نظر گرفته می‌شود:

$$C_{1n} = C_{1(n-a)} - m_{1n}$$

اگر شرایط IMF فراهم شود:  $C_{1n} = IMF1$

IMF1 اولین مؤلفه‌ای است که از داده‌ها محاسبه می‌شود و باید دقیق‌ترین یا کوتاه‌ترین مؤلفه‌های دوره‌ای داده‌ها را شامل شود.

۴- جداسازی مؤلفه اول (IMF1) از داده‌ها:

$$r_1 = x(t) - IMF1$$

$r_1$ : داده جدید محسوب می‌شود. گام ۱ تا ۴،  $m$  بار تکرار می‌شود تا این که IMF $_m$  خیلی کوچک شود یا بیش از یک اکستریم موضعی نداشته باشد. در نهایت سری اصلی به صورت

$$x(t) = \sum_{m=1}^m IMF_m + r_m$$

نمایش داده می‌شود.

IMF های حاصل از EMD شامل فرکانس‌های بالا تا پایین هستند. هر IMF به دلیل دامنه‌ها و فرکانس‌های وابسته به زمان، کلی‌تر از یک تابع هارمونیک ساده است (آقنیزکا<sup>۱</sup> و دیوید<sup>۲</sup>، ۲۰۲۲).

مدل‌های یادگیری عمیق به دلیل ظرفیت قدرتمندشان در گرفتن وابستگی دنباله، پیش‌بینی سری‌های زمانی را به تدریج پیشرفته کرده‌اند. با این وجود، به دلیل وجود غیر ایستایی در داده‌های دنیای واقعی، انجام پیش‌بینی‌های دقیق هنوز چالش برانگیز است، که نشان می‌دهد توزیع داده‌ها به سرعت در

<sup>3</sup> Liu

<sup>1</sup> Agnieszka

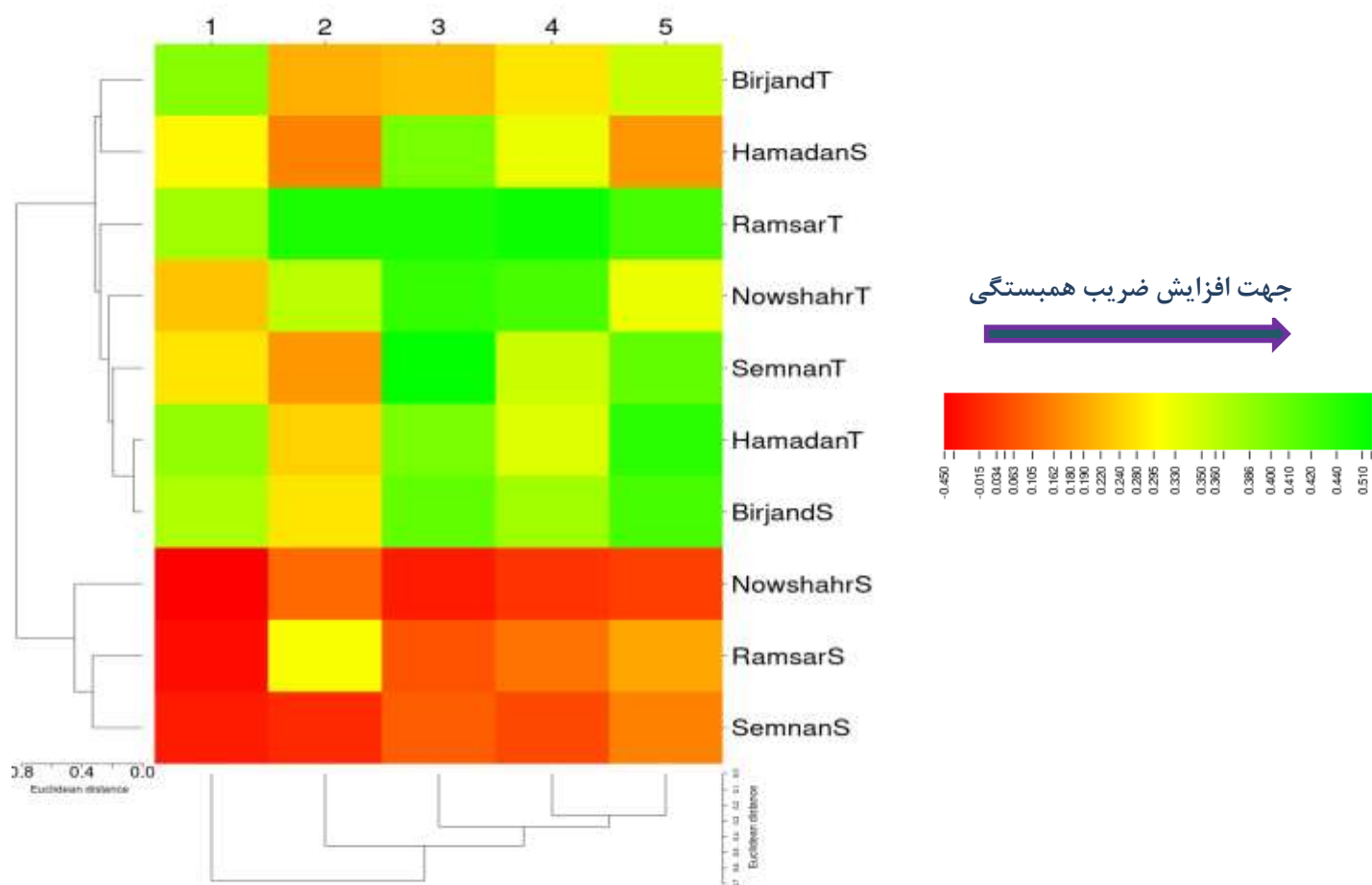
<sup>2</sup> Dawid

۶- بکارگیری آماره‌های ارزیابی در بررسی عملکرد گروه اول، دوم و سوم

### ۳. نتایج و بحث

داده‌های مورد استفاده تحقیق دمای بیشینه و مجموع ساعات آفتابی در بازه زمانی ۱۹۹۰ تا ۲۰۲۳ در نظر گرفته شدند. تقسیم‌بندی داده‌ها به دو مرحله واسنجی و صحت‌سنجی براساس نسبت ۸۰:۲۰ بود. بیشینه مقدار متوسط دمای بیشینه در بازه زمانی مورد بررسی مربوط به ایستگاه‌های بیرجند و سمنان و کمینه مربوط به ایستگاه‌های رامسر و همدان است. دلیل این که در دو حالت بیشینه و کمینه، دو ایستگاه انتخاب شده به خاطر تفاوت جزئی دمای بیشینه و کمینه بین دو

ایستگاه می‌باشد. بیشینه مقدار متوسط مجموع ساعات آفتابی در بازه زمانی مورد بررسی مربوط به ایستگاه‌های بیرجند و سمنان و کمینه مربوط به ایستگاه همدان است. در مرحله بعد اقدام به مدلسازی دمای بیشینه و مجموع ساعات آفتابی خواهد شد که شامل تعیین دقیق داده‌های ورودی به مدل و پارامترهای مدل است. یکی از عواملی که می‌تواند صحت مدلسازی را تحت تاثیر قرار دهد، نوع داده‌های ورودی به مدل است. بنابراین در مدلسازی دمای بیشینه از ۵ تاخیر زمانی استفاده شد و ضریب همبستگی داده‌های مشاهداتی با داده‌های تاخیری محاسبه شد و سه تاخیر با بیشترین ضریب همبستگی به عنوان داده‌های ورودی در نظر گرفته شدند. مقادیر ضریب همبستگی در شکل ۴ آورده شده است.



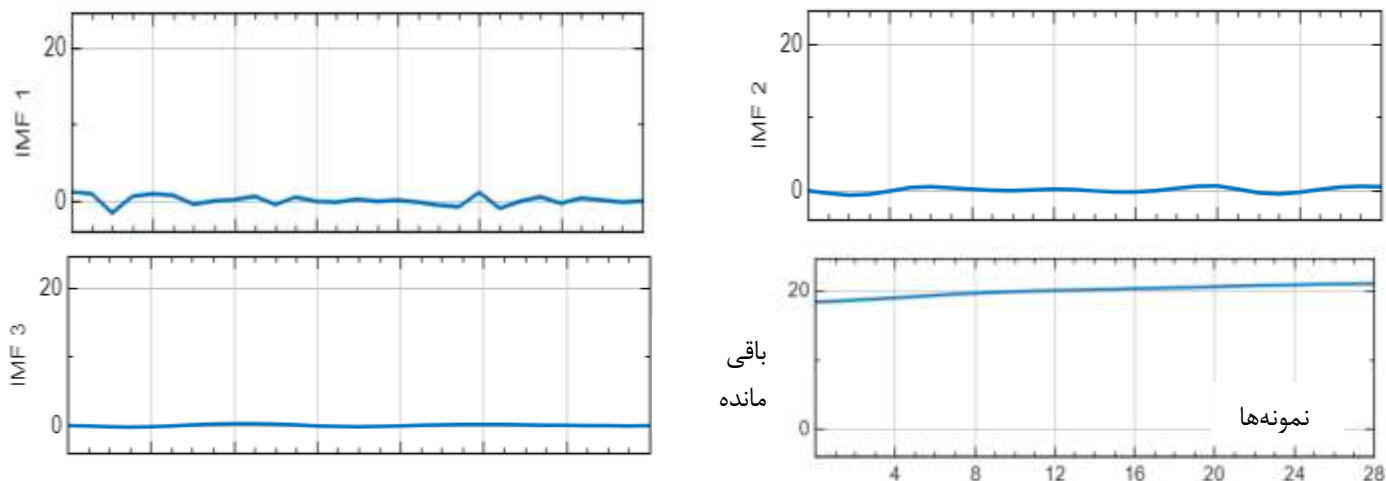
شکل ۴. نمودار heat map در نمایش ضریب همبستگی سری اصلی با ۵ تاخیر زمانی مختلف در داده‌های دمای بیشینه (T) و مجموع ساعات آفتابی (S)

مورد داده‌های دمای بیشینه در ایستگاه‌های رامسر و نوشهر تاخیرهای ۲، ۳ و ۴ در ایستگاه همدان تاخیرهای ۱، ۴ و ۵ در

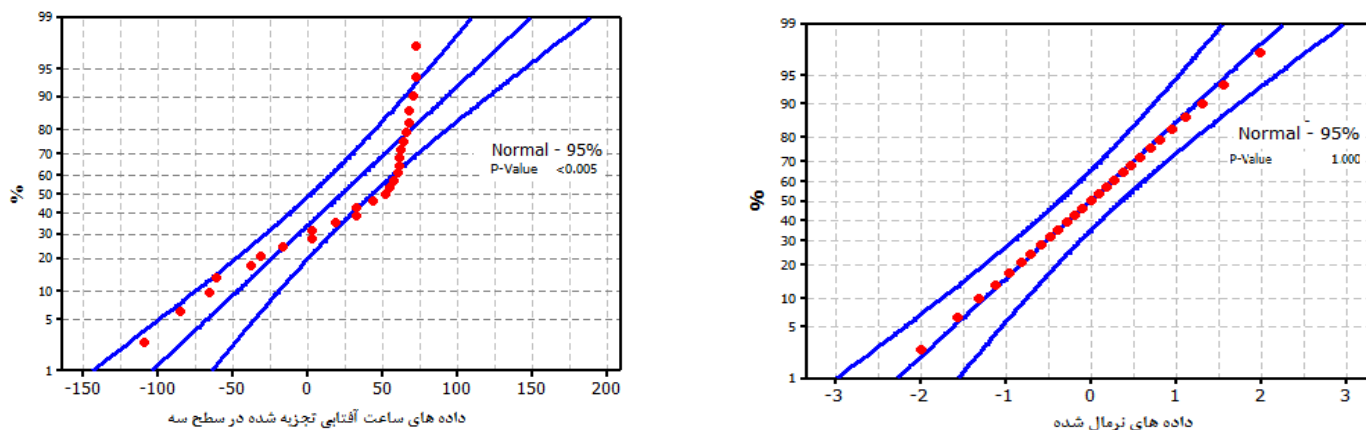
همان طور که در شکل ۴ مشاهده می‌شود، ضریب همبستگی تاخیرهای مختلف در ایستگاه‌ها با هم متفاوت هستند. در

ادامه با تاخیرهای انتخابی سه نوع مدل سازی با مدل رگرسیون بردار پشتیبان (تحلیل حساسیت مدل بر روی توابع کرنل و ضریب پنالتی انجام پذیرفت) انجام می گیرد: الف) داده های ورودی شامل سه سری زمانی با تاخیر انتخابی براساس بیشترین مقدار ضریب همبستگی ب) داده های ورودی شامل تجزیه سه سری با تاخیر زمانی غالب با EMD: تحلیل حساسیت بر روی نوع درون یابی، پیشینه مقدار IMF انجام گرفت (شکل ۵) ج) داده های ورودی شامل نرمال سازی داده های تجزیه شده گام ب با تابع Normal scores در نرم افزار Minitab16 (شکل ۶).

ایستگاه سمنان تاخیرهای ۳،۴ و ۵ در ایستگاه بیرجند تاخیرهای ۱،۳ و ۵ دارای بیشترین مقدار ضریب همبستگی هستند. این مساله تاثیر تاخیرهای مختلف را بر نوع مدل سازی نشان می دهد. در ایستگاه های با اقلیم بسیار مرطوب رامسر و نوشهر تاخیر ۵ در سه تاخیر انتخاب شده نیست و این تاخیر در ایستگاه های خشک و نیمه خشک وارد سه تاخیر اول می شود. در مورد داده های مجموع ساعات آفتابی نیز این مساله مشاهده می شود، به طوری که سه تاخیر اول در ایستگاه های رامسر و نوشهر ۲، ۳ و ۵ است، در حالی که در ایستگاه های بیرجند و سمنان سه تاخیر شامل ۳، ۴ و ۵ می باشد. این مساله اهمیت داده های ورودی را براساس اقلیم نشان می دهد. در



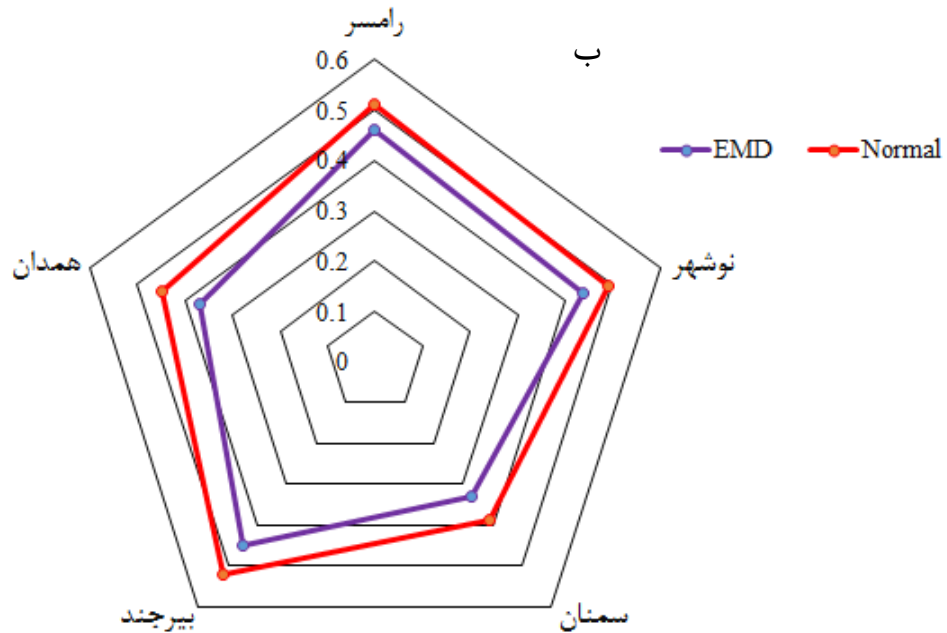
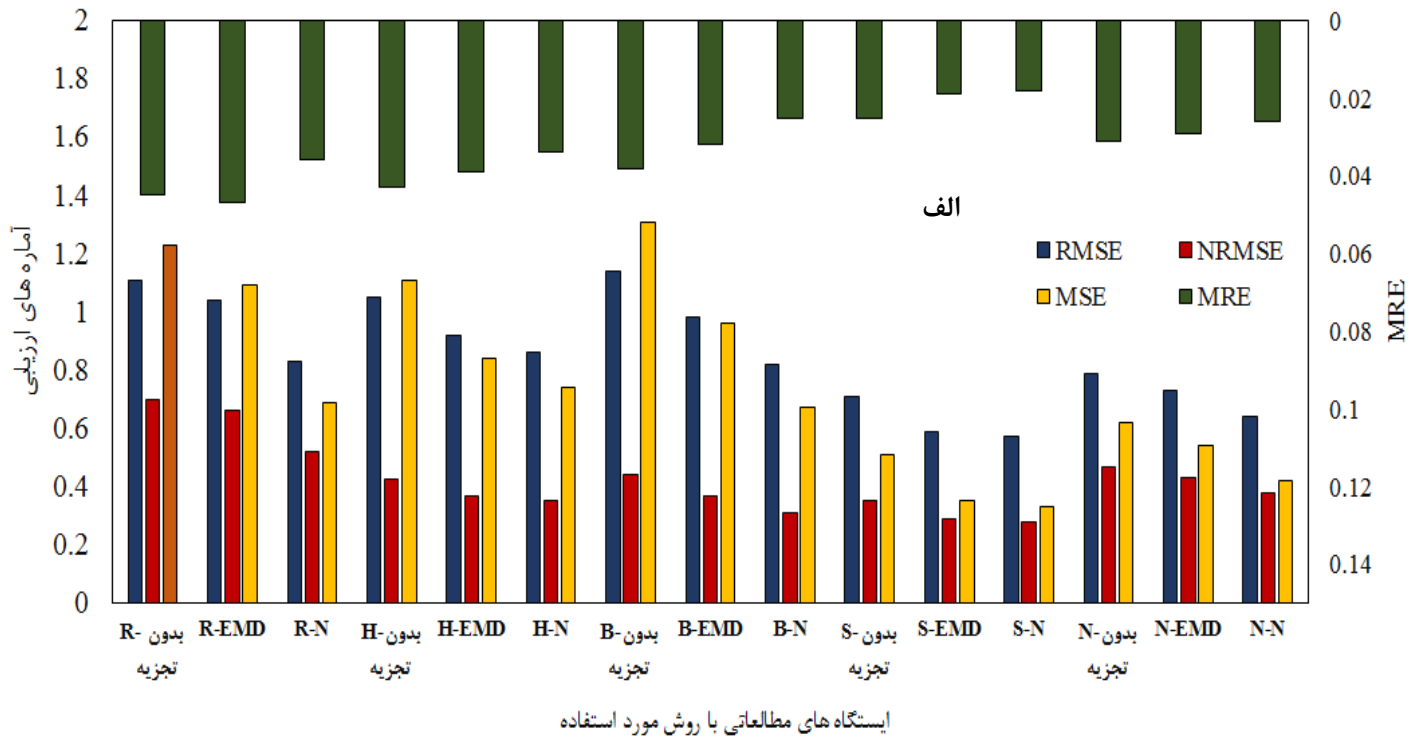
شکل ۵. زیر سری های حاصل از EMD برای دمای بیشینه در همدان برای تاخیر زمانی ۵



شکل ۶. نمودار احتمالاتی داده های تجزیه شده ساعت آفتابی سری اول، سطح تجزیه سه و داده های نرمال شده در ایستگاه رامسر

که با استفاده از نرمال‌سازی سری، این نقاط در محدوده توزیع نرمال قرار گرفته‌اند. در ادامه آماره‌های ارزیابی جهت مقایسه سه نوع مدل‌سازی در شکل ۷ آورده شده است.

براساس شکل ۵ با افزایش تعداد IMFها از نوسانات سری کاسته می‌شود تا حدی که به حالت خطی می‌رسد. این تجزیه شامل سه سطح تجزیه است. در شکل ۶ داده‌های نمودار احتمالاتی سری اصلی خارج از محدوده توزیع نرمال است



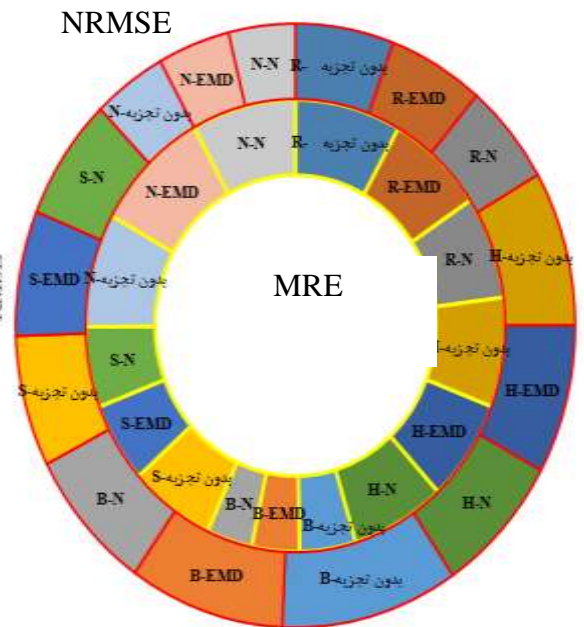
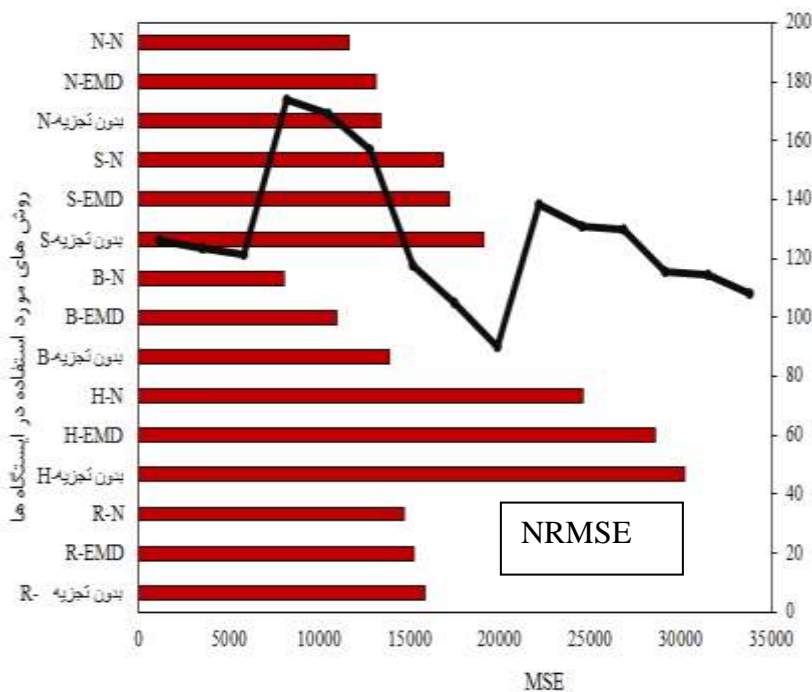
شکل ۷. آماره‌های ارزیابی سه روش مورد استفاده، بدون تجزیه، تجزیه با EMD و نرمال‌سازی در تمام ایستگاه‌ها (الف)، مقایسه مقدار S در بررسی عملکرد دو روش EMD و نرمال‌سازی (R: نماد رامسر، H: همدان، B: بیرجند، N: نوشهر، S: سمنان)

مدلسازی گروه اول به مدلسازی نرمال به ترتیب برابر با ۱۵/۷۸، ۱۸/۱۸، ۱۸/۴۲ و ۲۹/۷۷ بود. میزان کاهش آماره‌های خطا از مدلسازی گروه اول به گروه سوم بیشتر از مدلسازی گروه اول به مدلسازی نرمال سری ۱ می‌باشد و این مساله اهمیت نرمال کردن سری‌های زمانی تجزیه شده را نشان می‌دهد. در مورد ایستگاه همدان اختلاف آماره‌ها در حالت تجزیه و نرمال شده کم می‌باشد و بهبودی آنچنانی در سری‌های نرمال شده مشاهده نشد. به همین دلیل به بررسی نرمال بودن سری‌های تجزیه شده پرداخته شد که زیرسری‌های غیر نرمال شامل زیر سری ۴، ۷ و ۱۲ بودند. در گام بعدی ضریب همبستگی زیر سری‌های ۷، ۴ و ۱۲ در حالت تجزیه با داده‌های مشاهداتی (حالت الف) و سپس ضریب همبستگی زیر سری‌های نرمال شده ۷، ۴ و ۱۲ حالت (ب) با داده‌های مشاهداتی محاسبه شدند. ضریب همبستگی زیر سری‌های ۴ و ۷ در حالت ب نسبت به الف همراه با افزایش ۵ و ۸۰ درصد بود. به همین خاطر بجای نرمال‌سازی کل زیر سری‌ها از زیر سری نرمال ۷ (حالت ج) استفاده شد که میزان کاهش RMSE از مدلسازی ساده به مدلسازی با تجزیه و مدلسازی با نرمال‌سازی به ترتیب برابر با ۱۲/۳۸ و ۱۸/۰۹ درصد، در مورد NRMSE به ترتیب برابر با ۱۳/۵۵ و ۱۸/۲۲ درصد، در مورد MRE به ترتیب برابر با ۹/۳ و ۲۰/۹۳ درصد و در مورد MSE به ترتیب برابر با ۲۴/۳۲ و ۳۳/۳۳ درصد می‌باشد. در مورد دیگر از سری نرمال ۴ و ۷ (حالت د) استفاده شد که میزان کاهش RMSE، NRMSE، MRE و MSE از مدلسازی ساده به نرمال شده به ترتیب برابر با ۱۳/۳۳، ۱۵/۸۸، ۱۱/۶۲ و ۲۵/۲۲ درصد است. کامل مشخص است که میزان کاهش در حالت د نسبت به حالت ج کمتر است و این مساله اهمیت نرمال شدن سری ۷ که همبستگی بالایی با داده مشاهداتی داشت را نشان می‌دهد. در مورد ایستگاه سمنان نیز تفاوت فاحشی بین مدلسازی با تجزیه سری زمانی و مدلسازی با سری تجزیه شده نرمال مشاهده نشد. بنابراین نرمال بودن سری‌های تجزیه شده بررسی شد و زیر سری‌های ۳، ۵، ۸ و ۱۲ نرمال نبودند، به همین خاطر فقط زیر

براساس شکل ۷، در ایستگاه‌های رامسر، نوشهر و بیرجند تجزیه سری زمانی و نرمال‌سازی تمام سری‌های تجزیه شده دقت پیش‌بینی را بالا برده است، به طوری که در ایستگاه رامسر میزان کاهش RMSE از مدلسازی ساده به تجزیه سری زمانی و از مدلسازی ساده به نرمال‌سازی به ترتیب ۶/۳ و ۲۵/۲۲ درصد، در ایستگاه نوشهر برابر با ۷/۵۹ و ۱۸/۹۸ درصد و در ایستگاه بیرجند برابر با ۱۴/۰۳ و ۲۸/۰۷ درصد بود. میزان کاهش NRMSE در ایستگاه رامسر برابر با ۵/۷۱ و ۲۵/۷۱ درصد، در ایستگاه نوشهر برابر با ۸/۵۱ و ۱۹/۱۴ درصد و در ایستگاه بیرجند برابر با ۱۵/۹ و ۲۹/۵۴ درصد می‌باشد. میزان کاهش MSE در ایستگاه رامسر برابر با ۱۱/۳۸ و ۴۳/۹ درصد، در ایستگاه نوشهر برابر با ۱۲/۹ و ۳۲/۲۵ درصد و در ایستگاه بیرجند برابر با ۲۶/۷۱ و ۴۸/۸۵ درصد بود. نتایج بیان شده برای زمانی بود که تمام سری‌های تجزیه شده نرمال شوند. در بررسی بیشتر از آزمون کولموگروف-اسمیرنوف برای تعیین وضعیت نرمال بودن سری‌های تجزیه شده ایستگاه رامسر استفاده شد که سری ۶ (سطح معنی‌داری ۰/۰۱) و ۹ (۰/۰۲۶) از سری نرمال فاصله داشتند. با توجه به میزان سطح معنی‌داری در مدلسازی دیگر فقط از سری نرمال شده ۶ استفاده شده که میزان کاهش RMSE، NRMSE، MRE و MSE از حالت مدلسازی گروه اول به مدلسازی نرمال برابر با ۱۱/۷۱ و ۱۱/۴۲، ۲/۲۲ و ۲۱/۹۵ درصد بود که اگر مقایسه‌ای با میزان درصدهای کاهش با مدلسازی با تمام سری‌های تجزیه شده نرمال انجام گیرد مشاهده می‌شود میزان کاهش در حالت اخیر نسبت به نرمال تنها با زیرسری ۶ بیشتر است و این اهمیت نرمال‌سازی زیر سری ۹ را در روند مدلسازی نشان می‌دهد. در ایستگاه بیرجند نیز بررسی بیشتری بر روی روند نرمال‌سازی انجام گرفت، به این صورت که با آزمون کولموگروف-اسمیرنوف تمام سری‌های تجزیه شده بررسی نرمال بودن را انجام دادند که سری‌های ۱-۴-۶ و ۸ نرمال بودند. حال در مدلسازی فقط از زیر سری نرمال ۱ استفاده شده که میزان کاهش RMSE، NRMSE، MRE و MSE از حالت

ایستگاه‌ها از مدل‌سازی بدون تجزیه به مدل‌سازی با نرمال‌سازی سری‌های تجزیه شده برابر به ترتیب برابر با ۲۲/۰۱ و ۲۲/۵۲ است. مقایسه مقادیر کاهش RMSE و NRMSE در دو حالت اخیر در تمام ایستگاه‌ها نشان داد که نرمال‌سازی سری‌های تجزیه شده نسبت به فقط تجزیه سری زمانی افزایش دارد. همچنین میزان افزایش شاخص S از مدل‌سازی با تجزیه سری زمانی به مدل‌سازی با زیر سری‌های نرمال به ترتیب در ایستگاه‌های رامسر، همدان، بیرجند، سمنان و نوشهر برابر با ۱۰/۸۶، ۲۱/۶۲، ۱۵/۵۵، ۱۸/۱۸ و ۱۱/۳۶ درصد است. در ادامه شکل ۸ آماره‌های خطا را برای داده‌های مجموع ساعت آفتابی نشان می‌دهد.

سری ۳ نرمال شده در روند مدل‌سازی استفاده شد. میزان کاهش RMSE از مدل‌سازی با سری‌های اصلی به مدل‌سازی با سری‌های تجزیه شده و نرمال شده به ترتیب برابر با ۱۶/۹ و ۱۹/۷۱ درصد، میزان کاهش NRMSE برابر با ۱۷/۱۴ و ۲۰ درصد، میزان کاهش MRE برابر با ۲۴ و ۲۸ درصد و میزان کاهش MSE برابر با ۳۱/۳۷ و ۳۵/۲۹ درصد بود. دقت پیش‌بینی با تجزیه سری زمانی افزایش یافت، به طوری که متوسط کاهش RMSE و NRMSE در کل ایستگاه‌ها از مدل‌سازی بدون تجزیه به استفاده از تجزیه به ترتیب برابر با ۱۱/۴۴ و ۱۶/۱۲ درصد می‌باشد. علاوه بر تجزیه سری زمانی، نرمال‌سازی سری زمانی توانست دقت مدل‌سازی را افزایش دهد، میزان متوسط کاهش RMSE و NRMSE در کل



شکل ۸. بررسی تاثیر نرمال کردن سری‌های تجزیه شده ساعات آفتابی با آماره‌های ارزیابی در ایستگاه‌های مختلف

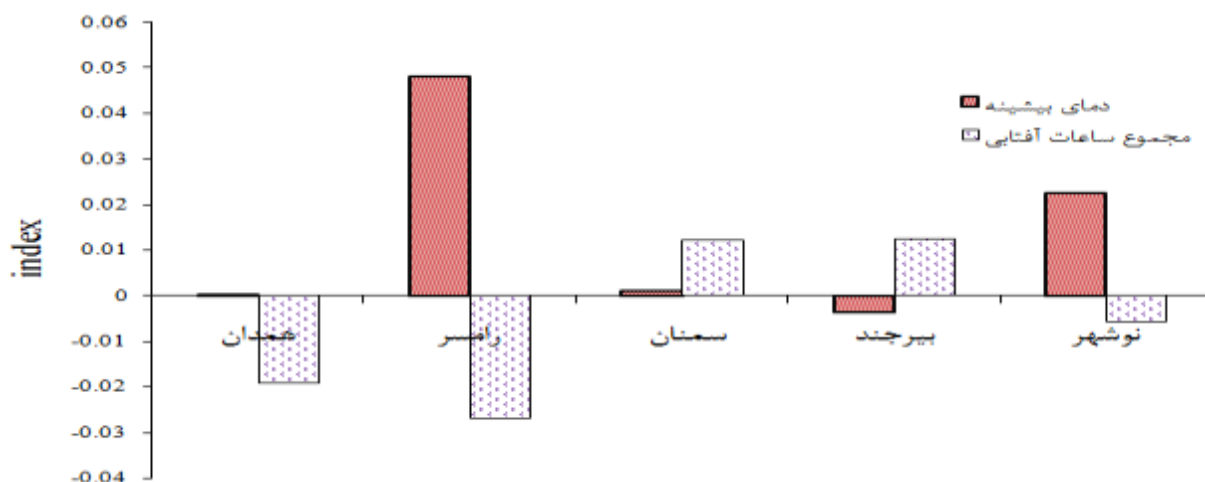
و نرمال‌سازی سری‌های تجزیه شده برابر با ۹/۷۹، ۹/۶۱ و ۱۸/۶۲، در ایستگاه نوشهر با مدل‌سازی بدون تجزیه به مدل‌سازی با تجزیه برابر با ۰/۸۷، ۲/۵۵ و ۱/۷۲ نرمال‌سازی سری‌های تجزیه شده برابر با ۶/۷، ۹/۷۷ و ۱۲/۹۵ است. در

در مورد داده‌های ساعات آفتابی هم نرمال‌سازی سری‌های تجزیه شده روند پیش‌بینی را توسعه داد. در ایستگاه همدان میزان کاهش RMSE، NRMSE و MSE از مدل‌سازی بدون تجزیه به مدل‌سازی با تجزیه برابر با ۲/۷۵، ۳/۸۴ و ۵/۴۲

زیرسری‌های غیر نرمال را با تابع مربوطه نرمال کرده و ترتیب جایگذاری در روند مدل‌سازی بررسی شد (۱: استفاده از زیر سری ۶ نرمال شده، ۲: استفاده از زیر سری ۸ نرمال شده، ۳: استفاده از زیرسری ۱ نرمال شده، ۴: استفاده از زیرسری ۲ نرمال شده، ۵- استفاده از نرمال زیرسری های ۱-۲-۶-۸). میزان کاهش RMSE از حالت اصلی به حالت‌های ۱، ۲، ۳، ۴ و ۵ به ترتیب برابر با ۵/۳۲، ۵/۹۴، ۵/۱۳، ۴/۸۸ و ۶/۱ درصد است. به طور کلی مدل‌سازی با تجزیه سری‌های اصلی توانست دقت مدل‌سازی را افزایش دهد که این نتیجه در تحقیق شاهی نژاد و دهقانی (۱۳۹۵) و لخواج و همکاران (۲۰۲۴) هم مشاهده شد. از سوی دیگر نرمال‌سازی داده‌های تجزیه شده دقت پیش‌بینی را افزایش داد که تاثیر نرمال‌سازی در روند تجزیه سری زمانی در تحقیق کانگ و همکاران (۲۰۲۳) در مورد سری رواناب مشاهده شد.

اگر متوسط NRMSE در مورد مدل‌سازی دمای بیشینه با نرمال‌سازی سری‌های تجزیه شده در تمام ایستگاه‌ها براساس اقلیم محاسبه شود، مقدار NRMSE در اقلیم خشک، نیمه-خشک و مرطوب به ترتیب برابر با ۰/۲۹۵، ۰/۴۵ و ۰/۳۵ است. در مورد مجموع ساعات آفتابی مقدار NRMSE در اقلیم خشک، نیمه‌خشک و مرطوب به ترتیب برابر با ۰/۴۲۵، ۰/۴۷ و ۰/۲۹ است. الگوریتم پیشنهادی در دمای بیشینه و مجموع ساعات آفتابی به ترتیب در اقلیم خشک و بسیار مرطوب دارای کمینه خطا بود. شکل ۹ شاخصی (index) را به صورت اختلاف بین یک و تقسیم مجموع مقادیر پیش‌بینی شده بر مجموع داده‌های مشاهداتی برای تمام ایستگاه‌ها نشان می‌دهد.

ایستگاه رامسر میزان کاهش MRE،NRMSE،RMSE و MSE از مدل‌سازی بدون تجزیه به مدل‌سازی با تجزیه به ترتیب برابر با ۳/۲، ۰۹/۰۶، ۷/۸۴ و ۴/۰۴ است و در صورت نرمال‌سازی تمامی زیرسری‌ها بهبودی مشاهده نشد. به همین دلیل آزمون نرمال بر روی سری‌های تجزیه شده انجام گرفت که زیر سری‌های ۳ و ۱۰ نرمال نبودند و در این حالت از نرمال شده این دو زیر سری در روند مدل‌سازی استفاده شد. میزان کاهش MRE،NRMSE،RMSE و MSE از مدل‌سازی بدون تجزیه به مدل‌سازی با نرمال‌سازی اخیر شامل ۳/۶۲، ۳/۶۶، ۹/۸ و ۷/۱۱ بود. در ایستگاه بیرجند میزان کاهش MRE،NRMSE،RMSE و MSE از مدل‌سازی بدون تجزیه به مدل‌سازی با تجزیه به ترتیب برابر با ۱۱/۲۲، ۱۲/۵، ۱۴/۸۱ و ۲۱/۱۹ و از مدل‌سازی بدون تجزیه به نرمال-سازی شامل ۲۳/۷۳، ۲۵، ۱۸/۵۱ و ۴۱/۸۲ بود. در بیرجند نرمال‌سازی زیر سری‌ها باعث بهبود عملکرد شد، ولی برای بررسی بیشتر تنها زیر سری ۶ که نرمال نبود برای نرمال‌سازی استفاده شد که میزان کاهش RMSE و NRMSE نسبت به حالت بدون تجزیه و نرمال برابر با ۳۰/۶۱ و ۳۱/۲۵ درصد است. بنابراین نوع زیرسری انتخابی برای نرمال‌سازی می‌تواند نتایج پیش‌بینی را تحت تاثیر قرار دهد. در ایستگاه سمنان میزان کاهش MRE،NRMSE،RMSE و MSE از مدل‌سازی بدون تجزیه به مدل‌سازی با تجزیه برابر با ۵/۱۹، ۴/۵۴، ۱۲/۱۹ و ۱۱/۱ می‌باشد. نرمال‌سازی سری‌های تجزیه شده نتوانست دقت پیش‌بینی سری زمانی را افزایش دهد. به همین خاطر آزمون نرمال بودن زیر سری‌ها انجام گرفت که زیر سری‌های ۱، ۲، ۶ و ۸ نرمال نبودند، به همین خاطر



شکل ۹. مقادیر شاخص جهت تعیین بیش و کم برآورد الگوریتم پیشنهادی برای تمام داده‌ها و ایستگاه‌ها

در مورد دمای بیشینه بجز ایستگاه بیرجند بقیه ایستگاه‌ها دارای کم برآورد هستند. در مورد داده‌های مجموع ساعت آفتابی فقط ایستگاه‌هایی که در اقلیم خشک (ایستگاه سمنان و بیرجند) قرار دارند، دارای کم برآورد هستند. نوع اقلیم با توجه به تعداد ایستگاه‌های دارای کم و بیش برآورد، می‌تواند عملکرد مدل‌سازی را تحت تاثیر قرار دهد.

#### ۴. نتیجه‌گیری

تغییرات اقلیمی حاکم در سال‌های اخیر و اثرات آن مانند کاهش منابع آبی و پوشش گیاهی، پایش و پیش‌بینی برخی داده‌های هواشناسی از جمله دمای بیشینه را ضروری ساخته است. در این راستا الگوریتمی بر مبنای تلفیق تجزیه سری زمانی با تبدیل نرمال ارائه شد. یکی از مواردی که می‌تواند دقت مدل را تحت تاثیر قرار دهد، نوع داده‌های ورودی است که براساس بیشینه مقدار ضریب همبستگی با داده مشاهداتی انجام گرفت. این مساله اهمیت داده‌های ورودی را براساس اقلیم نشان می‌دهد که داده‌های ورودی می‌توانند نوسانات داده‌های هواشناسی را نشان دهند. نرمال کردن زیر سری‌های تجزیه شده توانست دقت پیش‌بینی را نسبت به مدل‌سازی ساده و مدل‌سازی با تجزیه سری زمانی افزایش دهد. نرمال‌سازی تاثیرات مشخصه‌های چولگی بر پیش‌بینی را کاهش می‌دهد.

در واقع نرمال‌سازی می‌تواند بر خطای ناشی از چولگی داده‌های مطالعاتی غلبه کند. مساله مهم دیگر در روند نرمال‌سازی زیر سری‌های تجزیه شده، نوع زیرسری است که باید نرمال شود، به طوری که تغییر نوع زیر سری نرمال شده می‌تواند نتایج پیش‌بینی را تحت تاثیر قرار دهد. الگوریتم پیشنهادی در دمای متوسط و مجموع ساعات آفتابی به ترتیب در اقلیم خشک و بسیار مرطوب دارای کمترین خطا بود. در مورد دمای بیشینه بجز ایستگاه بیرجند بقیه ایستگاه‌ها دارای کم برآورد و در مورد داده‌های مجموع ساعت آفتابی فقط ایستگاه‌ها در اقلیم خشک دارای تخمین کم برآورد هستند. با توجه به اهمیت نرمال‌سازی زیر سری‌های تجزیه شده یکی از پیشنهادات تحقیق می‌تواند استفاده از انواع تبدیلات جهت نرمال‌سازی و بررسی تاثیر آن بر مقادیر پیش‌بینی باشد. به طور کلی الگوریتم پیشنهادی می‌تواند به عنوان رهیافتی قوی دقت پیش‌بینی داده‌های هواشناسی را افزایش دهد که منجر به تصمیم‌گیری‌های با اعتبار بالا در تحقیقات تغییرات آب و هوایی، بهداشت عمومی، کشاورزی، مصرف انرژی و برنامه‌ریزی شهری شود.

#### منابع

1. Pezeshki, Z., Geraylo, H., & Soleimani Ivvari, S. A. (2018). Noise removal of temperature using wavelet transform and temperature prediction with SVM and

- inverse wavelet method. *Scientific Journal of Research in Computer Sciences*, (10), 1–18. (In Persian)
2. Roshangar, K., & Abdolzadeh, S. (2023). River discharge prediction using a hybrid long short-term memory, wavelet transform, and empirical mode decomposition approach in semi-arid and humid climates. *Iranian Journal of Irrigation and Drainage*, 17(4), 703–717. (In Persian)
  3. Shahi Nejad, B., & Dehghani, R. (2016). Application of wavelet neural networks in estimating mean air temperature of Sari County. *Climatological Research*, 7(27), 75–86. (In Persian)
  4. Salehi, S. M., Radmanesh, F., Zarei, H., Mansouri, B., & Selgi, A. (2018). Groundwater level prediction using a hybrid time series–wavelet model (Case study: Firuzabad Plain). *Irrigation Sciences and Engineering*, 41(4), 1–16. (In Persian)
  5. Agnieszka, W., & Dawid, K. (2022). Modeling seasonal oscillations in GNSS time series with complementary ensemble empirical mode decomposition. *GPS Solutions*, 26(4), 101. <https://doi.org/10.1007/s10291-022-01300-1>
  6. Cho, D., Yoo, C., Son, B., Im, J., Yoon, D., & Cha, D. H. (2022). A novel ensemble learning for post-processing of NWP model's next-day maximum air temperature forecast in summer using deep learning and statistical approaches. *Weather and Climate Extremes*, 35, 100410. <https://doi.org/10.1016/j.wace.2021.100410>
  7. Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903–995. <https://doi.org/10.1098/rspa.1998.0193>
  8. Kang, Y., Cheng, X., Chen, P., Zhang, S., & Yang, Q. (2023). Monthly runoff prediction by a multivariate hybrid model based on decomposition-normality and Lasso regression. *Environmental Science and Pollution Research*, 30(10), 27743–27762. <https://doi.org/10.1007/s11356-022-24145-3>
  9. Licer, M., Belmajdoub, H., & Minaoui, K. (2025). Enhanced deep learning approach for improved maximum temperature forecasting: A case study in the Sahara region, Morocco. *Theoretical and Applied Climatology*, 156(1), 1–14. <https://doi.org/10.1007/s00704-024-04628-3>
  10. Lin, M. L., Tsai, C. W., & Chen, C. K. (2021). Daily maximum temperature forecasting in changing climate using a hybrid of multi-dimensional complementary ensemble empirical mode decomposition and radial basis function neural network. *Journal of Hydrology: Regional Studies*, 38, 100923. <https://doi.org/10.1016/j.ejrh.2021.100923>
  11. Liu, Z., Cheng, M., Li, Z., Huang, Z., Liu, Q., Xie, Y., & Chen, E. (2024). Adaptive normalization for non-stationary time series forecasting: A temporal slice perspective. *Advances in Neural Information Processing Systems*, 36.
  12. Muchtadi-Alamsyah, I., Viltoriano, R., Harjono, F., Nazaretha, M., Susilo, M., Bayu, A., Josaphat, B., Hakim, A., & Syuhada, K. (2024). Support vector regression–based heteroscedastic models for cryptocurrency risk forecasting. *Applied Soft Computing*, 111792. <https://doi.org/10.1016/j.asoc.2024.111792>
  13. Salarijazi, M., Ghorbani, K., Mohammadi, M., Ahmadianfar, I., Mohammadrezapour, O., Naser, M. H., & Yaseen, Z. M. (2023). Spatial–temporal estimation of maximum temperature high return periods for annual time series considering stationary and nonstationary approaches in Iranian urban areas. *Urban Climate*, 49, 101504. <https://doi.org/10.1016/j.uclim.2023.101504>
  14. Shamshirband, S., Mohammadi, K., Khorasanizadeh, H., Yee, L., Lee, M., Petković, D., & Zalnezhad, E. (2016). Estimating the diffuse solar radiation using

- a coupled support vector machine–wavelet transform model. *Renewable and Sustainable Energy Reviews*, 56, 428–435. <https://doi.org/10.1016/j.rser.2015.11.055>
15. Stallone, A., Cicone, A., & Materassi, M. (2020). New insights and best practices for the successful use of empirical mode decomposition, iterative filtering, and derived algorithms. *Scientific Reports*, 10(1), 15161. <https://doi.org/10.1038/s41598-020-72166-6>
  16. Velivelli, S., Satyanarayana, G. C., & Ali, M. M. (2024). Predicting daily maximum temperature over Andhra Pradesh using machine learning techniques. *Theoretical and Applied Climatology*, 155(9), 8567–8585. <https://doi.org/10.1007/s00704-023-04677-8>
  17. Volvach, A., Kurbasova, G., & Volvach, L. (2024). Wavelets in the analysis of local time series of the Earth's surface air temperature. *Heliyon*, 10(1), e23741. <https://doi.org/10.1016/j.heliyon.2024.e23741>