Han et al. Journal of Cloud Computing (2024) 13:20
https://doi.org/10.1186/s13677-023-00560-1

Journal of Cloud Computing: Advances, Systems and Applications

RESEARCH

Open Access

Advanced series decomposition with a gated recurrent unit and graph convolutional neural network for non-stationary data patterns

Huimin Han[1], Harold Neira-Molin[2], Asad Khan[3*], Meie Fang[3*], Haitham A. Mahmoud[4], Emad Mahrous Awwad[5], Bilal Ahmed[7] and Yazeed Yasin Ghadi[6]

Abstract

In this study, we present the EEG-GCN, a novel hybrid model for the prediction of time series data, adept at addressing the inherent challenges posed by the data's complex, non-linear, and periodic nature, as well as the noise that frequently accompanies it. This model synergizes signal decomposition techniques with a graph convolutional neural network (GCN) for enhanced analytical precision. The EEG-GCN approaches time series data as a one-dimensional temporal signal, applying a dual-layered signal decomposition using both Ensemble Empirical Mode Decomposition (EEMD) and GRU. This two-pronged decomposition process effectively eliminates noise interference and distills the complex signal into more tractable sub-signals. These sub-signals facilitate a more straightforward feature analysis and learning process. To capitalize on the decomposed data, a graph convolutional neural network (GCN) is employed to discern the intricate feature interplay within the sub-signals and to map the interdependencies among the data points. The predictive model then synthesizes the weighted outputs of the GCN to yield the final forecast. A key component of our approach is the integration of a Gated Recurrent Unit (GRU) with EEMD within the GCN framework, referred to as EEMD-GRU-GCN. This combination leverages the strengths of GRU in capturing temporal dependencies and the EEMD's capability in handling non-stationary data, thereby enriching the feature set available for the GCN and enhancing the overall predictive accuracy and stability of the model. Empirical evaluations demonstrate that the EEG-GCN model achieves superior performance metrics. Compared to the baseline GCN model, EEG-GCN shows an average R2 improvement of 60% to 90%, outperforming the other methods. These results substantiate the advanced predictive capability of our proposed model, underscoring its potential for robust and accurate time series forecasting.

Keywords Time series forecasting, EEMD, CEEMDAN, GCN

*Correspondence:
Asad Khan
asad@gzhu.edu.cn
Meie Fang
fme@gzhu.edu.cn
Full list of author information is available at the end of the article

Springer Open

© 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

In the era of rapid industrialization and informatization, people are increasingly relying on various sensors to obtain data [1]. Due to the large and constantly increasing number of sensors deployed by humans, the explosive growth of time series data in various fields has followed suit. Today, time series data has become one of the most common types of data, such as changes in air quality data in a region, traffic flow changes at a certain intersection in a road network, fluctuations in stock prices in the stock trading market, greenhouse gas emissions and agricultural effects [2–4], all of which are recorded and represented in a time series. Researchers can analyze these recorded data to further explore the hidden patterns behind these changes. The more accurate researchers' analysis of these patterns, the higher the accuracy of time series prediction. Accurately predicting time series data is beneficial for us to plan ahead and better allocate resources. The environmental protection department of the government can use historical air quality data in a region to predict the changes in air quality data in that region in the future, thereby making better arrangements for pollution prevention and control in that area [5]. In industrial production, managers can make better plans for the use of resources such as electricity, natural gas, and coal by predicting and analyzing time series data [6]. The traffic management department can use information on historical traffic flow to predict road congestion and remind people to arrange better travel routes in advance [7]. Time series data prediction has now become a very popular research direction, which can help people make corresponding plans in advance, reduce costs, improve efficiency, and is of great significance to improve social productivity [8–10].

In the field of traffic volume forecasting, models can broadly be categorized into parametric and non-parametric based on their structural foundation. Moreover, within the domain of deep learning methodologies, models are subclassified into generative, discriminative, and hybrid deep structures, each demonstrating its unique capabilities and advancements over time [11]. The evolution of research has seen a shift from traditional parametric statistical models towards non-parametric and subsequently to hybrid models, indicating a progression towards more complex and nuanced modeling techniques.

Early applications of parametric models often employed growth curves for forecasting metrics like rail transit passenger volumes (Yuan et al. [12]). Common among these parametric approaches are various time-series models and their derivatives, which are praised for their simplicity and interpretability [13–15]. Nonetheless, these models traditionally falter when addressing the non-linear nature of traffic flows, often leading to substantial prediction errors.

To mitigate the shortcomings of parametric models, non-parametric models like the support vector regression (SVR) algorithm have been introduced with notable success. Toan. T.D reported that SVR shows superior performance in forecasting traffic flow, particularly with small-sample, high-dimensional data sets characterized by non-linearity, offering a robust generalization capability that circumvents overfitting and thereby provides more accurate short-term urban traffic flow predictions [16].

Exploring the utility of recurrent neural networks, Yutian Liu. investigated three RNN variants applied to traffic data, concluding that RNNs offer commendable prediction capabilities, albeit with LSTM models showing slightly higher error rates [17]. Luo Xianglong. enhanced the training efficiency of support vector machines (SVM) by integrating the Discrete Fourier Transform (DFT) method, which helped to reduce the training scale and expedited the training process without compromising prediction accuracy [18].

In another innovative approach, Changxi Ma. leveraged a hybrid model combining Spatiotemporal Feature Selection Algorithm (STFSA) with a convolutional neural network (CNN) to create a two-dimensional matrix for short-term traffic flow prediction, yielding better accuracy than single models like SVR, SARIMA, KNN, ANN, or even combined models like STFSA-ANN [19]. Wang S. extended this hybrid model concept by integrating STFSA with a gated recurrent unit (GRU), which exhibited substantial improvements over standalone CNN and GRU models in both precision and reliability for short-term traffic forecasting [20].

Noreen Zaffer put forward a CNN-LSTM multi-step prediction model that incorporated feature data with an attention mechanism, showcasing an impressive accuracy rate of nearly 99%, with effective application across varying conditions such as peak and non-peak hours, and differentiating between working days and holidays [21]. Zhang W. proposed three hybrid deep learning models (CL-CN-G, CL-CNG, and G-CN-CL) integrating CNN, GRU, and ConvLSTM to specifically address the forecasting of traffic flow under distinctive conditions such as holidays and adverse weather scenarios. Case studies demonstrated the high accuracy and efficacy of these models, with the G-CN-CL model being particularly outstanding [22, 23].

This trajectory of research underscores a dynamic shift towards leveraging the strengths of various modeling techniques to enhance predictive performance in traffic volume forecasting. The integration of deep learning architectures and hybrid models exemplifies the innovative strides in the field, aiming to tackle the inherent non-linear and complex patterns observed in traffic data.

The research contribution of integrating Ensemble Empirical Mode Decomposition (EEMD), Gated Recurrent Unit (GRU), and Graph Convolutional Network (GCN) for prediction purposes lies in addressing the complexities of time-series data that are both spatially and temporally correlated. Each component of the EEMD-GRU-GCN method brings a unique strength to the prediction model, making the collective methodology robust and sophisticated for various forecasting tasks. Here's how each component contributes:

Ensemble Empirical Mode Decomposition (EEMD)

Data decomposition

EEMD effectively decomposes non-linear and non-stationary time series data into a finite number of intrinsic mode functions (IMFs), which simplifies the complexity of the original data.

Noise reduction

It helps in reducing noise and enhancing the signal-to-noise ratio, which is crucial for accurate forecasting.

Feature extraction

EEMD is an advanced feature extraction technique that identifies the underlying structures within the data, which can be critical for understanding complex patterns.

Gated Recurrent Unit (GRU)

Temporal relationships

GRU is a type of recurrent neural network that is particularly good at capturing temporal dependencies, even over long sequences, which is vital for time-series prediction.

Modeling dynamics

It allows the model to include the dynamics of the system being studied, learning when to forget previous inputs and when to update its beliefs with new data.

Efficiency

GRUs are computationally more efficient than other types of RNNs, like LSTMs, without compromising the performance, making them suitable for real-time prediction tasks.

Graph Convolutional Network (GCN)

Spatial correlation

GCN extends the utility of convolutional neural networks to graph-structured data, enabling the model to capture spatial correlations in data that cannot be represented in a Euclidean space.

Complex relationships

It is particularly useful for datasets where the relationships between entities are as important as the entities themselves, such as in traffic networks or social networks.

Scalability

GCNs are scalable to large datasets, making them applicable to complex systems with numerous interacting components.

Research Contribution of the EEMD-GRU-GCN Method

The combination of EEMD, GRU, and GCN in a single predictive framework leads to a powerful approach for tackling prediction problems:

Holistic analysis

The EEMD-GRU-GCN method can provide a holistic analysis of time-series data by taking into account both the temporal sequence and spatial connections between different parts of the data.

Enhanced accuracy

The multi-faceted nature of the approach leads to improved prediction accuracy, as it can deal with various types of irregularities in the data.

Versatility

This method can be adapted to a wide range of applications, from financial markets and energy load forecasting to environmental monitoring and traffic flow prediction.

Improved generalization

By combining EEMD's feature extraction, GRU's temporal dynamics learning, and GCN's spatial relationship understanding, the model is less likely to overfit and more likely to generalize well to unseen data.

Advanced insights

The method can also provide insights into the nature of the data being studied, revealing complex interdependencies that simpler models might miss.

Research backgroud

Research on time series data prediction can generally be divided into three directions, namely, research on prediction methods based on statistics, research on prediction methods based on machine learning, and research on prediction methods based on hybrid models. Among them, machine learning includes traditional machine learning and deep learning, and hybrid prediction models mainly consist of two parts: signal decomposition of time series data and time series prediction.

Traditional method

In the realm of energy system forecasting, considerable advancements have been made to identify efficacious methodologies suitable for real-world application. Such forecasting models are crucial in mitigating system failure risks and enhancing the reliability of energy systems through the projection of future scenarios [24].

Historically, an analog methodology was initially employed to project wind speed distributions, representing a nascent step in predictive modeling [25]. This was superseded by the advent of time series models, which aimed to forecast wind power several hours ahead, thereby facilitating more agile energy management strategies [26]. For short-term wind speed forecasting, the Kalman filter emerged as a dynamic tool that assimilated new data to refine predictions continually [27].

Traditional statistical methods have long been used to emulate the characteristics of time series data, such as ARIMA (Auto-Regressive Integrated Moving Average) and AR-ARCH (Auto-Regressive Conditional Heteroskedasticity), both of which have found applications in financial markets for predicting return rates [28]. The fractional-ARIMA model, which offers predictive capabilities for several days in advance, demonstrated superior accuracy compared to the persistence model in a case study involving a 750 kW wind turbine [29]. Moreover, the ARIMA model has been effectively adapted to forecast global solar irradiance, with modifications such as the combination of ARIMA and repeated wavelet transform yielding significant improvements in forecasting performance [30].

In an innovative step, Wang et al. incorporated an extreme learning model with ARIMA, validating its accuracy through various case studies for wind projection [31]. The synergy between Artificial Neural Networks (ANN) and ARIMA in a hybrid model developed by K R Nair underscored the potential for greater accuracy than when these models operate independently [32]. The integration of machine learning techniques with ARIMA has been suggested to further enhance the precision and consistency of wind speed forecasts (Liu et al. [33]). Additionally, Asim et al., introduced an ARIMA-based model designed to improve accuracy and manage the uncertainties inherent in wind speed prediction and carbon emission control [34, 35].

It is critical to acknowledge that ARIMA-based models exhibit optimal performance with stationary time series data. However, energy-related time series such as solar radiation and wind speed typically manifest seasonality and trends. To address these non-stationary characteristics, the Seasonal ARIMA (SARIMA) model has been employed, with Xianqi Z. demonstrating its high accuracy in predicting thermal energy requirements for district heating systems [36]. The SARIMA-RVFL (Random Vector Functional Link) model, designed for short-term solar photovoltaic generation predictions, and Wang H. et al.'s application of the SARIMA model for monthly wind velocity forecasting have both shown improved accuracy over traditional ARIMA-based approaches [37].

ANNs have seen widespread use due to their capacity to resolve complex nonlinear equations, thus enabling predictions across diverse future scenarios. Time series statistical methods coupled with ANNs have been extensively applied in the prediction of solar and wind energy patterns (Shuai Hu et al. [38]). The implementation of ANN techniques in solar irradiance prediction has yielded more accurate results compared to empirical regression models [39]. Diverse ANN architectures such as feed-forward propagation (FFBP), adaptive linear element (ADALINE), and radial basis function neural networks (RBFNN) have demonstrated varying levels of forecasting acuity, contingent upon their respective structures and parameterizations [40]. Feed-forward neural networks (FFNN) have been broadly applied to wind power prediction with satisfactory accuracy [41].

A novel approach using genetic neural networks (GNN), which apply a genetic algorithm for weight and bias optimization instead of the traditional backpropagation method, has shown promising results in wind velocity prediction [42, 43]. Enhancing ANN training with particle swarm optimization (PSO) has also been reported to produce superior outcomes compared to conventional training methods [44]. For instance, a study employing ANN to predict solar irradiance a day ahead in a grid-connected solar photovoltaic plant reported a mean absolute error (MAE) of 3.21% and a mean bias error (MBE) of 8.54% [45].

Support vector machines (SVMs), which are adept at modeling non-linear data patterns similar to ANN techniques, have exhibited improved prediction performance in multi-layer perception neural networks (Uncuoglu, et al. [46]). Additionally, wavelet networks—a hybrid of wavelet theory and neural network methodology—have been applied in solar irradiance prediction, with one particular study demonstrating their competitive performance against other neural network techniques [47]. Both ANNs and SVMs have demonstrated proficiency in capturing and modeling the complex non-linear trends in energy forecasting.

Series decomposition methods for prediction

In the realm of short-term load forecasting (STLF), several methodologies have been employed over the years to enhance prediction accuracy, such as traditional algorithms, Similar Day (SD) selection, Empirical Mode Decomposition (EMD) techniques, artificial intelligence (AI), and an amalgamation of different forecasting models [48, 49]. 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.). The ever-evolving energy grids have necessitated the incorporation of diverse variables in forecasting models, such as climatic conditions, seasonal holidays, and dynamic pricing structures [50], revealing the inadequacies of conventional forecasting approaches that often struggle with non-linear dynamics [51].

The SD selection method relies on the analysis of historical data, pinpointing past days with load patterns that resemble the target day's expected conditions. Attributes like the day of the week and meteorological conditions serve as a basis for prediction (Maxwell et al. [52]). This method has been refined through the integration of the XGB algorithm to determine attribute significance and calculate distances for optimal SD selection [53]. Despite its utility, the standalone SD method may not fully encapsulate the intricate nature of electrical load patterns, prompting researchers to suggest its combination with other predictive techniques for improved robustness [54].

AI and machine learning (ML) technologies are increasingly adopted by electric utility providers to tackle complex load forecasting. Despite significant research efforts, achieving high accuracy in STLF remains a complex endeavor due to the non-stationarity of electrical load data and the prediction of long-term dependencies [55]. Models such as Long Short-Term Memory (LSTM) networks and their bidirectional variants (BiLSTM) are used to forecast demand-side load across different time horizons (Ullah I, et al. [56]). Gated Recurrent Unit (GRU) models have found applications in forecasting short-term load for electric vehicle (EV) charging stations and battery state-of-charge predictions [57, 58]. Comparative assessments of LSTM, BiLSTM, and GRU models indicated the superior performance of BiLSTM in predicting the load for EV fleets, despite the challenges posed by the complexity of aggregate load data [59].

The EMD method has become a staple in diverse forecasting applications, ranging from energy consumption to renewable energy outputs and commodity prices [60]. It excels at distilling original datasets into intrinsic mode functions (IMFs), facilitating the analysis of unstable and non-stationary time series data [61]. Among the variations of EMD, the Complete Ensemble EMD with Adaptive Noise (CEEMDAN) stands out for its efficient spectral separation capabilities at a reduced computational load [62]. Recent advancements have seen the CEEMDAN method utilized to enhance the input/output data structures for electrical demand forecasting, yielding models with substantially improved accuracy [63].

The convergence of these advanced methodologies signifies a progressive stride in the field of STLF, highlighting a collective move towards intricate, multi-faceted approaches that address the complex nature of power consumption patterns. Integrating various models and techniques to compensate for individual limitations has become a key strategy in developing more reliable and precise forecasting systems.

In the early stages of research on time series prediction, researchers first used methods based on statistics to complete the task. Nepal B. et al. used an autoregressive moving average model (ARMA) to predict power load [64]. However, since most time series data have strong non-stationarity, ARMA does not have good predictive performance for non-stationary time series. In order to better handle non-stationary time series data, scholars have improved the ARMA model by adding differential terms to obtain the ARIMA model, which can analyze the periodicity and oscillations of time series data. Saglam M. used the ARIMA model to predict Turkey's energy demand [65]. Although statistical time series prediction models have achieved good predictive performance, when faced with time series data with increasing volume and complexity, models based on statistics are overwhelmed.

With the emergence of machine learning, researchers have seen new solutions. Brouno et al. 66] used the support vector machine (SVM) method to predict stock trends, while Gupta et al. used SVM to construct a time series prediction model [67]. The experiments showed that SVM has stronger feature extraction capabilities for nonlinear data compared to prediction models based on statistics and better robustness to noise in data. Ashfaq et al. used the KNN method to predict short-term power load [68]. KNN is a non-parametric unsupervised learning algorithm which is simple, easy to use and has strong applicability, while [69] used ANN to predict AQI time series data in the air. ANN is a combination of multiple neurons capable of non-linear output. Compared with classical machine learning methods such as SVM and KNN, it has stronger data fitting ability. Deep learning is an important branch of machine learning. With the increase of data volume and computing power, deep learning has become increasingly prominent. Deep learning can learn more complex data features [70]. Recurrent neural networks (RNN) can retain previously processed information and pass it to the next time step, making them very suitable for solving time series prediction tasks However, when the input sequence data is long, RNN may encounter the problems of vanishing or exploding gradients. To improve this problem, researchers have improved RNN and obtained the long short-term memory network (LSTM). (Zha et al. [71]) used the convolutional neural network (CNN) combined with LSTM to predict natural gas production, using CNN to extract data features and further improve the predictive accuracy of the LSTM network. Graph convolutional neural network (GCN) has strong learning ability for data relations. Zhang et al. [72] predicted traffic flow using a GCN-based model.

In recent years, more and more scholars have started to use hybrid models to complete time series prediction tasks. Hybrid prediction models generally consist of two parts: signal decomposition and signal prediction. Commonly used signal decomposition methods include EMD, EEMD, VMD, etc. Compared with single-structured prediction models, hybrid models often achieve better performance [73, 74] used EMD to decompose the original sequence data and then used SVM to predict to achieve short-term power load forecasting. Shu et al. [75] used EMD to decompose the original sequence data, then extracted features using CNN, and finally used LSTM neural network to model the extracted features and obtain predictive results. Experiments have shown that this model performs significantly better than single models. However, EMD lacks rigorous mathematical proof and may produce mode mixing in some cases. EEMD is a method improved from EMD to solve the problem of mode mixing. Wu et al. [76] used EEMD combined with LSTM to predict oil prices. Yin S. et al. [77] predicted international financial data using a combination model of VMD, ARIMA, and TEF. However, VMD cannot effectively decompose the non-periodic parts of non-stationary signals.

Proposed methods

Before introducing how the EEMD-CEEMDAN-GCN hybrid model predicts time series data, we first briefly describe the basic principles of the relevant theories used to construct this model, namely, the Ensemble Empirical Mode Decomposition (EEMD), the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), and the principle and application of Graph Convolutional Neural Network (GCN).

The Ensemble Empirical Mode Decomposition (EEMD) is an advanced time series analysis method used to process complex data. It is particularly useful for non-linear and non-stationary time series. EEMD is an improvement over the original Empirical Mode Decomposition (EMD) process, which was developed to decompose a signal into a finite set of intrinsic mode functions (IMFs) that are simple oscillatory modes.

Here's an overview of the EEMD method with a focus on the mathematical formulae involved:

Empirical Mode Decomposition (EMD)

The EMD method decomposes a signal $x(t)$ into a sum of oscillatory components called intrinsic mode functions (IMFs) and a residue $r(t)$:

$$x(t) = \sum_{i=1}^{n} IMF_i(t) + r(t)$$

The IMFs are functions that satisfy two conditions:

1. The number of extrema and the number of zero-crossings must either equal or differ at most by one.
2. At any point, the mean value of envelope defined by the local maxima and the envelope defined by the local minima is zero.

Ensemble Empirical Mode Decomposition (EEMD)

EEMD improves upon EMD by adding white noise to the signal to assist in the sifting process and to prevent mode mixing. The steps are as follows:

1. Add white noise:

Add a white noise series $W_n(t)$ to the signal:

$$x_n(t) = x(t) + w_n(t)$$

where n represents the ensemble number.

2. Decompose:

Decompose each noisy signal $W_n(t)$ using EMD to get IMFs:

$$x_n(t) = \sum_{i=1}^{N_n} IMF_{i,n}(t) + r_n(t)$$

where N_n is the number of IMFs obtained for the n-th ensemble.

3. Ensemble mean:

Repeat the above steps for N ensembles and take the ensemble mean of the.

corresponding IMFs to get the final set of IMFs:

$$IMF_i(t) = \frac{1}{N} \sum_{n=1}^{N} IMF_{i,n}(t)$$

where i indicates the i-th IMF.

4. Final decomposition:

The final decomposition of the original signal using EEMD is given by:

$$x(t) = \sum_{i=1}^{N_{IMF}} IMF_i(t) + r(t)$$

where *NIMFs* is the total number of IMFs averaged across all ensembles, and r(t) is the residual signal after subtracting all IMFs.

The addition of white noise in multiple ensembles serves to cancel out the noise in the averaging process, allowing for a more stable and robust extraction of the IMFs. Each IMF can then be analyzed to understand the underlying processes or used in forecasting models for prediction.

GCN

Graph Convolutional Networks (GCNs) are a powerful neural network architecture for processing data that is structured as graphs. They are used to capture the dependence of graphs via message passing between the nodes of graphs. Here's a basic overview of the GCN methodology along with mathematical formulae:

GCN Overview
In a GCN, every node

Let $G = (V, E)$ be a graph with nodes $v \in V$ and edges $e \in E$. Let X be the node feature matrix where each row represents the feature vector of a node. Let A be the adjacency matrix of G, and D be the diagonal degree matrix where Dii is the sum of the weights of all edges attached to node i.

The graph convolution operation is defined as follows:

$$H^{(l+1)} = \sigma(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)})$$

where:

- $H(l)$ is the matrix of activations in the l-th layer; $H(0) = X$.
- $W(l)$ is the weight matrix for the l-th layer.
- $\widetilde{A} = A + IN$ is the adjacency matrix of the graph G with added self-connections IN(identity matrix).
- $\widetilde{D}$ is the diagonal degree matrix of $\widetilde{A}$.
- $\sigma(\cdot)$ is the activation function, such as ReLU $\sigma(x) = \max(0, x)$.

Normalization

The normalization term $\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}$ is crucial as it prevents the scale of the features from increasing with the number of nodes.

Multi-layer GCN.

A multi-layer GCN can be constructed by stacking multiple graph convolution layers:

$$H^{(l+1)} = \sigma(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)})$$
$$\vdots$$
$$Z = H^{(L)}$$

where L is the number of layers, and Z is the output of the final layer which can be used for tasks like node classification, graph classification, or link prediction.

Feature learning

The GCN model learns to map nodes to a space where the graph structure is maximally informative about the nodes' final representations, making it effective for tasks that require capturing the dependencies in graph-structured data. This generalized method allows GCNs to be applied to any graph, providing a means for the nodes to effectively "communicate" with each other and thereby learn a representation that is informed by their local graph neighborhood.

Proposed EEG-GCN model

The EEMD-GRU-GCN (Ensemble Empirical Mode Decomposition—Gated Recurrent Unit—Graph Convolutional Network) prediction algorithm is a complex, hybrid model that combines signal processing, recurrent neural networks, and graph-based neural networks to predict time series data. Below is a conceptual outline of how you might implement such an algorithm, divided into stages for clarity:

Stage 1: Signal Decomposition with EEMD.

Signal Preprocessing

Prepare your time series data, handling any missing values, anomalies, and normalizing if necessary.

Apply EEMD

Use Ensemble Empirical Mode Decomposition to decompose the time series into a set of intrinsic mode functions (IMFs).

This step helps in handling non-stationary and non-linear properties of the time series.

Stage 2: feature learning with GRU

Prepare data for RNN

Transform the IMFs into sequences suitable for RNN processing.

Define a window size that represents how many past time steps are used to predict the future value.

Design GRU network

Construct a GRU architecture, which is particularly effective in capturing temporal dependencies.

Configure the network with an appropriate number of units and layers for your problem.

Train GRU model

Train the GRU on the sequences from the decomposed time series.

You may train individual GRU models for each IMF or a single GRU model on all IMFs combined, depending on the complexity and characteristics of the data.

Stage 3: graph-based learning with GCN

Feature extraction

Extract relevant features from the GRU model's outputs. These features represent learned temporal patterns in the data.

Construct graph

Build a graph where nodes represent different entities or time steps in your data.

Define edges based on the relationships or interactions between these entities/time steps.

Design GCN model

Set up a Graph Convolutional Network that can operate on the graph structure, taking the features extracted by the GRU as input.

Train GCN model

Train the GCN to learn the interdependencies represented in the graph structure.

This stage allows the model to capture complex patterns that are not just temporal but also structured in a non-Euclidean space (the graph).

Stage 4: prediction and model evaluation

Combine models for prediction

Integrate the outputs from both the GRU and the GCN models.

This could involve a simple concatenation of features, a weighted average, or a more complex fusion technique.

Make predictions

Use the combined model to make predictions on new data.

Post-process these predictions if necessary to ensure they are in the correct format or scale.

Evaluate Performance

Assess the model's accuracy, stability, and generalization using appropriate metrics (e.g., R^2, MAE, RMSE).

Stage 5: optimization and refinement

Hyperparameter tuning

Optimize the model by tuning hyperparameters such as learning rates, window sizes, and the number of units in the GRU and GCN.

Model refinement

Refine the model by incorporating domain-specific knowledge into the graph structure or by enhancing the signal decomposition step.

Experiment with different architectures or additional layers like attention mechanisms to improve performance.

Stage 6: deployment and monitoring

Deployment

Deploy the model for real-world prediction tasks.

Ensure there's a pipeline for feeding new data into the model and for handling real-time predictions if necessary.

Continuous monitoring

Regularly monitor the model's performance to detect any drift or performance degradation.

Update and retrain the model with new data as it becomes available. The overall structure diagram is shown in Fig. 1.

Experimental setting and results

Dataset description

In this section, the three standard datasets used in the experimental part of this paper are introduced: Air Quality, Energy and Traffic.

Air quality dataset

The Air Quality dataset contains air quality data recorded by sensors in Guangzhou, Guangdong Province, China from January 1, 2017 to August 14, 2021, with a sampling frequency of once a day.

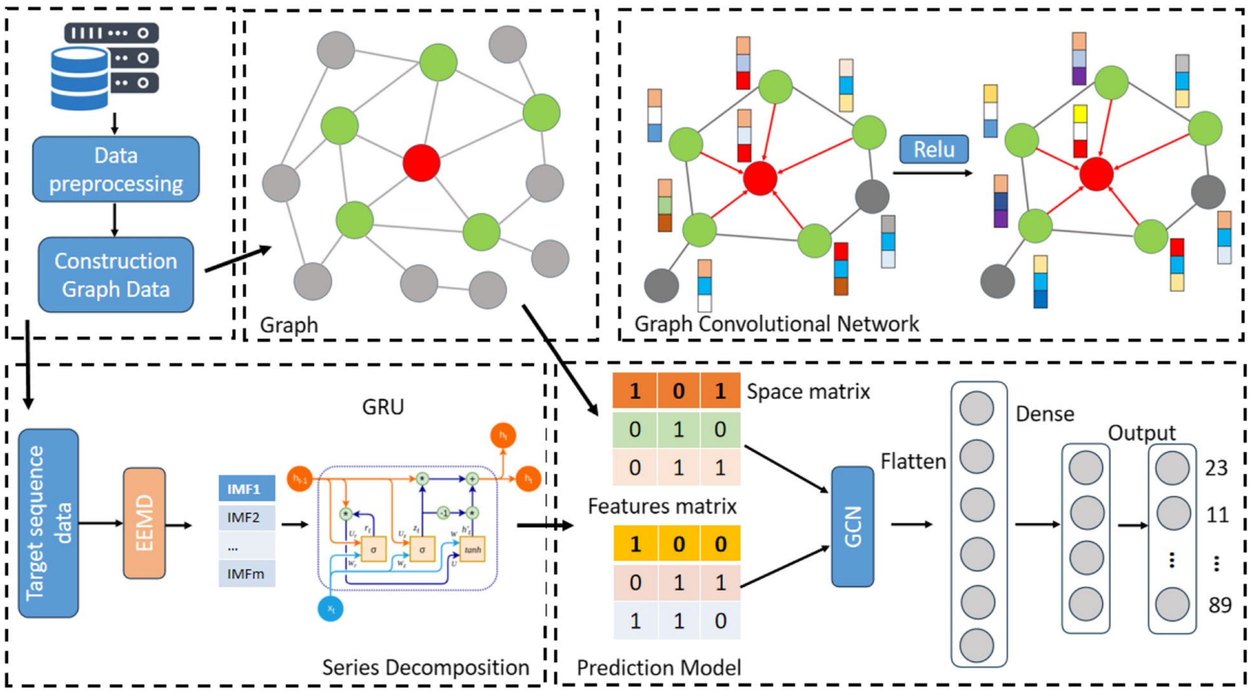

Fig. 1 Proposed model

Energy dataset

The Energy dataset contains energy data recorded by sensors in Netherlands between August 4, 2022 and April 23, 2023, with a sampling frequency of once every 15 min.

Traffic dataset

The Traffic dataset contains traffic flow data recorded by sensors on roads in London between November 1, 2015 and June 30, 2017, with a sampling frequency of once an hour.

Experimental settings

Python 3.8.5 and Pytorch1.7.0 are used to implement the proposed algorithm. The training hardware consists of an i7-10700K CPU and an NVIDIA GeForce RTX 3090 GPU. Table 2 shows the hyperparameter setting for all the models used in this study.

In order to compare the performance of various prediction algorithms, this paper selects four evaluation metrics, MAE, MSE, MAPE, and R2, to evaluate the prediction performance of the proposed model. They stand for Mean Absolute Error, Mean Square Error, Mean Absolute Percentage Error, and R Squared, respectively. Their formulas are as follows:

1) MAE refers to Mean Absolute Error in machine learning, which is a common metric used to evaluate the accuracy of prediction models. It reflects the degree of difference between the predicted values and actual values of the model, with the calculation formula being the absolute difference between the predicted and actual values divided by the total number of samples. A smaller MAE indicates better predictive ability of the model.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\widehat{y}_i - y_i\right| \tag{1}$$

2) MSE stands for Mean Squared Error. It is a common metric used in the evaluation of machine learning models and other prediction models. MSE measures the average of the squared differences between the predicted and actual values of a target variable in a dataset. A lower MSE score indicates that the model is better at making accurate predictions.

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

3) MAPE stands for Mean Absolute Percentage Error. It is a measure of accuracy used in forecasting and prediction models to evaluate the difference between actual and predicted values. It is calculated as the average of the absolute percentage differences

between the actual and predicted values, expressed as a percentage. MAPE values range from 0% (perfect accuracy) to 100% (complete inaccuracy).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\widehat{y}_i - y_i\right|}{y_i} \tag{3}$$

4) R2, also known as R squared, is a statistical measure that represents the proportion of the variance in the dependent variable that is explained by the independent variable(s) in a regression model. It is a value between 0 and 1, with higher values indicating a better fit of the model to the data. R2 is often used to evaluate the accuracy and usefulness of a regression model, and it can help to determine how well the model predicts the outcomes of interest.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\widehat{y}_i - y_i\right)^2}{\sum_{i=1}^{n}\left(\overline{y} - y_i\right)} \tag{4}$$

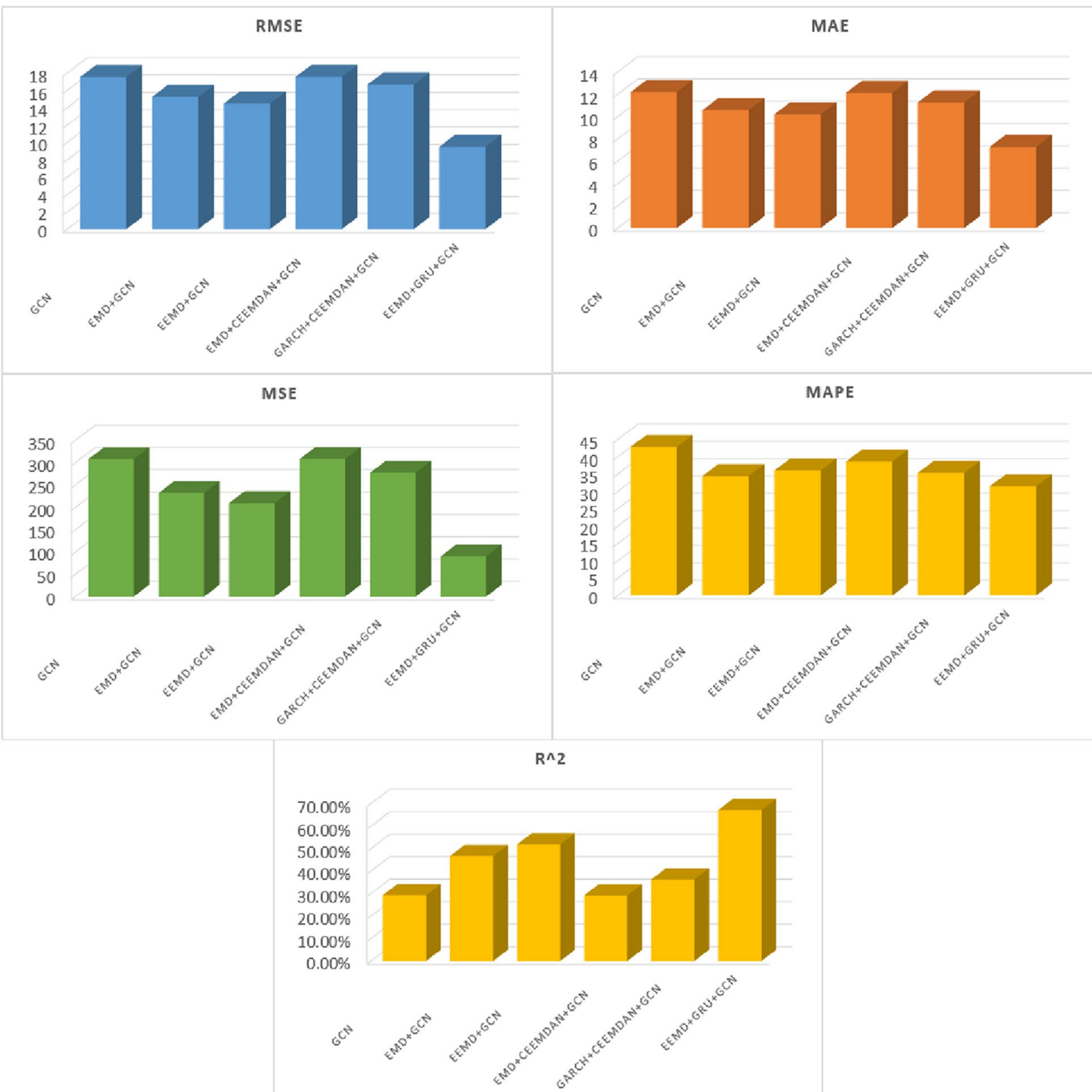

Fig. 2 Evaluation results of various models on the Air Quality dataset

In the formula for calculating the four evaluation metrics, MAE, MSE, MAPE, and R2, y_i represents the actual value of the input sample of the model, $\widehat{y}_i$ represents the predicted value output by the model, n represents the number of input samples, and i represents the sequence number of the sample.

Experimental results

We conducted experiments on three datasets, Air Quality, Energy, and Traffic, and compared the experimental results of five models, including GCN, EEMD-GCN, CEEMDAN-GCN, EMD-CEEMDAN-GCN, and the proposed EEG-GCN. The performance of these models on the Air Quality dataset is shown in Fig. 2, the performance on the Energy dataset is shown in Fig. 3, and the performance on the Traffic dataset is shown in Fig. 4.

Figure 2 shows the evaluation performances of different models on the Air Quality dataset. Predicting air quality is a complex task involving the analysis of data on pollutants like particulate matter and gases, along

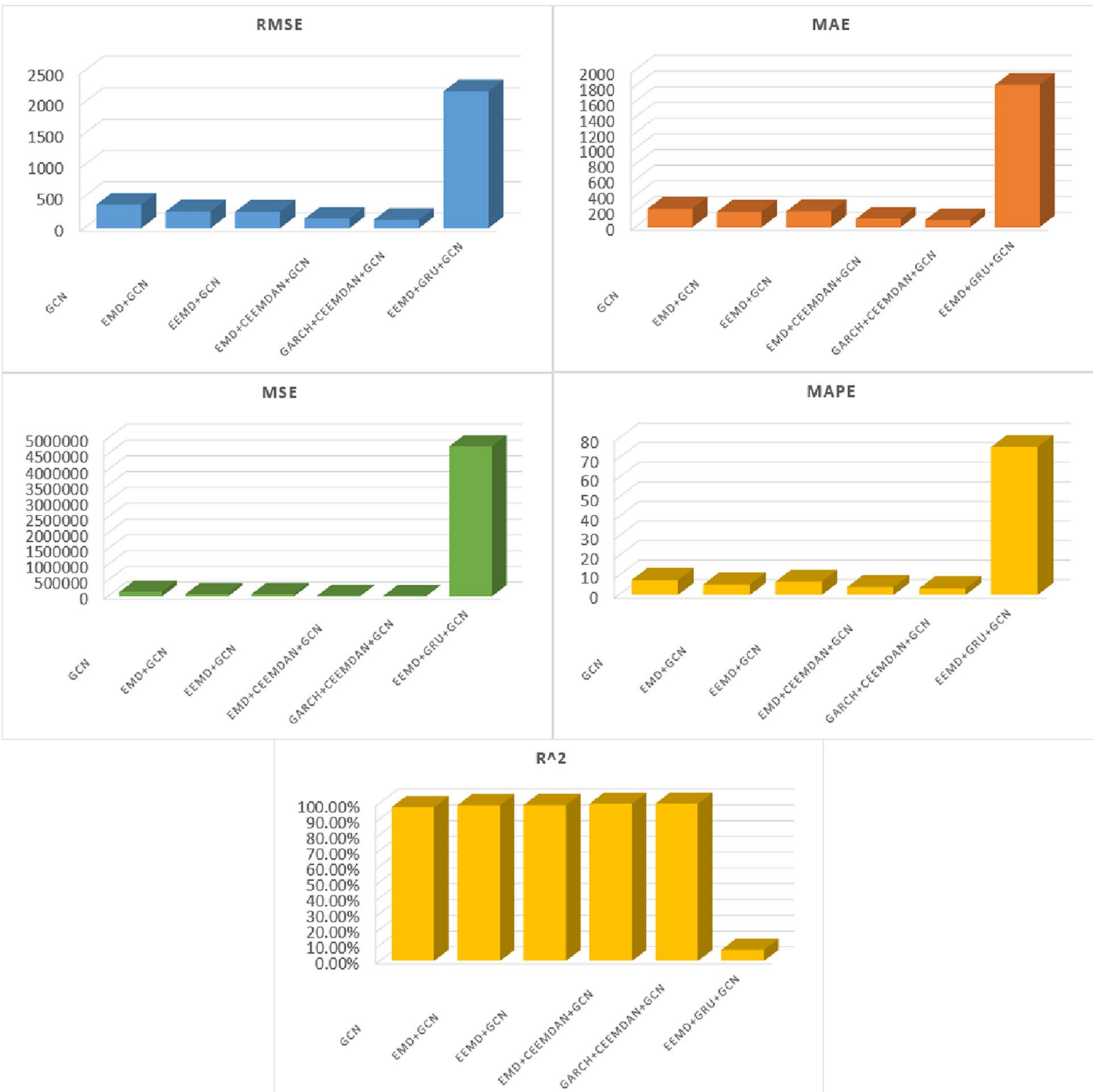

Fig. 3 Evaluation results of various models on the Energy dataset

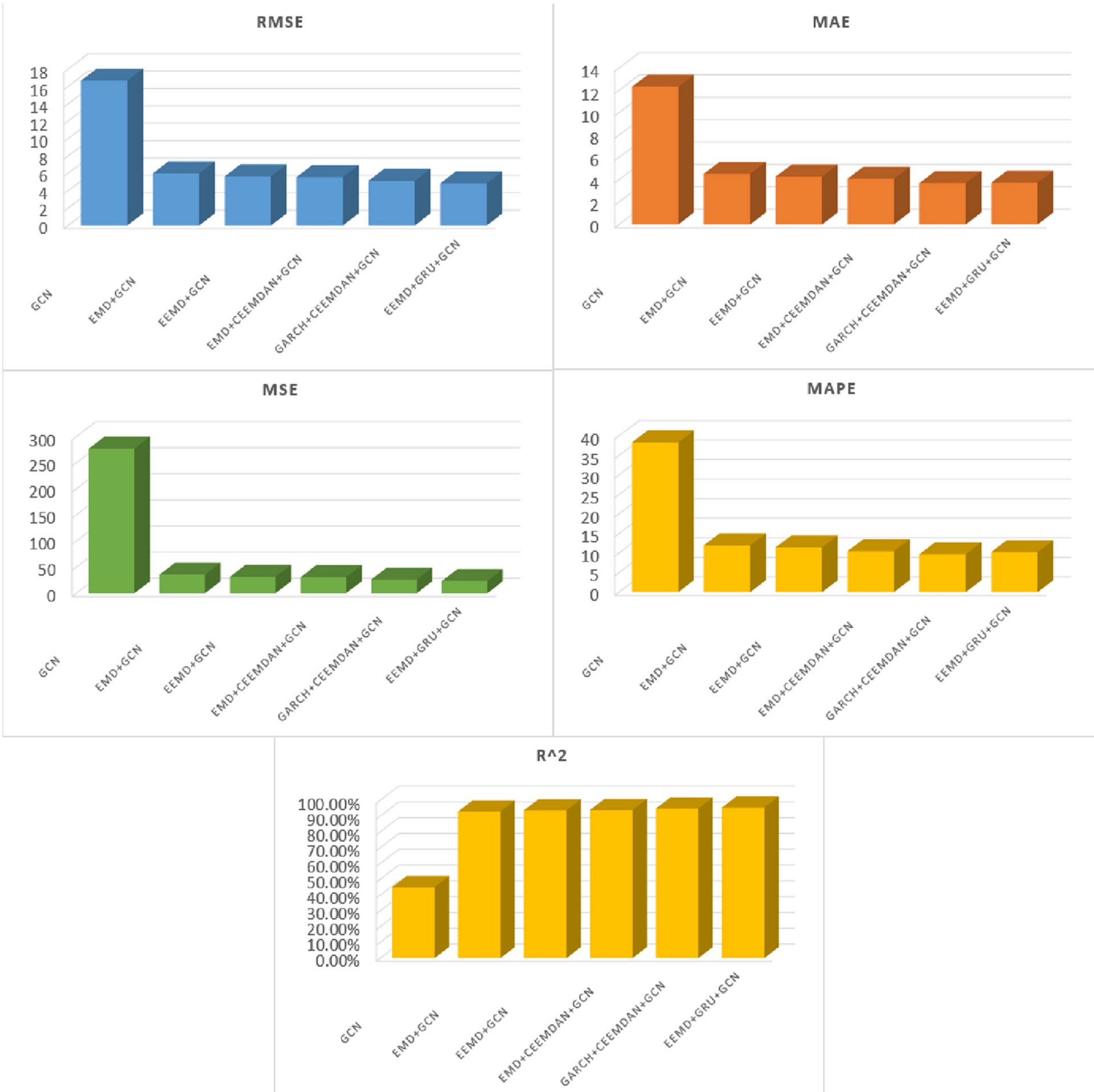

Fig. 4 Evaluation results of various models on the Traffic dataset

with environmental factors. This process is challenging due to the spatial and temporal variability of air quality, influenced by local pollution sources, weather, and seasonal changes. The complexity lies in understanding the interdependencies between different pollutants and environmental conditions. Moreover, predictions are critical for public health, as inaccuracies can have serious implications. Factors such as changing environmental policies, industrial activities, and data collection inconsistencies further complicate accurate prediction. Overall, effective air quality prediction requires managing variable, complex data while considering public health impact and data accuracy. RMSE of the proposed method is the lowest i.e., 9.48 while other algorithms are higher such as GARCH-CEEMDAN-GCN (16.64), EMD-CEEMDAN-GCN (17.54), EEMD-GCN (14.46), EMD-GCN (15.22) and GCN (17.52). Similarly, MAE is the lowest i.e., 7.27 for the proposed method while for other algorithms are higher GARCH-CEEMDAN-GCN (11.26), EMD-CEEMDAN-GCN (12.13), EEMD-GCN (10.18), EMD-GCN

(10.57) and GCN (12.18). For MSE, algorithms are GARCH-CEEMDAN-GCN (276.96), EMD-CEEMDAN-GCN (307.59), EEMD-GCN (209.2), EMD-GCN (231.62) and GCN (306.99) which is highest as compared to proposed method i.e. 90.01. MAPE is also the lowest for a proposed method which is 31.41 as compared to other methods i.e., GARCH-CEEMDAN-GCN (35.32), EMD-CEEMDAN-GCN (38.56), EEMD-GCN (35.98), EMD-GCN (34.41) and GCN (42.8). R2 of the proposed method is the highest among all the methods i.e., 67.10%.

Figure 3 shows the evaluation performances of different models on the Energy dataset. Predicting energy needs and production from datasets that track various sources like fossil fuels and renewables is complex due to the diversity of energy types and their unique characteristics. Challenges include managing demand fluctuations influenced by factors like weather and economic conditions, navigating infrastructure constraints of power grids and storage, and adapting to policy changes that affect energy markets. Environmental sustainability considerations and the rapid evolution of energy technologies, like renewables and energy-efficient devices, further complicate predictions. Thus, effective energy sector prediction requires sophisticated models capable of adapting to a dynamic landscape with varying demands, technological advancements, and regulatory environments. Results for all the parameters are high due to complex nature and proposed method accuracy is low due to nature of dataset.

Figure 4 shows the evaluation performances of different models on the Traffic dataset. Predicting traffic is a complex task specifically network traffic, which involves the analysis of vast and fast-generated data types such as packet counts, byte sizes, and IP addresses. This is crucial for managing network performance and security. Challenges in this field include the high volume and speed of data generation, requiring efficient processing techniques; the complexity and variability of traffic due to user behavior and external factors like cyber-attacks; the need for accurate anomaly detection in a dynamic environment; the presence of temporal dependencies where past patterns affect future ones; privacy concerns due to the sensitivity of the data; and the need for adaptability in predictive models due to evolving network technologies and usage patterns. Therefore, effectively predicting network traffic demands handling large-scale, complex data while maintaining accuracy, privacy, and adaptability in models. Compared with other models, RMSE for the EEMD-GRU-GCN model is lowest with a value of 4.84, while other methods GARCH-CEEMDAN-GCN (5.14), EMD-CEEMDAN-GCN (5.57), EEMD-GCN (5.64), EMD-GCN (5.98) and GCN (16.65) are higher than proposed. Similarly, MAE is the lowest for the proposed method i.e., 3.72 while GARCH-CEEMDAN-GCN (3.69), EMD-CEEMDAN-GCN (4.08), EEMD-GCN (4.27), EMD-GCN (4.52) and GCN (12.31) are more than proposed method. MSE is 23.488 for the proposed algorithm while other methods GARCH-CEEMDAN-GCN (26.39), EMD-CEEMDAN-GCN (31.01), EEMD-GCN (31.83), EMD-GCN (35.78) and GCN (277.14) is much higher. MAPE follows a different pattern for other algorithms i.e., GARCH-CEEMDAN-GCN (9.69), EMD-CEEMDAN-GCN (10.48), EEMD-GCN (11.45), EMD-GCN (11.97) and GCN (38.29) while proposed method 10.24 is second lowest. Since the R2 is the highest 95% for the proposed model as compared to GARCH-CEEMDAN-GCN (95%), EMD-CEEMDAN-GCN (94%), EEMD-GCN (94%), EMD-GCN(93%) and GCN(45%) which shows the proposed method is outstanding in traffic dataset.

From the experiments conducted on these three datasets, it can be observed that on the Air Quality dataset, the proposed hybrid model performs best in terms of all metrics.

Additionally, Fig. 5 shows the comparison between the real and predicted data of all algorithms on a selection of 150 consecutive data points from the test set of the Air Quality dataset. Figure 6 shows the same comparison on a selection of 150 consecutive data points from the test set of the Energy dataset. Finally, Fig. 7 shows the comparison on a selection of 150 consecutive data points from the test set of the Traffic dataset.

Discussion

The benefits arising from the EEG-GCN model presented in this study are extensive and can be observed in various domains that depend on the accurate analysis of time series data. Some of the key benefits are detailed below:

Enhanced Forecasting Accuracy: By integrating advanced signal decomposition with a graph convolutional neural network, the EEG-GCN model offers a marked improvement in forecasting accuracy. This is crucial for industries where precision in prediction can have significant economic implications, such as in stock market trading or energy supply planning.

Noise Reduction and Signal Clarity: The utilization of EEMD and CEEMDAN within the EEG-GCN model effectively filters out noise, thereby providing clearer signals for analysis. This is particularly beneficial in environments where data is heavily contaminated with noise, such as in medical signal processing or environmental monitoring.

Improved Decision-Making: With more reliable forecasts, decision-makers in businesses, governments, and other organizations can plan with greater confidence. This could mean better inventory manage-

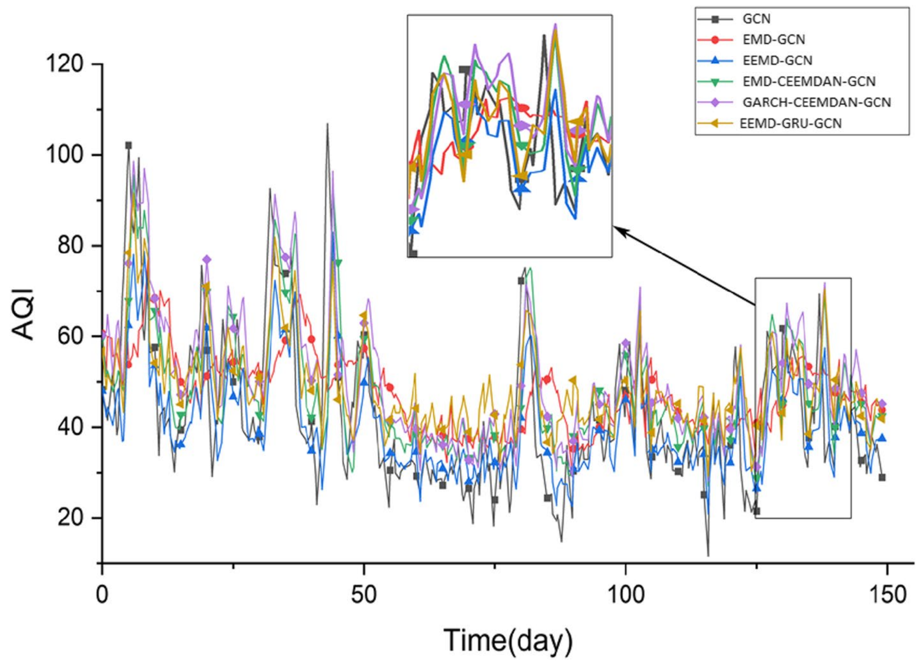

Fig. 5 Time series visual comparison of prediction results of 150 points observation of all the algorithms in the Air Quality dataset

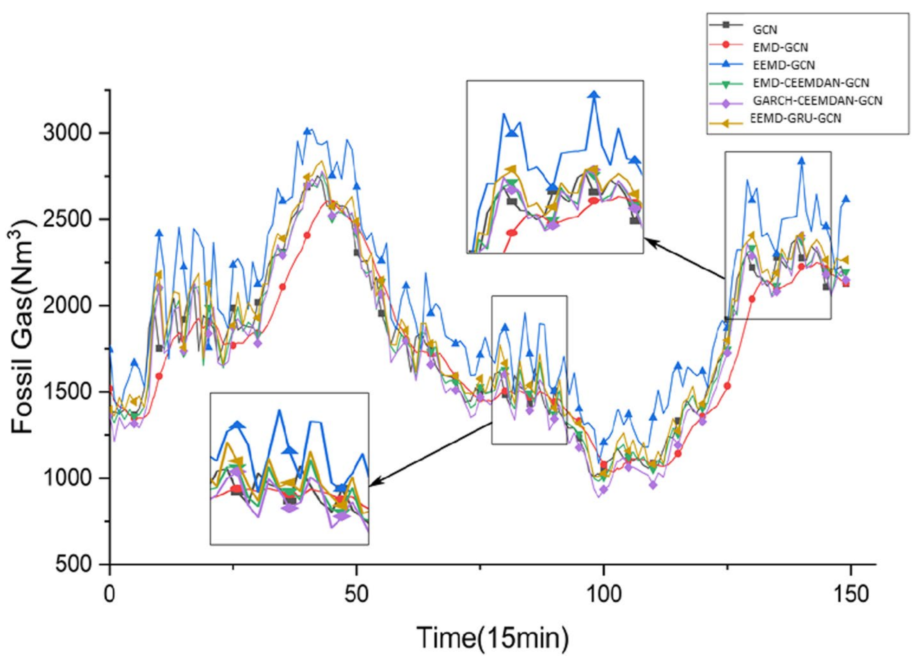

Fig. 6 Time series visual comparison of prediction results of 150 points observation of all the algorithms in the Energy dataset

ment in retail, more effective policy development in public health, or enhanced resource allocation in disaster management.

Operational Efficiency: In sectors like manufacturing and logistics, where time series predictions are used for demand forecasting, the EEG-GCN model can contribute to leaner operations by optimizing production schedules and supply chain operations, thereby reducing waste and improving customer satisfaction.

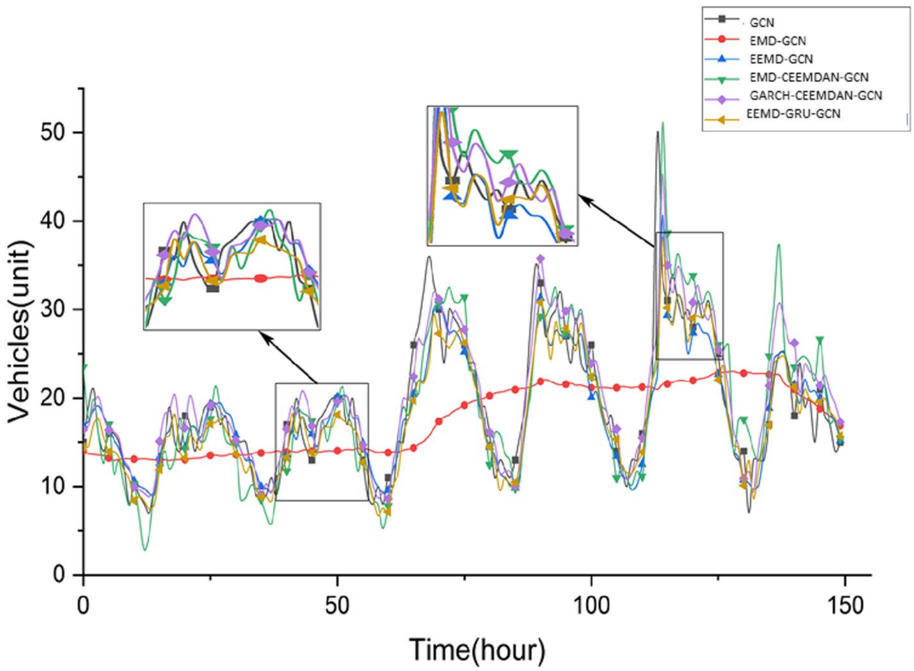

Fig. 7 Time series visual comparison of prediction results of 150 points observation of all the algorithms in the Traffic dataset

Energy Sector Advancement: The energy sector can greatly benefit from more accurate predictions of renewable energy outputs, leading to improved grid management and energy storage solutions. This could help in balancing supply and demand, thus facilitating a transition to greener energy sources.

Risk Mitigation: Financial institutions and insurance companies can use the model to better understand and predict market dynamics or claim trends, which can lead to more effective risk assessment and mitigation strategies.

Technological Innovation: The EEG-GCN model's approach encourages further innovation in machine learning and artificial intelligence by showcasing the effectiveness of hybrid models that can be tailored for specific complex data scenarios.

Cross-Disciplinary Applications: Given its flexibility and accuracy, the model has potential applications across a wide range of disciplines, from climate science and healthcare to urban planning and environmental protection.

Resource Management: For sectors like agriculture, where time series data can predict seasonal patterns and crop yields, the EEG-GCN model can lead to more efficient use of water, fertilizers, and other resources, contributing to sustainable practices.

Customizability and Scalability: The model's architecture allows for customization to suit the specific nuances of various types of time series data, which means it can be scaled and adapted for different industries and applications.

In essence, the EEG-GCN model's ability to deliver more accurate and reliable time series predictions translates into potential economic benefits, operational improvements, risk reduction, and the enabling of better strategic planning across diverse sectors.

The practical implications of this study are multifaceted and have considerable potential to impact various domains where time series data play a critical role. At the heart of the EEG-GCN model is its ability to manage the inherent complexity of temporal data, making it a valuable tool for industries and sectors that rely heavily on accurate forecasting. First and foremost, the EEG-GCN model's superior handling of noise and non-linearities makes it an exceptional candidate for deployment in financial markets, where time series data are notoriously volatile and noisy. The ability of the EEG-GCN to decompose these signals into more manageable components means that financial analysts could achieve more accurate forecasts of stock prices, market indices, and economic indicators. This increased accuracy could significantly reduce the risk of unforeseen market volatility and allow for better asset allocation and risk management strategies.

In the energy sector, particularly in renewable energy management, the EEG-GCN model can be leveraged to predict energy production from sources such as wind and solar power, which are inherently intermittent and unpredictable. The model's decomposition of complex weather-related data into simpler sub-signals could lead to more accurate predictions of energy availability. This could, in turn, facilitate more efficient grid management and energy storage, reduce wastage, and ensure a steadier supply of renewable energy to consumers. Another area where the EEG-GCN model shows promise is in environmental monitoring and climate science. Climate datasets are characteristically rich in non-linear trends and noise due to the myriad factors that affect weather systems. The EEG-GCN's enhanced capability to dissect and understand these datasets can assist in more reliable climate modeling and forecasting, which is essential for planning in agriculture, disaster management, and policy-making.

Healthcare could also benefit from this model, particularly in the analysis of medical time series data such as heart rate or glucose level monitoring. The EEG-GCN's ability to sift through the 'noise' of biological variability and other artifacts to predict patient-specific events could lead to more personalized and timely healthcare interventions. Moreover, the incorporation of GRU into the EEG-GCN framework, resulting in the EEMD-GRU-GCN, presents a methodological advancement for handling data across time with more nuanced interpretations. This aspect of the model is crucial for real-time monitoring systems, such as those used in industrial process control or traffic management, where understanding the temporal sequence of events is as important as recognizing patterns within them.

In summary, the EEG-GCN model holds significant practical utility across a wide array of fields that require the forecasting of complex time series data. Its empirical strength demonstrated through improved performance metrics, positions it as a potentially transformative tool for decision-makers seeking to derive actionable insights from challenging datasets. The ability to turn complex, noisy, and non-linear time series into accurate predictions can lead to more informed decisions, optimized operations, and a better understanding of future scenarios in various sectors.

Despite the notable advancements offered by the EEG-GCN model in time series data prediction, it is essential to acknowledge the limitations inherent in this work:

Computational Complexity: The EEG-GCN model incorporates complex algorithms such as EEMD and GCN, which could be computationally intensive. This may require significant computational resources and could be a limiting factor for real-time applications or for use in environments with limited computing infrastructure.

Data Requirement: The efficacy of the model is contingent upon the availability of high-quality, granular data. In cases where data is sparse, irregular, or of poor quality, the performance of the model might be compromised.

Overfitting Risk: As with many sophisticated models, there is a potential risk of overfitting, where the model becomes too closely fitted to the training data, impairing its generalization capabilities to unseen data.

Interpretability: Neural network-based models, including GCNs, are often considered 'black boxes' due to their complex nature, which can make it challenging to interpret the decision-making process or the significance of various inputs.

Dependency on Parameter Tuning: The performance of the EEG-GCN model heavily relies on the appropriate tuning of parameters. Finding the right configuration requires expertise and can be time-consuming, potentially limiting its accessibility to non-experts.

Generalizability: While the model has shown promising results, the extent to which it can be generalized across different domains and datasets without significant reconfiguration is unclear. Different types of time series data may require bespoke adjustments to the model.

Model Adaptation: As data evolves over time, the model may require retraining or updating to maintain accuracy, which could be a resource-intensive process.

Algorithmic Bias: Any predictive model is subject to the risk of bias, which can be introduced through the training data or the subjective choices in the model design process. Such bias could affect the fairness and reliability of predictions.

Transferability Across Domains: The adaptability of the model across various fields has yet to be thoroughly tested. Success in one domain, like energy forecasting, doesn't automatically ensure success in another, like financial markets.

Technology Integration: The integration of the EEG-GCN model into existing systems may pose challenges, as it might not be compatible with legacy systems or could require substantial changes to current workflows.

Training Time: Given the sophisticated nature of the model, the training time might be considerable, especially for very large datasets, which could be a bottleneck for time-sensitive applications.

Susceptibility to Dynamic Changes: Time series data can be influenced by sudden, unforeseen events (e.g., economic crashes, natural disasters). The model's ability to quickly adapt to such non-regular, abrupt changes is not fully established.

Recognizing these limitations is essential for the ongoing development and application of the EEG-GCN model. Addressing these challenges through continued research and development can lead to improved versions of the model that are more robust, efficient, and widely applicable.

Conclusion

In conclusion, the EEG-GCN model represents a significant advancement in the field of time series data prediction, demonstrating remarkable improvements in accuracy and stability over existing models. By intelligently integrating signal decomposition methods with the innovative graph convolutional neural network approach, the model adeptly navigates the complexities of non-linear and periodic data characteristics, while also effectively mitigating the influence of noise.

However, the study acknowledges the limitations, including computational demand, the necessity for high-quality data, the risk of overfitting, challenges with interpretability, and the critical need for meticulous parameter tuning. These constraints highlight the scope for further refinement and optimization of the model.

Future work could focus on several aspects:

Efficiency Optimization: Developing strategies to reduce the computational load of the EEG-GCN model without compromising prediction accuracy could make it more viable for a broader range of applications, including those with limited computational resources.
Data Quality Enhancement: Investigating methods to enhance the model's robustness to data quality, potentially through advanced data preprocessing techniques or robustness measures, could extend its applicability.
Interpretability Improvement: Efforts to increase the interpretability of the GCN component, such as through the development of visualization tools or the integration of explainable AI principles, would be beneficial.
Hyperparameter Tuning Automation: Implementing automated machine learning (AutoML) techniques for hyperparameter optimization could minimize the need for manual tuning and open the model's use to a wider audience.
Domain Adaptability: Conducting cross-domain studies to test the transferability of the model could provide insights into its versatility and adaptability to different types of time series data.
Dynamic Adaptation: Enhancing the model to better cope with abrupt changes in data patterns by incorporating real-time learning capabilities could greatly improve its utility in dynamic environments.
Bias Mitigation: Developing methodologies to detect and correct biases in both training data and model predictions is crucial to ensure fairness and reliability in different application scenarios.
System Integration: Addressing the challenges of integrating the EEG-GCN model into existing technological frameworks could accelerate its adoption in industry.
Training Time Reduction: Investigating methods to decrease model training time, possibly through parallel computing or more efficient algorithms, would make the model more practical for large datasets and real-time applications.
Model Generalization: Further research is needed to understand the conditions under which the model generalizes best and to develop guidelines for adapting the model to a variety of situations.

The EEG-GCN model's promising performance lays a solid foundation for future research and potential practical applications. It opens up new avenues for the predictive analysis of time series data across different sectors such as finance, weather forecasting, energy management, and beyond. As the model continues to evolve, it is poised to become an even more indispensable tool for analysts and decision-makers facing the challenge of extracting meaningful insights from complex temporal data streams.

Authors' contributions
All authors reviewed the manuscript.

Funding
This research was sponsored by the Key Laboratory of Philosophy and Social Sciences in Guangdong Province of Maritime Silk Road of Guangzhou University (GD22TWCXGC15) and the National Natural Science Foundation of China (Grant No. 622260-101). The authors present their appreciation to King Saud University for funding this research through the Researchers Supporting Program number (RSPD2024R1006), King Saud University, Riyadh, Saudi Arabia.

Declarations

Competing interests
The authors declare no competing interests.

Author details
[1]Mechanical and Electrical Engineering College, Hainan Vocational University of Science and Technology, Haikou 571126, China. [2]Department of Computer Science and Electronics, Universidad de La Costa, CUC, 080002 Barranquilla, Colombia. [3]Metaverse Research Institute, School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou, Guangdong 510006, People's Republic of China. [4]Department of Industrial Engineering, College of Engineering, King Saud University, P.O. Box 800, 11421 Riyadh,

Saudi Arabia. [5]Department of Electrical Engineering, College of Engineering, King Saud University, P.O. Box 800, 11421 Riyadh, Saudi Arabia. [6]Department of Computer Science, Al Ain University, Al Ain, UAE. [7]Department of Structural Engineering, Faculty of Civil Engineering, Doctoral School, Silesian University of Technology, Akademicka 2, 44-100 Gliwice, Poland.

Received: 2 November 2023 Accepted: 28 November 2023
Published online: 18 January 2024

References

1. Salles R, Pacitti E, Bezerra E, Porto F, Ogasawara E (2022) T S Pred: a framework for nonstationary time series prediction. Neurocomputing 467:197–202
2. Goudarzi G, Birgani YT, Assarehzadegan MA, Neisi A, Dastoorpoor M, Sorooshian A, Yazdani M (2022) Prediction of airborne pollen concentrations by artificial neural network and their relationship with meteorological parameters and air pollutants. J Environ Health Sci Eng 20(1):251–264
3. Méndez M, Merayo MG, Núñez M (2023) Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model. Eng Appl Artif Intell 121:106041
4. Bhatti U, Masud M, Bazai S, Tang H (2023). Editorial: Investigating AI-based smart precision agriculture techniques. Front Plant Sci 14 https://doi.org/10.3389/fpls.2023.1237783.
5. Fischer E, Barreca G, Greco A et al (2023) Seismic risk assessment of a large metropolitan area by means of simulated earthquakes. Nat Hazards 118:117–153
6. Mahmoud A, Mohammed A (2021) A survey on deep learning for time-series forecasting. In: Hassanien AE, Darwish A. (eds) Machine learning and big data analytics paradigms: analysis, applications and challenges. Studies in Big Data. vol 77. Springer, Cham.
7. Guo K, Yu X, Liu G, Tang S (2023) A Long-Term Traffic Flow Prediction Model Based on Variational Mode Decomposition and Auto-Correlation Mechanism. Appl Sci 13:7139
8. Hahn Y, Langer T, Meyes R, Meisen T (2023) Time Series Dataset Survey for Forecasting with Deep Learning. Forecasting 5:315–335
9. Ning Y, Kazemi H, Tahmasebi P (2022) A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet. Comput Geosci 164:105126
10. Dong S, Xiao J, Xiaolin Hu, Fang N, Liu L, Yao J (2023) Deep transfer learning based on Bi-LSTM and attention for remaining useful life prediction of rolling bearing. Reliab Eng Syst Saf 230:108914
11. He R, Zhang C, Xiao Y, Lu X, Zhang S, Yanbing Liu Y (2024) Deep spatio-temporal 3D dilated dense neural network for traffic flow prediction. Expert Syst Appl 237(Part A):121394.
12. Yuan Y, Shao C, Cao Z, Chen W, Yin A, Yue H, Xie B (2019) Urban rail transit passenger flow forecasting method based on the coupling of artificial fish swarm and improved particle swarm optimization algorithms. Sustainability 11:7230
13. Zheng H, Chen J, Huang Z, Yang K, Zhu J (2022) Short-Term Online Forecasting for Passenger Origin-Destination (OD) Flows of urban rail transit: a graph-temporal fused deep learning method. Mathematics 10:3664
14. Banerjee N, Morton A, Akartunal K (2020) Passenger demand forecasting in scheduled transportation. Eur J Oper Res 286(3):797–810
15. Li W, Sui L, Zhou M et al (2021) Short-term passenger flow forecast for urban rail transit based on multi-source data. J Wireless Com Network 2021:9
16. Toan TD, Truong V-H (2021) Support vector machine for short-term traffic flow prediction and improvement of its model training using nearest neighbor approach. Transp Res Rec 2675(4):362–373
17. Liu Y, Rasouli S, Wong M, Feng T, Huang T (2024) RT-GCN: Gaussian-based spatiotemporal graph convolutional network for robust traffic prediction. Inform Fusion 102:102078
18. Luo X, Li D, Zhang S (2019) Traffic Flow Prediction during the Holidays Based on DFT and SVR. Journal of Sensors 2019:1–10. https://doi.org/10.1155/2019/6461450
19. Ma C, Zhao Y, Dai G, Xu X, Wong SC (2022). A novel STFSA-CNN-GRU Hybrid model for short-term traffic speed prediction. IEEE Transact Intell Transport Systems. PP. 1–10. https://doi.org/10.1109/TITS.2021.3117835.
20. Wang S, Shao C, Zhang J, Zheng Y, Meng M (2022) Traffic flow prediction using bi-directional gated recurrent unit method. Urban Inform 1(1):16
21. Zafar N, Haq IU, Chughtai JU, Shafiq O (2022) Applying Hybrid Lstm-Gru Model Based on Heterogeneous Data Sources for Traffic Speed Prediction in Urban Areas. Sensors (Basel) 22(9):3348
22. Zhang W, Yao R, Du X, Ye J. (2021). Hybrid deep spatio-temporal models for traffic flow prediction on holidays and under adverse weather. IEEE Access. PP. 1–1. https://doi.org/10.1109/ACCESS.2021.3127584.
23. Bhatti UA, Huang M, Neira-Molina H, Marjan S, Baryalai M, Tang H, Wu G, Bazai SU (2023) MFFCG – Multi feature fusion for hyperspectral image classification using graph attention network. Expert Syst Appl 229 (Part A):120496
24. Fodstad M, del Granado PC, Hellemo L, Knudsen BR, Pisciella P, Silvast A, Bordin C, Schmidt S, Straus J (2022) Next frontiers in energy system modelling: a review on challenges and the state of the art. Renew Sustain Energy Rev 160:112246.
25. Joe P, Sun J, Yussouf N, Goodman S, Riemer M, Gouda KC, Golding B, Rogers R, Isaac G, Wilson J, Li PW, Wulfmeyer V, Elmore K, Onvlee J, Chong P and Ladue J (2022) Predicting the weather: a partnership of observation scientists and forecasters. In: Golding, B. (eds) Towards the "Perfect" Weather Warning. Springer, Cham.
26. Yan J, Möhrlen C, Göçmen T, Kelly M, Wessel A, Giebel G (2022) Uncovering wind power forecasting uncertainty sources and their propagation through the whole modelling chain. Renew Sustain Energy Rev 165:112519
27. Wang H (2023) Extreme learning Kalman filter for short-term wind speed prediction. Front Energy Res 10:1047381. https://doi.org/10.3389/fenrg.2022.1047381
28. Fattah J, Ezzine L, Aman Z, El Moussami H, Lachhab A 2018 Forecasting of demand using ARIMA model. Int J Eng Bus Manag 10.
29. Hanifi S, Lotfian S, Zare-Behtash H, Cammarano A (2022) Offshore wind power forecasting—A new hyperparameter optimisation algorithm for deep learning models. Energies 15:6919
30. Ospina R, Gondim JAM, Leiva V, Castro C (2023) An overview of forecast analysis with ARIMA Models during the COVID-19 Pandemic: methodology and case study in Brazil. Mathematics 11:3069
31. Wang S, Wang J, Haiyan Lu, Zhao W (2021) A novel combined model for wind speed prediction – Combination of linear model, shallow neural networks, and deep learning approaches. Energy 234:121275
32. Nair KR, Vanitha V and Jisma M (2017) Forecasting of wind speed using ANN, ARIMA and Hybrid models, 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kerala, India, 2017, pp. 170–175,
33. Liu H, Tian H-q, Li Y-F (2012) Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. Appl Energy 98:415–424
34. Aasim SN, Singh AM (2019) Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. Renew Energy 136:758–768.
35. Bhatti UA, 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
36. Zhang X, Wu X, Zhu G, Lu X, Wang K (2022) A seasonal ARIMA model based on the gravitational search algorithm (GSA) for runoff prediction. Water Supply 22(8): 6959–6977.
37. Wang H, Yan S, Ju D, Ma N, Fang J, Wang S, Li H, Zhang T, Xie Y, Wang J (2023) Short-term photovoltaic power forecasting based on a feature rise-dimensional two-layer ensemble learning model. Sustainability 15:15594
38. Shuai Hu, Xiang Y, Zhang H, Xie S, Li J, Chenghong Gu, Sun W, Liu J (2021) Hybrid forecasting method for wind power integrating spatial correlation and corrected numerical weather prediction. Appl Energy 293:116951
39. Nawab F, Abd Hamid AS, Ibrahim A, Sopian K, Fazlizan A, Fauzan MF (2023) Solar irradiation prediction using empirical and artificial intelligence methods: a comparative review. Heliyon 9(6)
40. Deng Y, Zhou X, Shen J, Xiao G, Hong H, Lin H, Wu F, Liao BQ (2021) New methods based on back propagation (BP) and radial basis function (RBF) artificial neural networks (ANNs) for predicting the occurrence of haloketones in tap water. Sci Total Environ 10(772):145534
41. Ellahi M, Usman MR, Arif W, Usman HF, Khan WA, Satrya GB, Daniel K, Shabbir N (2022) Forecasting of wind speed and power through FFNN

and CFNN using HPSOBA and MHPSO-BAACs techniques. Electronics 11:4193
42. Chen N, Xiong C, Du W, Wang C, Lin X, Chen Z (2019) An Improved Genetic Algorithm Coupling a Back-Propagation Neural Network model (IGA-BPNN) for water-level predictions. Water 11:1795
43. Bhatti UA, Tang H, Wu G, Marjan S, Hussain A (2023) Deep learning with graph convolutional networks: an overview and latest applications in computational intelligence. Int J Intell Syst 2023:1–28
44. Al-Majidi SD, Abbod MF, Al-Raweshidy HS (2020) A particle swarm optimisation-trained feedforward neural network for predicting the maximum power point of a photovoltaic array. Eng Appl Artif Intell 92:103688
45. Husein M, Chung IY. Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: A deep learning approach. Energies. 2019;12(10):1856.
46. Uncuoglu E, Citakoglu H, Latifoglu L, Bayram S, Laman M, Mucella Ilkentapar, Alper Oner AA (2022) Comparison of neural network, Gaussian regression, support vector machine, long short-term memory, multigene genetic programming, and M5 Trees methods for solving civil engineering problems. Appl Soft Comput 129:109623
47. Zhang C, Zhang M (2022) Wavelet-based neural network with genetic algorithm optimization for generation prediction of PV plants. Energy Rep 8:10976–10990
48. Park R-J, Song K-B, Kwon B-S (2020) Short-term load forecasting algorithm using a similar day selection method based on reinforcement learning. Energies 13:2640
49. 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. Compu Mat Continua 77:681–697. https://doi.org/10.32604/cmc.2023.037958.
50. Gagnon P, Cole W (2022) Planning for the evolution of the electric grid with a long-run marginal emission rate. iScience 25(3):103915.
51. Almani AA, Han X (2023) Real-time pricing-enabled demand response using long short-time memory deep learning. Energies 16:2410
52. Barton M, Lennox B (2022) Model stacking to improve prediction and variable importance robustness for soft sensor development. Digital Chem Eng 3
53. Zheng H, Yuan J, Chen L (2017) Short-term load forecasting using emd-lstm neural networks with a Xgboost algorithm for feature importance evaluation. Energies 10:1168
54. Vanting NB, Ma Z, Jørgensen BN (2021) A scoping review of deep neural networks for electric load forecasting. Energy Inform 4(Suppl 2):49
55. Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, Liu X, Yanjun Wu, Dong F, Qiu C-W, Qiu J, Hua K, Wentao Su, Jian Wu, Huiyu Xu, Han Y, Chenguang Fu, Yin Z, Liu M, Roepman R, Dietmann S, Virta M, Kengara F, Zhang Ze, Zhang L, Zhao T, Dai Ji, Yang J, Lan L, Luo M, Liu Z, An T, Zhang B, He X, Cong S, Liu X, Zhang W, Lewis JP, Tiedje JM, Wang Qi, An Z, Wang F, Zhang L, Huang T, Chuan Lu, Cai Z, Wang F, Zhang J (2021) Artificial intelligence: a powerful paradigm for scientific research. Innovation 2(4):100179
56. Ullah I, Muhammad Hasanat S, Aurangzeb K, Alhussein M, Rizwan M, Anwar MS (2023) Multi-horizon short-term load forecasting using hybrid of LSTM and modified split convolution. PeerJ Comput Sci 15(9):e1487
57. Muhammad Ahsan Zamee, Dongjun Han, Heejune Cha, Dongjun Won (2023) Self-supervised online learning algorithm for electric vehicle charging station demand and event prediction. J Energy Storage 71:108189.
58. Zheng J, Zhu J, Xi H (2023) Short-term energy consumption prediction of electric vehicle charging station using attentional feature engineering and multi-sequence stacked Gated Recurrent Unit. Comput Electr Eng 108:108694
59. He W, Li Z, Liu T, Liu Z, Guo X, Jinguang Du, Li X, Sun P, Ming W (2023) Research progress and application of deep learning in remaining useful life, state of health and battery thermal management of lithium batteries. J Energy Storage 70:107868
60. Bhatia K, Mittal R, Varanasi J, Tripathi MM (2021) An ensemble approach for electricity price forecasting in markets with renewable energy resources. Utilities Policy 70:101185.
61. Xinxin W, Xiaopan S, Xueyi A, Shijia L (2023) Short-term wind speed forecasting based on a hybrid model of ICEEMDAN, MFE, LSTM and informer. PLoS One 18(9):e0289161
62. Zhao L, Li Z, Zhang J, Teng B (2023) An integrated complete ensemble empirical mode decomposition with adaptive noise to optimize LSTM for significant wave height forecasting. J Mar Sci Eng 11:435
63. Liu H, Xiong X, Yang B, Cheng Z, Shao K, Tolba A (2023) A power load forecasting method based on intelligent data analysis. Electronics 12:3441
64. Nepal B, Yamaha M, Yokoe A, Yamaji T (2020) Electricity load forecasting using clustering and ARIMA model for energy management in buildings. Jpn Archit Rev 3:62–76
65. Saglam M, Spataru C, Karaman OA (2023) Forecasting electricity demand in Turkey using optimization and machine learning algorithms. Energies 16:4499
66. Henrique BM, Sobreiro VA, Kimura H (2018) Stock price prediction using support vector regression on daily and up to the minute prices. J Finance Data Sci 4(3):183–201.
67. Gupta D, Pratama M, Ma Z, Li J, Prasad M (2019) Financial time series forecasting using twin support vector regression. PLoS One 14(3)
68. Ashfaq T and Javaid N (2019) "Short-term electricity load and price forecasting using enhanced KNN," 2019 International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan, pp. 266–2665.
69. 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 Environ Policy 21(6):1341–1352
70. Sarker IH (2021) Deep Learning: a comprehensive overview on techniques, taxonomy, applications and research directions. Sn Comput Sci 2:420
71. Zha W, Liu Y, Wan Y, Luo R, Li D, Yang S, Yanmei Xu (2022) Forecasting monthly gas field production based on the CNN-LSTM model. Energy 260:124889
72. 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.
73. Yin H, Zuhong Ou, Huang S, Meng A (2019) A cascaded deep learning wind power prediction approach based on a two-layer of mode decomposition. Energy 189:116316
74. Zhang J, Siya W, Zhongfu T, Anli S (2023) An improved hybrid model for short term power load prediction. Energy 268:126561
75. Shu W, Gao Q (2020) Forecasting stock price based on frequency components by EMD and neural networks. Ieee Access 8:206388–206395
76. Wu YX, Wu QB, Zhu JQ (2019) Improved EEMD-based crude oil price forecasting using LSTM networks. Physica A 516:114–124
77. Yin C, Wang G, Liao J (2023) Application of VMD–SSA–BiLSTM algorithm to smart grid financial market time series forecasting and sustainable innovation management. Front Energy Res 11:1239542

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