

RESEARCH

Open Access



Time series forecasting model for non-stationary series pattern extraction using deep learning and GARCH modeling

Huimin Han¹, Zehua Liu^{2,3*}, Mauricio Barrios Barrios⁵, Jiu hao Li², Zhixiong Zeng⁴, Nadia Sarhan⁷ and Emad Mahrous Awwad⁶

Abstract

This paper presents a novel approach to time series forecasting, an area of significant importance across diverse fields such as finance, meteorology, and industrial production. Time series data, characterized by its complexity involving trends, cyclicity, and random fluctuations, necessitates sophisticated methods for accurate forecasting. Traditional forecasting methods, while valuable, often struggle with the non-linear and non-stationary nature of time series data. To address this challenge, we propose an innovative model that combines signal decomposition and deep learning techniques. Our model employs Generalized Autoregressive Conditional Heteroskedasticity (GARCH) for learning the volatility in time series changes, followed by Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) for data decomposition, significantly simplifying data complexity. We then apply Graph Convolutional Networks (GCN) to effectively learn the features of the decomposed data. The integration of these advanced techniques enables our model to fully capture and analyze the intricate features of time series data at various interval lengths. We have evaluated our model on multiple typical time-series datasets, demonstrating its enhanced predictive accuracy and stability compared to traditional methods. This research not only contributes to the field of time series forecasting but also opens avenues for the application of hybrid models in big data analysis, particularly in understanding and predicting the evolution of complex systems.

*Correspondence:

Zehua Liu
liuzehualp@126.com

¹ Mechanical and Electrical Engineering College, Hainan Vocational University of Science and Technology, Haikou 571126, China

² College of Water Conservancy and Civil Engineering, South China Agricultural University, Guangzhou 510642, China

³ Informationization Construction Center, Guangdong Industry Polytechnic Guangzhou, Guangzhou 510300, China

⁴ College of Engineering, South China Agricultural University, Guangzhou 510642, China

⁵ Department of Computer Science and Electronics, Universidad de La Costa, Barranquilla, Colombia

⁶ Department of Electrical Engineering, College of Engineering, King Saud University, P.O. Box 800, 11421 Riyadh, Saudi Arabia

⁷ Department of Quantitative Analysis, College of Business Administration, King Saud University, Riyadh, Saudi Arabia



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Introduction

Navigating the intricate seas of time series forecasting is no small feat, considering its vast applications in domains like traffic management, finance, hydrology, and more [1]. The fusion of cutting-edge sensor technology, computing prowess, and advanced communication channels has bestowed upon us a treasure trove of rich time series data, reshaping the very fabric of how we monitor and control complex real-world systems [2]. In the economic realm, economists ride the waves of stock price fluctuations, foreseeing market trends. Medical maestros employ biological time series data to decipher diseases and gauge patient well-being. Meteorologists, armed with atmospheric time series data, predict the capricious dance of climate changes. Even in the realm of industrial production, time series data orchestrates a symphony to monitor product quality.

Yet, peering into the future isn't a stroll in the park. Time series data, with its intricate dance of trends, cycles, seasons, and unpredictable whims, poses a formidable challenge. Unraveling these intricacies demands a profound understanding to accurately predict the twists and turns that lie ahead. The history of time series forecasting is a rich tapestry, woven with threads of linear regression, autoregressive and moving average models, and the contemporary marvels of deep learning and neural networks [2–4]. However, as the data at hand often boasts nonlinearity and non-stationarity, existing methods struggle to extract sufficient features for precise forecasting. Here steps in the unsung hero—signal decomposition. Signal decomposition, akin to an alchemical process, breaks down the complexity of time series data into digestible components—trends, cycles, and random noise. Classical methods like Fourier transform, wavelet transform, and empirical mode decomposition have peeled back the layers, offering glimpses into the intrinsic structure of time series data, becoming indispensable references for forecasting.

Enter the disruptor, deep learning—a juggernaut capable of plumbing the depths of input data with its nonlinear modules [5, 6]. As each module transforms representation levels, deep learning unfurls a multi-level tapestry of understanding, applied with resounding success in realms like computer vision [7, 8], speech recognition [9], and natural language processing [10]. However, the challenge persists, especially with nonlinear and non-stationary datasets. To surmount this, our paper unveils a novel time series forecasting model, seamlessly integrating signal processing and deep learning. A unique network structure emerges, designed to pluck features from sequence data at varying intervals.

Here lie our contributions:

- We propose a groundbreaking time series forecasting model, marrying the realms of signal processing and deep learning. Our custom-designed network structure is the maestro orchestrating the symphony of feature extraction.
- We harness GARCH to unravel the volatility nuances of time series changes. The processed data undergoes a transformative dance with CEEMDAN, a signal decomposition virtuoso, significantly decluttering the complexity. Finally, the baton passes to GCN, orchestrating the learning of data features.
- Our model faces the crucible of assessment against several typical time-series datasets, emerging with enhanced predictive accuracy and heightened stability. The results affirm the transformative potential of our proposed fusion model in the intricate realm of time series forecasting.

Related work

Embarking on the realm of time series data prediction unveils a rich tapestry woven with three distinctive approaches—one delving into statistical forecasts, another navigating the terrain of machine learning, and the third boldly venturing into the uncharted waters of hybrid models. Within the expansive domain of machine learning, a dual landscape emerges, encompassing both traditional methods and the sophisticated depths of deep learning. Simultaneously, hybrid predictive models take center stage, harmonizing the symphony of signal decomposition within time series data and the orchestration of predictive prowess.

In the early stages of time series forecasting exploration, intrepid researchers turned to the tried-and-true path of statistical principles. Burlando P and team unfurled the Autoregressive Moving Average (ARMA) model, forecasting short-term rainfall with commendable success [11]. However, grappling with the tumultuous seas of strongly non-stationary data, the ARMA model faced limitations. To surmount this challenge, a strategic evolution unfolded, giving rise to the Autoregressive Integrated Moving Average (ARIMA) model. This model, showcased by Meyler A in predicting inflation in Ireland, exhibited prowess in dissecting the periodicity and volatility of time series data [12]. Further refinements, such as the Seasonal ARIMA (SARIMA) model by Williams et al., adorned with a seasonal pattern, paved the way for predicting single variable automobile traffic data [13]. Yet, statistical-based models, despite their achievements, found their capabilities somewhat constrained in the face of increasingly complex time series data tasks.

The rise of machine learning heralded a new era, offering a diverse arsenal of solutions. Following this trajectory, Chen et al. and Tay FEH, Cao L introduced

Support Vector Machines (SVMs) [14, 15] to forecast financial time series data, showcasing superior performance in nonlinear feature extraction and resilience to data noise [16]. LV et al. embraced the simplicity and adaptability of the K-Nearest Neighbors (KNN) method for short-term power load prediction [17]. Meanwhile, the realm of Artificial Neural Networks (ANNs) unfolded as a powerful tool, with Hung N Q successfully predicting rainfall in Bangkok and Maleki and Goudarzi using ANNs for air AQI sequence data [18–20]. The prowess of deep learning, riding the waves of big data and computational power, unfolded with Recurrent Neural Networks (RNNs). Tokgöz A and Ünal G demonstrated its effectiveness in predicting Turkey’s power load [21]. However, challenges of vanishing or exploding gradients led to the optimization of RNNs, birthing the Long Short-Term Memory networks (LSTM). Chang Y S et al. harnessed LSTM’s capabilities in predicting air pollution data [22, 23]. Integration with Convolutional Neural Networks (CNN) further enhanced predictions, exemplified by Zha et al.’s work in predicting natural gas production [24]. The artistry of Graph Convolutional Networks (GCN) found expression in predicting road

traffic speed, as demonstrated by Yu B et al. [25]. Bhatti et al. identifies importance of GCN and its variants in the field of science [26].

In recent years, a cadre of scholars has embraced the avant-garde—hybrid models. These models, fusing signal analysis and prediction, promise superior performance. Mainstream signal decomposition methods such as Empirical Mode Decomposition (EMD), Ensemble EMD (EEMD) [27], and Variational Mode Decomposition (VMD) [28] took the stage. Scholar Zhang W et al. wielded EMD coupled with SVM for short-term power load prediction [29]. Shu and team, innovators, blended EMD with CNN and LSTM, achieving superior predictive results [30]. Experimental validations substantiate the supremacy of such hybrid models over their singular counterparts. Addressing the mathematical gaps in EMD, an improved version emerged Ensemble EMD with Adaptive Noise (EEMDAN). Yan Y et al. harnessed EEMDAN with LSTM to predict wind speed with precision [31]. Zhu Q et al. navigated the challenges of non-stationary signals, combining Variational Mode Decomposition (VMD) with Bi-GRU for rubber futures time series prediction [32]. A further evolution, Complementary Ensemble Empirical

Table 1 Advantages and disadvantages of several common sequence decomposition methods

Series Decomposition Methods		
Method	Advantage	Disadvantage
VMD	Featuring rigorous theoretical derivation, widespread application, and relatively low computational complexity [34].	the decomposed results are heavily restricted to the selection of the penalty parameter α and the number of sub-modes K [35].
EMD	Able to decompose signals adaptively [36].	mode mixing, end effect, poor noise immunity [37].
EEMD	Has strong adaptability, effectively overcoming the phenomenon of mode mixing [38].	its reconstruction error is large and its integrity is poor [39].
CEEMD	Ensure decomposition effectiveness, while reducing reconstruction errors caused by white noise [40].	There is a loss of information for high frequency components [41].
CEEMDAN	Successfully resolved the issue of white noise transmission from high frequency to low frequency [42].	There exist noise after decomposing complex sequential data [43].

Table 2 Advantages and disadvantages of several commonly used sequence prediction methods

Prediction Methods		
Method	Advantage	Disadvantage
SVM	Possessing good generalization ability and being insensitive to noisy data [44].	difficulty occurs in selecting suitable kernel [45].
KNN	The principle is simple and there are few influencing factors [46].	The computational complexity is high, and the algorithm performance is easily affected by the samples [17].
ANN	Able to learn non-linear features of data [47].	Easily trapped in local optima [48].
LSTM	Has the ability to capture long-term sequence features [49].	difficult to balance the depth and operation time complexity of its model [50].
CNN-LSTM	Reducing the risk of overfitting can uncover more feature information [51, 52].	There is a problem of neural network model degradation [53].
GCN	High efficiency, possessing the ability to effectively learn data relationships, and providing interpretability [54].	There is a problem of excessive smoothing [55].

Mode Decomposition with Adaptive Noise (CEEMDAN), resolved the white noise transmission conundrum. Zhou F et al. showcased the power of CEEMDAN and LSTM, painting a portrait of carbon price forecasts [33]. We have compiled and summarized the above methods, and Table 1 and 2 display these summaries.

Considering the advantages and disadvantages of several methods, this study opts for CEEMDAN for the decomposition of time series. This is because, compared to VMD and EEMD, CEEMDAN has stronger adaptability and stability. Compared with models such as SVM, KNN, and ANN, the GCN has a stronger ability to learn data relationships and higher computational efficiency. Compared with LSTM and CNN-LSTM, the GCN has fewer parameters and a simpler structure. Considering all this, we chose the GCN model as the predictive model.

Proposed methodology

Prior to discussing the forecasting method of the GARCH-CEEMDAN-GCN combination model for time series data, our first step is to provide a succinct account of the foundational theories applied in the establishment of this aforementioned hybrid model. This involves detailing the principles and practicalities of the Autoregressive Conditional Heteroskedasticity model (GARCH), the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), and the Graph Convolutional Neural Network (GCN).

GARCH

Bollerslev, building on the idea of ARMA modeling, constructed the Generalized Autoregressive Conditional Heteroscedasticity model (GARCH). The GARCH model is commonly utilized for modeling financial time series in investment and financial management. It's specifically designed for varying volatility, meaning we can predict different degrees of volatility for different points in time.

The basic form of the GARCH model is as follows:

Assume that the disturbance series u_t has the following structure:

$$u_t = \sqrt{h_t} \epsilon_t \tag{1}$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \epsilon_t \sim N(0, 1) \tag{2}$$

In which, $\sigma_\epsilon^2 = 1$, and $\alpha_0 > 0, \alpha_i \geq 0, \beta_i \geq 0$. Since $\{\epsilon_t\}$ is a white noise process and is independent of the past values of u_{t-i} , both the conditional and unconditional mean of u_t are zero. However, the conditional variance of u_t equals $E_{t-1} u_t^2$, which is h_t . This model, which allows for conditional heteroscedasticity with both autoregressive terms and moving average terms,

is called the Generalized Autoregressive Conditional Heteroskedasticity model, denoted as GARCH(p, q). Clearly, if $p=0, q=1$, then GARCH(0, 1) is the same as the ARCH(1) model. If all β_t are zero, then the GARCH(p, q) is equivalent to the ARCH(q) model. Therefore, the GARCH model can be seen as a generalization of the ARCH model, or the ARCH model can be seen as a special case of the GARCH model.

Introduce the lag operator polynomial:

$$\alpha(B) = \alpha_1 B + \dots + \alpha_q B^q \tag{3}$$

$$\beta(B) = \beta_1 B + \dots + \beta_p B^p \tag{4}$$

Thus, the GARCH (p, q) model can be represented as:

$$h_t = \alpha_0 + \alpha(B)u_t^2 + \beta(B)h_t \tag{5}$$

If the roots of $1-\beta(B)=0$ are all outside the unit circle, then

$$h_t = \frac{\alpha_0}{1-\beta(1)} + \frac{\alpha(B)}{1-\beta(B)}u_t^2 = a_0^* + \sum_{j=1}^{\infty} \delta_j u_{t-j}^2 \tag{6}$$

where B is the one-step lag operator at time t, $a_0^* = \frac{\alpha_0}{1-\beta(1)}$, δ_j is the coefficient of B^j in the expansion of $\frac{\alpha(B)}{1-\beta(B)}$.

Therefore, it can be seen that if a series follows a GARCH(p, q) process, under certain conditions, it can be represented by an infinite order ARCH process with a reasonable lag structure. Thus, in practical applications, for a high-degree ARCH model, it can be represented by a relatively simple GARCH model, which reduces the estimation parameters and facilitates the identification and estimation of the model.

CEEMDAN

CEEMDAN, standing for Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, is an enhanced version of the EEMD algorithm. It employs a partitioned approach to signal processing and gives rise to the intrinsic mode functions (IMFs) of the signal by way of several decompositions, as demonstrated in Algorithm 1. The unique feature of CEEMDAN, as compared to EEMD, is its adaptive noise generation during the decomposition phase, allowing the process to be steadier and more precise. Furthermore, it capably addresses the endpoint effect and mode mixing issues inherent in conventional EMD. The sequence of operations in the CEEMDAN algorithm is presented in Fig. 1. The procedures for the CEEMDAN algorithm are outlined as follows:

1. Incorporate Gaussian white noise into the initial signal $f(t)$, as demonstrated in Eq. (7):

$$f_i(t) = y(t) + \varepsilon_0 n_i(t), i = 1, 2, \dots, N \tag{7}$$

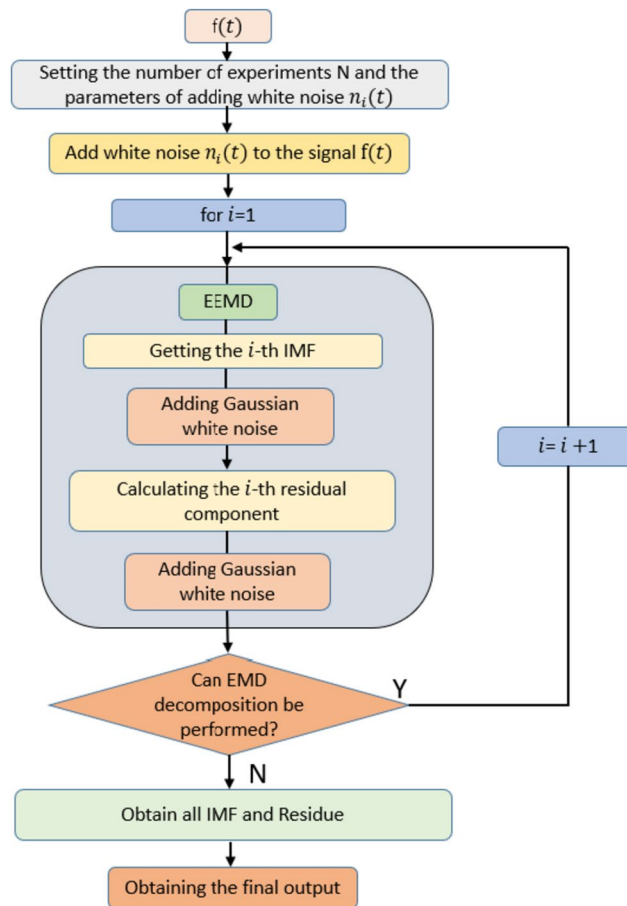


Fig. 1 The flowchart of CEEMDAN

The symbol ε_0 denotes the ratio of signal-to-noise, $n_i(t)$ symbolizes the Gaussian white noise introduced during the i -th cycle, and N stands for the count of cycles.

2. Carry out EMD decomposition on the newly added signal $f_i(t)$, which is interleaved with Gaussian white noise, to ascertain the first intrinsic mode function, designated as IMF_1 , along with its associated residual component r_1 . These are represented in Eqs. (8) and (9).

$$IMF_1 = \frac{1}{N} \sum_{i=1}^N E_1[f_i(t)] \tag{8}$$

$$r_1(t) = x(t) - IMF_1 \tag{9}$$

In formula (6), E represents the EMD decomposition operation.

3. Add $\varepsilon_1 E_1[n_i(t)]$ to $r_1(t)$, perform EMD decomposition, and then perform summation and averaging to obtain IMF_2 , as shown in formula (10):

$$IMF_2 = \frac{1}{N} \sum_{i=1}^N E_1\{r_1(t) + \varepsilon_1 E_1[n_i(t)]\} \tag{10}$$

4. When $k = 2, 3, \dots, K$, use the corresponding intrinsic mode function to compute the k -th residual component $r_k(t)$, as depicted in Eq. (11):

$$r_k(t) = r_{k-1}(t) - IMF_k \tag{11}$$

5. At every stage, introduce Gaussian white noise to produce a fresh signal, subsequently determine its primary intrinsic mode component, and employ this as a novel mode component for the original signal. Consequently, the k -th mode component in this stage is IMF_{k+1} , as shown in formula (12):

$$IMF_{k+1} = \frac{1}{N} \sum_{i=1}^N E_1\{r_k(t) + \varepsilon_k E_k[n_i(t)]\} \tag{12}$$

6. Carry out the procedures detailed in steps 4 and 5 repeatedly. Once the signal is no longer susceptible to EMD decomposition, the ultimate k mode components and the final residual component of the signal will be realized, as depicted in Eq. (13):

$$R(t) = f(t) - \sum_{k=1}^K IMF_k \tag{13}$$

7. The ultimate result of the CEEMDAN decomposition of the signal $f(t)$ is represented in Eq. (14):

```

Determining N as the number of experiments and  $n_i(t)$  as the
parameters for white noise.
Add  $n_i(t)$  to the signal  $f(t)$ 
For  $i=1; i \leq N; i++$ ;
 $f_i(t) = f(t) + n_i(t)$ ;
EEMD decomposition of  $f_i(t)$ ;
    Getting the  $i$ -th IMF ;
    Add Gaussian white noise;
    Calculating the residual component;
    Add Gaussian white noise to residual component;
    Until the signal can no longer be decomposed by EMD:
    Break;
End For
Obtain all IMF and Residue.
End
    
```

Algorithm 1. CEEMDAN

$$f(t) = \sum_{i=k}^K IMF_{k+1} + R(t) \tag{14}$$

$$H^{(l+1)} = \sigma(D^{-1/2}(A + I)D^{-1/2}H^{(l)}W^{(l)}) \tag{18}$$

GCN

A graph can be defined as: $G=(V, E)$, where V represents the collection of all nodes within the graph, and E represents the set of all edges connecting nodes within the graph. The fundamental purpose of a Graph Convolutional Network (GCN) is to discern and extract specific features within graph data. This distinguishes it from traditional Convolutional Neural Networks (CNNs) because a GCN has the capacity to identify and extract spatial features from non-Euclidean topological graphs, determining correlations among the different data points within the graph structure. The principle at the heart of the GCN is to build a system akin to a message-passing mechanism. It starts by drawing from the original graph-structured data, continuously identifying features and transmitting data, with ongoing updates made to the information related to both the target node itself and its associated spatial domain. The equation for GCN is as follows:

$$D_{ii} = \sum_j A_{ij} \tag{15}$$

$$H^{(0)} = X \tag{16}$$

$$H^{(1)} = \sigma(D^{-1/2}(A + I)D^{-1/2}XW^{(0)}) \tag{17}$$

In the provided equation, A stands for the adjacency matrix, which outlines the spatial relations of the original data points. I denote the identity matrix. Furthermore, $A + I$ represents an augmented adjacency matrix where self-loops or self-connections have been introduced. D is the degree matrix that signifies the relationship between each node and its neighbors within the graph. The value corresponding to a particular node in the degree matrix is higher if the node is connected to more nodes in the graph. The quantity $D^{-1/2}(A + I)D^{-1/2}$ pertains to the normalization procedure applied to the adjacency matrix. The utility of this operation is to alleviate issues frequently encountered in deep learning, such as the gradient vanishing or exploding problem. The matrix X is a compilation of vertex feature data, embodying the feature matrix. Lastly, W signifies the weight of every edge connection in the graph; $H^{(1)}$ symbolizes the outcome derived from the initial message propagation round, while $H^{(l+1)}$ denotes the results accumulated from the $(l+1)$ -th round of updates. $W^{(l)}$ corresponds to the connection weight parameter that's been updated and aggregated at the current state. The initialization of the W parameter within the context of the Graph Convolutional Network (GCN) algorithm isn't overly stringent. In contrast with other deep learning models, a GCN allows effective node feature updates through the stacking of shallow layers. This results in a lower volume of parameters and reduces the computational time complexity.

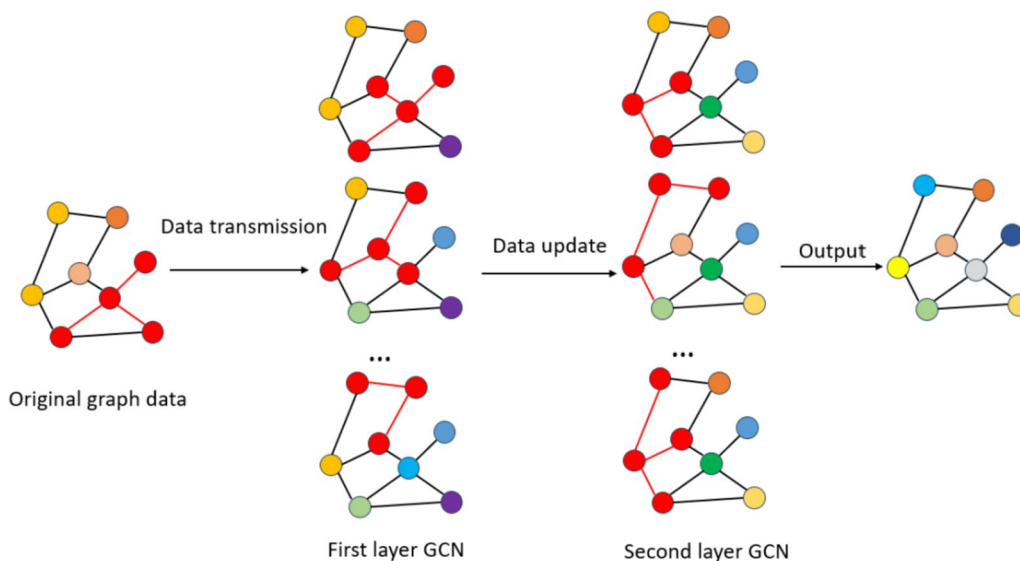


Fig. 2 Structure diagram of GCN

In the context of the raw graph-structured data, continuous feature extraction and data relay processes act on the information within the target node and its adjacent nodes, resulting in dynamic updates. This concept is depicted in Fig. 2.

Proposed methodology

The model introduced in this paper achieves highly accurate prediction of time series data by integrating the GARCH model, the CEEMDAN signal decomposition method, and the GCN deep learning method. Time series data has complex characteristics such as high nonlinearity, non-stationarity, and inherent noise in the data itself. The effect is often less than ideal when using a single model for prediction. The GARCH model possesses the ability to capture the dynamic volatility of the time series data. The inputted time series data will become easier to analyze after processing with the GARCH model. The residuals of the GARCH model are further analyzed using the CEEMDAN method, which decomposes the complex signal into a series of intrinsic mode functions (IMFs), contributing to unveiling various frequency components hidden behind the data and can further reduce the complexity of the time series data. Finally, feature extraction from the decomposed IMFs is performed using the Graph Convolutional Neural Network model and predictions are made based on this. The detailed modeling steps are as follows:

- Preprocess the raw data, fill the missing values in the data with the average value, standardize the data, and arrange all serial data in a uniform chronological order.

- Process the sequence data to be predicted with the GARCH model.
- Decompose the data processed by the GARCH model using the CEEMDAN algorithm to get their modal components, IMFs.
- Use the modal components obtained from the decomposition to build input features for the GCN, then use the GCN network model for prediction to get the final prediction results. The overall structure diagram is shown in the Fig. 3.

Dataset and evaluation metrics

Dataset

This research utilizes three comprehensive public datasets—Energy, Traffic, and Air Quality—to conduct a thorough analysis. The Air Quality dataset, in particular, plays a pivotal role in this study. Sourced from the China National Environmental Monitoring Center’s website (www.cnemc.cn), it provides an extensive record of daily air quality measurements from Guangzhou, Guangdong Province, China. This dataset covers a significant period, from January 1, 2017, to August 14, 2021, offering a rich temporal range for analysis. The dataset is equipped with sensor-based measurements that capture various environmental parameters, providing insights into the air quality trends and patterns over the years. Figure 4 in the study offers a detailed visualization of this dataset, presenting the data in a format that is both informative and accessible. To complement this, Table 3 systematically presents the statistical information derived from the dataset, offering a quantitative overview that includes metrics such as mean values, standard deviations, and other relevant statistical measures.

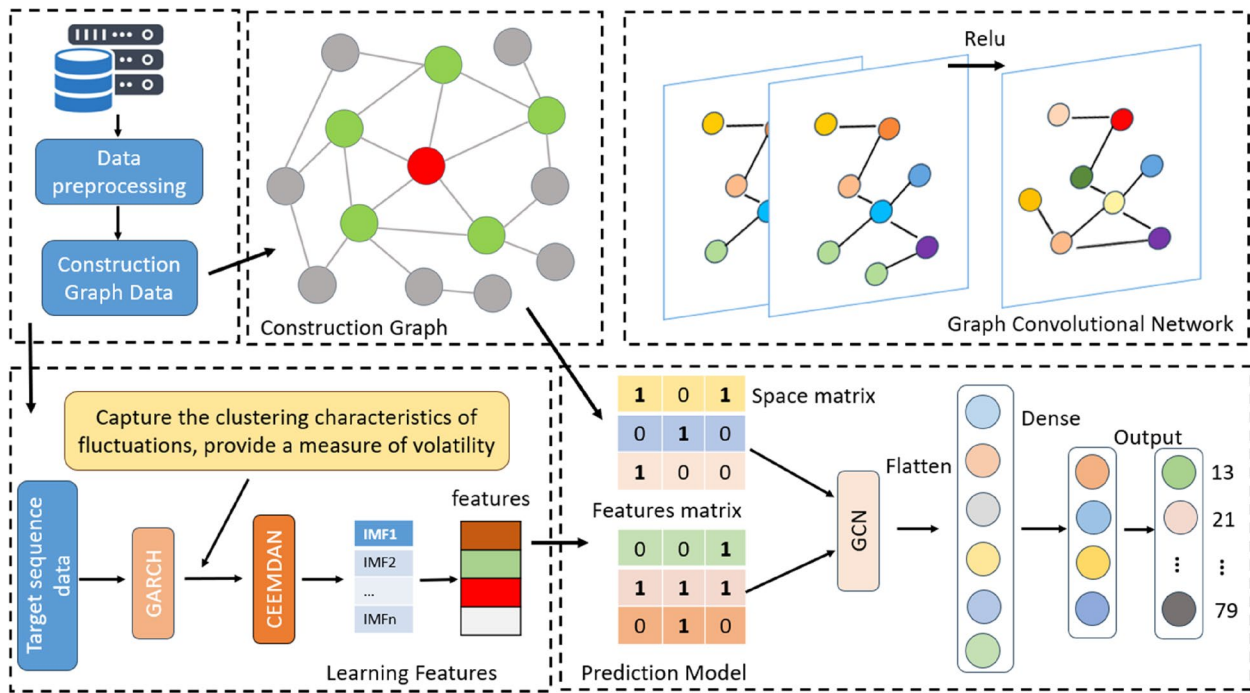


Fig. 3 The flow chart of the proposed model

These datasets collectively facilitate a multi-faceted exploration of environmental factors, allowing for a comprehensive understanding of the interplay between air quality, energy consumption, and traffic patterns. The Air Quality dataset, with its extensive coverage and detailed measurements, is particularly instrumental in shedding light on the environmental dynamics of Guangzhou, thereby contributing significantly to the broader objectives of the study.

The Energy dataset, sourced from Kaggle, encompasses energy data logged by sensors from August 4, 2022, through April 23, 2023, at a quarter-hourly sampling frequency. A visual representation of this dataset is provided in Fig. 5, while its statistical overview can be found in Table 4.

The Traffic dataset, also obtained from Kaggle, comprises data related to traffic flow captured hourly through road sensors between November 1, 2015, and June 30, 2017. A graphic representation of the Traffic dataset is displayed in Fig. 6, with comprehensive statistics provided in Table 5.

Figure 7 illustrates box plots for all three datasets: Air Quality, Energy, and Traffic. The parameter configurations used in this study are furnished in Table 6.

Evaluation metrics

This study makes use of four distinct evaluation indices, namely MAE, MSE, MAPE, and R², to assess the efficacy of the introduced prediction model. The acronyms represent Mean Absolute Error, Mean Square Error, Mean Absolute Percentage Error, and R Squared in

order. The mathematical representations of these metrics are provided below:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \tag{19}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \tag{20}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i} \tag{21}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2} \tag{22}$$

In the computation of the four appraisal measures, namely, Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R²), y_i stands for the true value of the model’s input sample, while \hat{y}_i indicates the prediction rendered by the model. The term n denotes the quantity of input samples and i signifies the index of each sample. For the indicators MAE, MSE, and MAPE, a lower value denotes superior model performance. As for R², its range is between 0 and 1, inclusive. An R² value closer to 1 signifies a superior performance attributed to the model.

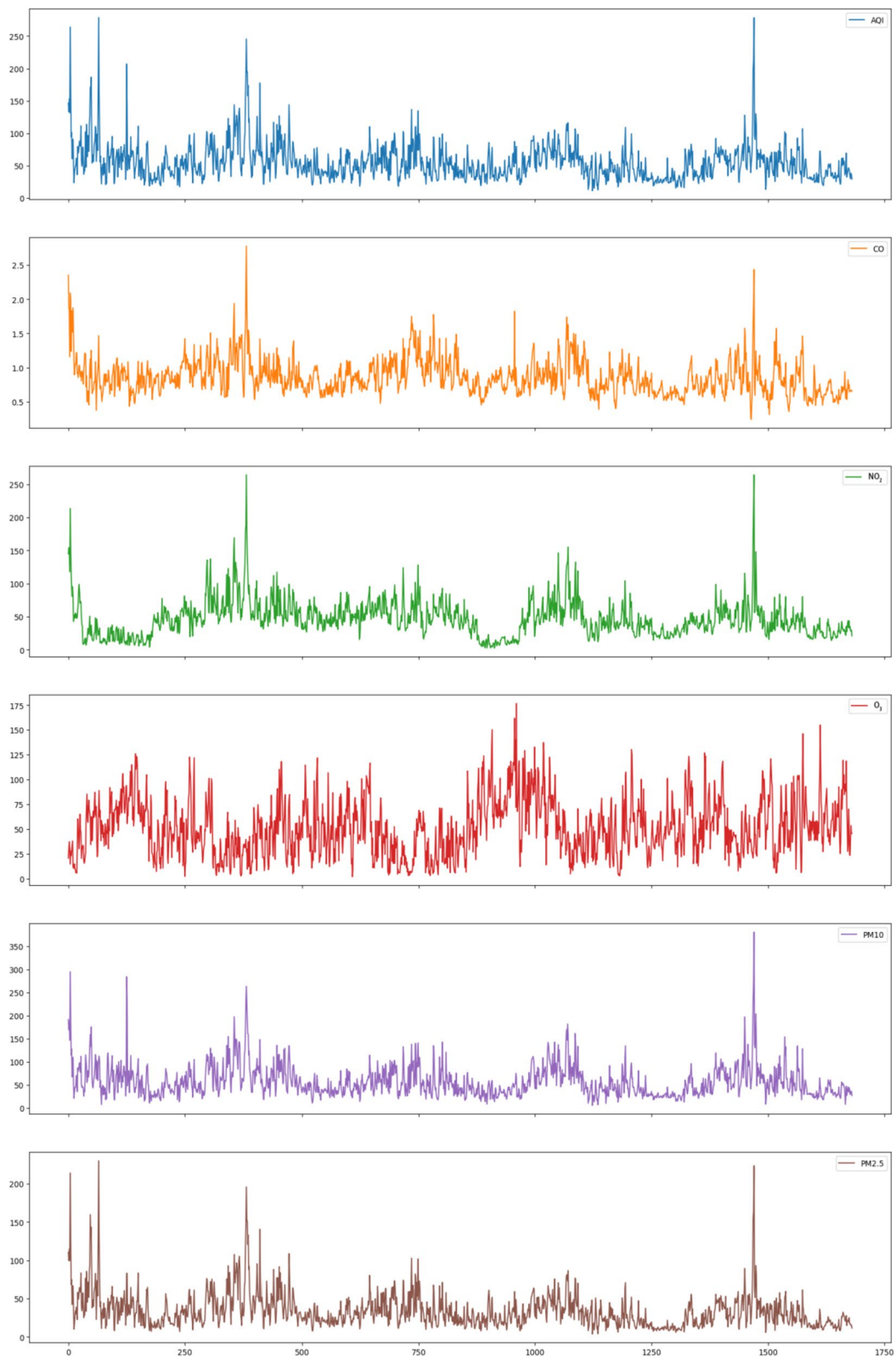


Fig. 4 Visualization of air quality data set display

Table 3 Statistical description of AQI and air pollutant patterns 2017–2021

Year	Analysis	Statistical analysis of air pollutants						Mean change from previous year (%)					
		AQI	CO	NO ₂	O ₃	PM10	PM2.5	AQI	CO	NO ₂	O ₃	PM10	PM2.5
2021	max	278.42	2.44	264.33	155.04	380.71	223.54						
	min	13.23	0.21	14.08	5.88	8.00	5.92						
	mean	51.57	0.74	40.76	58.47	55.57	28.90	15.81%	-5.13%	0.41%	17.96%	18.83%	17.22%
	std	29.09	0.27	27.90	27.21	41.59	23.21						
	median	44.65	0.68	34.88	55.00	44.03	22.29						
2020	max	127.83	1.56	119.25	130.42	197.21	89.38						
	min	11.13	0.39	12.17	2.91	5.63	3.83						
	mean	44.52	0.78	40.61	49.61	46.76	24.66	-18.97%	-17.03%	-11.93%	-13.12%	-19.74%	-27.02%
	std	18.90	0.20	17.67	25.25	25.80	14.00						
	median	41.44	0.74	37.81	47.21	40.44	21.72						
2019	max	136.71	1.83	155.33	176.8	181.78	102.79						
	min	18.86	0.45	2.70	3.14	8.42	7.92						
	mean	54.87	0.94	46.08	57.08	58.20	33.80	-7.33%	6.83%	-23.05%	26.85%	-3.81%	-13.17%
	std	21.97	0.26	28.49	34.11	32.38	17.29						
	median	50.42	0.88	45.53	54.63	50.00	30.54						
2018	max	245.5	2.78	264.43	121.92	263.54	195.63						
	min	18.17	0.48	15.71	2.00	10.94	7.96						
	mean	59.21	0.88	59.84	45.02	60.51	38.89	-6.06%	-4.36%	32.79%	-10.21%	-8.27%	-6.79%
	std	29.78	0.25	25.16	26.24	33.40	24.33						
	median	53.20	0.85	54.84	42.10	51.96	32.67						
2017	max	278.71	2.35	213.46	126.00	294.75	229.92						
	min	17.42	0.37	4.13	2.33	7.67	7.13						
	mean	63.00	0.92	45.09	50.15	65.95	41.67						
	std	33.90	0.27	32.70	27.26	39.32	27.96						
	median	57.04	0.90	39.92	48.46	58.06	35.70						

Results

We have carried out experimental investigations on three distinct datasets, namely, Air Quality, Energy, and Traffic, comparing the output of five diverse models – GCN, EMD-GCN, EEMD-GCN, EMD-CEEMDAN-GCN, and the newly proposed model GARTCH-CEEMDAN-GCN. Figure 8 illustrates the performance of the models when applied to the Air Quality dataset. Similarly, the models’ performance against Energy and Traffic datasets can be observed in Figs. 9 and 10 respectively.

Figure 8 shows the evaluation performances of different models on the AQI dataset. Compared with other models, the MAE metric of GARCH-CEEMDAN-GCN is 7.5% better than GCN, 6.5% worse than EMD-GCN, 10.6% worse than EEMD-GCN, and 7.2% better than EMD-CEEMDAN-GCN. The MSE metric of GARCH-CEEMDAN-GCN is 9.9% better than GCN, 19.6% worse than EMD-GCN, 32.4% worse than EEMD-GCN, and

9.9% better than EMD-CEEMDAN-GCN. The MAPE metric of GARCH-CEEMDAN-GCN is 17.5% better than GCN, 2.6% worse than EMD-GCN, 1.8% better than EEMD-GCN, and 8.4% better than EMD-CEEMDAN-GCN. The R2 metric of GARCH-CEEMDAN-GCN is 24.1% better than GCN, 23.4% worse than EMD-GCN, 30.8% worse than EEMD-GCN, and 24.1% better than EMD-CEEMDAN-GCN.

Figure 9 shows the evaluation performances of different models on the ENERGY dataset. Compared with other models, the MAE metric of GARCH-CEEMDAN-GCN is 61.1% better than GCN, 53.3% better than EMD-GCN, 54.9% better than EEMD-GCN, and 13.6% better than EMD-CEEMDAN-GCN. The MSE metric of GARCH-CEEMDAN-GCN is 86.9% better than GCN, 74.5% better than EMD-GCN, 73.4% better than EEMD-GCN, and 19.7% better than EMD-CEEMDAN-GCN. The MAPE metric of GARCH-CEEMDAN-GCN is 56.2% better than

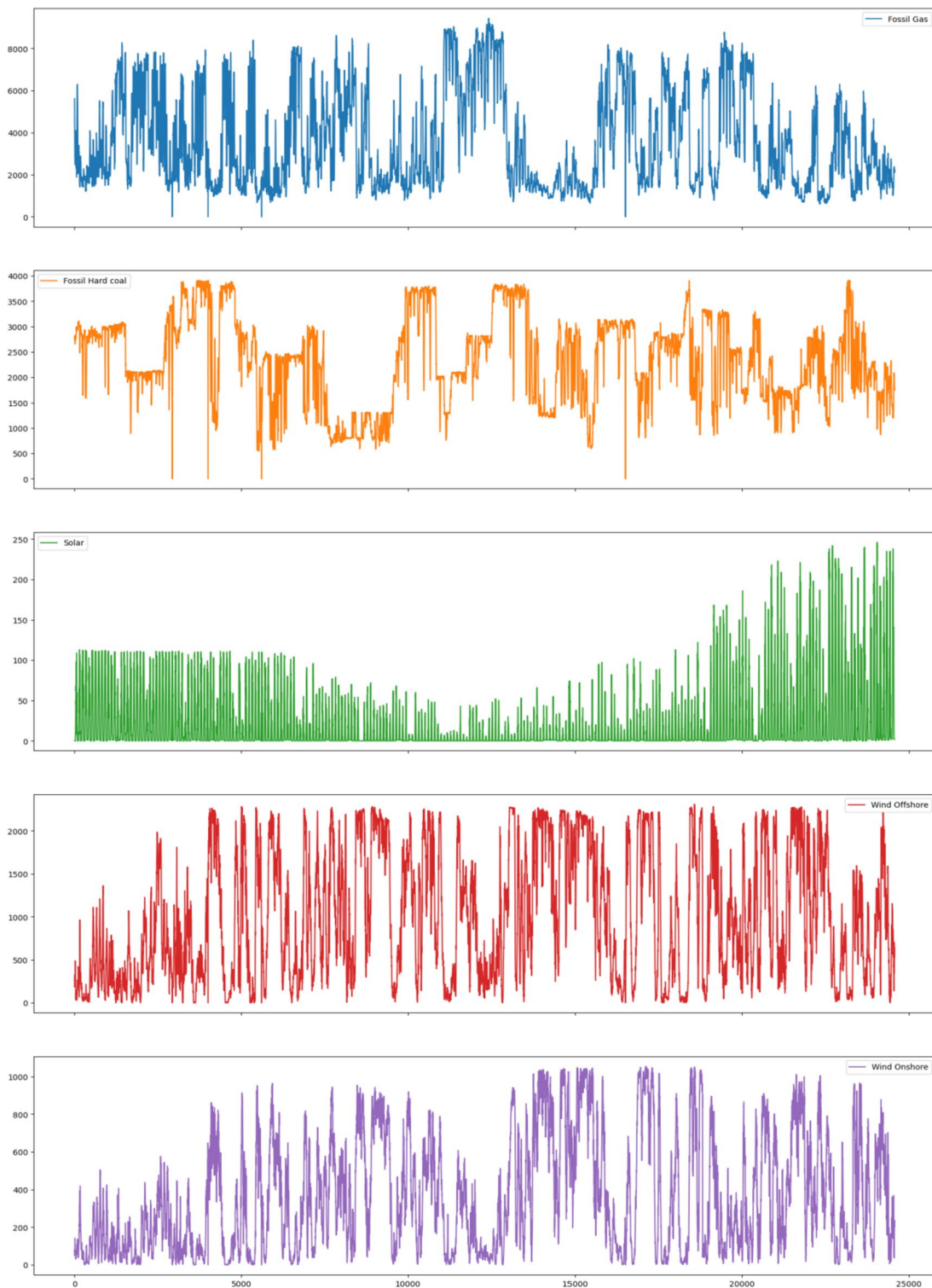


Fig. 5 Visualization of energy data set display

GCN, 38.2% better than EMD-GCN, 50.5% better than EEMD-GCN, and 17.4% better than EMD-CEEMDAN-GCN. The R2 metric of GARCH-CEEMDAN-GCN is

3.1% better than GCN, 1.0% better than EMD-GCN, 1.0% better than EEMD-GCN, and is the same as EMD-CEEMDAN-GCN.

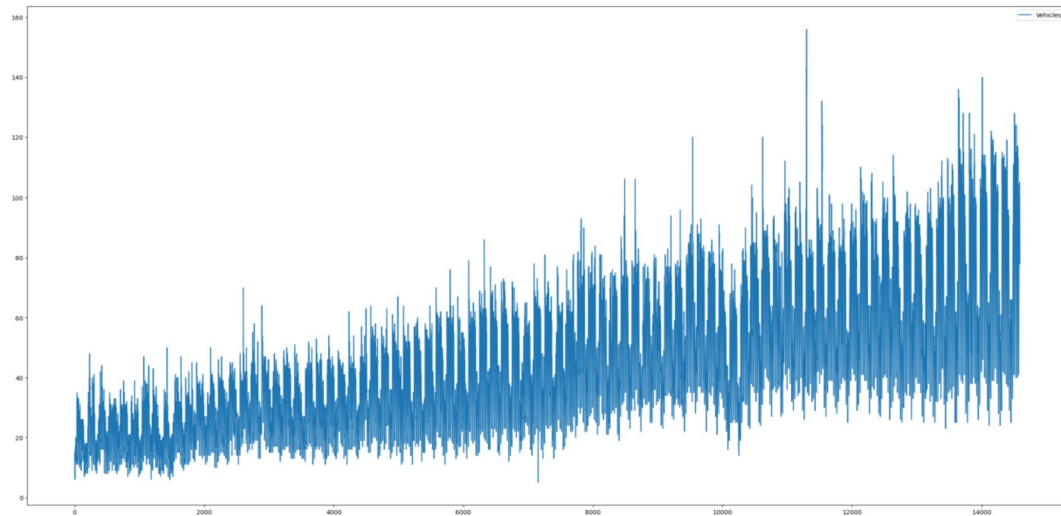


Fig. 6 Visualization of transportation data set display

Figure 10 shows the evaluation performances of different models on the TRAFFIC dataset. Compared with other models, the MAE (Mean Absolute Error) metric of GARCH-CEEMDAN-GCN is 70.0% better than GCN, 18.4% better than EMD-GCN, 13.6% better than EEMD-GCN, and 9.6% better than EMD-CEEMDAN-GCN. The MSE (Mean Squared Error) metric of GARCH-CEEMDAN-GCN is 90.5% better than GCN, 26.2% better than EMD-GCN, 17.1% better than EEMD-GCN, and

14.9% better than EMD-CEEMDAN-GCN. The MAPE (Mean Absolute Percentage Error) metric of GARCH-CEEMDAN-GCN is 74.7% better than GCN, 19.0% better than EMD-GCN, 15.4% better than EEMD-GCN, and 7.5% better than EMD-CEEMDAN-GCN. The R2 (R Squared) metric of GARCH-CEEMDAN-GCN is 111.1% better than GCN, 2.2% better than EMD-GCN, 1.1% better than EEMD-GCN, and 1.1% better than EMD-CEEMDAN-GCN.

Table 5 Statistical description of vehicles 2015–2017

		Statistical analysis of vehicles	Mean change from previous year (%)
Year	Analysis	Vehicles	Vehicles
2017	max	156	65.28%
	min	14	
	mean	64.94	
	std	22.85	
	median	61	
2016	max	120	90.35%
	min	5	
	mean	39.29	
	std	16.82	
	median	36	
2015	max	50	
	min	6	
	mean	20.64	
	std	7.78	
	median	19	

The experiments on these three datasets show that the proposed model performs poorly on the Air Quality dataset in terms of MAE, MSE, MAPE, and R2 metrics. However, on the Energy and Traffic datasets, the proposed model performs optimally in all indicators. Taken together, the performance of the proposed GARCH-CEEMDAN-GCN hybrid forecasting model is superior to GCN, EMD-GCN, EEMD-GCN, EMD-CEEMDAN-GCN, and GARCH-CEEMDAN-GCN models. This also verifies that using the GARCH model to further preprocess the data can enhance the analytical power of subsequent parts of the model.

Discussion

In our study on time series forecasting using a hybrid model of deep learning and GARCH for non-stationary series, we found that the integration of deep learning significantly enhances pattern recognition capabilities, particularly in complex and evolving datasets where traditional models fall short. The inclusion of GARCH modeling proved invaluable for handling the inherent variability and volatility, a feature especially pertinent in financial markets. While the model outperforms traditional methods in certain aspects, challenges such

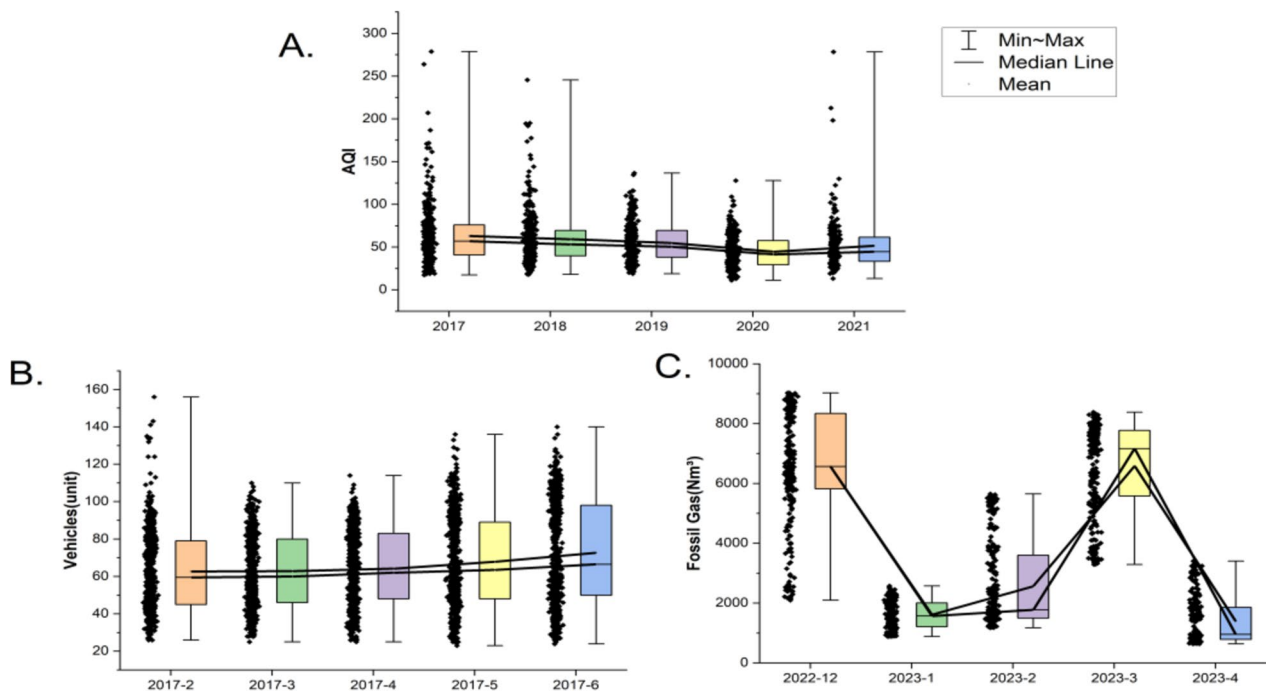


Fig. 7 Shows the statistical description of three datasets: **A** AQI, **B** Fossil Gas, and **C** Vehicles

as model complexity, computational demands, and the risk of overfitting were notable. These limitations underscore the need for careful data preprocessing and model tuning. The model’s practical applications are broad, offering potential benefits in sectors like finance, energy, and retail for improved forecasting and decision-making. However, its practical deployment necessitates balancing its sophisticated capabilities with its computational and expertise requirements. Future research should focus on enhancing interpretability, reducing computational overhead, and exploring adaptability to various types of non-stationary data, considering the ethical and societal implications in sensitive applications like economic forecasting. This study opens avenues for more nuanced and robust forecasting methods in the face of increasingly complex data landscapes.

The practical implications of a time series-forecasting model that combines deep learning with GARCH modeling for non-stationary series are significant

and diverse. This approach offers a powerful tool for extracting intricate patterns from complex and evolving data, which is highly relevant in areas like finance, where market volatility and economic indicators show non-stationary behaviors. By leveraging deep learning, the model can identify and learn from underlying patterns and relationships in the data that traditional models might miss. This leads to potentially more accurate and nuanced forecasts, essential for risk management, investment strategies, and economic planning. The use of GARCH modeling adds the ability to understand and predict the variability and volatility in the time series, a crucial aspect in financial forecasting and decision-making processes. Moreover, this hybrid modeling technique can be beneficial in other sectors like weather forecasting, energy consumption analysis, and demand forecasting in retail, where understanding and predicting complex, non-linear patterns is crucial. However, its practical application requires careful consideration of its complexity, computational demands,

Table 6 Parameters setting for the algorithms used in this study

Parameters	GCN	EMD-GCN	EEMD-GCN	EMD-CEEMDAN-GCN	GARCH-CEEMDAN-GCN
Batch size	64	64	64	64	64
Learning rate	0.01	0.01	0.01	0.01	0.01
Epochs	100	100	100	100	100

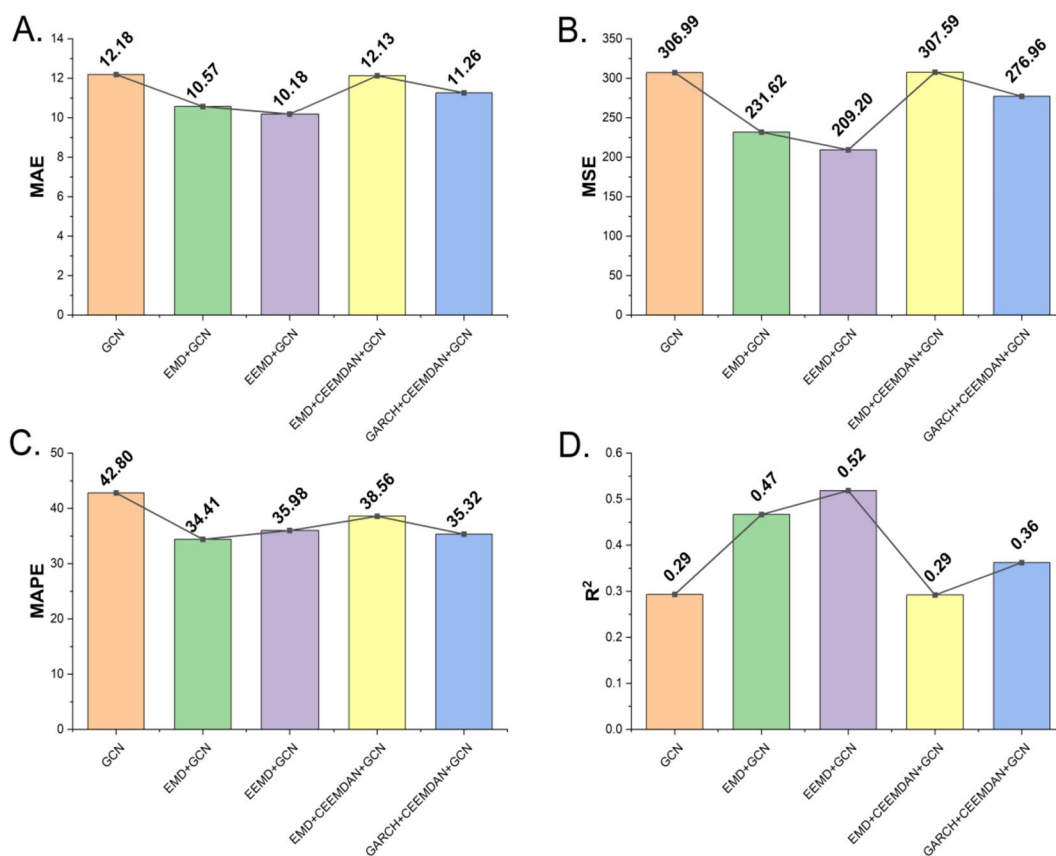


Fig. 8 Evaluation results of various models on the Air Quality dataset

and the need for expert knowledge in both deep learning and time series analysis.

The proposed method of time series forecasting for non-stationary series using deep learning and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) modeling, while innovative and potentially powerful, does have several limitations:

- **Complexity and Computation Requirements:** The integration of deep learning with GARCH modeling results in a complex model structure. This complexity requires significant computational resources for training and inference, especially with large datasets, making it less accessible for smaller organizations or individuals with limited computational capacity.
- **Overfitting Risks:** Deep learning models, due to their high parameter count and flexibility, are prone to overfitting, especially when the available data is not sufficiently large or diverse. This can lead to models that perform well on training data but poorly on unseen data.

- **Data Sensitivity and Preprocessing:** The performance of this approach heavily depends on the quality and preprocessing of the input data. Non-stationary time series data often require careful preprocessing, such as detrending and deseasonalization, to ensure meaningful patterns are learned.
- **Model Interpretability:** Deep learning models, particularly those with complex architectures, suffer from a lack of interpretability. Understanding the internal workings and decision processes of these models can be challenging, which may be a critical drawback in fields where explainability is crucial.
- **Parameter Tuning and Model Selection:** The success of the model heavily depends on the choice of architecture for the deep learning component and the parameters of the GARCH model. Finding the optimal configuration requires extensive experimentation and expertise, which can be time-consuming and resource-intensive.
- **Sensitivity to Market Conditions (for financial time series):** In financial applications, the model's performance can be highly sensitive to market conditions.

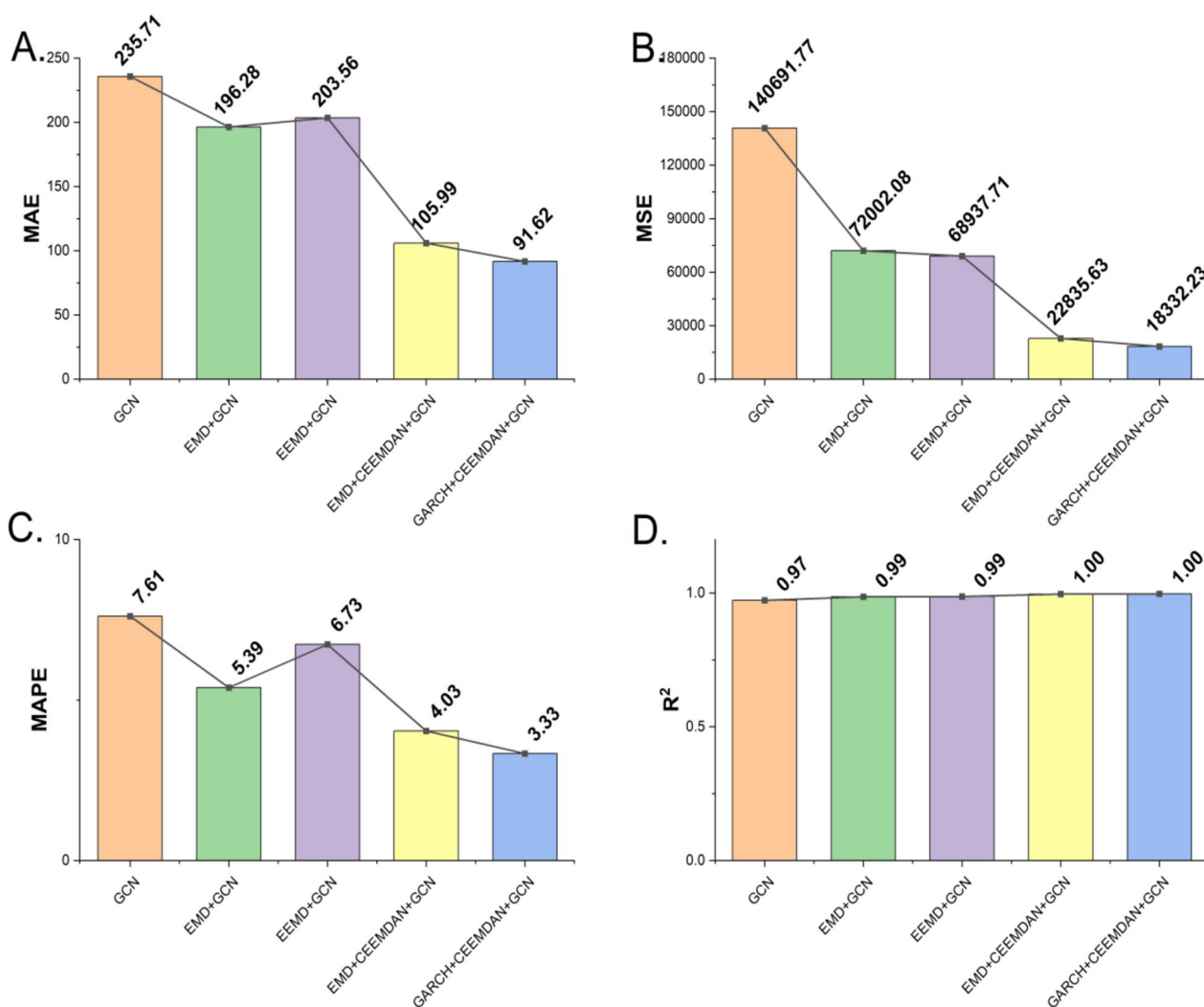


Fig. 9 Evaluation results of various models on the Energy dataset

Non-stationary series in finance are often influenced by unpredictable external factors, which can lead to significant forecast errors.

- **Generalization Across Different Domains:** While the model might be effective for specific types of non-stationary time series, its performance may not generalize well across different domains or types of time series data.
- **Dependency on Historical Data:** GARCH models and deep learning both rely heavily on historical data. In situations where historical data may not fully capture future trends or patterns (such as unprecedented economic events), the model’s predictions may be less accurate.

- **Training Time:** Due to the complex nature of combining deep learning with GARCH models, the training time can be significantly longer than more traditional time series forecasting methods.
- **Requirement for Expert Knowledge:** Implementing and tuning this hybrid model requires a deep understanding of both deep learning and econometric models like GARCH. This expertise may not be readily available in all organizations.
- These limitations highlight the need for careful consideration and possibly the integration of additional techniques or methodologies to address specific challenges when applying this model to real-world time series forecasting problems.

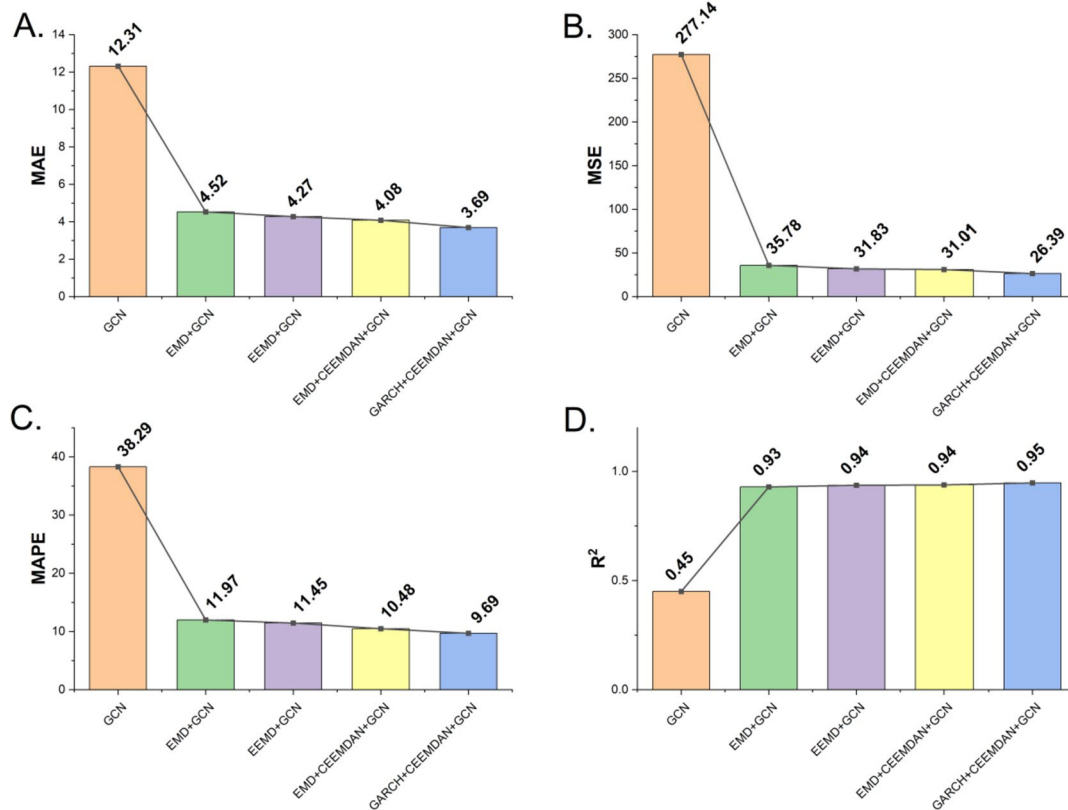


Fig. 10 Evaluation results of various models on the Traffic dataset

Conclusion

This study explores a hybrid time series prediction model that combines the GARCH-CEEMDAN-GCN model. The model uses a novel approach to preprocess input data for better analysis, then utilizes CEEMDAN for further data decomposition, reducing the complexity of the sequence data and enhancing the ability of the predictive model to learn data correlations and characteristics. In selecting the predictive model, we chose the GCN for its fewer parameters and excellent data learning abilities. Through observing the experimental results, we can find out the following:

- In predicting time series data, the hybrid prediction model with the GARCH data handling and signal decomposition steps produces better experimental results than single prediction models.
- When input data is preprocessed through the GARCH model, the convenience of data analysis is further enhanced.
- The application of the signal decomposition method allows the predictive model to handle input data with reduced complexity without having to

deal directly with complex raw data, making it easier for the predictive model to learn data relationships and extract features.

- Comparative experimental results show that our proposed GARCH-CEEMDAN-GCN based hybrid model offers the best prediction accuracy, is stable, and has strong data fitting capabilities. Furthermore, this model that we proposed can be used to predict time series data in other practical applications, such as forecasting stock prices and temperature changes, demonstrating its significant practical value.

Authors’ contributions

Huimin Han: Conceptualization, Methodology, Data curation, Writing - original draft preparation. Zehua Liu: Formal analysis, Investigation, Writing - review and editing. Mauricio Barrios Barrios: Software, Validation, Visualization. Jiuhao Li: Resources, Project administration, Funding acquisition. Zhixiong Zeng: Methodology, Data curation, Writing - original draft preparation. Emad Mahrous Awwad: Conceptualization, Writing - review and editing, Supervision. Nadia Sarhan: Formal analysis, Investigation, Visualization.

Funding

The authors present their appreciation to King Saud University for funding this research through Researchers Supporting Program number (RSPD2024R1052), King Saud University, Riyadh, Saudi Arabia.

No datasets were generated or analysed during the current study.

Declarations

Competing interests

The authors declare no competing interests.

Received: 19 November 2023 Accepted: 13 December 2023

Published online: 02 January 2024

References

- Li H, Jin K, Sun S, Jia X, Li Y (2022) Metro passenger flow forecasting through multi-source time-series fusion: an ensemble deep learning approach. *Appl Soft Comput* 120:108644. <https://doi.org/10.1016/j.asoc.2022.108644>
- Ogliari E, Dolara A, Manzolini G et al (2017) Physical and hybrid methods comparison for the day ahead PV output power forecast[J]. *Renewable Energy* 113:11–21
- Shafie-Khah M, Moghaddam MP, Sheikh-El-Eslami MK (2011) Price forecasting of day-ahead electricity markets using a hybrid forecast method[J]. *Energy Convers Manage* 52(5):2165–2169
- Qian Z, Pei Y, Zareipour H et al (2019) A review and discussion of decomposition-based hybrid models for wind energy forecasting applications[J]. *Appl Energy* 235:939–953
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning[J]. *Nature* 521(7553):436–444
- Kamilaris A, Prenafeta-Boldú FX (2018) Deep learning in agriculture: A survey[J]. *Comput Electron Agric* 147:70–90
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9
- Bhatti U, Tang H, Wu S (2023) Mangrove decline puts Pakistan's coasts at risk. *Science* 382:654–655. <https://doi.org/10.1126/science.adl3073>
- Dahl GE, Yu D, Deng L, Acero A (2022) Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition, *IEEE Trans. Audio, Speech, Lang. Process* 20(1):30–42
- Box GEP, Jenkins GM (1970) Time series analysis: forecasting and control. *J Am Stat Assoc* 68(342):199–201
- Burlando P, Rosso R, Cadavid LG et al (1993) Forecasting of short-term rainfall using ARMA models[J]. *J Hydrol* 144(1–4):193–211
- Meyler A, Kenny G, Quinn T (1998) Forecasting Irish inflation using ARIMA models[J]
- Williams BM, Hoel LA (2003) Modeling and forecasting vehicular traffic flow as a seasonal arima process: theoretical basis and empirical results[J]. *J Transport Eng* 129(6):664–672
- Chen TT, Lee SJ (2015) A weighted LS-SVM based learning system for time series forecasting[J]. *Inf Sci* 299:99–116
- Tay FEH, Cao L (2001) Application of support vector machines in financial time series forecasting[J]. *Omega* 29(4):309–317
- Bhatti U, Mengxing H, Neira-Molin H, Marjan S, Baryalai M, Hao T, Wu G, Bazai S (2023) MFFCG – Multi feature fusion for hyperspectral image classification using graph attention network. *Expert Syst Appl* 229:120496. <https://doi.org/10.1016/j.eswa.2023.120496>
- Lv X, Cheng X, Tang Y (2018) Short-term power load forecasting based on balanced KNN[C]//IOP Conference series: materials science and engineering. IOP Publishing 322(7):072058
- Hung NQ, Babel MS, Weesakul S et al (2009) An artificial neural network model for rainfall forecasting in Bangkok, Thailand[J]. *Hydrol Earth Syst Sci* 13(8):1413–1425
- Maleki H, Sorooshian A, Goudarzi G, Baboli Z, Birgani YT, Rahmati M (2019) Air pollution prediction by using an artificial neural network model. *Clean Technol Envir* 21(6):1341–1352. <https://doi.org/10.1007/s10098-019-01709-w>
- Goudarzi G, Hopke PK, Yazdani M (2021) Forecasting PM2.5 concentration using artificial neural network and its health effects in Ahvaz, Iran. *Chemosphere* 283:131285. <https://doi.org/10.1016/j.chemosphere.2021.131285>
- Tokgöz A, Ünal G (2018) A RNN based time series approach for forecasting turkish electricity load[C]//2018 26th Signal processing and communications applications conference (SIU). IEEE, 1–4
- Chang YS, Chiao HT, Abimannan S et al (2022) An LSTM-based aggregated model for air pollution forecasting[J]. *Atmos Pollut Res* 11(8):1451–1463
- Bhatti U, Hashmi MZ, Sun Y, Masud M, Nizamani MM (2023) Editorial: Artificial intelligence applications in reduction of carbon emissions: Step towards sustainable environment. *Front Environ Sci* 11:1183620. <https://doi.org/10.3389/fenvs.2023.1183620>
- Zha W, Liu Y, Wan Y, et al (2022) Forecasting monthly gas field production based on the CNN-LSTM model[J]. *Energy* 124889
- Yu B, Lee Y, Sohn K (2020) Forecasting road traffic speeds by considering area-wide spatio-temporal dependencies based on a graph convolutional neural network (GCN)[J]. *Transport Res Part C Emerg Technol* 114:189–204
- Bhatti, Uzair & Tang, Hao & Wu, Guilu & Marjan, Shah & Hussain, Aamir. (2023). Deep learning with graph convolutional networks: an overview and latest applications in computational intelligence. *Int J Intell Syst*. 2023. <https://doi.org/10.1155/2023/8342104>
- Tang H, Bhatti U, Li J, Marjan S, Baryalai M, As M, Ghadi Y, Mohamed H (2023). A new hybrid forecasting model based on dual series decomposition with long-term short-term memory. *Int J Intell Syst*. 2023. <https://doi.org/10.1155/2023/9407104>
- Gendeel M, Zhang YX, Han AQ (2018) Performance comparison of ANNs model with VMD for short-term wind speed forecasting. *IET Renew Power Gener* 12(12):1424–1430. <https://doi.org/10.1049/iet-rpg.2018.5203>
- Zhang W, Liu F, Zheng X et al (2015) A hybrid EMD-SVM based short-term wind power forecasting model[C]//2015 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC). IEEE, 1–5
- Shu W, Gao Q (2020) Forecasting stock price based on frequency components by EMD and neural networks[J]. *IEEE Access* 8:206388–206395
- Yan Y, Wang X, Ren F et al (2022) Wind speed prediction using a hybrid model of EEMD and LSTM considering seasonal features[J]. *Energy Rep* 8:8965–8980
- Zhu Q, Zhang F, Liu S et al (2019) A hybrid VMD-BiGRU model for rubber futures time series forecasting[J]. *Appl Soft Comput* 84:105739
- Zhou F, Huang Z, Zhang C (2022) Carbon price forecasting based on CEEMDAN and LSTM[J]. *Appl Energy* 311:118601
- Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. *IEEE Trans Signal Process* 62:531–544. <https://doi.org/10.1109/TSP.2013.2288675>
- Liu C, Zhu L, Ni C (2018) Chatter detection in milling process based on VMD and energy entropy. *Mech Syst Signal Process* 105:169–182. <https://doi.org/10.1016/j.ymssp.2017.11.046>
- Gu G, Wang K, Wang Y et al (2016) Synchronous triple-optical-path digital speckle pattern interferometry with fast discrete curvelet transform for measuring three-dimensional displacements. *Opt Laser Technol* 80:104–111
- Maheshwari S, Pachori RB, Kanhangad V et al (2017) Iterative variational mode decomposition based automated detection of glaucoma using fundus images. *Comput Biol Med* 88:142–149
- Chen Z, Yuan C, Wu H et al (2022) An improved method based on EEMD-LSTM to predict missing measured data of structural sensors[J]. *Appl Sci* 12(18):9027
- Sun M, Li Z, Li Z, Li Q, Liu Y, Wang J (2020) A noise attenuation method for weak seismic signals based on compressed sensing and CEEMD. *IEEE Access* 8:71951–71964
- Zhang Q, Lou L (2021) Research on partial discharge signal denoising of transformer based on improved CEEMD and adaptive wavelet threshold[C]//2021 3rd International Academic Exchange Conference on Science and Technology Innovation (IAECST). IEEE, 1708–1712
- Mou Z, Niu X, Wang C (2020) A precise feature extraction method for shock wave signal with improved CEEMD-HHT[J]. *Journal of Ambient Intelligence and Humanized Computing* 1–12
- Torres ME, Colominas MA, Schlotthauer G, Flandrin P (2021) A complete ensemble empirical mode decomposition with adaptive noise. In *Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing*, Prague, Czech Republic, pp. 4144–4147

43. Colominas MA, Schlotthauer G, Torres ME (2014) Improved complete ensemble EMD: a suitable tool for biomedical signal processing. *Biomed Signal Process Control* 14:19–29 [CrossRef]
44. Zhang L, Li C, Chen L et al (2023) A Hybrid forecasting method for anticipating stock market trends via a soft-thresholding de-noise model and support vector machine (SVM)[J]. *World Basic Appl Sci J* 2023(13):597–602
45. Mohandes M (2002) Support vector machines for short-term electrical load forecasting. *Int J Energy Res* 26:335–345 [CrossRef]
46. Wang H, Xu P, Zhao J (2022) Improved KNN algorithms of spherical regions based on clustering and region division[J]. *Alex Eng J* 61(5):3571–3585
47. Kurani A, Doshi P, Vakharia A et al (2023) A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting[J]. *Ann Data Sci* 10(1):183–208
48. Jadav K, Panchal M (2012) Optimizing weights of artificial neural networks using genetic algorithms. *Int J Adv Res Comput Sci Electron Eng* 1:47–51
49. Zhang J, Qu S, Zhang Z et al (2022) Improved genetic algorithm optimized LSTM model and its application in short-term traffic flow prediction[J]. *PeerJ Comput Sci* 8:e1048
50. Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* 35(8):1798–1828
51. Zha W, Liu Y, Wan Y, et al (2022) Forecasting monthly gas field production based on the CNN-LSTM model[J]. *Energy* 124889
52. Bhatti U, Bazai S, Hussain S, Fakhar S, Ku C, Marjan S, Por Y, Jing L. (2023). Deep Learning-Based Trees Disease Recognition and Classification Using Hyperspectral Data. *Computers, Materials & Continua*. 77:681–697. <https://doi.org/10.32604/cmc.2023.037958>
53. Wang J, Peng B, Zhang X (2018) Using a stacked residual LSTM model for sentiment intensity prediction. *Neurocomputing* 322:93–101
54. Zhang Q, Jin Q, Chang J, et al (2018) Kernel-weighted graph convolutional network: A deep learning approach for traffic forecasting[C]//2018 24th International Conference on Pattern Recognition (ICPR). IEEE, 1018–1023
55. Miao X, Gürel N M, Zhang W, et al (2021) Degnn: Improving graph neural networks with graph decomposition[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining 1223–1233

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
