Pol. J. Environ. Stud. Vol. 33, No. 2 (2024), 1121-1134
DOI: *10.15244/pjoes/172756* ONLINE PUBLICATION DATE: 2023-11-27

Original Research

Hybrid Climate Forecasting: Variational Mode Decomposition and Convolutional Neural Network with Long-Term Short Memory

Huimin Han[1], Sibghat Ullah Bazai[2], Mughair Aslam Bhatti[2]*, Abdul Basit[5], Abdul Wahid[6], Uzair Aslam Bhatti[4], Yazeed Yasin Ghadi[7], Abdulmohsen Algarni[8]**

[1]Mechanical and Electrical Engineering College, Hainan Vocational University of Science and Technology, Haikou, 571126, China
[2]School of Geography, Nanjing normal university, Nanjing 210023, China
3Department of Computer Engineering, Balochistan University of Information Technology, Engineering, and Management Sciences (BUITEMS), Quetta, Pakistan
[4]School of information and Communication Engineering Hainan University, Haikou, China
[5]Department of Computer Science and IT University of Baluchistan, Quetta, Pakistan. 87300
[6]Department of Electronic Engineering, Balochistan University of Information Technology, Engineering, and Management Sciences (BUITEMS), Quetta, Pakistan
[7]Department of Computer Science, Al Ain University, UAE
[8]Department of Computer Science, King Khalid University, Abha 61421, Saudi Arabia

Received: 14 July 2023
Accepted: 21 September 2023

Abstract

Ozone (O_3) pollution has surfaced as a significant threat to urban air quality in contemporary years. The precise and efficient forecast of ozone levels is fundamental in the mitigation and management of ozone pollution. Even though the air quality monitoring network offers useful multi-source pollutant concentration data for predicting ozone levels, existing models still grapple with issues arising from outlier and redundant sites influencing prediction precision, and cross-contamination between different pollutants. Also, the non-linear and volatile nature of monthly runoff makes accurate prediction more complex, provide a more granular and timely view of atmospheric flow variations. In this research, we introduce a hybrid model that unites Variational Modal Decomposition (VMD), particularly useful for separating mixed signals or extracting meaningful patterns from noisy or complex data, Convolutional Long Short-Term Memory Neural Network (CNN-LSTM) is designed for processing sequences of data with grid-like structures, such as images or video frames. CNN-LSTMs use convolutional operations to capture spatial patterns and LSTM units to model temporal dependencies, making them effective for tasks like video analysis, image sequence prediction, and spatiotemporal data processing, and VMD-CNN-LSTM to counter these issues. We commence by deconstructing the historical data series from the Nanjing air quality monitoring stations using VMD. Then, the Ensemble Empirical Mode

*e-mail: mughairbhatti@nnu.edu.cn
**e-mail: a.algarni@kku.edu.sa

Decomposition (EEMD) algorithm is applied to the VMD residual to acquire characteristic components or Intrinsic Mode Functions (IMFs). Each IMF is independently trained via LSTM to produce predictions for each component. Ultimately, we secure the final prediction by linearly superimposing the predictions from all components. The LSTM's adaptive learning ability and memory function make it ideal for managing long-term data, leading to more precise predictions. To evaluate the prediction performance on the test set, our VMD-CNN-LSTM model is compared with other models such as EMD-LSTM, EMD-CNN-LSTM, and VMD-LSTM using root mean square error (RMSE), mean absolute error (MAE), and Nash coefficient (NSE). Our findings reveal that the VMD-CNN-LSTM model surpasses the other models, displaying higher prediction precision and lower errors. Importantly, the model shows enhanced fitting of peak and valley values, thus providing a promising strategy for monthly runoff prediction. In this research, we've put forth a unique hybrid model, VMD-CNN-LSTM, for monthly ozone prediction. By amalgamating VMD, CNN, and LSTM, our model effectively tackles challenges associated with outlier and redundant sites, cross-pollution between pollutants, and nonlinearity makes it hard to model the intricate runoff relationships accurately, while instability results in unpredictable fluctuations, both of which impact the accuracy and reliability of monthly runoff predictions and make it more impactful in Environmental Management, Energy Optimization, Agriculture, Urban Planning, Climate Resilience

Keywords: ozone prediction model, LSTM, series decomposition, VMD

Introduction

The epoch of the Industrial Revolution catalyzed a massive leap in global industries and economies, unfortunately, along with an amplification of air pollutants from various industrial sources. The expansion and severity of worldwide pollution escalated, drawing increased focus [1-4]. An increase in air pollutant levels, coupled with the deteriorating state of the global environment, significantly impacts human health and overall quality of life. This underscores the urgent imperative to reduce pollutant concentrations and improve air quality [5-8]. Although numerous environmental protection policies have been introduced, alongside growing global awareness of atmospheric pollution reduction, excessive atmospheric pollutant concentrations persist, and environmental degradation remains a critical issue [9-10].

Air quality is principally impacted by pollutants like carbon monoxide (CO), carbon dioxide (CO_2), ozone (O_3), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and particulate matter ($PM_{2.5}$, PM_{10}). Among these, ozone plays a crucial role in the human health and ecological environment. Ozone is beneficial in small amounts, with a safe concentration at 0.2 mg/m^3 [11], while a concentration above 100 μg/m3 is considered unsafe. The increase in ozone concentrations over the years can be attributed to industrial plant emissions and increased solar intensity [12, 13]. Elevated ozone concentrations can severely harm human health, particularly the respiratory system, and negatively impact the ecological environment, causing a decrease in crop yields and damaging agriculture and forestry [13].

The importance of ozone pollution prediction spans several areas:

– Public Health: Ozone pollution, a detrimental air pollutant, can harm human health, causing respiratory problems, aggravating asthma and other respiratory diseases, and increasing the risk of cardiovascular issues. Accurate ozone prediction enables timely warnings and necessary measures to protect public health.
– Environmental Impact: Ozone pollution affects not just human health but the environment as well. It can harm vegetation, reducing crop yields and disrupting ecosystems. Predicting ozone levels allows for proactive measures to mitigate these environmental impacts.
– Policy and Regulations: Ozone pollution prediction aids in shaping policies and regulations aimed at lowering pollution levels. It helps policymakers identify sources of pollution and develop targeted interventions.
– Urban Planning and Infrastructure: Ozone prediction informs decisions regarding urban planning and infrastructure development. By understanding the spatial and temporal patterns of ozone pollution, city planners can minimize exposure to high ozone levels.
– Economic Implications: Ozone pollution can have significant economic consequences, impacting agricultural productivity and incurring health-related costs. Accurate prediction of ozone pollution enables better resource allocation and planning.

It is essential to protect the ecological environment while developing material civilization [14] as the sustainable coexistence of both ensures long-term prosperity for present and future generations. As awareness of environmental protection improves and concerns about air pollution status grow, establishing a scientific and effective atmospheric pollutants control mechanism becomes critical [15-16]. Air pollution prediction can effectively prevent serious pollution events, reduce the damage from heavily polluted weather, and minimize economic losses caused by severe air pollution [17-18].

Research has focused on ozone (O_3) forecasting [19]. Efforts have been made to enhance the accuracy of O_3 forecasting using classic methods such as regression analysis, time series analysis, autoregressive moving average models, and gray models [20, 21]. Despite their advantages, these methods often fail to capture nonlinear patterns prevalent in ozone data [22-24]. The advent of artificial intelligence (AI) technology has introduced intelligent methods for air quality and ozone forecasting [25]. Techniques such as expert systems, machine learning, and fuzzy reasoning have been utilized in this domain [26-30]. Among them, artificial neural networks (ANNs), capable of handling nonlinear factors, are often chosen for ozone forecasting [31]. Artificial Neural Networks (ANNs) mirror the human brain's structure, comprised of several layers of interconnected neurons, each governed by activation functions. This design permits ANNs to navigate intricate data, notably nonlinear data, that traditional ozone forecasting techniques struggle to handle [32-35].

Back-propagation algorithm (BP), the dominant method among ANN models, grapples with sluggish convergence and protracted iteration cycles [36], When discussing the challenges of back-propagation algorithms in Artificial Neural Network (ANN) models, it's important to note that 'convergence speed' refers to how quickly the algorithm reaches a stable and accurate solution during training. This impacts the efficiency of the learning process. Additionally, 'iteration cycles' indicate the number of times the algorithm updates its weights and biases by iteratively adjusting them during training. Efficiently managing these aspects is crucial for optimizing the training process in ANNs. Various refinements have been proposed to counter these limitations. For instance, Huang et al. employed Particle Swarm Optimization (PSO) to enhance the BP algorithm, boosting the convergence speed in air quality forecasting [37]. Likewise, Yang et al. constructed a BP-ANN model for PM2.5 forecasting, yielding satisfactory outcomes [38]. Qiao et al. employed a refined ant colony algorithm in conjunction with BPNN to predict peak load in Chengdu, showcasing a marked improvement in prediction accuracy compared to conventional regression models [39]. Ya et al. underscored the prowess of Convolutional Neural Networks (CNNs) in extracting nonlinear data features, emphasizing their advanced feature extraction capabilities [40], Convolutional Neural Networks (CNNs) are a class of deep learning models designed for processing grid-like data such as images. They excel at feature extraction through the use of specialized layers called 'convolutions.' These convolutions systematically scan the input data, identifying patterns, edges, textures, and more. This hierarchical feature extraction allows CNNs to automatically learn and represent complex features, making them highly effective for image analysis, object recognition, and other tasks. Their ability to capture spatial relationships and hierarchies in data has made CNNs a go-to choice for computer vision applications. Deep learning approaches, such as Recurrent Neural Networks (RNNs) and their variants like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have attracted interest for their performance in time series prediction tasks, particularly in deciphering underlying patterns in larger, more intricate datasets [41-43]. While RNNs have traditionally been employed, they're hampered by the vanishing gradient problem, which LSTM effectively resolves [44]. LSTM employs gated neural units output gates, input gates and forgetting gates, to retain contextual data and extract long-term features from time series data [45].

Numerous studies have harnessed LSTM for ozone forecasting, demonstrating its efficacy. For instance, Zhang et al. developed an ozone prediction methodology based on LSTM, delivering encouraging forecasting results [46]. Pak et al. presented a blend of CNN and LSTM for short-term ozone forecasting, using CNN to draw out high-level features and enhance accuracy when the LSTM input sequence is long [47]. This synergy of CNN and LSTM has been shown to be a superior model for predicting air quality time series data [48, 49]. The amalgamation of AI methods, particularly ANNs and deep learning models like LSTM, has progressed the ozone forecasting field. These models excel at deciphering the nonlinear characteristics of ozone data, enhancing prediction accuracy against traditional linear analysis methods. Further investigation into hybrid models, merging techniques such as CNN and LSTM, might yield improved results in forecasting air quality and ozone levels. The ongoing advancement and honing of AI-based techniques will contribute to more precise and reliable ozone forecasting, aiding in the control and reduction of air pollution.

To overcome the restrictions of traditional artificial intelligence models, series decomposition methods are utilized to process time series signals. Signal processing methods play a vital role in combined models and directly impact the model's predictive performance. Standard sequence preprocessing techniques include Wavelet Decomposition (WD) [50] and Empirical Mode Decomposition (EMD) [51]. The EMD decomposition algorithm doesn't rely on any basis function and differs significantly from wavelet decomposition. It's particularly efficient at handling non-stationary and nonlinear complex signals. Li et al. [52] put forward the AQI forecasting algorithm using Variational Mode Decomposition (VMD) and improved the VMD results by employing bald eagle search to propose a BVMD refined model for better outcomes. The research conducted in Xi'an, Beijing, and Shanghai shows that the proposed method is superior to other decomposition techniques.

Consequently, VMD dynamically dissects the relevant components associated with each center frequency within the frequency spectrum, demonstrating superior precision in decomposition. The VMD method's ability to decompose data has proven particularly beneficial in feature selection for predictive models and has seen successful applications in forecasting trends in asset prices within financial and energy markets [53]. Therefore, this study adopts the VMD decomposition approach as the main tool for decomposition in our modeling process.

In this paper, we use CNN-LSTM to directly process deformation monitoring time series data in the time domain and extract the data dependence information therein. Combined with VMD, we propose a deformation prediction model based on VMD-CNN-LSTM. Initially, VMD decomposes the deformation monitoring data sequence into a set of relatively stable intrinsic modal components with different frequency scale characteristics, reducing the influence of nonlinearity and non-stationarity, and easing prediction; the optimal number of modal components in VMD is determined by calculating the permutation entropy of the decomposition margin; then CNN-LSTM is directly employed to predict each modal component, and mine the multi-scale features in the original data according to the characteristics of different modal components; finally, the predicted values of each component are overlaid to obtain the final result.

The primary contributions of this paper are:

• Complexity and environmental variation cause the Ozone (O_3) data series to have a high degree of variation. The VMD-CNN-LSTM model is used to analyze complex time series data of Ozone (O_3) by decomposing it into multiple sub-signal components in the frequency domain using the Variational Mode Decomposition (VMD) technique. The VMD decomposition results in the original sequence, various VMD components, and residual items. The residual series is further decomposed using the Ensemble Empirical Mode Decomposition (EEMD) technique and combined with the LSTM model for predictive analysis.

In a study comparing different machine learning models using Ozone data from four Chinese provinces, the proposed VMD-CNN-LSTM model outperformed other models. The performance was evaluated using various indicators such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE), indicating the superior predictive capability of the VMD-CNN-LSTM model for Ozone prediction in this particular context.

The paper is structured as follows: Section 2 provides a brief introduction to the hybrid model construction. Section 3 presents results and analysis for Nanjing city. Section 4 provides discussion. The final section draws conclusions.

Materials and Methods

Before we delve into the construction of the VMD-CNN-LSTM hybrid model for predicting changes in ozone (O_3) levels, it's crucial to briefly describe the components of this combined model: Variational Mode Decomposition (VMD) technology and the CNN combined with LSTM neural network.

Variational Mode Decomposition (VMD)

Variational Mode Decomposition (VMD) is an advanced signal processing methodology that has garnered significant interest recently. It's a data-oriented strategy utilized to break down time series data into multiple modal functions or components, each exhibiting distinct frequency behaviors. VMD integrates the principles of variational optimization and mode decomposition, allowing it to effectively identify the fundamental oscillatory components within a given time series. The VMD algorithm commences by iteratively breaking down the input signal into multiple modes through the resolution of an optimization problem. This optimization process seeks the optimal decomposition that reduces both the likeness between adjacent modes and the departure from the original signal. By progressively updating the modes, VMD systematically segregates the signal into different components, each symbolizing a specific oscillatory mode.

One of VMD's primary strengths is its capability to process non-stationary and multi-component time series signals. Unlike conventional decomposition methodologies, such as Fourier or wavelet analysis, VMD adapts to the local traits of the data, offering a more precise representation of the underlying oscillatory components. Additionally, VMD can manage signals with irregular or non-uniform sampling rates, making it an apt fit for a broad spectrum of time series applications.

The decomposed modes obtained from VMD can be individually analyzed or amalgamated to reconstruct the original signal. This versatility enables researchers and practitioners to study individual frequency components, recognize hidden patterns, and execute advanced analyses on the decomposed modes. Additionally, VMD has been effectively merged into various applications, such as signal denoising, anomaly detection, and time series forecasting, highlighting its efficacy in augmenting our understanding and prediction of complex dynamic systems. The process of signal decomposition through VMD is also the solution to the variational constraint problem. The model expression for this variational constraint problem is illustrated in Equation (1).

$$\left\{ \begin{array}{l} \min\limits_{\langle u_k \rangle, \langle w_k \rangle} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\} \\ s.t. \sum_k u_k = f \end{array} \right\} \tag{1}$$

In the given equation, {uk}: = {u1,…,uk} represents the modal component VMF obtained after the decomposition process; {wk}: = {w1,…,wk} are the center frequencies that correspond to each VMF respectively. * denotes the convolution symbol; ∂_t signifies the partial derivative with respect to time t, δ(t) stands for the impulse function; f is the original input signal. By incorporating the Lagrange multiplier λ(t) and the quadratic penalty factor α, the constraint variational problem is transformed into an unconstrained variational problem, which can be represented in the following format:

$$L(\{u_k\},\{w_k\},\lambda) := \alpha \sum_k \parallel \partial_t \left[\left(\delta(t)+\frac{j}{\pi t}\right) u_k(t)\right] e^{-jw_k t} \| e_2^2 + \| f(t) - \sum_k u_k(t) \parallel \Sigma_2^2 + \left(\lambda(t), f(t) - \sum_k u_k(t)\right) \quad (2)$$

The quadratic penalty factor α is utilized in the equation to ensure the precision of signal reconstruction, especially when there is Gaussian noise present. By incorporating the quadratic penalty factor, the variational problem can effectively control the impact of noise and enhance the accuracy of reconstructing the original signal. This regularization term plays a crucial role in balancing the fidelity of the reconstructed signal and the complexity of the decomposition components, ultimately improving the overall quality of the signal reconstruction process. It helps to minimize the impact of noise on the reconstructed signal. Additionally, the Lagrangian operator is utilized to enforce strict constraints on the optimization problem.

To solve the constrained variational problem described in equation (2), an iterative search method called the Alternate Direction Method of Multipliers (ADMM) can be applied. ADMM aims to find the saddle point of the Lagrangian function by iteratively updating the variables. This iterative process allows for the determination of the optimal solution for the variational problem, including the VMFuk (variational mode function) and center frequency. By employing ADMM, the optimization problem can be effectively solved, and the desired solution that satisfies the given constraints can be obtained. This approach enhances the accuracy of the variational problem solution and facilitates the extraction of VMFuk and the corresponding center frequency. The expressions of wk are:

$$\hat{u}_k^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i\neq k} \hat{u}_i(w) + \frac{\hat{\lambda}(w)}{2}}{1+2\alpha(w-ww_k)^2}$$
$$\widehat{w}_k^{n+1} = \frac{\int_0^\infty w|\hat{u}_k(w)|^2 dw}{\int_0^\infty \hat{u}_k|(w)|^2 dw} \quad (3)$$

CNN and Long Short-Term Memory (LSTM) Neural Network

The CNN model comprises two 1-dimensional convolutional layers. The initial convolution layer extracts valuable new insights from the original time series data, enabling the subsequent LSTM model to retain more pertinent information while omitting redundant data. Both convolutional layers feature an identical number of convolution kernels, meaning that the dimensions of their output features are the same. The second convolution layer delves deeper into the relationships among different features, thereby fortifying the connections among the extracted features and enhancing the precision of the forecast results.

The Long Short-Term Memory (LSTM) module, a type of recurrent neural network (RNN), is explicitly designed to manage long-term dependencies and detect sequential patterns in time series data. Its application spans a variety of domains, including natural language processing, speech recognition, and time series forecasting.

The LSTM module comprises multiple memory cells, storing information over time. This enables the network to retain crucial context and effectively handle elongated sequences. Each memory cell consists of three primary components:

Cell State (Ct): The cell state serves as the LSTM module's "memory". It carries information throughout the sequence and selectively retains or discards data via gates.

Input Gate (i): The input gate governs the influx of information into the cell state. It identifies which parts of the input should be stored in the cell state, updating it with new relevant information.

Forget Gate (f): The forget gate decides which parts of the previous cell state's information should be discarded. It regulates how much the previous cell state should influence the current one.

Moreover, the LSTM module includes an output gate (o) that manages the flow of information from the cell state to the module's output. It determines which parts of the cell state should be utilized to generate the module's output. The operations within an LSTM module are executed using various activation functions and weight matrices. These computations involve element-wise operations, such as the sigmoid function (σ) and hyperbolic tangent function (tanh), which infuse non-linearity and allow the LSTM to model intricate relationships in the data.

The LSTM module is designed to alleviate the vanishing gradient issue, which can arise in traditional RNNs when gradients decrease as they propagate through long sequences. By selectively retaining and updating information through the gates, LSTM can learn long-term dependencies more efficiently and prevent the loss of relevant data. The forget gate, denoted as ft, controls the network's memory function and can be expressed as follows:

$$f_t = \sigma(W_f[h_t - 1, X_t] + b_f) \quad (4)$$

where σ represents the sigmoid function, which can be written as:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{5}$$

The input gate i_t is another important gate, and has a similar form to the forget gate f_t:

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i) \tag{6}$$

In the formula, W_i and b_i represent the weight and bias values, but these values are different from the values of the forget gate.

LSTM (Long Short-Term Memory) is a type of recurrent neural network (RNN) architecture that offers several benefits for prediction tasks:

Capturing Long-Term Dependencies: LSTM is specifically designed to handle long-term dependencies in sequential data. It overcomes the vanishing gradient problem of traditional RNNs by using a memory cell that can store and retrieve information over long time lags. This makes LSTM well-suited for capturing complex patterns and dependencies in time series data, allowing it to make accurate predictions.

Handling Variable-Length Sequences: LSTM can efficiently handle variable-length input sequences. It automatically learns to adapt its memory cell to different time steps, making it flexible for predicting sequences of different lengths. This is particularly useful for applications where the length of the input sequence may vary, such as natural language processing or financial time series analysis.

Memory and Context Retention: The memory cell in LSTM allows it to retain information from earlier time steps, enabling it to capture context and dependencies over extended periods. This ability is essential for accurate prediction in time series data, as historical context often plays a crucial role in determining future patterns.

Learning Non-Linear Relationships: LSTM models have non-linear activation functions that enable them to learn complex non-linear relationships in the data. This makes them effective for capturing intricate patterns and trends that may not be easily captured by traditional linear models.

Handling Noisy Data: LSTM models can handle noisy data by learning to filter out irrelevant information and focus on relevant patterns. The memory cell allows the model to selectively retain important features while disregarding noisy or irrelevant inputs, leading to improved prediction accuracy in the presence of noise.

Proposed VMD-CNN-LSTM Model

Ozone behavior is characterized by its non-stationary nature, nonlinearity, and other intricate properties. Predicting ozone levels accurately using a single technique proves to be a complex task. Nonetheless, the technique known as Variational Mode Decomposition (VMD) has shown promise in effectively breaking down complicated ozone signals into numerous simplified modal components. By applying prevalent prediction methodologies to each modal component extracted via VMD decomposition, a notable improvement in prediction accuracy can be observed. Prior studies have mainly concentrated on the modeling of modal components estimated through VMD decomposition, often overlooking the intricate information within the residual terms post-decomposition.

To overcome this issue, a hybrid model of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) neural networks can be deployed. Both CNN and LSTM networks are adept at capturing autocorrelation within time series data and possess the ability to remember long-term patterns. By merging the VMD decomposition method with the

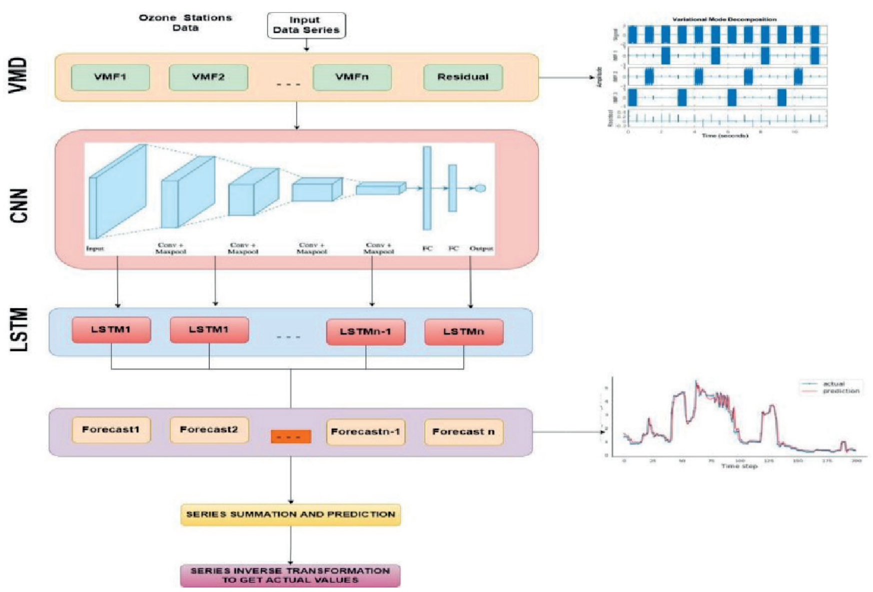

Fig. 1. Hybrid Forecasting Model: VMD-CNN-LSTM for Improved Time Series Prediction.

CNN-LSTM network structure, the prediction accuracy of ozone levels can be substantially amplified. This integrated approach enables separate modeling and prediction of each modal component, harnessing the uniformity extracted by VMD. Moreover, the complex information within the residual terms is processed by the CNN-LSTM network, effectively recognizing temporal dependencies and long-term patterns.

Here's an outline of the comprehensive modeling steps of the proposed methodology:

Step 1: VMD Decomposition The initial ozone sequence is deconstructed using the VMD decomposition technique. This yields individual modal components, referred to as Variational Mode Functions (VMFs). Furthermore, the total VMF data is deducted from the original time series data, resulting in the residual term post-VMD decomposition.

Step 2: Normalization and Data Splitting The disassembled VMF components are normalized, and suitable training and testing samples are chosen. The normalized VMFs then serve as the input to the CNN-LSTM model. The CNN component of the model enhances the series prediction via convolutional operations, and the LSTM module is utilized to predict each Intrinsic Mode Function (IMF) subsequence within the VMFs individually. The predictions for each VMF component subsequence are derived.

Step 3: Residual Term Prediction The residual term obtained from VMD decomposition undergoes further prediction. Initially, CNN is employed to enhance the series prediction by applying convolutional operations to the residual term. Subsequently, LSTM is used to predict each subsequence within the residual term. The prediction outcomes of these subsequence predictions are combined to yield the final prediction result for the residual item.

Step 4: Aggregation of Prediction Results To obtain the final prediction result for the original sequence, the predictions derived from each VMD component and the residual item are aggregated. This involves summing up the predictions of each VMD component and the residual item obtained after VMD decomposition. By overlaying these predictions, the ultimate prediction result for the original sequence is obtained. By adhering to these modeling steps, the proposed methodology aims to achieve accurate ozone prediction by effectively deconstructing the original sequence, individually training each component using CNN-LSTM, and amalgamating the predictions to generate the final result.

Complete flow of the implementation is shown in Fig. 1.

Results

In this part, the areas selected for data collection and the results of the suggested approach's execution will be discussed.

Area of Study

Our research was conducted in Nanjing, a historical yet modern sub-provincial city in China, serving as the capital of Jiangsu Province. The city, as of 2019, is divided into 11 administrative districts and spans a total area of 6,587 square kilometers, comprising 95 streets and 6 towns. Home to roughly 9,314,685 people by the end of 2020, Nanjing pulses with the energy of a thriving urban hub. The climate here is subtropical monsoon with average annual rainfall of 1200 mm, leading to four distinct and unique seasons.

The seasons in Nanjing have their specific attributes. Spring welcomes bright sunny days, monsoons come with plenty of rainfall, summers are generally hot, and autumn is dry with mild temperatures. Winters in the city are notably cold and dry. Each season brings its own allure for tourists, who are also drawn by the city's historical significance and its various cultural and natural attractions, including ancient landmarks and local cuisine.

Data on Ozone

This paper pays particular attention to the daily average datasets from nine ozone monitoring stations in Nanjing. The data covers the period from January 2018 to December 2021 for each station. The assumption is made that the ozone concentration at each station follows a normal distribution. A suite of statistical measures, including minimum and maximum average values for each station, mean, median, and standard deviation (SD) of ozone concentrations, were computed to get a better picture of air pollution levels.

To illustrate the regional variation in air pollution levels, a geographic information system (GIS) tool named ArcGIS (version 10.5) was employed. This software, using spatial data, facilitates the creation of graphical maps. By including the geographical coordinates of the monitoring stations and corresponding ozone concentration values, maps demonstrating the distribution of air pollution levels across the Nanjing region were generated. For the modeling and prediction stages, data from 2018 to 2020 was used to train the prediction model. The remaining data was utilized for testing and validation to evaluate the developed prediction model's accuracy and performance.

Validation Techniques and Comparative Algorithms

To evaluate the effectiveness of the model's predictions, three evaluation indicators have been selected: root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). These indicators will provide valuable insights into the accuracy and performance of the forecasting model. The calculation is as follows:

$$e_{MSE} = \frac{1}{n\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

$$e_{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{8}$$

$$e_{MAPE} = \frac{\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|}{y_i} \tag{9}$$

$$R^2 = \frac{\sum_{i=1}^{k}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{k}(y_i - \bar{y})^2} \tag{10}$$

In the formula, y_i represents the actual value of the station ozone, and $\hat{y}_i$ represents the predicted value of the station ozone. The variable n represents the size of the test sample, and i denotes the sequential number of the test sample point.

To evaluate the benefits of the proposed model, a comparative analysis is conducted using four direct prediction models, namely LSTM, GRU, BiLSTM, and Bi-GRU. In addition, three time series models, including ARIMA, SARIMA, and Prophet, are employed. Furthermore, an ablation study of the proposed approach, EEMD-LSTM, is performed by excluding VMD from the model.

Series Decomposition and Forecasting Results

The Variational Mode Decomposition (VMD) technique is used to break down the original yield sequence into various VMF components and residual elements. Subsequently, the residual component and the series undergo a second round of decomposition using the Ensemble Empirical Mode Decomposition (EEMD). This decomposed data is then integrated with the Long Short-Term Memory (LSTM) model for predictive analysis. To evaluate the proposed method, we contrast it with the EEMD-LSTM model, which combines EEMD with LSTM technology.

To compare and analyze the effectiveness of the various combination models, we take into account the Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2) values. Fig. 2 presents the average outcomes obtained from all the Nanjing stations and provides a comparative analysis of various models. The MAE values for the models show the lowest error for the VMD-CNN-LSTM model, affirming its exceptional prediction accuracy.

Similar patterns are observed for the MSE, MAPE, and R^2 values, with the VMD-CNN-LSTM model consistently outperforming the other models. In the VMD-CNN-LSTM model, the MAE and MSE values are reduced by 46% and 50% respectively, and the MAPE value is reduced by 4% compared to the EEMD-LSTM model. Furthermore, the R^2 value increased by 4%, indicating an excellent match between true and predicted values. The results highlight the exceptional performance of the proposed VMD-CNN-LSTM model in accurately forecasting ozone data, surpassing other models, including EEMD-LSTM.

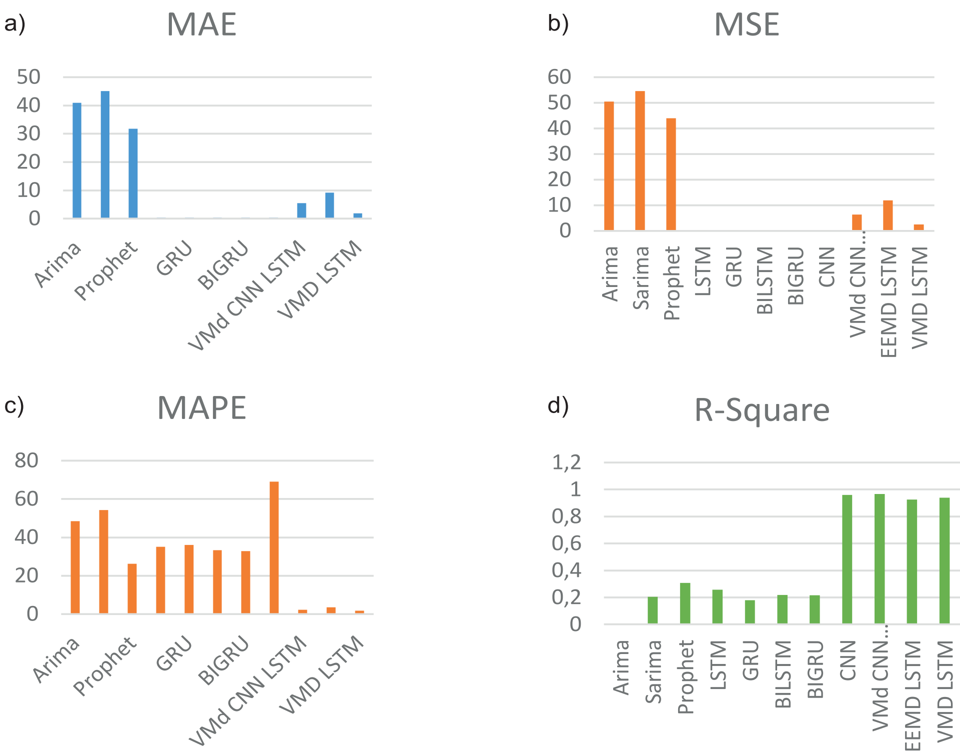

Fig 2. Comparison of different algorithms with proposed model of Nanjing a) MAE, b) MSE, c) MAPE, d) R^2.

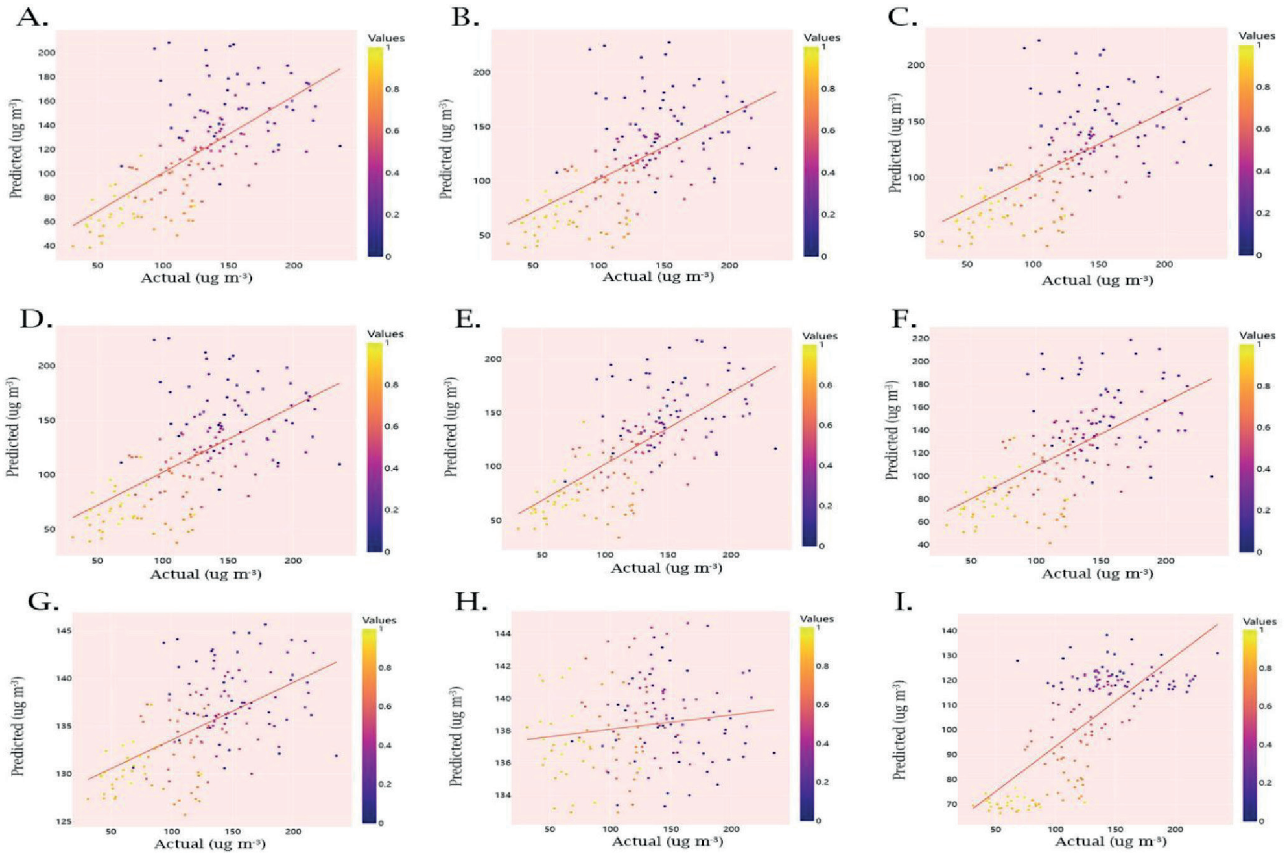

Fig. 3. Comparing Actual and Predicted Values for All Algorithms a) ARIMA, b) SARIMA, c) Prophet, d) LSTM, e) GRU, f) BI-LSTM, g) BI-GRU, h) EEMD-LSTM, and i) VMD-CNN-LSTM.

The value of R^2 is a reliability coefficient between zero and one hundred (or 0 and 1.0). A higher R^2 indicates a more reliable model. Due to the significance of both model stability and flexibility, optimizing R^2 is not the goal. To achieve optimal results in comparing the adjusted R^2 with the original R^2 value, it is desirable for the two numbers to be relatively close. Upon comparing the R^2 values of all prediction models, it is evident from Fig. 3 that the VMD-CNN-LSTM approach yielded the highest value ($R^2 = 0.98$).

Fig. 4 offers a visual comparison of 150 days of observational data, underscoring the areas where our prediction model aligns with the actual values. Given the data's non-linear and dynamic nature, our prediction model's accuracy improves over time, increasingly aligning with the actual figures. The results further show that while LSTM models are adept at memorizing long-term patterns and generally provide accurate predictions, they fall short when it comes to complex Ozone data. Decomposing complex time-series data into subseries with varied frequencies via EEMD enhances prediction accuracy across all stations. When comparing performance across stations, LSTM appeared to underperform compared to GRU when inappropriate settings were used. The accuracy of the predictions was further improved by applying VMD to discern denoising

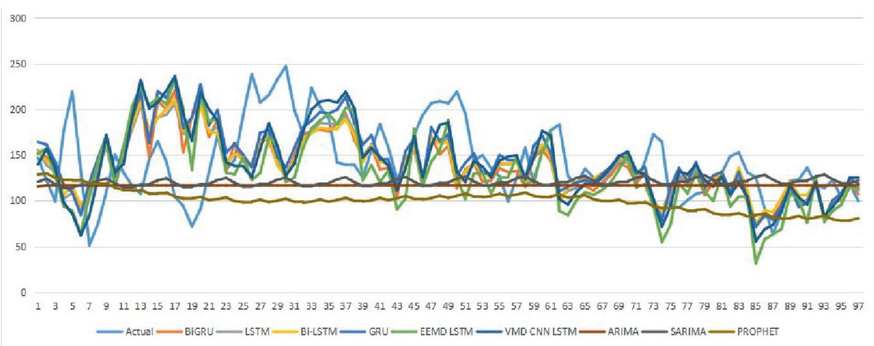

Fig. 4. Visual comparison of 100 days prediction results of all the methods.

patterns within the data for LSTM. The VMD-CNN-LSTM model excelled in predicting short-term Ozone levels, demonstrating its applicability in various other scenarios.

Discussion

This investigation delves into enhancing the accuracy of ozone predictions through the application of decomposition techniques. The complexity of ozone series can hinder the efficiency of direct prediction models. Several decomposition methods, such as Ensemble EMD (EEMD), Empirical Mode Decomposition (EMD), and Variational Mode Decomposition (VMD), have been tried, but their effectiveness is often limited due to modal aliasing and performance inefficiencies.

Variational Modal Decomposition (VMD) has proven to be a promising method for successful decomposition of ozone series, thereby mitigating issues present in previous techniques. After the initial decomposition, the Intrinsic Mode Function (IMF) components that remain complex can be further broken down, leading to a reduction in the complexity of the ozone series. Nevertheless, identifying which IMF components have high complexity is a daunting task.

Research conducted by Wang et al. [54] performed a straightforward decomposition of the IMF components, revealing that the primary component displayed the most complexity. In this research, VMD is used to measure the complexity of each IMF component and set quantitative criteria for choosing complex components. While VMD adeptly resolves modal aliasing and performance inefficiencies, it is crucial to properly predetermine the decomposition level and penalty factor for achieving optimal decomposition results.

Studies by Wu et al. [55] have shown that the use of series decomposition in comprehensive frameworks can greatly improve model predictive performance. For instance, an EEMD-LSTM model applied to pollution data from Anyang demonstrated marked improvements in terms of MAE, RMSE, and MAPE over the LSTM model. Other models, like VMD-SE-LSTM and EEMD-LSTM, also exhibited good prediction performance, prediction stability, and early warning accuracy across various datasets. These observations are consistent with this research, which also found EEMD-LSTM superior to LSTM following series decomposition.

An EMD decomposition model proposed by Huang et al. [56, 57] was utilized to filter noise from air quality data, leading to the extraction of IMF components. Using an EMD-IPSO-LSTM air quality prediction model, each IMF component was then modeled, resulting in improved prediction accuracy and superior model fitting compared to LSTM and EMD-LSTM models. Our research proposes a similar strategy, using EEMD, which also yields improved results compared to LSTM and EEMD-LSTM.

Our study's methodology effectively predicts ozone levels, albeit with some limitations that warrant attention. The available experimental data was inadequate due to constraints in the experimental setup, and further refinement is needed. Additionally, our study lacked data on meteorological factors near the monitoring stations, which could have significantly bolstered our model's performance.

The newly proposed VMD-CNN-LSTM model in this research brings several advantages to the prediction of monthly ozone levels:

- Enhanced Prediction Accuracy: By decomposing historical data series with the VMD technique, the VMD-CNN-LSTM model seeks to improve the prediction accuracy of monthly ozone levels.
- Management of Outliers and Redundant Sites: Ozone prediction often suffers from outliers and redundant data from multi-source pollutant concentration monitoring. The VMD-CNN-LSTM model minimizes these influences by decomposing the data and training each component individually.
- Dealing with Cross-Interference: Cross-interference between different pollutants needs consideration when predicting ozone levels based on multi-source data. This model effectively manages this interference by separately training each decomposed component.
- Long-Term Memory and Adaptive Learning: The LSTM part of the model demonstrates strong adaptive learning and memory capabilities. This allows the model to learn effectively from long-term data and accurately predict ozone levels.
- Comparative Performance: A comparison with other models such as EMD-LSTM, EMD-CNN-LSTM, and VMD-LSTM was carried out. The VMD-CNN-LSTM model outperformed these models in prediction accuracy and had lower error rates, proving its effectiveness in capturing extreme ozone levels.

Conclusions

While soft computing-based prediction models strive to improve accuracy, many often neglect the critical step of data preprocessing. Raw data is often filled with noise and superfluous information, emphasizing the need for effective data preprocessing in prediction models. This study introduces a decomposition algorithm for preprocessing, designed to reduce data dimensionality and extract key features from raw input. Despite significant advancements in natural language processing and computer vision via decomposition algorithms and deep learning, air quality time-series forecasting has seen limited progress in this regard. The primary aim of this paper is to present a novel predictive model that harnesses the benefits of EEMD, VMD, and LSTM. Decomposition algorithms can help disentangle complex air quality data into meaningful components, aiding in the identification of long-term trends, seasonal variations, and irregular patterns. Meanwhile, deep learning techniques, like recurrent neural networks (RNNs) or long short-term memory (LSTM) networks, can capture intricate temporal dependencies in the data, enabling more accurate and timely predictions of air quality levels. This combination of approaches can significantly enhance our ability to forecast air quality, benefiting public health and environmental management. Our analysis of Ozone data from nine stations in Nanjing has led to the following conclusions:

The integration of VMD and EEMD decomposition into various frequency components notably improved ozone prediction accuracy when utilized as input for the LSTM model.

- Based on historical data, LSTM's hidden layer neural units were selected automatically, facilitating predictions of both short-term and long-term trends. Hidden layer neural units in Long Short-Term Memory (LSTM) networks are automatically determined through training on historical data. The network adjusts the number of units in these layers during training to capture the relevant patterns in the data, enabling it to predict both short-term and long-term trends effectively. This dynamic adjustment is a key feature of LSTMs, as it allows the network to adapt to the complexity of the underlying temporal dependencies in the dataset.
- The hybrid VMD-CNN-LSTM model put forth in this study outperforms other models, showing a high degree of fit between real and predicted values. This approach has proven effective in accurately forecasting AQI (Air Quality Index) in practical scenarios.
- Predicting future ozone levels, affected by intricate factors like humidity and weather conditions, poses a challenge. Future research may benefit from integrating the proposed model with these multidimensional complex influencing factors to improve overall forecasting effectiveness.
- Future work could consider exploring optimal ensemble models for the decomposed modes, optimal ensemble model refers to a combination of multiple prediction models, each trained on a specific mode obtained through data decomposition. These ensemble models are designed to leverage the strengths of individual mode-specific models, enhancing predictive accuracy and robustness by effectively blending their forecasts. The goal is to harness the collective predictive power of these models to improve the overall quality of predictions, particularly in scenarios involving multifaceted and complex data patterns, rather than using a simple addition approach. Moreover, developing an intelligent forecasting system and smart decision-making system for ozone monitoring could help in devising appropriate management policies based on forecasted outcomes. However, while the proposed VMD-CNN-LSTM method holds potential in enhancing ozone prediction accuracy, several limitations must be acknowledged:
- Data Availability: The method's efficacy is closely tied to the availability of accurate and comprehensive ozone data. Insufficient or inconsistent data can limit the model's performance and applicability.
- Parameter Selection: Proper selection of parameters like the decomposition level, penalty factor, and network architecture is crucial in the VMD-CNN-LSTM method. Inappropriate parameter selection can lead to inferior results and may necessitate manual tuning. Robust optimization techniques can be used for automated parameter selection. These techniques play a critical role in fine-tuning model parameters to achieve optimal predictive performance. By automating this process, we ensure that our models adapt effectively to the data, resulting in more accurate and reliable predictions.
- Interpretation Complexity: The multistage process of decomposition, prediction, and recombination in the VMD-CNN-LSTM method can complicate interpretation and understanding of inherent data patterns and relationships. Decomposition separates the data into distinct components, prediction models operate on these components individually, and recombination integrates their predictions. This complexity can make it challenging to directly interpret the final prediction in terms of the original data, as the influence of each component on the overall prediction may not be immediately transparent. Therefore, understanding the intricate relationships among these stages becomes crucial for meaningful interpretation.
- Sensitivity to Outliers and Noise: Like any predictive model, the VMD-CNN-LSTM method could be sensitive to data outliers and noise, possibly affecting prediction accuracy and necessitating additional preprocessing or outlier detection techniques.
- Generalization to Other Locations: The method's performance may vary when applied to ozone data from different geographical locations or cities with distinct environmental conditions, potentially requiring custom adjustments for optimal performance.
- Computational Complexity: The implementation of the VMD-CNN-LSTM method could be computationally intensive, particularly with larger datasets or lengthy time series, requiring adequate computational resources and efficient strategies for practical application. By streamlining the computational demands of the modeling process, we can achieve faster predictions and reduce resource requirements, making the approach more feasible for real-time or large-scale applications. This optimization enhances the model's efficiency and usability, which is particularly valuable in operational settings, such as air quality management or environmental monitoring.

– Limited Consideration of External Factors: The VMD-CNN-LSTM method mainly focuses on decomposing the ozone time series and capturing its inherent features, often not considering external factors like meteorological conditions, air pollution sources, or human activities. Including such factors may improve the model's predictive capabilities.

Furthermore, the method suggested in this research paper can be expanded to encompass additional areas of energy prediction, such as forecasting crude oil prices and wind speeds. By refining this approach and broadening its applications, we can enhance predictions and decision-making processes in various sectors, thereby leading to improved management strategies and better resource utilization.

Funding

This research was financially supported by the Deanship of Scientific Research at King Khalid University under research grant number (R.G.P.2/549/44)

Acknowledgments

Thanks for reviewers and editors for providing valuable suggestion during review process.

Conflicts of Interest

The authors declare no conflict of interest.

Disclaimer/Publisher's Note

The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

References

1. FAHIM A., ADDAE B., OFOSU-ADARKWA J., QINGMEI T., BHATTI U. Industry 4.0 and Higher Education: An Evaluation of Barriers Affecting Master's in Business Administration Enrolments Using a Grey Incidence Analysis. IEEE Access. **PP**, 1, **2021**.
2. GALVAN L.P.C., BHATTI U.A., CAMPO C.C., TRUJILLO R.A.S. The Nexus Between CO_2 Emission, Economic Growth, Trade Openness: Evidences From Middle-Income Trap Countries. Frontiers in Environmental Science. **10**, **2022**.
3. YUAN C., LI Q., NIE W., YE C. A depth information-based method to enhance rainfall-induced landslide deformation area identification. Measurement. **219**, 113288, **2023**.
4. QIAN L., ZHENG Y., LI L., MA Y., ZHOU C., ZHANG D. A New Method of Inland Water Ship Trajectory Prediction Based on Long Short-Term Memory Network Optimized by Genetic Algorithm. Applied Sciences. **12** (8), 4073, **2022**.
5. HASNAIN A., HASHMI M.Z., BHATTI U.A., NADEEM B., WEI G., ZHA Y., SHENG Y. Assessment of Air Pollution before, during and after the COVID-19 Pandemic Lockdown in Nanjing, China. Atmosphere. **12** (6), 743, **2021**.
6. BHATTI U., MENGXING H., NEIRA-MOLIN H., MARJAN S., BARYALAI M., HAO T., WU G., BAZAI S. MFFCG – Multi feature fusion for hyperspectral image classification using graph attention network. Expert Systems with Applications. **229**, 120496, **2023**.
7. ZHU W., CHEN J., SUN Q., LI Z., TAN W., WEI Y. Reconstructing of High-Spatial-Resolution Three-Dimensional Electron Density by Ingesting SAR-derived VTEC into IRI model. IEEE Geoscience and Remote Sensing Letters. **19**, 1, **2022**.
8. YANG Y., LIU L., ZHANG P., WU F., WANG Y., XU C., ZHANG L., AN S., KUZYAKOV Y. Large-scale ecosystem carbon stocks and their driving factors across Loess Plateau. Carbon Neutrality. **2** (1), 5, **2023**.
9. HAO Y., GAI Z., YAN G., WU H., IRFAN M. The spatial spillover effect and nonlinear relationship analysis between environmental decentralization, government corruption and air pollution: Evidence from China. Science of The Total Environment. **763**, 144183, **2021**.
10. TIAN H., HUANG N., NIU Z., QIN Y., PEI J., WANG J. Mapping Winter Crops in China with Multi-Source Satellite Imagery and Phenology-Based Algorithm. Remote Sensing. **11**, **2019**.
11. BOTONDI R., LEMBO M., CARBONI C., ERAMO V. The Use of Ozone Technology: An Eco–Friendly Method for the Sanitization of the Dairy Supply Chain. Foods. **12** (5), 987, **2023**.
12. KUROKAWA J., OHARA T. Long-term historical trends in air pollutant emissions in Asia: Regional Emission inventory in ASia (REAS) version 3. Atmospheric Chemistry and Physics. **20** (21), 12761, **2020**.
13. LU S., WANG N., WANG C. Oxidation and biotoxicity assessment of microcystin-LR using different AOPs based on UV, O_3 and H_2O_2. Frontiers of Environmental Science & Engineering. **12**, **2018**.
14. ARORA N. K., FATIMA T., MISHRA I., VERMA M. K., MISHRA J., MISHRA V. Environmental sustainability: challenges and viable solutions. Environmental Sustainability. **1**, 309, **2018**.
15. LI Y., HUANG S., YIN C., SUN G., GE C. Construction and Countermeasure Discussion on Government Performance Evaluation Model of Air Pollution Control: A Case Study from Beijing-Tianjin-Hebei Region. Journal of Cleaner Production. **254**, 120072, **2020**.
16. LIU Q.Y., LI D.Q., TANG X.S., DU W. Predictive Models for Seismic Source Parameters Based on Machine Learning and General Orthogonal Regression Approaches. Bulletin of the Seismological Society of America. **2023**.
17. HASNAIN A., SHENG Y., HASHMI M.Z., BHATTI U.A., HUSSAIN A., HAMEED M., MARJAN S., BAZAI S.U., HOSSAIN M.A., SAHABUDDIN M., WAGAN R.A., ZHA Y. Time Series Analysis and Forecasting of Air Pollutants Based on Prophet Forecasting Model in Jiangsu

Province, China. Frontiers in Environmental Science. **10**, **2022**.

18. BHATTI U.A., YAN Y., ZHOU M., ALI S., HUSSAIN A., QINGSONG H., YU Z., YUAN L. Time Series Analysis and Forecasting of Air Pollution Particulate Matter (PM2.5): An SARIMA and Factor Analysis Approach. IEEE Access. **9**, 41019, **2021**.
19. ZHU S., WANG X., SHI N., LU M. CEEMD-subset-OASVR-GRNN for ozone forecasting: Xiamen and Harbin as cases. Atmospheric Pollution Research. **11**, **2020**.
20. RYU Y.H., HODZIC A., DESCOMBES G., HU M., BARRÉ J. Toward a Better Regional Ozone Forecast Over CONUS Using Rapid Data Assimilation of Clouds and Meteorology in WRF – Chem. Journal of Geophysical Research: Atmospheres. **124** (23), 13576, **2019**.
21. YAFOUZ A., NAJAH A.-M., ZAINI A., EL-SHAFIE A., YAFOUZ A., ZAINI N., AHMED A., EL-SHAFIE A. Ozone Concentration Forecasting Based on Artificial Intelligence Techniques: A Systematic Review. Water, Air, & Soil Pollution: Focus. **232**, **2021**.
22. JIA P., CAO N., YANG S. Real-time hourly ozone prediction system for Yangtze River Delta area using attention based on a sequence to sequence model. Atmospheric Environment. **244**, 117917, **2021**.
23. GONG S., BAI X., LUO G., LI C., WU L., CHEN F., RAN C., XI H., ZHANG S. Climate change has enhanced the positive contribution of rock weathering to the major ions in riverine transport. Global and Planetary Change. **228**, 104203, **2023**.
24. CHENG B., WANG M., ZHAO S., ZHAI Z., ZHU D., CHEN J. Situation-Aware Dynamic Service Coordination in an IoT Environment. IEEE/ACM Trans. Netw. **25** (4), 2082, **2017**.
25. JIN N., ZENG Y., YAN K., JI Z. Multivariate Air Quality Forecasting with Nested LSTM Neural Network. IEEE Transactions on Industrial Informatics. **PP**, 1, **2021**.
26. YAN R., LIAO J., YANG J., SUN W., NONG M., LI F. Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. Expert Systems with Applications. **169**, 114513, **2020**.
27. LIU D.-R., LEE S.-J., HUANG Y., CHIU C.-J. Air pollution forecasting based on attention-based LSTM neural network and ensemble learning. Expert Systems. **37** (3), e12511, **2020**.
28. WANG J., LI H., YANG H., WANG Y. Intelligent multivariable air-quality forecasting system based on feature selection and modified evolving interval type-2 quantum fuzzy neural network. Environmental Pollution. **274**, 116429, **2021**.
29. YIN L., WANG L., HUANG W., LIU S., YANG B., ZHENG W. Spatiotemporal Analysis of Haze in Beijing Based on the Multi-Convolution Model. Atmosphere. **12** (11), 1408, **2021**.
30. YIN L., WANG L., HUANG W., TIAN J., LIU S., YANG B., ZHENG W. Haze Grading Using the Convolutional Neural Networks. Atmosphere. **13** (4), 522, **2022**.
31. WEN L., YUAN X. Forecasting CO_2 emissions in Chinas commercial department, through BP neural network based on random forest and PSO. The Science of the total environment. **718**, 137194, **2020**.
32. FREEMAN B.S., TAYLOR G., GHARABAGHI B., THÉ J. Forecasting air quality time series using deep learning. Journal of the Air & Waste Management Association. **68** (8), 866, **2018**.
33. CABANEROS S.M., CALAUTIT J.K., HUGHES B.R. A review of artificial neural network models for ambient air pollution prediction. Environmental Modelling & Software. **119**, 285, **2019**.
34. YIN L., WANG L., ZHENG W., GE L., TIAN J., LIU Y., YANG B., LIU S. Evaluation of Empirical Atmospheric Models Using Swarm-C Satellite Data. Atmosphere. **13** (2), 294, **2022**.
35. BAN Y., LIU X., YIN Z., LI X., YIN L., ZHENG W. Effect of urbanization on aerosol optical depth over Beijing: Land use and surface temperature analysis. Urban Climate. **2023**, 101655, **2023**.
36. WANG B., KONG W., GUAN H. Air quality forcasting based on gated recurrent long short-term memory model. Association for Computing Machinery, Chengdu, China, **2019**.
37. HUANG Y., XIANG Y., ZHAO R., CHENG Z. Air Quality Prediction using Improved PSO-BP Neural Network. IEEE Access. **PP**, 1, **2020**.
38. YANG K., SHI Y., LUO Y., LIU R., SUN W., SUN M. Assessing spatiotemporal air environment degradation and improvement represented by PM2.5 in China using two-phase hybrid model. Sustainable Cities and Society. **59**, 102180, **2020**.
39. QIAO D., YAO J., ZHANG J.-W., LI X.-L., MI T., ZENG W. Short-term air quality forecasting model based on hybrid RF-IACA-BPNN algorithm. Environmental Science and Pollution Research. **29**, **2022**.
40. YAMASHITA R., NISHIO M., DO R.K.G., TOGASHI K. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. **9** (4), 611, **2018**.
41. QI L., ZHU M., WU L., PAN X., TANG X., WANG Z. Deep Learning for Air Quality Forecasts: a Review. Current Pollution Reports. **6**, 1, **2020**.
42. PAWAR K., JALEM R., TIWARI V. Stock Market Price Prediction Using LSTM RNN: Proceedings of ICETEAS 2018. **2019**.
43. CHEN J., LIU Z., YIN Z., LIU X., LI X., YIN L., ZHENG W. Predict the effect of meteorological factors on haze using BP neural network. Urban Climate. **51**, 101630, **2023**.
44. WANG Y., LIU Y., WANG M., LIU R. LSTM Model Optimization on Stock Price Forecasting. 2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES). 173, **2018**.
45. YANG D., CHEN K., YANG M., ZHAO X. Urban rail transit passenger flow forecast based on LSTM with enhanced long-term features. IET Intelligent Transport Systems. **13**, **2019**.
46. ZHANG X., ZHANG Y., LU X., BAI L., CHEN L., TAO J., WANG Z., ZHU L. Estimation of Lower-Stratosphere-to-Troposphere Ozone Profile Using Long Short-Term Memory (LSTM). Remote Sensing. **13** (7), 1374, **2021**.
47. PAK U., KIM C., RYU U., SOK K., PAK S.-N. A hybrid model based on convolutional neural networks and long short-term memory for ozone concentration prediction. Air Quality, Atmosphere & Health. **11**, 883, **2018**.
48. DENG T., MANDERS A., SEGERS A.J., BAI Y., LIN H. Temporal Transfer Learning for Ozone Prediction based on CNN-LSTM Model. **2021**.
49. ESLAMI E., CHOI Y., LOPS Y., SAYEED A. A real-time hourly ozone prediction system using deep convolutional neural network. **2019**.
50. ZHANG Z., CHEN H., HUANG X. Prediction of Air Quality Combining Wavelet Transform, DCCA Correlation

Analysis and LSTM Model. Applied Sciences. **13** (5), 2796, **2023**.
51. JIANG W., ZHU G., SHEN Y., XIE Q., JI M., YU Y. An Empirical Mode Decomposition Fuzzy Forecast Model for Air Quality. Entropy. **24** (12), 1803, **2022**.
52. LI G., TANG Y., YANG H. A new hybrid prediction model of air quality index based on secondary decomposition and improved kernel extreme learning machine. Chemosphere. **305**, 135348, **2022**.
53. LI Y., WANG S., WEI Y., ZHU Q. A New Hybrid VMD-ICSS-BiGRU Approach for Gold Futures Price Forecasting and Algorithmic Trading. IEEE Transactions on Computational Social Systems. **PP**, 1, **2021**.
54. WANG S., HAO J. Air quality management in China: Issues, challenges, and options. Journal of environmental sciences (China). **24**, 2, **2012**.
55. WU H., TANG X., WANG Z., WU L., LU M., WEI L., ZHU J. Probabilistic Automatic Outlier Detection for Surface Air Quality Measurements from the China National Environmental Monitoring Network. Advances in Atmospheric Sciences. **35** (12), 1522, **2018**.
56. HUANG L., HU J., CHEN M., ZHANG H. Impacts of power generation on air quality in China – part I: An overview. Resources, Conservation and Recycling. **121**, **2016**.
57. MA S., QIU H., YANG D., WANG J., ZHU Y., TANG B., SUN K., CAO M. Surface multi-hazard effect of underground coal mining. Landslides. **20** (1), 39, **2023**.