



Performance evaluation of signal decomposition methods in monthly precipitation estimation (Case study: Telezang station)

Nazanin Alivand¹ | Ali Mohammad Akhoond-Ali^{2✉} | Farshad Ahmadi³

1. Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran. E-mail: anazanin.alivand1974@gmail.com
2. Corresponding Author, Water Engineering Department, Shahid Chamran University of Ahvaz, Ahvaz, Iran. E-mail: aliakh@scu.ac.ir
3. Water Engineering Department, Shahid Chamran University of Ahvaz, Ahvaz, Iran. E-mail: F.ahmadi@scu.ac.ir

Article Info	ABSTRACT
Article type: Research Article Article history: Received 7 June 2025 Received in revised form 26 July 2025 Accepted 4 September 2025 Published online 10 December 2025 Keywords: <i>Data preprocessing</i> <i>Rainfall forecasting</i> <i>Variable Mode Decomposition</i> <i>Complete Ensemble Empirical Mode Decomposition</i>	Precipitation forecasting is of great importance due to its impact on agriculture, natural disaster management, and water supply. Therefore, in this study, the monthly precipitation of the Telezang station from 1966 to 2020 was modeled using the Variable Mode Decomposition (VMD) and Complete Ensemble Empirical Mode Decomposition (CEEMD) methods. Input data were defined for the Support Vector Machine (SVM) model based on four scenarios. In the first scenario, monthly precipitation values with up to four lags were considered as model inputs. In the second scenario, in addition to the lagged precipitation data, a periodic term was added to the input patterns of the model. In the third and fourth scenarios, the monthly precipitation data were decomposed using CEEMD and VMD, respectively, and provided to the model. The findings of this research indicated that adding the periodic term slightly improved the model's performance. Additionally, a comparison of the results from the data preprocessing methods using VMD and CEEMD showed that the VMD-SVM model outperformed the CEEMD-SVM model significantly, reducing the MAE index by an average of approximately 35.25 mm compared to the standalone model and 13.77 mm compared to the CEEMD-SVM model, while also achieving greater accuracy.

Cite this article: Alivand, N., Akhoond-Ali, A. M., & Ahmadi, F. (2025). Performance evaluation of signal decomposition methods in monthly precipitation estimation (Case study: Telezang station). *Journal of Water and Irrigation Management*, 15 (2), 525-542. DOI: <https://doi.org/10.22059/jwim.2025.396662.1239>





ارزیابی عملکرد روش‌های تجزیه سیگنال در برآورد بارش ماهانه (مطالعه موردی: ایستگاه تله‌زنگ)

نازنین عالی‌وند^۱ | علی محمد آخوند علی^۲ | فرشاد احمدی^۳

۱. دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: anazanin.alivand1974@gmail.com

۲. نویسنده مسئول، گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: aliakh@scu.ac.ir

۳. گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: F.ahmadi@scu.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی	پیش‌بینی بارش به دلیل تأثیر آن بر کشاورزی، مدیریت بلایای طبیعی و نیز تأمین آب از اهمیت زیادی برخوردار است. در مطالعه حاضر، بارش ماهانه ایستگاه تله‌زنگ در دوره آماری ۱۳۴۵ تا ۱۳۹۹ با استفاده از روش تجزیه مد متغیر (VMD) و روش تجزیه مد تجربی گروهی کامل (CEEMD) مدل‌سازی شد. داده‌های ورودی براساس چهار سناریو به مدل ماشین بردار پشتیبان (SVM) تعریف شد. در سناریوی اول مقادیر بارش ماهانه تا چهار تأخیر به عنوان ورودی مدل در نظر گرفته شدند. در سناریوی دوم علاوه بر داده‌های بارش تأخیریافته، ترم پربودیک (تناوبی) نیز به الگوهای ورودی مدل افزوده شد. در سناریوی سوم و چهارم به ترتیب داده‌های بارش ماهانه توسط CEEMD و VMD تجزیه و در اختیار مدل قرار گرفتند. رهیافت این پژوهش نشان داد افزودن ترم تناوبی عملکرد مدل را به مقدار کمی بهبود می‌بخشد. همچنین، مقایسه نتایج روش‌های پیش‌پردازش داده‌ها توسط VMD و CEEMD نشان داد که مدل VMD-SVM از عملکرد بسیار بهتری نسبت به مدل CEEMD-SVM برخوردار بوده است. به‌طوری‌که این روش توانسته مقدار شاخص MAE را به‌طور متوسط حدود ۳۵/۲۵ میلی‌متر نسبت به مدل منفرد و ۱۳/۷۷ میلی‌متر نسبت به مدل CEEMD-SVM کاهش دهد و دقت بیش‌تری نیز داشته است.
تاریخ دریافت: ۱۴۰۴/۰۳/۱۷	
تاریخ بازنگری: ۱۴۰۴/۰۵/۰۴	
تاریخ پذیرش: ۱۴۰۴/۰۶/۱۳	
تاریخ انتشار: ۱۴۰۴/۰۹/۱۹	
کلیدواژه‌ها: پیش‌پردازش داده پیش‌بینی بارندگی تجزیه مد متغیر تجزیه مد تجربی گروهی کامل	

استناد: عالی‌وند، نازنین؛ آخوند علی، علی محمد و احمدی، فرشاد (۱۴۰۴). ارزیابی عملکرد روش‌های تجزیه سیگنال در برآورد بارش ماهانه (مطالعه موردی: ایستگاه تله‌زنگ). نشریه مدیریت آب و آبیاری، ۱۵ (۲)، ۵۲۵-۵۴۲. DOI: <https://doi.org/10.22059/jwim.2025.396662.1239>



۱. مقدمه

بارش یکی از مهم‌ترین پدیده‌های هواشناسی است که بر چندین بخش اقتصادی مانند کشاورزی، تولید نیروی برق‌آبی و مدیریت منابع آب تأثیر می‌گذارد (Trinh, 2018; Hadad, 2011; Hartmann *et al.*, 2016). پیش‌بینی دقیق و به‌موقع بارش به‌منظور پیشگیری و کاهش پیامدهای بلایای طبیعی مانند رانش زمین، سیل و خشکسالی ضروری است (Pham *et al.*, 2020). پیش‌بینی بارش به‌دلیل فرایندهای پیچیده‌ای که در این پدیده هواشناسی نقش دارند، یک چالش دشوار محسوب می‌شود. تغییرات بسیار زیاد مکانی-زمانی در توزیع بارندگی نیز، پیش‌بینی آن را به امری فوق‌العاده سخت تبدیل کرده است. با این حال، پیش‌بینی قابل‌اعتماد این متغیر هواشناسی به‌ویژه باتوجه‌به تهدیدهای جهانی ناشی از تغییرات آب‌وهوایی، از اهمیت حیاتی برخوردار است (Dotse *et al.*, 2024). تاکنون روش‌های متعددی به‌منظور پیش‌بینی بارش توسعه یافته‌اند، از جمله پیش‌بینی عددی وضع هوا (NWP)، روش‌های آماری و یادگیری ماشین (ML) یا هوش مصنوعی (AI). این روش‌ها عموماً به دو دسته اصلی روش‌های مبتنی بر فیزیک و روش‌های داده‌محور تقسیم می‌شوند. مدل‌های NWP که مبتنی بر فیزیک هستند، رویکرد اصلی مورد استفاده توسط سازمان‌های هواشناسی در سراسر جهان محسوب می‌شوند. اگرچه این مدل‌ها توانایی بالایی در پیش‌بینی وضع هوا نشان داده‌اند، اما گاهی با عدم قطعیت‌های زیادی در ورودی‌های مدل مواجه هستند (Bauer *et al.*, 2015; Schultz *et al.*, 2021). در سال‌های اخیر، چندین تکنیک پیشرفته به‌عنوان ابزارهای جدید برای مدل‌سازی معرفی شده‌اند. مدل ماشین بردار پشتیبان^۱ (SVM) نمونه‌ای از این مدل‌ها می‌باشد (Essam *et al.*, 2022; Samantaray *et al.*, 2022).

مدل ماشین بردار پشتیبان یکی از تکنیک‌های محبوب در یادگیری ماشین و تحلیل داده‌ها می‌باشد که برای دسته‌بندی و رگرسیون استفاده می‌شود. ماشین بردار پشتیبان بر مبنای مفهوم «مرز جداساز» عمل می‌کند و هدف مدل، جداکردن داده‌ها به بهترین شکل ممکن می‌باشد (Vapnik, 1998). از این روش در مطالعات متعددی به‌منظور پیش‌بینی بارش استفاده شده است. در مطالعه Yao *et al.* (2024)، مدل ماشین بردار پشتیبان (SVM) را به جهت بهبود دقت پیش‌بینی بارش زمستانی کشور چین به‌کار بردند. نتایج نهایی بیانگر عملکرد مناسب این مدل در پیش‌بینی بارش بوده است. در پژوهشی، Tricha and Moussaid (2024) با بهره‌گیری از مدل‌های رگرسیون خطی، رگرسیون چندجمله‌ای، K نزدیک‌ترین همسایه (KNN)، ماشین بردار پشتیبان (SVM)، درخت تصمیم (DT)، جنگل تصادفی (RF) و استفاده از داده‌های هواشناسی مثل دما، رطوبت، سرعت باد و فشار، اقدام به پیش‌بینی بارش شهر کازابلانکا مراكش کردند و نشان دادند که استفاده از این رویکردها، دقت پیش‌بینی بارش را بهبود بخشیده است. در پژوهشی دیگر، Putri (2024)، مدل ماشین بردار پشتیبان (SVM) را جهت پیش‌بینی پتانسیل بارندگی در جاکارتا مورد استفاده قرار داد. بدین‌منظور داده‌های آب‌وهوای روزانه به‌کار گرفته شدند و نتایج حاکی از عملکرد قابل قبول مدل ماشین بردار پشتیبان بوده است.

تجزیه داده‌های بارش به مؤلفه‌های ساده‌تر، راه‌کاری مؤثر برای مقابله با چالش‌های ناشی از نویز، تغییرات فصلی و عوامل محیطی پیچیده محسوب می‌شود، چرا که این روش با جداسازی سیگنال‌های اصلی از تداخلات محیطی و شناسایی مؤلفه‌های کلیدی حساس به تغییرات آب‌وهوایی، نه تنها دقت مدل‌های پیش‌بینی را به‌طور چشم‌گیری افزایش می‌دهند، بلکه با حذف مؤلفه‌های نامرتب و نامرتب مانند نویز اندازه‌گیری و اثرات تصادفی، امکان تمرکز بر الگوهای معنی‌دار بارش را فراهم می‌سازند و در نهایت بستر مناسبی برای توسعه مدل‌های پیش‌بینی‌کننده مبتنی بر یادگیری ماشین و هوش مصنوعی ایجاد می‌کند، بنابراین، تجزیه داده‌های بارش به مؤلفه‌های ساده‌تر، امکان تحلیل دقیق‌تر الگوهای بارش را فراهم می‌سازد (Zhang *et al.*, 2018; Liu *et al.*, 2023).

تجزیه حالت تجربی گروهی کامل (CEEMD) که شکل پیشرفته‌ای از تجزیه حالت تجربی است (Huang *et al.*, 1998)،

می‌تواند به‌عنوان ابزاری مفید در بهبود دقتی که در مدل‌های موجود وجود ندارد، عمل کند. CEEMD قادر است سیگنال‌های نوسانی بزرگ را به اجزای فرکانسی کوچک‌تر تفکیک و تجزیه نماید. پس از آغاز به‌کار روش CEEMD، به‌دلیل دو ویژگی کلیدی وابستگی کامل به داده و خودسازگاری ذاتی موردتوجه پژوهش‌گران قرار گرفت (Alvanitopoulos *et al.*, 2014). بنابراین، CEEMD قادر است ضمن حفظ ساختار فیزیکی داده‌های ورودی، ویژگی‌های مرتبط و معنی‌دار را نیز استخراج نماید (Yeh *et al.*, 2010). این روش در مطالعات متعددی همچون Ali *et al.* (2020)، Shen and Ban (2023)، Zhang *et al.* (2023)، Dong (2024) و Xu *et al.* (2024) استفاده شده و عملکرد مناسب آن مورد تأیید بوده است.

تجزیه مد متغیر (VMD)، به‌عنوان یک روش تجزیه سیگنال مانند EMD، در مقایسه با روش‌های مشابه دارای مقاومت نویز قوی‌تر و پایه نظری ریاضی محکم‌تری است (Dragomiretskiy and Zosso, 2013). این روش هم‌زمان یک سیگنال را به مجموعه‌ای از توابع حالت ذاتی (IMF) با خاصیت پراکندگی مشخص تجزیه می‌کند. VMD یک تکنیک پردازش سیگنال عالی است که برای تجزیه و تحلیل سیگنال‌های غیرایستا و غیرخطی مناسب می‌باشد. از زمان پیشنهاد این روش در سال ۲۰۱۴، به‌طور گسترده‌ای در زمینه‌های مختلفی مانند مهندسی پزشکی، تضعیف امواج سطحی، تحلیل زمان-فرکانس لرزه‌ای، پردازش تصویر، تحلیل داده‌های مالی و ... مورد استفاده قرار گرفته است (Lahmiri and Boukadoum, 2014; Liu *et al.*, 2015; Liu *et al.*, 2016; Lahmiri and Shmuel, 2017; Lahmiri, 2016). از این روش نیز در پژوهش‌های متعددی همچون Parsaie *et al.* (2024)، Dong *et al.* (2024) و Fardi Rad and Hosseini (2025) استفاده شده و عملکرد مطلوب آن گزارش گردیده است.

نظر به پیشینه پژوهش ارائه‌شده، مشاهده می‌شود که برآورد بارش همواره موردتوجه پژوهش‌گران بوده و بدین‌منظور از روش‌های یادگیری ماشین و به‌ویژه مدل SVM استفاده شده است. در اکثر مطالعات انجام‌شده روش SVM به‌صورت منفرد مورد استفاده قرار گرفته و ترکیب آن با روش‌های پیش‌پردازش‌کننده نظیر VMD و CEEMD کم‌تر موردتوجه بوده است. از این‌رو، هدف از مطالعه حاضر توسعه مدل‌های VMD-SVM و CEEMD-SVM برای پیش‌بینی بارش و ارزیابی عملکرد آن‌ها می‌باشد. علاوه بر این در توسعه الگوهای ورودی خاصیت تناوبی با دو رویکرد خطی و غیرخطی در بهبود عملکرد مدل ماشین بردار پشتیبان مورد بحث و بررسی قرار خواهد گرفت.

۲. مواد و روش‌ها

۲.۱. منطقه مورد مطالعه و داده‌ها

در مطالعه حاضر، از اطلاعات بارش ایستگاه تله‌زنگ واقع در حوضه آبریز دز در طی دوره آماری ۱۳۴۵ تا ۱۳۹۹ استفاده شده است. حوضه آبریز دز یکی از مهم‌ترین حوضه‌های آبریز کشور محسوب می‌شود که در منطقه زاگرس مرکزی و جنوب‌غربی ایران واقع شده است. این حوضه به‌دلیل برخورداری از بارش‌های نسبتاً مناسب و شرایط توپوگرافی خاص، از پتانسیل آبی قابل‌توجهی برخوردار است. اهمیت این حوضه آبریز به‌دلیل نقش حیاتی آن در تغذیه سفره‌های آب زیرزمینی، حفظ تعادل اکولوژیکی و تأمین آب برای بخش‌های مختلف کشاورزی، صنعت و شرب است. وجود اکوسیستم‌های متنوع شامل جنگل‌های بلوط و مراتع غنی، بر ارزش زیست‌محیطی این منطقه می‌افزاید. این حوضه به‌دلیل موقعیت جغرافیایی و ویژگی‌های هیدرولوژیکی منحصربه‌فرد، همواره موردتوجه پژوهش‌گران و برنامه‌ریزان منابع آب بوده است. ایستگاه تله‌زنگ در موقعیت جغرافیایی ۴۸ درجه و ۳۲ دقیقه شمالی و ۱۵ درجه و ۴۸ دقیقه شرقی با ارتفاع ۴۴۰ متر از سطح دریا در حوضه آبریز دز واقع شده است. این ایستگاه با مساحتی بالغ بر ۱۶۲۰۱ کیلومتر مربع، یکی

از ایستگاه‌های کلیدی برای بارش‌های غربی- جنوب‌غربی در استان خوزستان محسوب می‌شود. میانگین بارش سالانه این ایستگاه طبق داده‌های ثبت‌شده در دوره آماری ۵۵ ساله، ۸۳۱/۸ میلی‌متر بوده است. مشخصات ایستگاه تله‌زنگ در جدول (۱) ذکر گردیده است. شکل (۱) نیز موقعیت حوضه آبریز دز در ایران و موقعیت ایستگاه تله‌زنگ را نشان می‌دهد.

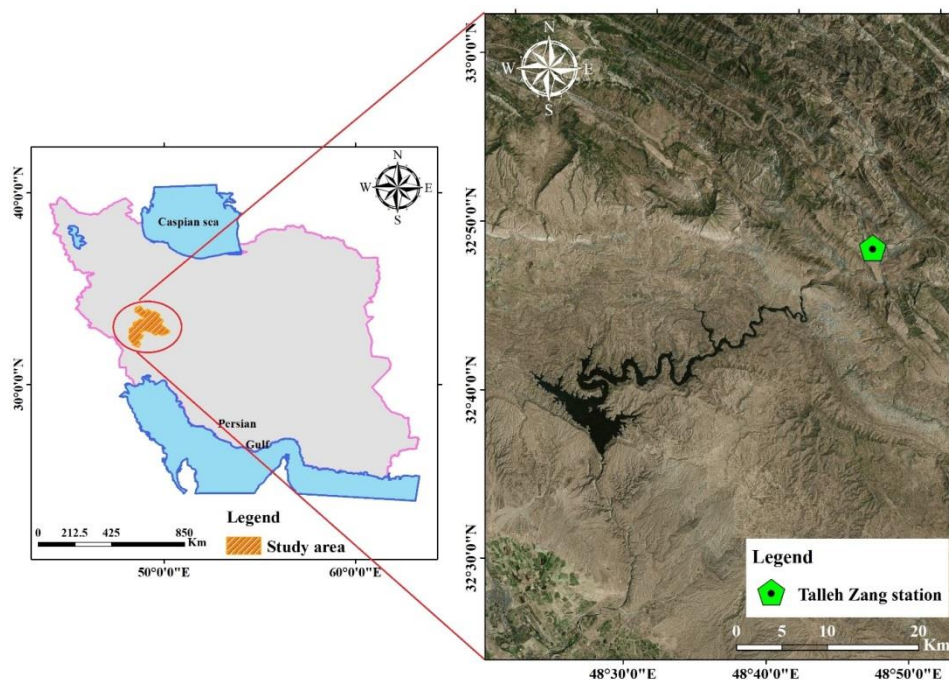


Figure 1. Location of the study area along Dez Basin

Table 1. The statistical specifications of the Talleh Zang Rain gauge station during the statistical period of 1966-2020

Time Scale	Minimum (mm)	Mean (mm)	Maximum (mm)	Standard Deviation (mm)	Coefficient of Variation (%)
October	0	8.1	87.0	17.9	2.1
November	0	85.86	331.5	79.4	0.9
December	0	149.8	479.0	98.8	0.6
January	0	151.4	378.5	94.7	0.6
February	0	157.3	424.5	91.0	0.5
March	0	121.1	379.5	90.7	0.7
April	0	110.7	333.0	96.0	0.8
May	0	44.1	192.5	45.9	1.0
June	0	1.36	27.0	4.4	3.2
July	0	0	0	0	0
August	0	1.0	49.0	6.5	6.5
September	0	0.7	24.5	3.5	4.5
Annual	0	831.8	1546.5	306.3	0.3

۲.۲. مدل ماشین بردار پشتیبان (SVM)

مدل ماشین بردار پشتیبان (SVM) اولین بار توسط Vapnik (1995) معرفی شد. این روش یک الگوریتم یادگیری تحت‌نظارت قدرتمند است که برای مسائل طبقه‌بندی و رگرسیون استفاده می‌شود (Du et al., 2017). SVM با یافتن صفحه بهینه‌ای که نقاط داده متعلق به کلاس‌های مختلف را جدا می‌کند، کار می‌کند. هدف آن حداکثر کردن فاصله بین کلاس‌ها می‌باشد و از بردارهای پشتیبان (نقاط داده نزدیک به صفحه) برای تعریف مرز بهینه استفاده می‌کند. SVM می‌تواند به مسائل طبقه‌بندی خطی و غیرخطی رسیدگی کند و با استفاده از توابع کرنل مختلف مانند خطی، چندجمله‌ای

یا شعاعی (RBF) این کار را انجام می‌دهد (Imran et al., 2025; Basak et al., 2007). رگرسیون بردار پشتیبان را به شکل معادله (۱) در مجموعه‌ای از توابع خطی که داده‌ها را منعکس می‌کنند، در نظر بگیرید (Zakaria and shabri, 2012):

$$f(x) = \sum_{k=1}^N \alpha_k + \alpha_k^* x_k^T x + b \quad \text{رابطه ۱}$$

در معادله فوق، N تعداد مقادیر آموزشی است، درحالی‌که $x_k \in R^n$ و $y_k \in R$ به ترتیب مقادیر ورودی و خروجی هستند. سپس می‌توان از مسئله اولیه زیر برای تعریف مسئله بهینه‌سازی استفاده کرد.

$$\text{Min} J \quad w, \xi, \xi^* = \frac{1}{2} w^T w + c \sum_{k=1}^N \xi_k + \xi_k^* \quad \text{رابطه ۲}$$

$$\text{Such that} \begin{cases} y_k - w^T x_k - b \leq \varepsilon + \xi_k \\ w^T x_k + b - y_k \leq \varepsilon + \xi_k^* \\ \xi_k, \xi_k^* \geq 0 \end{cases} \quad \text{رابطه ۳}$$

در رابطه (۲)، ثابت c سطح تحمل انحراف از دقت موردنظر را تعیین می‌کند و با متغیرهای کمکی ξ_k, ξ_k^* برای $k = 1, \dots, N$ مرتبط است. مسئله ابتدا باید به فرم لاگرانژ بیان شود، پس از آن برنامه‌ریزی درجه دوم مسئله دوگانه باید آن را حل کند. تابع خطی در فضای دوگانه تبدیل می‌شود.

$$f(x) = \sum_{k=1}^N \alpha_k + \alpha_k^* x_k^T x + b \quad \text{رابطه ۴}$$

پیش‌بینی‌های SVM برای یک سناریوی غیرخطی در نظر گرفته شده است. ضرایب لاگرانژ هستند. مدل فضای وزنی اساسی که در ادامه آمده است، برای تسهیل

$$f(x) = w^T \phi(x) + b \quad \text{رابطه ۵}$$

که در آن، $\phi(x): R^n \rightarrow R^{nh}$ نمایانگر نگاشت داده‌های ورودی به یک فضای با ابعاد بالاتر می‌باشد. هدف نهایی در این روش، یافتن تابع $f(x)$ است که بهترین تطابق را با داده‌های آموزشی داشته باشد و درعین‌حال، از پیچیدگی زیاد جلوگیری کند. در این مورد از تابع هسته استفاده شده است که منجر به $K(x_k, x_l) = \phi(x_k)^T \phi(x_l)$ برای $k = 1, \dots, N$ می‌شود. هسته استفاده‌شده با توجه به نوع داده‌ها انتخاب می‌شود. تابع کرنل شعاعی (RBF) به دلیل کاربرد گسترده در این مطالعه مورد استفاده قرار گرفته است (Mahmood et al., 2024).

۳.۲. روش تجزیه مد تجربی گروهی کامل (CEEMD)

روش تجزیه مد تجربی گروهی کامل (Complete ensemble empirical mode decomposition) اولین بار توسط Torres et al. (2011) ارائه شد. تکنیک تجزیه داده CEEMD به غیرایستایی و غیرخطی بودن داده‌های ورودی می‌پردازد. براساس تجزیه حالت تجربی (EMD)، CEEMD نه تنها می‌تواند مشکل سردرگمی حالت EMD را حل کند، بلکه می‌تواند با افزودن نویز سفید مثبت و منفی به سیگنال اصلی، باقیمانده نویز سفید را نیز حذف کند. مراحل اصلی این روش به شرح زیر می‌باشد (Zhang et al., 2020).

مرحله ۱- با اضافه کردن نویز سفید مثبت و منفی I_i و $-I_i$ به سیگنال اصلی X_i ، سیگنال‌های سنتز شده P_i و N_i به دست می‌آیند.

$$\begin{cases} P_i = X_i + I_i \\ N_i = X_i - I_i \end{cases} \quad \text{رابطه ۶}$$

مرحله ۲- از EMD برای تجزیه سیگنال مصنوعی به دست آمده از رابطه فوق استفاده می‌شود (Tian and Chen, 2021).

$$\begin{cases} C_{1j}^+, C_{2j}^+, \dots, C_{Mj}^+ \\ C_{1j}^-, C_{2j}^-, \dots, C_{Mj}^- \end{cases} \quad \text{رابطه ۷}$$

که در آن، C_{ij}^+ برابر با z امین IMF یا باقیمانده سیگنال سنتز شده پس از اضافه کردن نویز سفید مثبت در آزمون i ام است، C_{ij}^- برابر با z امین IMF یا باقیمانده سیگنال سنتز شده پس از اضافه کردن نویز سفید منفی در آزمون i ام است، M مجموع IMF و باقیمانده است. مرحله ۳- مراحل (۱) و (۲)، M بار تکرار شوند تا IMF M و باقیمانده به دست آید.

$$\begin{cases} P_i = X_i + I_i \\ N_i = X_i - I_i \end{cases} \quad \text{رابطه ۸}$$

مرحله ۴- در این مرحله، مقادیر میانگین کلی برای تمامی مؤلفه‌های مد ذاتی (IMF) و باقیمانده محاسبه می‌گردد که حاصل تجزیه سری زمانی بارش با روش CEEMD است (Lian, 2022).

۴.۲. الگوریتم تجزیه مد متغیر (VMD)

تجزیه مد متغیر (Variational Mode Decomposition یا VMD) یک تکنیک پیشرفته به منظور تحلیل و تجزیه سیگنال‌ها و سری‌های زمانی است. این روش به‌ویژه برای داده‌های غیرایستا و پیچیده طراحی شده و می‌تواند به استخراج توابع حالت ذاتی (Intrinsic Mode Functions یا IMFs) کمک کند. هدف VMD تجزیه یک سیگنال چندمؤلفه‌ای به چندین سیگنال تک‌مؤلفه‌ای است که فرکانس لحظه‌ای آن‌ها دارای معنای فیزیکی مشخصی باشد (Dragomiretskiy and Zosso, 2013; Xue et al., 2019). در اینجا، الگوریتم جهت متناوب ضرب‌کننده‌ها (ADMM) برای رفع محدودیت‌های مختلف برای تکرار مد، فرکانس مرکزی و ضریب لاگرانژ به کار گرفته شده است (Hestenes, 1969; Rockafellar, 1973; Bertsekas, 2014).

در نهایت، مدهای تفکیک شده پس از برآورده شدن شرط توقف به دست می‌آیند. شرح مختصری از این روش در زیر ارائه شده است (Liu et al., 2023):

الف- برای به دست آوردن طیف فرکانسی یک طرفه، سیگنال تحلیلی هر مد توسط تبدیل هیلبرت محاسبه می‌شود. رابطه ۹

$$(\delta(t) + \frac{j}{\pi t}) * u_k(t),$$

ب- طیف مد از طریق ترکیب نمایی با فرکانس مرکزی تخمینی متناظر، به باند پایه منتقل می‌شود.

$$\left[(\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-j\omega_k t}, \quad \text{رابطه ۱۰}$$

ج- پهنای باند هر مؤلفه از طریق تحلیل هموارسازی گاوسی سیگنال دمودوله شده یعنی نرم L^2 مربع گرادیان تخمین زده می‌شود.

$$\left\{ \min_{u_k, \omega_k} \left\{ \sum_k \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \right. \\ \left. s.t. \sum_k u_k(t) = f(t) \right\} \quad \text{رابطه ۱۱}$$

u_k و ω_k به ترتیب مجموعه هر مد و فرکانس مرکزی آن هستند.

برای اطمینان از مقاومت در برابر نویز و صحت سیگنال بازسازی شده، از یک ضریب جریمه درجه دوم و ضریب لاگرانژی استفاده می‌شود (Bertsekas, 1976; Nocedal and Wright, 1999). سپس، مسئله بدون قید را می‌توان به صورت زیر نوشت:

$$L(u_k, \omega_k, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \|f(t) - \sum_k u_k(t)\|_2^2 + \langle \lambda(t), f(t) - \sum_k u_k(t) \rangle, \quad (12)$$

روش ADMM برای حل مسئله بدون قید استفاده می‌شود.

$$\begin{cases} u_k^{n+1} = \arg \min L(u_k, \omega_k, \lambda^n), \\ \omega_k^{n+1} = \arg \min L(u_k^n, \omega_k, \lambda^n), \\ \lambda^{n+1} = \lambda^n + \tau (f - \sum_k u_k^{n+1}). \end{cases} \quad (13)$$

که در آن، n تعداد تکرار است و بین عبارات منظم و درست تعادل برقرار می‌کند. اکنون، u_k^{n+1} را می‌توان به صورت زیر بازنویسی کرد:

$$u_k^{n+1}(t) = \arg \min \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) + \frac{\lambda(t)}{2} \right\|_2^2, \quad (14)$$

با اعمال تساوی پارسوال و نرم L_2 می‌توان مسئله فوق را در حوزه فرکانس حل کرد.

$$\hat{u}_k^{n+1}(\omega) = \arg \min \left\{ \int_0^\infty 4\alpha |\omega - \omega_k|^2 |\hat{u}_k(\omega)|^2 + 2 \left| \hat{f}(\omega) - \sum_k \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right|^2 d\omega \right\}, \quad (15)$$

با انجام تغییر متغیر $w \leftarrow w - \omega_k$ معادلات به فرم ساده‌تر زیر قابل ارائه می‌باشد:

$$u_k^{n+1}(w) = \frac{f(\omega) - \sum_{i \neq k} u_i(\omega) + \frac{\lambda(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2}, \quad (16)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |u_k(\omega)|^2 d\omega}{\int_0^\infty |u_k(\omega)|^2 d\omega}, \quad (17)$$

$$\lambda^{n+1}(\omega) = \lambda^n(\omega) + \tau (f(\omega) - \sum_k \hat{u}_k^{n+1}(w)). \quad (18)$$

که در آن، $\sum_k \|u_k^{n+1} - u_k^n\|_2^2 / \|u_k^n\|_2^2 < \epsilon$ شرط همگرایی و $\in$ سطح تلورانس همگرایی است که از قبل تعیین شده است.

۵.۲. توسعه مدل

روش‌های داده‌محور مانند مدل‌های یادگیری ماشین برای پیش‌بینی خروجی به مجموعه‌ای از داده‌های ورودی نیاز دارند. انتخاب ورودی‌های مناسب در پیش‌بینی بارندگی، به‌ویژه هنگام استفاده از رویکردهای هوش مصنوعی بسیار مهم است، زیرا عملکرد و دقت مدل به متغیرهای ورودی بستگی دارد. بنابراین، انتخاب کارآمدترین پارامترهای ورودی و بررسی کامل آن‌ها در طول فرایند مدل‌سازی ضروری است. از آنجایی که بارندگی از نظر زمانی و مکانی به‌طور قابل‌توجهی تغییر می‌کند، انتخاب تعداد و نوع مناسب متغیرهای ورودی برای توسعه مدل‌ها در پیش‌بینی بارندگی ضروری است. اگرچه تعداد نامحدودی ورودی ممکن وجود دارد، اما لازم است زیرمجموعه کوچک‌تری از متغیرهای بالقوه مرتبط با بارندگی برای مکان مورد مطالعه انتخاب شود. این امر با به حداقل رساندن زمان اجرا و عدم قطعیت مرتبط با تعمیم، کارایی فرایند آموزش را بهبود می‌بخشد (Dotse et al., 2024). در این مطالعه، سعی گردید الگوهای ورودی با سناریوهای مختلف توسعه یابد. به همین جهت چهار سناریو تعریف گردید. در سناریوی اول، تأخیر مناسب برای مدل‌سازی بارش با استفاده از روش تابع خودهمبستگی جزئی وزن‌دهی شده و پس از انتخاب تأخیرهای مؤثر، الگوهای تأخیر پایه (LB) ایجاد شدند

(شکل ۲). معمولاً در تحلیل سری‌های زمانی، تأخیرهایی انتخاب می‌شوند که مقدار خودهمبستگی جزئی آن‌ها از حد آستانه معینی (مثلاً خطوط زرد در نمودار) فراتر رفته و از نظر آماری معنی‌دار باشند. همان‌گونه که در شکل (۲) مشخص است، اگرچه تا شش تأخیر قابلیت مدل‌سازی وجود دارد، اما افزایش تعداد تأخیرها ممکن است مدل را پیچیده‌تر کند بدون آنکه بهبود قابل‌توجهی در دقت مدل ایجاد شود (Ahmadi, 2020). در نتیجه تأخیر زمانی چهار ماهه برای انجام مدل‌سازی انتخاب شد. شایان ذکر است رکوردهای بارندگی با تأخیر (گذشته) به‌طور گسترده به‌عنوان ورودی‌های مدل برای پیش‌بینی بارندگی استفاده شده‌اند (Jiang and wu, 2013; El-Shafie *et al.*, 2011; Kalteh, 2017).

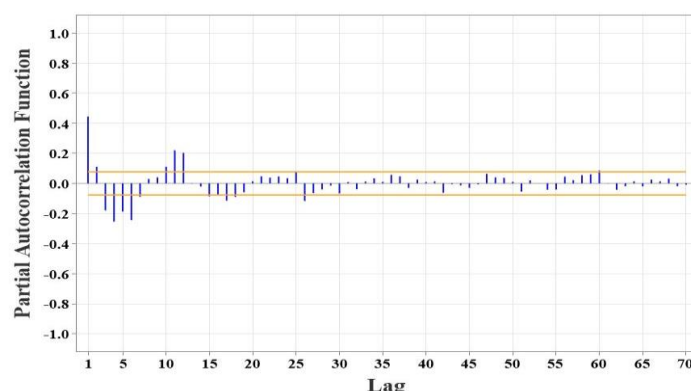


Figure 2. Partial autocorrelation function (PACF) of precipitation at Talleh Zang rain gauge station

در سناریوی دوم، علاوه بر متغیرهای منتخب، خاصیت فصلی خطی و غیرخطی به‌کار گرفته‌شد. بدین‌منظور از عبارت‌های $\cos[2\pi.i/12]$ و $\sin[2\pi.i/12]$ ($i = 1, \dots, 12$) و ضرایب ساده ماهانه تحت عنوان α ($\alpha = 1, \dots, 12$) به‌منظور بسط اثر خاصیت پریودیک خطی (LPT)^۳ و غیرخطی (NLPT)^۴ استفاده شده و به الگوهای ورودی افزوده شد. در سناریوی سوم، روش تجزیه مد تجربی گروهی کامل (CEEMD) به‌منظور پیش‌پردازش سری زمانی بارش به‌کار گرفته شد. در سناریوی چهارم، از روش تجزیه مد متغیر (VMD) به‌منظور انتقال بهتر اطلاعات به مدل منفرد SVM و حذف اثر نوفه‌ها بهره برده شد. به بیانی دیگر، ابتدا الگوهای ورودی (LB) تجزیه و سپس به مدل معرفی شدند و روش تلفیقی VMD-SVM توسعه داده شد. جدول زیر الگوهای ورودی توسعه یافته برای هر یک از مدل‌ها را نشان می‌دهد.

Table 2. Scenario-based input parameter variations in the simulation framework

Scenario	Row	Pattern	Input pattern layout
Time-lagged precipitation data (S1)	1	S1M1	$R_t = f(R_{t-1})$
	2	S1M2	$R_t = f(R_{t-1}, R_{t-2})$
	3	S1M3	$R_t = f(R_{t-1}, R_{t-2}, R_{t-3})$
	4	S1M4	$R_t = f(R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4})$
Time-lagged precipitation data with periodic terms (S2)	5	S2M1	$R_t = f(LPT, NLPT, R_{t-1})$
	6	S2M2	$R_t = f(LPT, NLPT, R_{t-1}, R_{t-2})$
	7	S2M3	$R_t = f(LPT, NLPT, R_{t-1}, R_{t-2}, R_{t-3})$
	8	S2M4	$R_t = f(LPT, NLPT, R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4})$
Precipitation data decomposition using the CEEMD method (S3)	9	S3M1	$R_t = f(CEEMD(R_{t-1}))$
	10	S3M2	$R_t = f(CEEMD(R_{t-1}, R_{t-2}))$
	11	S3M3	$R_t = f(CEEMD(R_{t-1}, R_{t-2}, R_{t-3}))$
	12	S3M4	$R_t = f(CEEMD(R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}))$
Precipitation data decomposition using the VMD method (S4)	13	S4M1	$R_t = f(VMD(R_{t-1}))$
	14	S4M2	$R_t = f(VMD(R_{t-1}, R_{t-2}))$
	15	S4M3	$R_t = f(VMD(R_{t-1}, R_{t-2}, R_{t-3}))$
	16	S4M3	$R_t = f(VMD(R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}))$

۲.۶. ارزیابی مدل‌ها

در این پژوهش، به منظور ارزیابی مدل‌های به کار گرفته شده در پیش‌بینی بارش در محل ایستگاه تله‌زنگ از معیارهای ضریب همبستگی (R)، میانگین خطای مطلق (MAE)، جذر میانگین مربعات خطا (RMSE) و ضریب نش- ساتکلیف (NS) استفاده شده است:

$$R = \frac{\sum_{i=1}^n Q_i - \bar{Q}_i \quad \bar{Q}_i - \bar{Q}}{\sqrt{\sum_{i=1}^n Q_i - \bar{Q}_i^2} \sqrt{\sum_{i=1}^n \bar{Q}_i - \bar{Q}^2}} \quad \text{رابطه (۱۱)}$$

$$MAE = \frac{\sum_{i=1}^n |Q_i - \bar{Q}_i|}{n} \quad \text{رابطه (۱۲)}$$

$$RMSE = \left(\frac{\sum_{i=1}^n Q_i - \bar{Q}_i^2}{n} \right)^{0.5} \quad \text{رابطه (۱۳)}$$

$$NS = 1 - \frac{\sum_{i=1}^n Q_i - \bar{Q}_i^2}{\sum_{i=1}^n Q_i - \bar{Q}^2} \quad \text{رابطه (۱۴)}$$

در روابط فوق، Q_i مقدار مشاهده شده در گام زمانی i ام، $\bar{Q}_i$ مقدار محاسبه شده در همان زمان، n تعداد داده‌ها و $\bar{Q}$ میانگین مقادیر مشاهداتی می‌باشد. شاخص کارایی نش- ساتکلیف (NS) از پرکاربردترین معیارها در ارزیابی مدل‌های هیدرولوژیکی محسوب می‌شود (Nash and Sutcliffe, 1970). به طور کلی، چنانچه مقدار شاخص نش- ساتکلیف بیش‌تر از ۰/۷ باشد، مدل از عملکرد عالی برخوردار است. مقادیر بین ۰/۳ تا ۰/۷ نشان‌دهنده عملکرد رضایت‌بخش بوده، درحالی‌که اعداد کمتر از ۰/۳ حاکی از عملکرد غیرقابل قبول مدل هستند (Roodaki and Azizian, 2020). قابل ذکر است مدلی که دارای کمترین مقدار خطا (RMSE و MAE) و نیز بیش‌ترین مقدار دقت (R و NS) باشد، مناسب‌ترین مدل خواهد بود.

۳. نتایج و بحث

۳.۱. پیش‌بینی بارش ماهانه با استفاده از مدل مستقل و تناوبی

در مطالعه حاضر، به منظور مدل‌سازی و پیش‌بینی بارش ماهانه ایستگاه تله‌زنگ، از الگوریتم ماشین بردار پشتیبان (SVM) استفاده گردید. داده‌های مورد استفاده در این پژوهش به دو بخش مجزای داده‌های آموزش (۸۰ درصد، معادل ۵۲۸ ماه) و داده‌های آزمون (۲۰ درصد، معادل ۱۳۲ ماه) تقسیم‌بندی و فرایند مدل‌سازی براساس هشت الگوی ورودی (ارائه شده در جدول ۲) طراحی گردید که هر یک از این الگوها با در نظر گرفتن پارامترهای بارش تأخیریافته و خاصیت تناوبی تنظیم گردیدند تا امکان ارزیابی دقیق عملکرد مدل فراهم گردد.

ضریب تناوبی (Periodic Coefficient) یک پارامتر ریاضی ساده برای شناسایی و مدل‌سازی الگوهای چرخه‌ای در داده‌های سری زمانی است. این ضریب با نرمال‌سازی داده‌ها در چرخه‌های مشخص مثلاً ماهانه یا سالانه، تغییرات تکراری مانند الگوهای فصلی بارش را استخراج می‌کند. محاسبه آن براساس انحراف استاندارد شده از میانگین دوره‌ای انجام می‌شود که به جداسازی سیگنال‌های دوره‌ای از نویز کمک می‌کند. هنگامی که این ضریب به همراه مدل‌های

یادگیری ماشین مانند SVM استفاده می‌شود، دقت پیش‌بینی را بهبود می‌بخشد. سادگی و شفافیت این ضریب، آن را به یک ابزار مطلوب در مدل‌سازی با مدل‌های داده‌محور تبدیل کرده است (Ahmadi, 2020).

انتخاب تابع کرنل بهینه از جمله مراحل اساسی در پیاده‌سازی مدل‌های یادگیری ماشین محسوب می‌شود که تأثیر معنی‌داری بر دقت پیش‌بینی‌ها دارد. مطابق یافته‌های پژوهش Ahmadi (2020)، تابع کرنل پایه شعاعی (RBF) به دلیل توانایی بالای آن در مدل‌سازی روابط غیرخطی پیچیده، عملکرد مطلوب‌تری نسبت به سایر توابع کرنل ارائه می‌نماید. ویژگی‌های منحصربه‌فرد این تابع شامل ظرفیت بالای شناسایی الگوهای پیچیده و قابلیت تعمیم‌پذیری مناسب، موجب گردید تا در این پژوهش، تابع RBF به عنوان هسته اصلی مدل انتخاب شود. در جدول (۳) نتایج شاخص‌های ارزیابی RMSE، MAE، R و NS برای مدل SVM ارائه شده است. با توجه به این جدول مشاهده می‌شود که الگوهای ورودی منتخب در دقت و عملکرد مدل منفرد تأثیر مستقیمی داشته است. به عنوان مثال، کاربرد ترم تناوبی کارایی مدل را براساس شاخص MAE بهبود بخشیده است، به طوری که این شاخص بدون در نظر گرفتن ترم تناوبی برای الگوهای ورودی S1M1 تا S1M4 در مرحله آزمون به طور متوسط برابر با ۵۳/۱ میلی‌متر بوده و این درحالی است که برای الگوهای S2M1 تا S2M4 این مقدار به ۴۳/۲ میلی‌متر کاهش یافته است. بهبود در شاخص NS نیز با در نظر گرفتن خاصیت تناوبی در مرحله آزمون وجود داشته و حدود ۰/۱۳۳ بهبود مشاهده می‌شود.

نظر به نتایج حاصل از مدل‌سازی بارش با استفاده از ترم تناوبی و نیز مطالعات Ahmadi (2020) و Montaseri and Ghavidel (2017) مشاهده می‌شود که افزودن ترم تناوبی به الگوهای ورودی مدل‌های یادگیری ماشین، سبب بهبود عملکرد مدل‌ها خواهد شد. شکل (۳) نمودار جعبه‌ای برای مدل منفرد و تناوبی را نشان می‌دهد. این نمودار به منظور مقایسه توزیع داده‌ها بین گروه‌های مختلف (در اینجا مدل ماشین بردار پشتیبان و تناوبی) استفاده می‌گردد. با توجه به این نمودار مدل ماشین بردار پشتیبان منفرد از جنبه نزدیکی بیش‌تر میانه و گستردگی پیش‌بینی‌های این مدل به داده‌های واقعی بهتر بوده است. لیکن نمودارهای جعبه‌ای فقط توزیع داده‌ها را نشان می‌دهند. به عبارت دیگر، این نمودارها دید خوبی از توزیع کلی داده‌ها، میانه و پراکندگی آن‌ها ارائه می‌دهند اما حساسیت زیادی به تک‌تک مقادیر پیش‌بینی‌شده در طول زمان ندارند و شاخص‌های ارزیابی عملکرد کلی مدل را در طول کل دوره ارزیابی می‌کنند و حساسیت بیش‌تری به خطاهای فردی در پیش‌بینی‌ها دارند. برتری مدل ضریب تناوبی در شاخص‌ها نشان می‌دهد که این مدل در مجموع، پیش‌بینی‌های دقیق‌تری با خطای کم‌تر و همبستگی بالاتری با داده‌های واقعی ارائه داده است، حتی اگر توزیع کلی پیش‌بینی‌های آن در نمودار جعبه‌ای کمی متفاوت به نظر برسد. به عبارت دیگر، ممکن است مدل SVM در پیش‌بینی مقادیر میانی بارش عملکرد نسبتاً خوبی داشته باشد (که در نمودار جعبه‌ای دیده می‌شود)، اما در پیش‌بینی مقادیر بالا و پایین یا در طول زمان، خطاهای بزرگ‌تری دارد که در شاخص‌های ارزیابی، خود را نشان می‌دهد. در مقابل، مدل ضریب تناوبی ممکن است توزیع پیش‌بینی‌های آن کمی متفاوت باشد، اما به طور کلی پیش‌بینی‌های دقیق‌تری در طول زمان ارائه می‌دهد.

Table 3. Results of RMSE, MAE, R and NS evaluation indices for the SVM model

Model	Training phase					Testing phase				
	Pattern	RMSE (mm)	MAE (mm)	R	NS	Pattern	RMSE (mm)	MAE (mm)	R	NS
SVM	S1M1	79.61	57.75	0.54	0.29	S1M1	75.44	52.50	0.43	0.15
	S1M2	78.06	56.98	0.57	0.32	S1M2	74.31	52.93	0.45	0.18
	S1M3	70.67	51.30	0.67	0.44	S1M3	76.00	54.08	0.45	0.14
	S1M4	70.94	49.53	0.66	0.44	S1M4	75.19	52.94	0.46	0.16
	S2M1	65.33	41.31	0.72	0.52	S2M1	68.06	42.27	0.60	0.31
	S2M2	63.89	40.10	0.74	0.54	S2M2	69.41	43.06	0.57	0.28
	S2M3	62.74	39.09	0.75	0.56	S2M3	69.65	43.35	0.57	0.28
	S2M4	64.95	41.53	0.73	0.53	S2M4	69.27	44.03	0.57	0.29

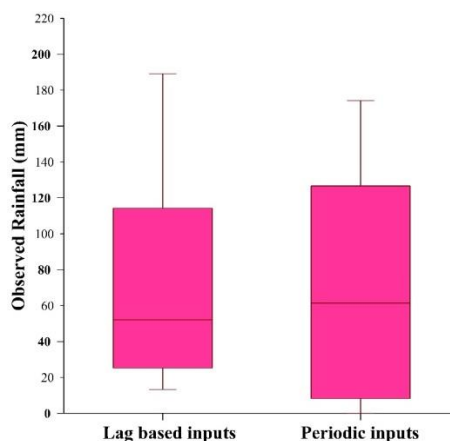


Figure 3. Box plot of observed and estimated rainfall obtained from the standalone SVM model with lag based and periodic inputs

۲.۳. نتایج حاصل از مدل CEEMD-SVM

به دلیل عدم توانایی مدل های منفرد و تناوبی در استفاده بهینه از اطلاعات موجود، راه کار دیگری تحت عنوان روش تجزیه مد تجزیه گروھی کامل (CEEMD) مورد استفاده قرار گرفت. بدین منظور بهترین الگوهای ورودی در هر سناریو انتخاب و با روش CEEMD تجزیه شدند و سپس به عنوان ورودی به مدل معرفی شدند. اساس کار این روش به تجزیه سری زمانی داده های بارش به مؤلفه های مد ذاتی (IMF) یا اجزای ساده تر می باشد. نتایج آماره های حاصل از روش CEEMD-SVM در جدول (۴) ارائه گردیده است. با توجه به این جدول، مشاهده می شود که تجزیه داده ها تأثیر بسیار مثبتی در بهبود کارایی مدل ماشین بردار پشتیبان داشته است. به عنوان مثال، در الگوی S3M3، مقدار RMSE از ۷۴/۳۱ در مدل SVM منفرد، به ۴۳/۶۴ در مدل CEEMD-SVM کاهش یافته است. بهبود دقت مدل سازی بارش با روش تجزیه CEEMD نیز مشهود بوده است. این امر نشان دهنده تأثیر مثبت پیش پردازش داده ها در فرایند مدل سازی می باشد. هم چنین می توان مشاهده نمود بهترین عملکرد مدل CEEMD-SVM با الگوی S3M3 به دست آمده است. در شکل (۴) نمودار گرافیکی مدل CEEMD-SVM برای بهترین عملکرد در مرحله آزمون ارائه شده است. با توجه به نمودار مشاهده می شود که مدل در پیش بینی مقادیر کمینه عملکرد قابل قبولی داشته، اما در پیش بینی مقادیر بیشینه (بارش های شدید یا پیک ها) با خطای بیش تری مواجه بوده است. این خطا می تواند ناشی از ماهیت غیر خطی و ناگهانی رخ داده های بارش شدید باشد. همبستگی بالای این مدل ($R=0.86$)، نشان می دهد که مدل به طور کلی در پیش بینی روند بارش عملکرد خوبی داشته است، به گونه ای که وقتی بارش به طور متوسط افزایش یا کاهش پیدا کرده، مدل نیز به طور کلی همین روند را دنبال کرده است. نتایج نشان می دهد که مدل CEEMD-SVM تشخیص و پیش بینی تغییرات کلی بارش موفق بوده لیکن اختلاف هایی بین مقادیر واقعی و پیش بینی شده وجود دارد. این خطاها می تواند ناشی از پیچیدگی های ذاتی پدیده بارش و محدودیت های مدل در پیش بینی دقیق جزئیات و مقادیر دقیق پیک ها باشد. به طور کلی مدل CEEMD-SVM در این زمینه عملکرد قابل قبولی داشته است.

Table 4. Results of RMSE, MAE, R and NS evaluation indices for the CEEMD-SVM model

Model	Training phase					Testing phase				
	Pattern	RMSE (mm)	MAE (mm)	R	NS	Pattern	RMSE (mm)	MAE (mm)	R	NS
SVM	S3M1	58.98	40.24	0.78	0.61	S3M1	53.03	35.42	0.77	0.58
	S3M2	19.69	13.33	0.98	0.96	S3M2	46.03	27.32	0.85	0.68
	S3M3	22.58	15.45	0.97	0.94	S3M3	43.64	31.46	0.87	0.72
	S3M4	22.81	15.65	0.97	0.94	S3M4	45.12	32.31	0.86	0.70

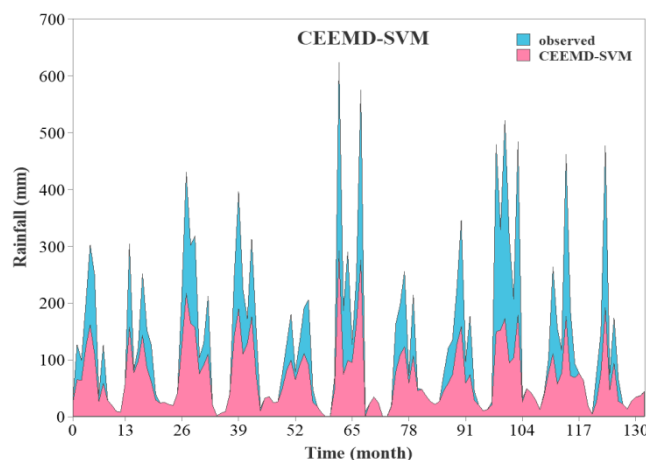


Figure 4. CEEMD-SVM model performance: Observed versus predicted values

۳.۳. نتایج حاصل از مدل تلفیقی VMD-SVM

با توجه به نتایج مدل تجزیه مد تجربی گروهی کامل (CEEMD) و اهمیت اساسی درک دقیق‌تر از تأثیرات متغیرهای خاص بر پیش‌بینی بارش، راه‌کار دیگر تجزیه داده‌ها تحت عنوان روش تجزیه مد متغیر (VMD) به جهت دستیابی به دقت و اطمینان بالاتر در پیش‌بینی‌ها و نیز مقایسه عملکرد آن با مدل تجزیه مد تجربی گروهی کامل موردتوجه قرار گرفت. در این مطالعه، داده‌های بارش ماهانه به‌منظور تجزیه سیگنال‌های پیچیده به اجزای ساده‌تر (IMFها) با استفاده از روش تجزیه مد متغیر تفکیک و به مدل ماشین بردار پشتیبان معرفی شد (Dragomiretskiy and Zosso, 2013). شکل (۵)، IMFهای تجزیه‌شده توسط VMD را برای داده‌های بارش ماهانه در ایستگاه تله‌زنگ نشان می‌دهد. با توجه به این شکل مشاهده می‌شود که تا سطح تجزیه سه، فرکانس‌های مرکزی به یکدیگر نزدیک شده و هیچ اختلاط مدی اتفاق نمی‌افتد. اما در سطح تجزیه چهار اختلاط مد رخ می‌دهد. بنابراین مقدار مناسب سطح تجزیه برای تحلیل داده‌های بارش ماهانه با یک تأخیر برابر با سه می‌باشد.

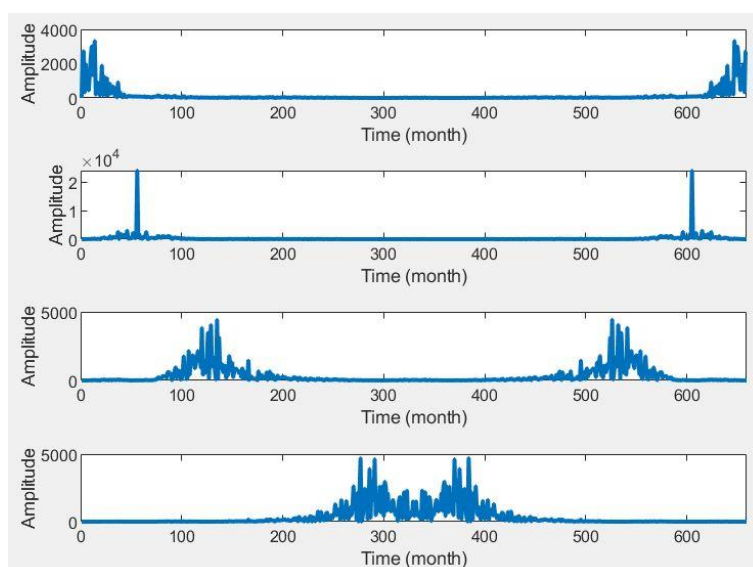


Figure 5. Frequency components obtained from VMD decomposition (levels 1–4)

همان گونه که بیان شد، هدف این مطالعه بررسی عملکرد مدل های تلفیقی مبتنی بر تجزیه سیگنال برای مدل سازی بارش ماهانه و مقایسه آن ها با یکدیگر است. عملکرد مدل با استفاده از شاخص های ارزیابی که کارایی و اثربخشی مدل را اندازه گیری می کنند، بررسی شد. نتایج در جدول (۵) خلاصه شده است. لازم به ذکر است، هنگامی که مقادیر NS و R نزدیک به یک و مقادیر RMSE و MAE نزدیک به صفر باشند، نشان می دهد که یک مدل در مقایسه با سایر مدل ها، کارایی بهتری دارد.

Table 5. Results of RMSE, MAE, R and NS evaluation indices for the VMD-SVM model

Model	Training phase					Testing phase				
	Pattern	RMSE (mm)	MAE (mm)	R	NS	Pattern	RMSE (mm)	MAE (mm)	R	NS
SVM	S4M1	56.949	38.462	0.799	0.637	S4M1	51.034	34.952	0.783	0.612
	S4M2	11.605	8.112	0.993	0.985	S4M2	20.448	15.037	0.930	0.910
	S4M3	7.892	5.545	0.997	0.993	S4M3	14.335	12.590	0.941	0.958
	S4M4	5.418	3.568	0.998	0.997	S4M4	9.404	8.858	0.968	0.966

در مقایسه عملکرد مدل های منفرد، تناوبی و مدل VMD، مدل ماشین بردار پشتیبان مبتنی بر VMD شاخص های کارایی و اثربخشی بهتری نسبت به دیگر مدل ها ارائه داد. از این نتیجه می توان دریافت که ترکیب VMD و مدل های منفرد مانند SVM می تواند عملکرد مدل منفرد را از نظر کارایی و اثربخشی بهبود بخشد. همچنین، عملکرد مدل SVM مبتنی بر VMD با مدل مبتنی بر CEEMD مقایسه شد. در نتیجه، مدل VMD-SVM عملکرد بهتری نسبت به مدل CEEMD-SVM ارائه داد، اما به طور کلی هنگامی که کارایی و اثربخشی مدل های مبتنی بر تجزیه با مدل های منفرد و تناوبی در نظر گرفته شد، مدل های VMD-SVM و CEEMD-SVM عملکرد بهتری داشتند. این نتایج نشان داد که تجزیه داده ها قادر به افزایش عملکرد مدل SVM است، درحالی که VMD برای بهبود عملکرد مدل ماشین بردار پشتیبان نسبت به CEEMD برای پیش بینی بارش ماهانه مناسب تر بود. در میان تمام مدل ها، چهار مدل برتر با بهترین کارایی و اثربخشی را می توان به ترتیب مدل های CEEMD-SVM، VMD-SVM، مدل ماشین بردار پشتیبان تناوبی و مدل منفرد SVM شناسایی کرد. شکل (۶) نمودار پراکندگی برای بهترین عملکرد مدل های مذکور را نشان می دهد. نمودارهای پراکندگی نشان دهنده درجه همبستگی و پراکندگی بین مقادیر تخمینی و مشاهده شده هستند.

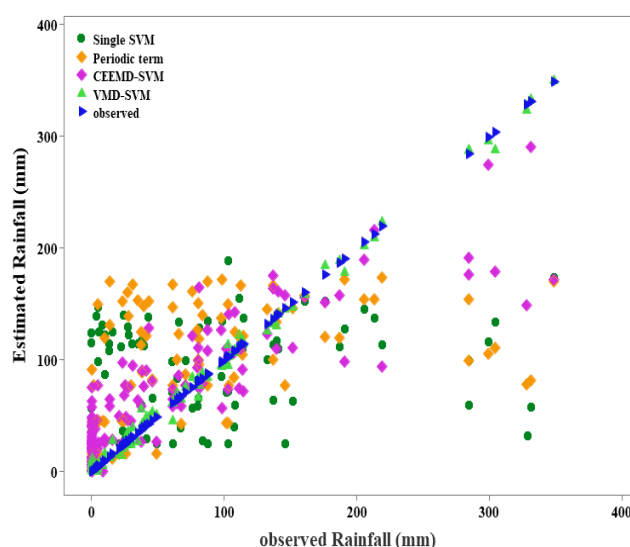


Figure 6. Observed versus modeled values: Scatter plot representation

از شکل (۶) مشاهده می‌شود که نقاط پراکندگی مدل‌های مبتنی بر VMD و CEEMD در مقایسه با مدل‌های منفرد و تناوبی، به خط شیب ۱:۱ نزدیک‌تر هستند. این نشان می‌دهد که مدل‌های یادگیری ماشین مبتنی بر VMD و CEEMD دقت بهتری نسبت به مدل‌های منفرد و تناوبی ارائه می‌دهند. از شکل می‌توان دریافت که مدل‌های منفرد و تناوبی نتایج بدتری نسبت به مدل‌های مبتنی بر VMD و CEEMD برای بارش شدید (کم و زیاد) ارائه دادند، حتی اگر همه مدل‌ها نتایج دقیقی برای داده‌های بارش ارائه دهند. به‌ویژه، همان‌طور که در این شکل مشاهده می‌شود، مدل‌های مبتنی بر تجزیه در مقایسه با مدل‌های منفرد و تناوبی برای بارش‌های بالاتر، بهترین مقادیر دقیق بارش را ارائه دادند. از آنجایی که مدل‌های یادگیری ماشین از داده‌های آموزشی با ساختار ساده‌تر برای هر مقیاسی که از تجزیه سری‌های زمانی در فرایند آموزش به‌دست می‌آید، استفاده می‌کنند، می‌توانند الگوی داده‌ها را به‌طور مؤثرتری یاد بگیرند. این ویژگی‌ها می‌توانند اثرات غیرخطی بودن و غیرایستابودن نهفته در سری‌های زمانی اصلی را کاهش دهند. بنابراین، مدل‌های تلفیقی مبتنی بر VMD و CEEMD می‌توانند به‌طور مؤثرتری با سری‌های زمانی غیرخطی و غیرایستا برخورد کنند و همچنین عملکرد مدل را بهبود بخشند.

۴. نتیجه‌گیری

در این مطالعه، دو مدل ترکیبی VMD-SVM و CEEMD-SVM، به‌منظور مدل‌سازی بارش ماهانه ایستگاه تله‌زنگ توسعه داده شدند و عملکرد مدل‌ها با استفاده از شاخص‌های کمی ارزیابی شد. نتایج با نتایج حاصل از مدل‌های منفرد و تناوبی مقایسه شد. نتایج نشان داد که افزودن ترم تناوبی به الگوهای ورودی مدل SVM می‌تواند دقت برآورد بارش را نسبت به حالت ساده بهبود بخشد لیکن مدل‌های یادگیری ماشین مبتنی بر VMD و CEEMD عملکرد بهتری نسبت به مدل‌های منفرد و تناوبی نشان دادند. همچنین برتری مدل VMD-SVM نسبت به CEEMD-SVM در پیش‌بینی بارش ماهانه کاملاً مشهود بوده است. به‌گونه‌ای که به‌کارگیری این روش توانسته مقدار خطای (RMSE) مدل منفرد را در بهترین عملکرد از ۷۴/۳۱ میلی‌متر به ۹/۴۰ میلی‌متر در مدل تجزیه مد متغیر کاهش دهد. درحالی‌که همین مقدار خطا در مدل منفرد ماشین بردار پشتیبان به مقدار ۴۳/۶۴ میلی‌متر در مدل CEEMD-SVM کاهش پیدا کرده است که نمایانگر عملکرد مطلوب‌تر مدل VMD-SVM در مدل‌سازی بارش ماهانه بوده است. مدل هیبریدی VMD-SVM با الگوی ورودی S4M4 مقدار شاخص‌های ارزیابی MAE و NS را نیز به‌ترتیب ۸/۸۵ میلی‌متر و ۰/۹۶ برآورد نمود. این نتایج تأیید می‌کند که VMD می‌تواند عملکرد مدل‌های یادگیری ماشین را برای مدل‌سازی بارش ماهانه بهبود بخشد و نیز بسیار بهتر از سایر روش‌های پیش‌پردازش داده‌ها نظیر CEEMD عمل نماید. بنابراین، VMD می‌تواند یک تکنیک جایگزین جدید برای مدل‌سازی بارش ماهانه براساس تجزیه سری‌های زمانی باشد.

۵. پی‌نوشت‌ها

1. Support Vector Machine (SVM)
2. Lag Based
3. Linear Periodic Term
4. Non Linear Periodic Term

۶. تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد.

۷. منابع

- Ahmadi. (2020). Performance evaluation of support vector machine methods and adaptive neurofuzzy inference system in predicting monthly river flow (Case study of Nazlu and Sezar rivers). *Iranian Soil and Water Research*, 51(3), 673-686. (In Persian).
- Ali, M., Prasad, R., Xiang, Y., & Yaseen, Z. M. (2020). Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. *Journal of Hydrology*, 584, 124647.
- Alvanitopoulos, P. F., Andreadis, I., Georgoulas, N., Zervakis, M., & Nikolaidis, N. (2014). Solar radiation time-series prediction based on empirical mode decomposition and artificial neural networks. In *Artificial Intelligence Applications and Innovations: 10th IFIP WG 12.5 International Conference, AIAI 2014, Rhodes, Greece, September 19-21, 2014. Proceedings 10* (pp. 447-455). Springer Berlin Heidelberg.
- Basak, D., Pal, S., & Patranabis, D. C. (2007). Support vector regression. *Neural Information Processing-Letters and Reviews*, 11(10), 203-224.
- Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567), 47-55.
- Bertsekas, D. P. (1976). Multiplier methods: A survey. *Automatica*, 12(2), 133-145.
- Bertsekas, D. P. (2014). *Constrained optimization and Lagrange multiplier methods*. Academic press.
- Dong, J., Wang, Z., Wu, J., Cui, X., & Pei, R. (2024). A novel runoff prediction model based on support vector machine and gate recurrent unit with secondary mode decomposition. *Water Resources Management*, 38(5), 1655-1674.
- Dong, S. (2024). Enhanced Prediction of Extreme Monthly Precipitation in Yinchuan City Using CEEMDAN-Decomposed LSTM and BP Neural Networks. In *2024 International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)* (pp. 491-496). IEEE.
- Dotse, S. Q., Larbi, I., Limantol, A. M., & De Silva, L. C. (2024). A review of the application of hybrid machine learning models to improve rainfall prediction. *Modeling Earth Systems and Environment*, 10(1), 19-44.
- Dragomiretskiy, K., & Zosso, D. (2013). Variational mode decomposition. *IEEE transactions on signal processing*, 62(3), 531-544.
- Du, J., Liu, Y., Yu, Y., & Yan, W. (2017). A prediction of precipitation data based on support vector machine and particle swarm optimization (PSO-SVM) algorithms. *Algorithms*, 10(2), 57.
- El-Shafie, A., Jaafer, O., & Seyed, A. (2011). Adaptive neuro-fuzzy inference system based model for rainfall forecasting in Klang River, Malaysia. *International Journal of Physical Sciences*, 6(12), 2875-2888.
- Essam, Y., Huang, Y. F., Ng, J. L., Birima, A. H., Ahmed, A. N., & El-Shafie, A. (2022). Predicting streamflow in Peninsular Malaysia using support vector machine and deep learning algorithms. *Scientific Reports*, 12(1), 38-83.
- Fardi Rad & Hosseini. (2025). Evaluation of a hybrid model for estimating monthly runoff using support vector machine based on variable mode analysis with whale optimization algorithm (Mashhad study basin). *Iranian Water Research*, 12(1), 39-53. (In Persian).
- Haddad, M. S. (2011). Capacity choice and water management in hydroelectricity systems. *Energy Economics*, 33(2), 168-177.
- Hartmann, H., Snow, J. A., Stein, S., Su, B., Zhai, J., Jiang, T., & Kundzewicz, Z. W. (2016). Predictors of precipitation for improved water resources management in the Tarim River basin: Creating a seasonal forecast model. *Journal of Arid Environments*, 125, 31-42.
- Hestenes, M. R. (1969). Multiplier and gradient methods. *Journal of optimization theory and applications*, 4(5), 303-320.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., & 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.
- Imran, M., Jannat Mishu, N. E., Khaliq, H., & Shahzad, F. (2025). Assessing Machine Learning Models for Precipitation Prediction in the Upper Indus Basin: A Comparative Analysis. *Environmental Science and Ecology*, 6(1), 10108.

- Jiang, L., & Wu, J. (2013). Hybrid PSO and GA for neural network evolutionary in monthly rainfall forecasting. In *Intelligent Information and Database Systems: 5th Asian Conference, ACIIDS 2013, Kuala Lumpur, Malaysia, March 18-20, 2013, Proceedings, Part I 5* (pp. 79-88). Springer Berlin Heidelberg.
- Kalteh, A. M. (2017). Enhanced monthly precipitation forecasting using artificial neural network and singular spectrum analysis conjunction models. *INAE Letters*, 2, 73-81.
- Lahmiri, S. (2016). Intraday stock price forecasting based on variational mode decomposition. *Journal of Computational Science*, 12, 23-27.
- Lahmiri, S., & Boukadoum, M. (2014, October). Biomedical image denoising using variational mode decomposition. In *2014 IEEE biomedical circuits and systems conference (BioCAS) proceedings* (pp. 340-343). IEEE.
- Lahmiri, S., & Shmuel, A. (2017). Variational mode decomposition based approach for accurate classification of color fundus images with hemorrhages. *Optics & Laser Technology*, 96, 243-248.
- Lian, L. (2022). Runoff forecasting model based on CEEMD and combination model: a case study in the Manasi River, China. *Water Supply*, 22(4), 3921-3940.
- Liu, W., Cao, S. Y., & He, Y. (2015). Ground roll attenuation using variational mode decomposition. In *77th EAGE Conference and Exhibition 2015* (Vol. 2015, No. 1, pp. 1-5). European Association of Geoscientists & Engineers.
- Liu, W., Cao, S., & Chen, Y. (2016). Applications of variational mode decomposition in seismic time-frequency analysis. *Geophysics*, 81(5), V365-V378.
- Liu, W., Liu, Y., Li, S., & Chen, Y. (2023). A review of variational mode decomposition in seismic data analysis. *Surveys in Geophysics*, 44(2), 323-355.
- Mahmood, O. A., Sulaiman, S. O., & Al-Jumeily, D. (2024). Forecasting for Haditha reservoir inflow in the West of Iraq using Support Vector Machine (SVM). *PloS one*, 19(9), e0308266.
- Montaseri & Ghavidel. (2017). Comparison of the performance of artificial intelligence models in estimating river water quality parameters during periods of drought and flood. *Water and Soil*, 30(6), 1733-1747. (In Persian).
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282-290.
- Nocedal, J., & Wright, S. J. (Eds.). (1999). *Numerical optimization*. New York, NY: Springer New York.
- Parsaie, A., Ghasemlounia, R., Gharehbaghi, A., Haghiabi, A., Chadee, A. A., & Nou, M. R. G. (2024). Novel hybrid intelligence predictive model based on successive variational mode decomposition algorithm for monthly runoff series. *Journal of Hydrology*, 634, 131041.
- Pham, B. T., Le, L. M., Le, T. T., Bui, K. T. T., Le, V. M., Ly, H. B., & Prakash, I. (2020). Development of advanced artificial intelligence models for daily rainfall prediction. *Atmospheric Research*, 237, 104845.
- Putri, A. L. R. (2024). Performance Analysis of the Support Vector Machine Algorithm in Predicting Rain Potential in DKI Jakarta. *Journal of Artificial Intelligence and Engineering Applications*, 3(3), 671-676.
- Rockafellar, R. T. (1973). A dual approach to solving nonlinear programming problems by unconstrained optimization. *Mathematical programming*, 5(1), 354-373.
- Roodaki, S., & Azizian, A. (2020). Uncertainty analysis due to the application of different infiltration methods on the performance of HEC-HMS model using GLUE algorithm. *Iran-Water Resources Research*, 16(2), 50-66. (In Persian).
- Samantaray, S., Das, S. S., Sahoo, A., & Satapathy, D. P. (2022). Monthly runoff prediction at Baitarani river basin by support vector machine based on Salp swarm algorithm. *Ain Shams Engineering Journal*, 13(5), 101732.
- Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., ... & Stadler, S. (2021). Can deep learning beat numerical weather prediction?. *Philosophical Transactions of the Royal Society*, 379(2194), 20200097.
- Shen, Z. Y., & Ban, W. C. (2023). Machine learning model combined with CEEMDAN algorithm for monthly precipitation prediction. *Earth Science Informatics*, 16(2), 1821-1833.
- Tian, Z., & Chen, H. (2021). Multi-step short-term wind speed prediction based on integrated multi-model fusion. *Applied Energy*, 298, 117248.
- Torres, M. E., Colominas, M. A., Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. In *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 4144-4147). IEEE.

- Tricha, A., & Moussaid, L. (2024). Evaluating machine learning models for precipitation prediction in Casablanca City. *Indonesian Journal of Electrical Engineering and Computer Science*, 35(2), 1325-1332.
- Trinh, T. A. (2018). The impact of climate change on agriculture: findings from households in Vietnam. *Environmental and resource economics*, 71(4), 897-921.
- Vapnik, V. (1998). The support vector method of function estimation. In *Nonlinear modeling: Advanced black-box techniques* (pp. 55-85). Boston, MA: springer us.
- Vapnik, V. (1995). *The nature of statistical learning theory* Springer. New York. 10: 978-1.
- Xu, H., Guo, Z., Cao, Y., Cheng, X., Zhang, Q., & Chen, D. (2024). Research on short-term precipitation forecasting method based on CEEMDAN-GRU algorithm. *Scientific Reports*, 14(1), 31885.
- Xue, Y. J., Cao, J. X., Wang, X. J., Li, Y. X., & Du, J. (2019). Recent developments in local wave decomposition methods for understanding seismic data: Application to seismic interpretation. *Surveys in Geophysics*, 40, 1185-1210.
- Yao, C., Yang, Z., Bai, H., Wu, Y., Gong, Z., & Feng, G. (2024). Ensemble prediction of winter precipitation in China based on Support Vector Machine Method. *Chinese Journal of Geophysics*, 67(10), 3670-3685.
- Yeh, J. R., Shieh, J. S., & Huang, N. E. (2010). Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Advances in adaptive data analysis*, 2(02), 135-156.
- Zakaria, Z. A., & Shabri, A. (2012). Streamflow forecasting at ungaged sites using support vector machines. *Applied Mathematical Sciences*, 6(60), 3003-3014.
- Zhang, S., Zhu, K., & Wang, C. (2025). A Novel Monthly Runoff Prediction Model Based on KVMD and KTCN-LSTM-SA. *Water*, 17(3), 460.
- Zhang, X., Shi, J., Chen, H., Xiao, Y., & Zhang, M. (2023). Precipitation prediction based on CEEMDAN-VMD-BILSTM combined quadratic decomposition model. *Water Supply*, 23(9), 3597-3613.
- Zhang, X., Zhang, Q., Zhang, G., Nie, Z., Gui, Z., & Que, H. (2018). A novel hybrid data-driven model for daily land surface temperature forecasting using long short-term memory neural network based on ensemble empirical mode decomposition. *International journal of environmental research and public health*, 15(5), 1032.
- Zhang, Y. A., Yan, B., & Aasma, M. (2020). A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM. *Expert systems with applications*, 159, 113609.