



ACBiGRU-DAO: Attention Convolutional Bidirectional Gated Recurrent Unit-based Dynamic Arithmetic Optimization for Air Quality Prediction

Vinoth Panneerselvam¹ · Revathi Thiagarajan²

Received: 6 December 2022 / Accepted: 28 May 2023 / Published online: 6 July 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Over the past decades, air pollution has turned out to be a major cause of environmental degradation and health effects, particularly in developing countries like India. Various measures are taken by scholars and governments to control or mitigate air pollution. The air quality prediction model triggers an alarm when the quality of air changes to hazardous or when the pollutant concentration surpasses the defined limit. Accurate air quality assessment becomes an indispensable step in many urban and industrial areas to monitor and preserve the quality of air. To accomplish this goal, this paper proposes a novel Attention Convolutional Bidirectional Gated Recurrent Unit based Dynamic Arithmetic Optimization (ACBiGRU-DAO) approach. The Attention Convolutional Bidirectional Gated Recurrent Unit (ACBiGRU) model is determined in which the fine-tuning parameters are used to enhance the proposed method by Dynamic Arithmetic Optimization (DAO) algorithm. The air quality data of India was acquired from the Kaggle website. From the dataset, the most-influencing features such as Air Quality Index (AQI), particulate matter namely $PM_{2.5}$ and PM_{10} , carbon monoxide (CO) concentration, nitrogen dioxide (NO_2) concentration, sulfur dioxide (SO_2) concentration, and ozone (O_3) concentration are taken as input data. Initially, they are preprocessed through two different pipelines namely imputation of missing values and data transformation. Finally, the proposed ACBiGRU-DAO approach predicts air quality and classifies based on their severities into six AQI stages. The efficiency of the proposed ACBiGRU-DAO approach is examined using diverse evaluation indicators namely Accuracy, Maximum Prediction Error (MPE), Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), and Correlation Coefficient (CC). The simulation result inherits that the proposed ACBiGRU-DAO approach achieves a greater percentage of accuracy of about 95.34% than other compared methods.

Keywords Air Quality · Prediction · Attention Convolutional Bidirectional Gated Recurrent Unit · Dynamic Arithmetic Optimization Algorithm · Accuracy

Introduction

Air pollution is a major problem human beings face every day all over the world. The main common factor of air pollution is urbanization and industrialization. At present,

this process is continuing and this is growing globally. The World Health Organization expresses that 9/10 of the people are undergoing this air pollution. $PM_{2.5}$, PM_{10} , CO, O_3 , and SO_2 are the frequent air pollutants that cause haze, soil acidification, and fog. Air pollution can lead to various health issues such as heart attacks and lung diseases. To enhance the accuracy of predictions in different scenarios, Ma et al. (2019) proposed integrating transfer learning techniques and a Bi-directional Long Short-Term Memory (BiLSTM) neural network as a potential solution to overcome these drawbacks. The expansion of industrialization and urbanization has resulted in the emergence of air pollution, leading to health complications as noted by Bekkar et al. (2021). Consequently, the well-being of developed nations has been adversely affected by the burden of pollution, as highlighted by Liu et al. (2021). The association of morbidity and mortality

Responsible Editor: Marcus Schulz

✉ Vinoth Panneerselvam
vinoth.ttk@gmail.com; vinoth@mepcoeng.ac.in
Revathi Thiagarajan
revathi@mepcoeng.ac.in

¹ Department of Computer Science and Engineering, Mepco Schlenk Engineering College, Sivakasi, India

² Department of Information Technology, Mepco Schlenk Engineering College, Sivakasi, India

with a mass of pollutants in the air and the estimation of air quality is done earlier (Tiwari et al. 2021). A dimensionless indicator, Air Quality Index (AQI), numerically expresses the consequences and remarks of the air quality. AQI grows the air quality management and measures the sulfur dioxide (SO_2), nitrogen dioxide (NO_2), airborne particles ($\text{PM}_{2.5}$, PM_{10}), and ground-level ozone (O_3) for the welfare of human and public health. Adaptation of bidirectional LSTM (Bi-LSTM) has forward and backward data for the good performance of RNN and regression of $\text{PM}_{2.5}$ prediction. The sequence model of time series air quality helps in analyzing the exploration of air quality forecasting by the BI-LSTM model (Zhang et al. 2020). Long and short-term is a powerful tool that helps in predicting air pollution predictions and power grid operations (Dairi et al. 2020). The photochemical process creates tropospheric ozone O_3 pollutants that influenced structural materials, human health, and plants. This O_3 pollutant shows negative consequences in industries and even in air managers (Guo et al. 2023a).

To measure air quality, it is necessary to gather time series data in a specific order, and often O_3 is included as a parameter in this process, as noted by Freeman et al. (2018). Poor environmental air quality has detrimental effects on human health. Thus, taking steps to prohibit environmental air pollution is crucial for the continuous development of a healthy nation. Air quality modeling and air quality prevention are substantial steps for preventing the quality of air. PM_{10} is the most dangerous pollutant which affects children and aged ones (Samal et al. 2020). Many air pollution components affect $\text{PM}_{2.5}$, and the information on air quality accumulated in the experiment contains PM_{10} , SO_2 , NO_x , $\text{PM}_{2.5}$, NO_2 , CO , O_3 , and NO . In this LSTM is employed for training the machine which develops the gradient explosion and gradient disappearance experienced by RNN. The forecasting result is combined with a warning system when the forecasting accuracy $\text{PM}_{2.5}$ develops considerably. Kristiani et al. (2022) suggest that among various methods, LSTM is the most suitable model for managing time series datasets. On the other hand, CNN has distinct advantages in feature capture. Also, the authors propose utilizing LSTM deep learning methods to implement comparison and optimization methods in machine learning. This paper develops a novel prediction model to predict the quality of air very accurately. The key contributions of this paper are described as follows:

- A novel air quality prediction model, Attention Convolutional Bidirectional Gated Recurrent Unit based Dynamic Arithmetic Optimization (ACBiGRU-DAO), is proposed to accurately predict the short-term effect of air quality in one health and classify the health risk into different types such as moderate, risky for sensitive people, unhealthy, very unhealthy, and hazardous.

- The elimination of irrelevant features and misclassification issues is overcome by the Attention Convolution Bidirectional Gated Recurrent Unit (ACBiGRU) method while performing complex tasks.
- The Dynamic Arithmetic Optimization (DAO) approach is integrated with ACBiGRU to enhance the prediction model and also tune the hyperparameters of the architecture to minimize the error rate for severity analysis.
- The air quality data obtained from the Kaggle dataset is preprocessed through the imputation of missing values using the adaptive sliding window technique and data transformation pipeline.

The section left in this article is arranged as follows: various literature works based on AQI are reviewed in the “Literature Survey” section. The background techniques are elaborated on in the “Background” section. The proposed methodology for predicting the Air quality index is presented in the “Proposed Methodology” section. The experimental results and analysis are discussed in the “Experimental Results and Analysis” section. Finally, the conclusion of the article and its future directions are discussed in the “Conclusion” section.

Literature Survey

Benhaddi and Ouarzazi (2021) illustrated a multivariate time series (MTS) forecasting using WaveNet-temporal-convolutional neural network (WTCNN) method. The WTCNN model was compared with long short-term memory (LSTM) and gated recurrent unit (GRU) for identifying its memory consumption, flexibility, stability capabilities, flow control, structure, and robustness. For experimental purposes, Marrakesh city of Morocco was selected for urban air quality prediction by using multivariate time series datasets and six types of real-world multi-sensors. The WTCNN method outperformed compared to other state-of-art methods. But multi-step time series forecasting is complex compared to single-step forecasting. Xu and Yoneda (2019) illustrated a long short-term memory auto-encoder multitask learning (LSTM-AML) model for predicting $\text{PM}_{2.5}$ time series. Guo and He (2021) employed the Beijing dataset for monitoring the air quality concerning wind speed, temperature, wind direction, weather conditions, humidity, and pressure. The performance metrics like root mean square error (RMSE), mean absolute error (MAE), and symmetric mean absolute percentage error (SMAPE) were evaluated to attain a better performance rate. The information about gas emissions and economic factors was not included in the prediction of the $\text{PM}_{2.5}$ time series.

Lin et al. (2020) discussed the time series prediction by air quality prediction system-based neuro-fuzzy modelling (AQPS-NFM). The neuro-fuzzy model is trained by Steepest Descent Backpropagation (SDB). The Taiwan air quality hourly datasets were utilized by using a few parameters like

wind speed, temperature, humidity, and wind direction. The result showed that the attained fuzzy rules have high quality, and the parameters were optimized very effectively. Ma et al. (2019) described the prediction of air quality based on transfer learning-based bi-directional long short-term memory (TL-BiLSTM). The developed method was used to enhance the prediction accuracy of air pollutants at larger resolutions. Various machine learning methods were compared to the TL-BiLSTM model in performance evaluation. The result showed that the TL-BiLSTM model has small error values and large temporal resolutions.

The prediction of AQI can be challenging due to the hindrances caused by component inter-correlation and volatile AQI patterns. To overcome these issues, Jin et al. (2021) introduced the multi-task multi-channel nested long short-term memory (MTMC-NLSTM) network. The authors implemented the root means squared propagation (RMSprop) optimizer to minimize the loss function of mean square error (MSE). The MTMC-NLSTM approach decomposes data into high- and low-frequency parts to facilitate effective learning. However, the effectiveness of this method depends on the quality of the preprocessing module. Ge et al. (2021) presented multi-scale spatiotemporal graph convolution network (MST-GCN) to promote strong spatial correlation and design long-term temporal dependencies in forecasting air quality. Based on domain categories, the trends and features of air quality were partitioned into numerous groups and encoded the correlation across regions utilizing distance and similarity graphs. Then, it used a fusion block tensor to combine the separated groups. The drawback of this technique was the high computation cost produced during exploring dynamic spatial correlation.

Zhang et al. (2021) elaborated to overcome the difficulty obtained while predicting the unstable variations of air quality by the variational mode decomposition-bidirectional long short-term memory (VMD-BiLSTM) method. Here, the established method disintegrates actual time series data into several components based on frequency domain. The BiLSTM model helps to capture extensive feature details and thus enhances prediction accuracy but requires more training time. To enhance the air quality prediction performance, Mao et al. (2021) developed a temporal sliding long short-term memory extended (TS-LSTME) approach. Based on the circumstance, the TS-LSTME approach selects appropriate influencing factors to forecast the targeted variables. The $PM_{2.5}$ concentration in the air was measured along with the utilization of temporal and meteorological data. The underestimation and overestimation results of the TS-LSTME approach affect air quality prediction accuracy.

Zhao et al. (2019) illustrated the determination of air quality by spatiotemporal collaborative convolutional neural network combined with a long-short term memory (STCNN-LSTM). At first, regional air quality prediction

problems helped with the analysis of closely connected areas and the prediction of several locations. Next, the utilization of time-sliding windows, correlation analysis of factors, and clustering algorithm supported in constructing Relevance data cube (3-dimensional data structure). At last, STCNN-LSTM was constructed with the combination of CNN and LSTM by dealing with other neural networks. Therefore, this model was efficient and appropriate for the air quality field and predicted accurately. But, this method led to emitting pollution and caused dispersion.

Zhang et al. (2019) elaborated the LightGBM model to predict $PM_{2.5}$ concentration and to overcome the issue of high-dimensional large-scale data. A sliding window mechanism was applied for increasing training dimensions. Hence, this method was appropriate for excavating the characteristics with a powerful connection. But, this model's performance accuracy was limited. Guo et al. (2023b) studied the climate change has a significant impact on air quality and human health, emphasizing the need for accurate predictions. Their results indicate that air temperature in the Sahara region will experience a sustained increase throughout the twenty-first century. These predictions hold crucial implications for climate policy-making and underscore the urgent requirement for proactive measures to address the multifaceted challenges posed by climate change.

Ma et al. (2020) established a prediction of air quality using transfer learning-based stacked bidirectional long short-term memory (TLS-BLSTM). The main subjective of this work is to control air pollution and minimize its impacts on human beings. Using various machine learning methods to provide an accurate prediction of air quality meanwhile, it requires several data to train the model. The experimental analyses were conducted and evaluated that TLS-BLSTM diminished the average RMSE rate upto 35.21% which improved the performance. On the other hand, it is easy to overfill, and dropout is much harder to implement.

Maleki et al. (2019) illustrated a prediction of air pollution by using an artificial neural network (ANN) model. To mitigate the adverse health effects of air pollution, forecasting air quality has become crucial. Decision-makers and practitioners in urban air quality rely on ANNs to estimate the spatiotemporal distribution of air pollutants and indices of air quality. However, after the ANN training, the data may produce output even with incomplete information.

Wang and Song (2018) discussed the deep spatial ensemble design for predicting air quality. The data-driven method utilized the meteorological data and historical air quality for predicting the air quality. The deep LSTM was utilized for learning the short-term and long-term dependencies of the air quality. The experimental results showed that the scheme improved the predicting accuracy. In order to improve the prediction accuracy, some additional information such as overseas air pollution, climate satellite information, and microsensor information was required.

Zou et al. (2021) discussed the air quality prediction for the spatiotemporal attention process. The predicted air quality area gathers the influences of relative sites by spatial attention process. The experimentation results showed that the scheme achieved higher performance when compared to other methods. This spatiotemporal attention process was only used to predict single-city air pollution.

Li et al. (2019) established a modified least squares support vector machine–based multi-objective multi-verse optimization (LSSVM-MOMVO) algorithm for addressing air quality monitoring issues. The data pre-processing for forecasting AQI was carried out by integrating the decomposition method with the feature selection model, and the MOMVO method was employed to eliminate noise and choose optimal input structures. The LSSVM-MOMVO algorithm was then utilized for AQI forecasting, which exhibited strong stability and high accuracy, as indicated by comparative analysis. Moreover, it was found to be a reliable and efficient approach for air quality monitoring. However, it is worth noting that the LSSVM-MOMVO algorithm has limitations in solving optimization problems.

Zhu et al. (2018) developed a machine learning approach for air quality forecasting, which utilized an optimization algorithm to solve the problem of multi-task learning (MTL) and improve convergence speed. The results showed improved performance in predicting air pollution; however, nearby meteorology stations needed to exhibit similar characteristics to enhance the prediction performance.

While numerous machine learning and deep learning techniques have been established for predicting and mitigating air pollution, these methods have some limitations such as high computational cost, poor prediction performance, minimum accuracy, and high error rate. To overcome these limitations, the present study proposes a novel approach called attention convolutional bidirectional gated recurrent unit based dynamic arithmetic optimization (ACBiGRU-DAO).

Background

This section provides a comprehensive description of the proposed ACBiGRU-based DAO approach. The proposed approach is formed by the integration of Dynamic Arithmetic Optimization (DAO) with an Attention Convolutional Bidirectional Gated Recurrent Unit (ACBiGRU) model. The detailed elaboration of the technique is listed as follows.

Attention Convolutional Bidirectional Gated Recurrent Unit (ACBiGRU)

Various deep learning models possess their own merits and demerits for predicting air quality. A convolution neural network (CNN) provides quick and accurate results in extracting

features (Liu et al. 2019). The BiGRU learns data sequences in frontward and backward modes. The attention strategy efficiently computes the significant features of the data. But when extracting key features, the CNN eliminates the data features corresponding to positions that make the network face misclassification results while handling complex tasks. This complexity is avoided by the enhanced feature prediction activities of BiGRU and attention mechanism, making the convolution network more capable to extract main data features. So, considering all these advantages, the ACBiGRU model is formed. The AQI, PM_{2.5}, PM₁₀, CO, NO₂, SO₂, and O₃ are considered inputs in which each layer is determined by the BiGRU layer. Each layer generates an individual output based on air quality prediction. The input layer is estimated by a hidden layer in both forward and backward directions which is performed by BiGRU and is directed to the attention layer which emphasized each output layer for predicting and validating the control of air pollution.

In the input layer, the most significant data features used to analyze and predict air quality are inputted into the framework. In the BiGRU layer, the inputted data patterns are learned in both frontward and backward directions. A gated recurrent unit is a modified form of recurrent neural network that helps to design the data sequences. However, its vanishing gradient problem creates certain issues in analyzing data sequences. Therefore, the BiGRU model is introduced that learns the data patterns in a two-way direction and initializes weights to resolve the gradient issues. Moreover, the GRU measures all the passing data features and outputs a vector with a constant dimension and its architecture is shown in Fig. 1a. The major operations performed by GRU are as follows: (i) using the reset gate, the GRU decides which data to be discarded from the preceding moment; (ii) using the update gate, GRU chooses the data to be uploaded from the current moment. Fig. 1b. shows the block diagram of ACBiGRU model. The numerical expression for the functioning of the reset gate is given by

$$S_t = \gamma(U_r R_t + W_r I_{t-1} + C_r) \tag{1}$$

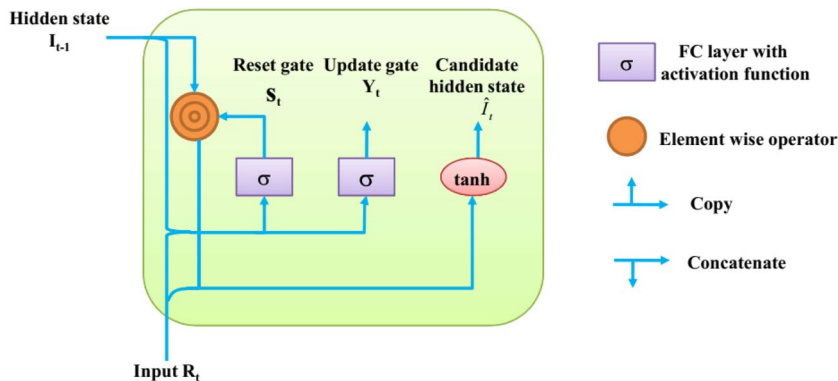
From the above equation, I_{t-1} represents the candidate activation vector, R_t is the input vector, U_r and W_r represents the weight information, C_r represents bias, and S_t represents the reset gate. The numerical expression for the functioning of the update gate is given by

$$Y_t = \gamma(U_z R_t + W_z I_{t-1} + C_z) \tag{2}$$

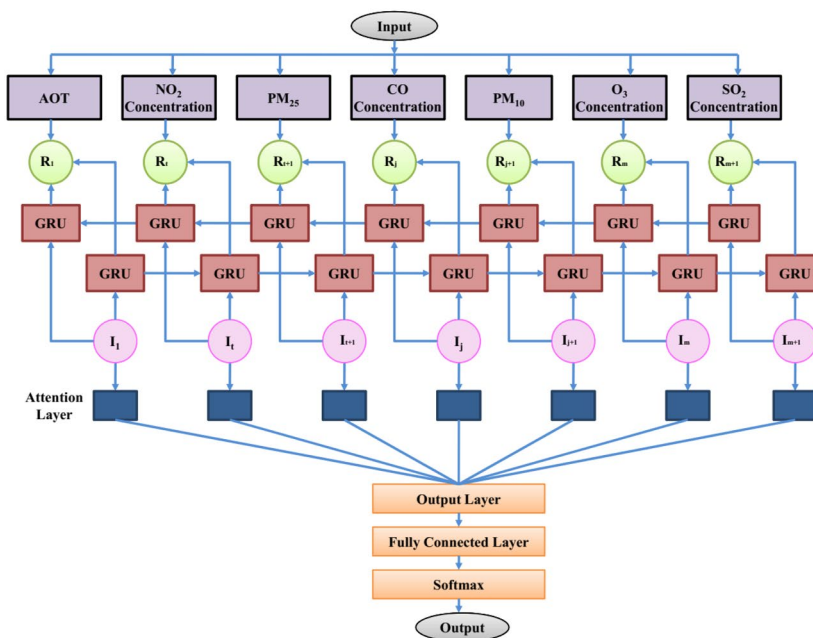
From the above equation, GRU computes the candidate memory content, which is the main step for calculating the current moment output, Y_t denotes the update gate, C_z represents bias, and U_z and W_z denote weight information.

$$\hat{I}_t = \tanh(U_r I_t + W_r I_{t-1} + C) \tag{3}$$

Fig. 1 (a) GRU Architecture and (b) ACBiGRU Architecture



(a) GRU architecture



(b) Block diagram of ACBiGRU model

From the above equation, C represents bias, U and W represents the weight of information.

$$I_t = (1 - Y_t)\hat{I}_t + Y_t I_{t-1} \tag{4}$$

The output I_t obtained through the BiGRU network is allotted with weights by attention strategy in the attention layer. The attention strategy is utilized to emphasize the effects related to data patterns. The weight assigned to each feature vector is represented by a value, with higher weights indicating more significant feature vectors. Therefore, more attention is given to feature vectors with higher weights, as these play a crucial role in enhancing the air quality prediction efficiency of the model. The mathematical formulation of the attention mechanism is modeled as

$$B_t = \sum_{t=1}^{t_h} b_{pq} I_t \tag{5}$$

$$b_{ij} = \text{soft} (U_{b2} \tanh (U_{b1} I_t))_{\max} \tag{6}$$

where b_{pq} represents attention weight's size, I_t represents the length of sequence data; U_{b1} and U_{b2} are weights. To select very deep key features in air quality data, different dimensioned convolution kernels are utilized. This work utilized three diverse convolution kernels such as the pooling layer, dropout layer, and fully connected layer for the extraction of significant features. The intermediary semantic data measured using the attention mechanism is provided as input to the convolution network. The term $B_{p:q}$ represents the splicing of p^{th} to the feature vector.

$$B_{p:q} = B_p \oplus B_{p+1} \oplus \dots \oplus B_j \tag{7}$$

$$t_p = g(UB_{p:p+n-1} + c) \tag{8}$$

From the above equation, g represents the hyperbolic tangent function, U represents weight data, n is the breadth of weight data, and c represents bias. By means of maximum pooling, the deep key feature vectors are extracted. At last, the prediction results obtained by the pooling layer are connected concurrently as an output of the convolution layer. The computations are made as follows:

$$t_q = [t_1, t_2, \dots \dots t_{m-1}, t_m] \tag{9}$$

$$\hat{t} = \max\{t_q\} \tag{10}$$

$$\hat{t} = [\hat{t}_1, \hat{t}_2, \hat{t}_3] \tag{11}$$

The term q represents the number of convolution kernels and m implies the number of convolution results. The output obtained by the convolutional layer is fed into a fully connected layer and then to the SoftMax layer to measure the probability of classified output.

$$R_v = U_D \hat{t} + D_C \tag{12}$$

$$S_i = \frac{\exp(S_i)}{\sum k \exp(S_i)} \tag{13}$$

From the above equation, U_D represent the weight of the network, D_C which is biased. The block diagram of the ACBiGRU model is portrayed in Fig. 1 b.

Dynamic Arithmetic Optimization (DAO) Algorithm

Khodadadi et al. (2022) proposed the development of a function by integrating a new accelerator with two dynamic features in the fundamental arithmetic optimization. The investigation and candidate solution phases are changed during the optimization process due to the dynamic version's characteristics of investigation and utilization. The DAO algorithm reduces the effort required to improve parameters using metaheuristic, which is a significant advantage of the algorithm.

Dynamic Accelerated Function

During the search period, Dynamic Accelerated Function (DAF) plays a vital role as a unique function. Balancing the maximum and minimum basic variables are essential for the arithmetic optimization algorithm. Additionally, achieving better internal parameters in this algorithm is influenced by the new descending functions that are obtained from DAO. The numerical expression of the modification component in this algorithm is given below.

$$D_a \text{function} = \left(\frac{I_{\max}}{I} \right)^\omega \tag{14}$$

From the above equation, the maximum and current iterations are indicated by I_{\max} and I , ω representing the constant value. While repeating every algorithm the performance is reduced.

Dynamic Candidate Solution

Investigation and utilization are the important two stages of metaheuristic algorithms which have the quality of balancing between these two stages. For accentuating the investigation and utilization, each solution needs to be upgraded during the processing of optimization from a good, attained solution. The function of the dynamic candidate solution is given in mathematical expression.

$$P_{x,y}(C_1 + 1) \begin{cases} B(p_y) \div (D_{cs} + \mu) \times ((UL_y - LL_y) \times \gamma + LL_y), \\ S_2 < 0.5 \\ B(p_y) \times D_{cs} \times ((UL_y - LL_y) \times \gamma + LL_y), \\ \text{otherwise} \end{cases} \tag{15}$$

$$P_{x,y}(C_1 + 1) = \begin{cases} B(p_y) - (D_{cs} \times ((UL_y - LL_y) \times \gamma + LL_y), \\ S_3 < 0.5 \\ B(p_y) + D_{cs} \times ((UL_y - LL_y) \times \gamma + LL_y), \\ \text{otherwise} \end{cases} \tag{16}$$

From the above equation, D_{cs} represents a dynamic candidate solution, $P_{x,y}$ represents x^{th} position of the y^{th} position, $B(p_y)$ denotes y^{th} position in the best-obtained phase, UL_y denotes the upper limit value of the y^{th} position, LL_y denotes the lower limit of the y^{th} position, γ is a controlling parameter, μ is a small number to avoid division by 0, S_2 and S_3 denotes random numbers in (0,1). The performance of D_{cs} is proposed because of the reduction in the percentage of candidate solutions while iterating and the value is decreased.

$$D_{cs}(0) = 1 - \sqrt{\frac{I}{I_{\max}}} \tag{17}$$

$$D_{cs}(n + 1) = D_{cs}(n) \times 0.99 \tag{18}$$

The speed of the AO algorithm is raised by utilizing DAO algorithm's candidate solutions which are explored in iterations and investigations. This Algorithm can function without parameters that are supported for metaheuristic algorithms (Guo et al. 2020). The leftover approach is similar to the AO algorithm while DAO algorithm applies dynamic functions is the dissimilarity between AO algorithm and DAO algorithm. Several parameters need to be adjusted because the

DAO algorithm receives an advantage from it. But it is always against rival algorithms due to the demand for parameter adjustment for various issues. The modified mechanism-based iteration opposition and non-fitness improvement are the main disadvantages of this algorithm.

Algorithmic Steps of ACBiGRU-based DAO Approach

Figure 2 illustrates the flow diagram of the proposed ACBiGRU-DAO approach. To gain better prediction results, the data learning process of the ACBiGRU model is enhanced by modifying or setting optimal parameter values using the DAO algorithm, and this modified version of the technique is formally represented as ACBiGRU-DAO approach. The algorithmic steps are given as follows:

Step 1: The BiGRU layer, attention layer, and convolution layer are the three significant layers determined in the ACBiGRU model.

Step 2: At the initial BiGRU layer, the input data patterns are learned in both forward and backward directions and also it makes use of both future and historical information to analyze the long-term data representations.

Step 3: Then, the effects associated with the data patterns are highlighted using the attention strategy. In addition, in this attention layer, the output obtained from the BiGRU layer is assigned with weights.

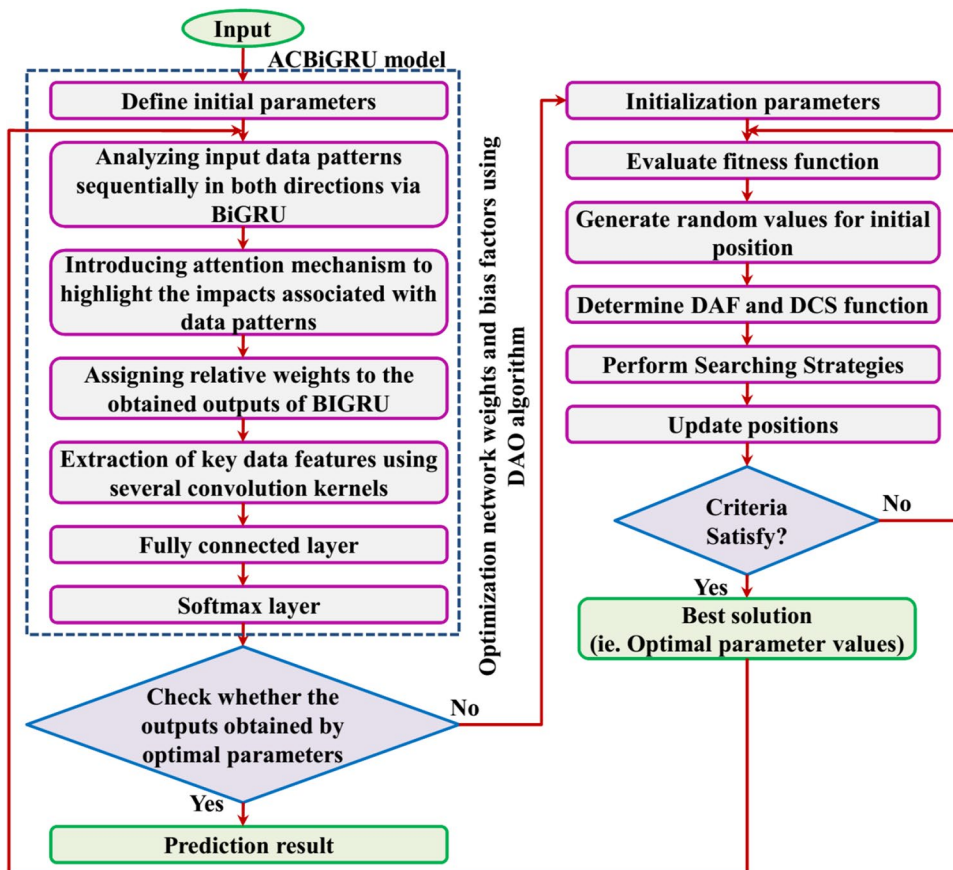
Step 4: The higher value of weight represents the significant feature vectors, so the attention weight is given more consideration else it will influence the air quality prediction efficiency of the model.

Step 5: Moreover, the deep key features in the data are extracted through convolution kernels. The randomly selected parameter value at the initial iteration does not provide more accuracy in the prediction process.

Step 6: Therefore, to attain a greater prediction accuracy rate, the parameters of the ACBiGRU model namely network weights and bias factors are tuned using the DAO algorithm.

Step 7: This algorithm incorporates dynamic accelerated function and dynamic candidate solution strategies to enhance the standard performance of the arithmetic optimization algorithm.

Fig. 2 Workflow of the proposed ACBiGRU-DAO approach



Step 8: The DAO algorithm not only selects the best solution (i.e., optimal parameter values) for the model but also accelerates the convergence ability of the model.

Step 9: Thus, the proposed ACBiGRU-DAO approach effectively predicts the data features and classifies them based on the severity levels into different stages.

Proposed Methodology

The air quality assessment becomes an indispensable step in many urban and industrial areas to monitor and preserve the quality of air. The increasing concentration of pollution in the environment creates negative impacts on the quality of human life. It requires an air quality monitoring system to predict the quality of air by gathering details of pollutant concentrations. To fulfill the requirement, this paper proposes a novel automated prediction model ACBiGRU-DAO for air quality prediction and data is taken from the Kaggle dataset. To enhance the results, the data utilized is initially preprocessed through the steps namely imputation of missing values and data transformation. The preprocessed

outcome is then fed into the proposed model that accurately predicts the air quality and classifies them based on severity into 6 AQI stages. Figure 3 shows the structure of the air quality prediction model.

Experimental Data Source

The dataset utilized in this work is taken from two Kaggle websites namely India Air Quality data (62.54 MB) and Air Quality data in India (2015–2020) (2.57 MB). Each of these datasets contains AQI ranges calculated on a yearly and daily basis respectively from various stations across India.

Data Preprocessing Phase

Data preprocessing contains two processes: imputation of missing values and data transformation (Huang et al. 2022).

Imputation of missing values

Imputation refers to the process of replacing missing data with substitute values, while preserving the most important information of the dataset. In regression methods, lagged variables are estimated using robust estimates to reduce negative bias coefficients generated during the regression process. Lagged variables

Fig. 3 Structure of air quality prediction model



enable the detection of possible associations between current and past variables in sequential time, and this process is also known as the adaptive sliding window technique.

Adaptive sliding window The adaptive sliding window extended the downfall determination under the sample downfall point. The outliers are organized in ascending order throughout the problem of window sample estimation. It is more powerful in interrupting the estimator rather than gradually being placed amidst data (Guo et al. 2021). These can be done by sequencing the outliers at the due time in modification with the window coefficient.

- The sequential monitors are divided into step lengths and window breadth for segmenting the series into mutually autonomous windows.
- The normalized interquartile range (NIQR) and median in every window are approximated as the standard divergence of the window sample and mean.
- Several windows with apparent outliers are chosen for computing the mean value of $t(y)$ and the threefold standard deviation as d_o .

$$c(\epsilon, y, t) = s_{up} |t(y') - t(y)| \tag{19}$$

From the above equation, the supremum is denoted by s_{up} , the corrupted samples are indicated by ϵ , y' , $t(y)$ represents standard deviation, ϵ^* represents determining the breakdown point is given in the form of numerical expression

$$\epsilon^*(y, t) = i_{nf} \{ \epsilon | c(\epsilon; y, t) \} \tag{20}$$

From the above equation, i_{nf} represents the infimum. The breakdown point of the window sample is formulated as

$$\epsilon^*(y, t) = i_{nf} \{ \epsilon | c(\epsilon; y, t) \geq r_o \} \tag{21}$$

From the above equation, the standard deviation of three or fourfold is indicated by r_o .

- The NIQR is utilized as $t(y')$ an estimated window sample to compute the bias, whereas the initial window is higher than the break-down point window.
- After the breakdown point, the monitoring sequence is extracted with step length and window breadth.
- The pertaining processing of the monitoring series is segmented into the dependent distortion characteristics. The above 4 steps are repeated in this adaptive processing.

Data Transformation

The climatic condition contains categorical information, $PM_{2.5}$, PM_{10} , CO, NO_2 , SO_2 , and O_3 which is necessary to

be calculated. These climatic conditions provided as input are not statistical data and it has to be transformed into statistical data (He et al. 2022, 2014). This is carried out by utilizing the statistical values in substituting for the meaning of information. Table 1 depicts various stages of the AQI and its ranges.

The data transformation for the proposed ACBiGRU-DAO method is performed by the min–max normalization approach. In the min–max normalization scheme, the mapping range of the attribute based on the upper and lower bound is indicated by $[dv, hv]$ with newly generated values $[dv_{NEW}, hv_{NEW}]$. The interval for the target range for the normalization is determined as $[0, 1]$ or $[-1, 1]$. The expression for the normalized range y' is formulated as

$$y' = \frac{y - dv}{hv - dv} (hv_{NEW} - dv_{NEW}) + dv_{NEW} \tag{22}$$

ACBiGRU-DAO Algorithm for Air Quality Prediction

The air quality prediction using the proposed ACBiGRU model is presented in Fig. 4. The following steps illustrate the step-by-step procedures carried out to determine air quality prediction.

(i) Input data

The air quality data collected from the Kaggle dataset is comprised of (i) India’s air pollution level over the years and (ii) air quality index data with hourly pollutant concentration levels from the year 2015 to 2020. From the dataset, the seven major air pollutant parameters are taken as input to determine the air quality by the proposed model.

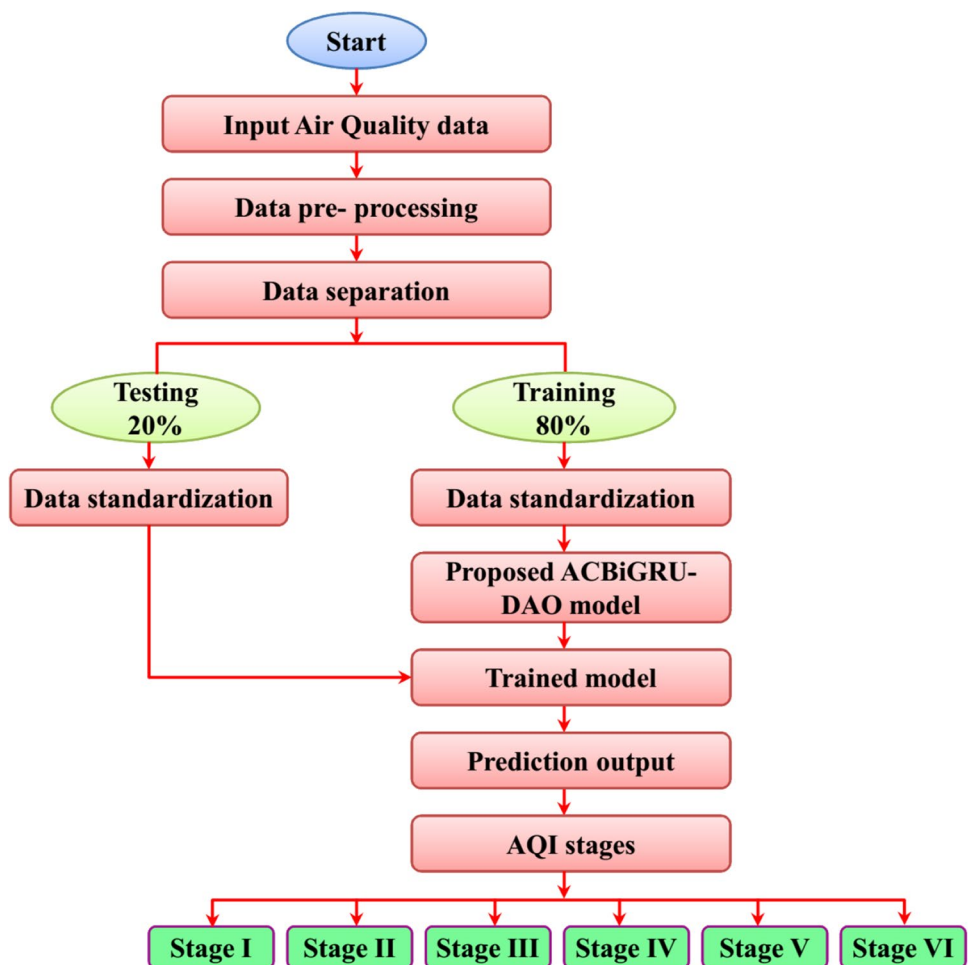
(ii) Data preprocessing

The drastic fluctuations in the concentration level of actual air quality data interrupt the prediction performance of the model. So, to get better prediction results in dealing with the air quality dataset, the inputted data is initially preprocessed. The air quality data is initially transformed into the statistical format

Table 1 Various AQI stages

AQI stages	Ranges	Air Pollution Level
Stage 1	0–50	Good
Stage 2	51–100	Moderate
Stage 3	101–150	Unhealthy for sensitive groups
Stage 4	151–200	Unhealthy
Stage 5	201–300	Very unhealthy
Stage 6	> 300	Hazardous

Fig. 4 Steps of the proposed ACBiGRU-DAO prediction model



by the data transformation process and then the missing values were imputed using the adaptive sliding window technique.

(iii) Data separation and standardization

After preprocessing, the dataset is partitioned into two divisions for testing (20%) and training (80%). Then, both the training and testing sets are standardized through the z-score standardization technique to ensure uniformity among the data. This standardization process helps to improve the prediction accuracy of the model. The numerical expression of z-score standardization is given by

$$x_k = \frac{y_k - \bar{y}}{\sigma} \tag{23}$$

The term x_k indicates standardized data, y_k represents input data, \bar{y} signifies the mean value of input data, and σ implies standard deviation. Moreover, the mean value \bar{y} and standard deviation σ of input data are measured using the following equation:

$$\bar{y} = \frac{1}{N} \sum_{k=1}^N y_k \tag{24}$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k - \bar{y})^2} \tag{25}$$

(iv) Training phase

The standardized data is provided as input to the proposed ACBiGRU-DAO model for training. At this phase, the optimal parameter values for the ACBiGRU model are assigned by an adaptive tuning process using the DAO algorithm. By using the optimized parameters, the features in the data are effectively trained by the ACBiGRU-DAO model.

(xxii) Testing phase

The standardized testing data are provided to the trained network model to formulate further predictions.

(vi) Prediction

The air quality is accurately predicted and classified by the ACBiGRU-DAO model according to pollutant concentration severities into 6 AQI stages. These stages help to map the quality of air. The air quality index values vary between the range of 0 and 500. The AQI range from 0 to 50 indicates normal/good air quality, 51 to 100 represents acceptable air quality, greater than 100 implies unhealthy air which might cause health effects for sensitive people and the range above 300 indicates hazardous air.

This research article determines the air quality data obtained all over the world to validate the variations that happen in the air quality during pandemic situations. The validation of air quality is performed based on comparing various air-polluting input features such as particulate matters namely $PM_{2.5}$ and PM_{10} , carbon monoxide (CO) concentration, nitrogen dioxide (NO_2) concentration, sulfur dioxide (SO_2) concentration, and ozone (O_3) concentration. The fundamental effects of air quality are predicted by quasi-experimental approaches to eject the distracting features. Various existing WTCNN, LSTM-AML, AQPS-NFM, TL-BiLSTM, MTMC-NLSTM, STCNN-LSTM, and TLS-BLSTM methods are employed to predict air quality index. However the prediction performance is diminished due to the utilization of very limited feature parameters, the crucial factors such as economic factor details as well as the gas emission warning was also not given importance. The exploration of dynamic spatial correlation is often performed with a high computational cost. The air pollution emission is increased which leads to generating dispersion and creating negative impacts in human life. The existing techniques predicted the pollution mostly within the limited temporal and spatial scale.

To overcome this problem and handle the high multicollinearity present in the input samples, the ACBiGRU-DAO method is proposed which improves the performance and accurately classifies the severity level of the air quality and protects the life of humans affected by air pollution. The optimal parameter values are determined which helps to monitor and predict the air quality in urban and industrial locations. The validation to enhance the air quality is offered in policy intervention of air quality using the comprehensive data from January 1 to July 5, 2020, which recovered the AQI independently in 597 significant cities by $PM_{2.5}$, PM_{10} , CO, NO_2 , SO_2 , and O_3 . The envelope calculation for air quality is contributed which determines the variations while predicting the AQI which reduces the air pollution. The prediction is performed using the Kaggle dataset which enhanced the effectiveness and prediction accuracy.

Experimental Results and Analysis

To analyze the effectiveness, the various parameter metrics namely Accuracy, Maximum Prediction Error (MPE), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Correlation Coefficients (CC) are utilized. The different methods such as WTCNN, LSTM-AML, AQPS-NFM, TL-BiLSTM, and the proposed ACBiGRU-DAO algorithm are applied for the comparative analysis to get high performance in the proposed ACBiGRU-DAO algorithm.

Parameter Setting

The parameter setting of the proposed ACBiGRU-DAO algorithm is explained in Table 2.

Performance Measures

The performance rate evaluation for the proposed method is validated by MPE, MAE, MSE, RMSE and Correlation Coefficient are employed to evaluate the performance rate.

Accuracy: Accuracy is defined as the ratio of accurate prediction of the true value to the total prediction and is formulated as

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (26)$$

The terms $p(i)$ and $a(i)$ indicate the predicted air concentration and actual air concentration, respectively.

MPE: It is the ratio of the maximum difference between the actual error and predicted error to the actual error value.

$$\text{MPE} = \max \left\{ \frac{|p(i) - a(i)|}{a(i)} \times 100 \right\} \quad (27)$$

Table 2 Parameter setting

Methods	Parameters	Values
ACBiGRU	Optimizer	Adam
	Batch size	128
	Loss function	Cross entropy
	Epoch	50
DAO	Exploration	1
	Exploitation	1.5
	Population size	50
	Number of iterations	100

MAE: The magnitude difference between observed air concentrations to the predicted air concentration is referred to as mean absolute error.

$$MAE = \frac{1}{n} \sum_{i=1}^n |p(i) - a(i)| \tag{28}$$

Here, *n* represents the maximum number of data samples.

MSE: It is the measurement of the amount of error obtained in estimating the statistical samples.

$$MSE = \frac{1}{n} \sum_{i=1}^n (p(i) - a(i))^2 \tag{29}$$

RMSE: The root mean square error signifies the square root of residual variance.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (p(i) - a(i))^2} \tag{30}$$

Correlation Coefficient (CC): It computes the statistical relationship between two independent data variables *y* and *z*.

$$CC = \frac{n(\sum yz) - (\sum y)(\sum z)}{[n \sum y^2 - (\sum y)^2][n \sum z^2 - (\sum z)^2]} \tag{31}$$

Performance Analysis

Table 3 represents the set of data features provided to the proposed ACBiGRU-DAO framework to determine the air quality. The most influencing factors in the air quality data are taken as input data which includes the features such as current PM_{2.5} and PM₁₀, AQI, CO, NO₂, SO₂, and O₃ concentration.

Table 4 illustrates the statistical results achieved using the proposed ACBiGRU-DAO approach and various other existing approaches in terms of different evaluation indicators namely accuracy, precision, recall, specificity, F-measure, and total no. of days the AQI stages predicted accurately. The proposed ACBiGRU-DAO approach attained air quality prediction accuracy of about 95.34% with a minimal error rate (4.66%).

Figure 5 represents the graphical analysis of prediction accuracy analysis concerning different methods. From the graph, it came to know that the proposed ACBiGRU-DAO approach gained more accuracy than other compared methods. The accuracy rate of the proposed ACBiGRU-DAO approach is 95.34%.

Figure 6 describes the graphical illustration of the error rate (%) concerning different methods. The error rate achieved by the proposed ACBiGRU-DAO approach is 4.6% whereas the other methods produced a greater error rate than

Table 3 Training features are taken from the air quality dataset

Sl. no	Features	Description
1	AQI	Current Air Quality Index value
2	PM _{2.5} and PM ₁₀	Particulate matters at an hour ago
3	CO	Carbon monoxide (CO) concentration was recorded at an hour ago
4	NO ₂	Nitrogen dioxide (NO ₂) concentration was recorded at an hour ago
5	SO ₂	Sulfur dioxide (SO ₂) concentration was recorded at an hour ago
6	O ₃	Ozone (O ₃) concentration was recorded at an hour ago

Table 4 Statistical results of proposed ACBiGRU-DAO and another approach

Evaluation Indicators	WTCNN	LSTM-AML	AQPS-NFM	TL-BiLSTM	ACBiGRU-DAO
Accuracy	86.03%	91.51%	85.21%	87.67%	95.34%
Precision	94.37%	93.67%	91.33%	87.83%	95.31%
Recall	76.67%	79.56%	80.05%	89.56%	90.59%
Specificity	89.49%	87.21%	91.24%	88.94%	93.58%
F-measure	88.94%	87.83%	81.43%	89.56%	90.33%
Error rate	13.97%	8.49%	14.79%	12.33%	4.66%
Maximum prediction error	4.3	3	2.4	3.4	1
Mean absolute error	0.3	0.4	0.76	0.3	0.059
Mean square error	0.6	0.4	0.81	0.3	0.057
Root mean square error	0.72	0.84	0.94	0.54	0.238
Correlation coefficient	0.54	0.62	0.69	0.71	0.8
The total no. of days the AQI stages predicted accurately	314	334	311	320	348

Fig. 5 Accuracy analysis

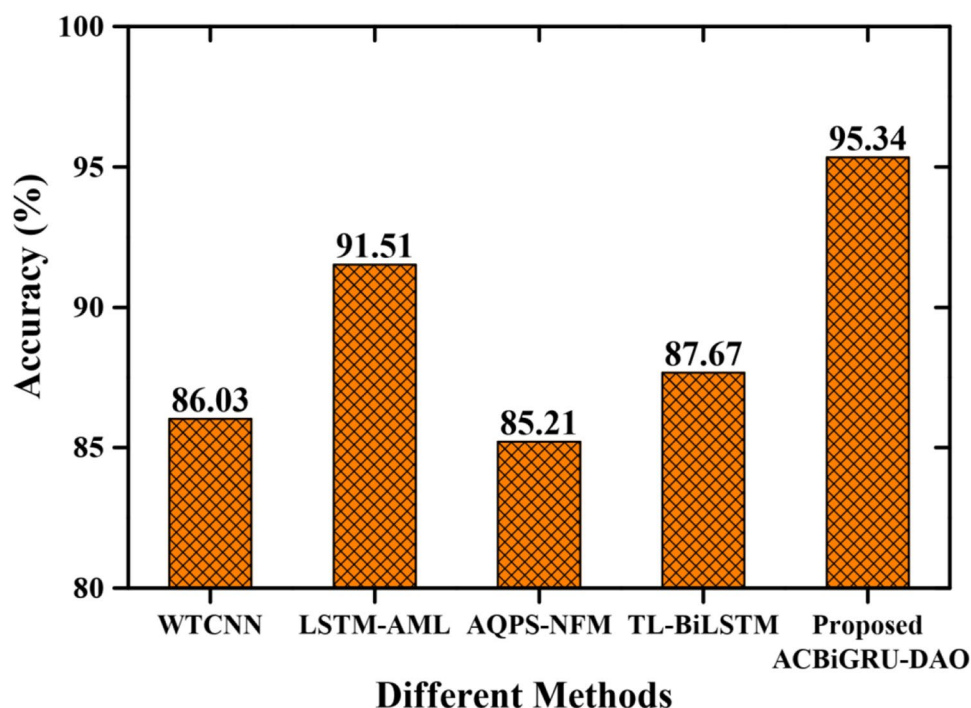
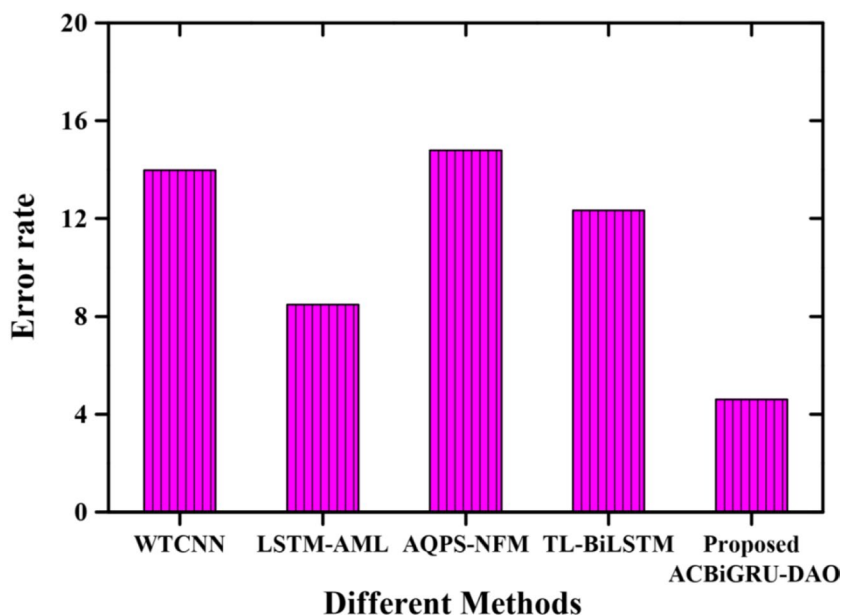


Fig. 6 Error rate evaluation



the proposed approach. This indicates the proposed approach is well effective than other compared methods with a very less error rate in air quality prediction. Figure 7 represents the MPE rate analysis of proposed and existing methods. The maximum prediction error achieved by the proposed ACBiGRU-DAO approach is 1 (very low) as compared to existing methods. The lower the prediction error, the greater will be prediction accuracy.

The correct prediction of air quality index stages concerning a total number of days is shown in Fig. 8. The graph

displays that the proposed ACBiGRU-DAO approach correctly predicted the AQI stages for about 348 days while the other compared methods namely WTCNN, LSTM-AML, AQPS-NFM, and TL-BiLSTM correctly predicted for 314 days, 334 days, 311 days, and 320 days, respectively. This illustrates the proposed method accurately predicts the AQI stages in more number of days than other compared methods. Figure 9 portrays the graphical results of MAE for different methods. The validation of the proposed approach is 0.059 which is less compared to MAE values of existing

Fig. 7 MPE analysis

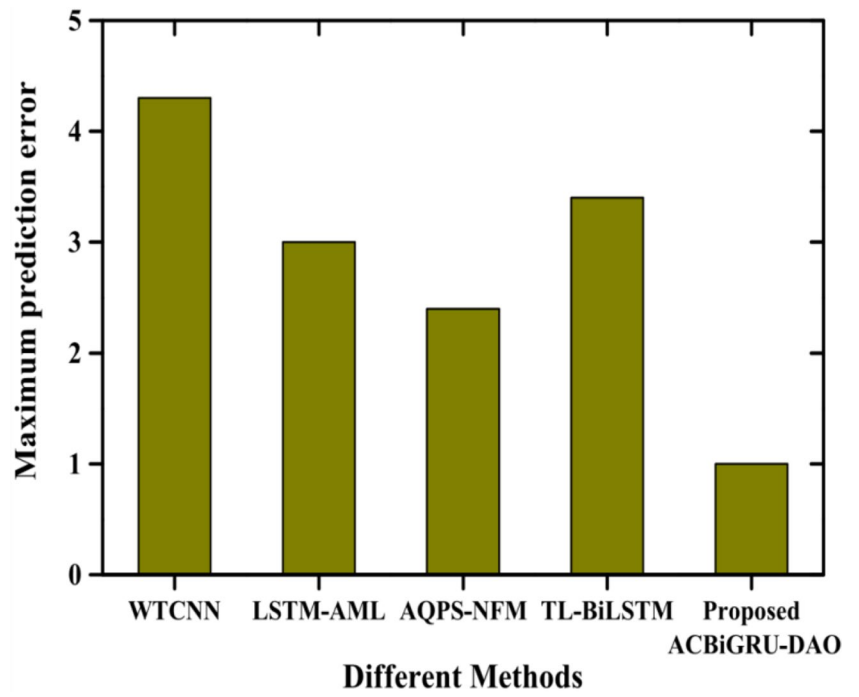


Fig. 8 Estimation of correctly predicted days of AQI

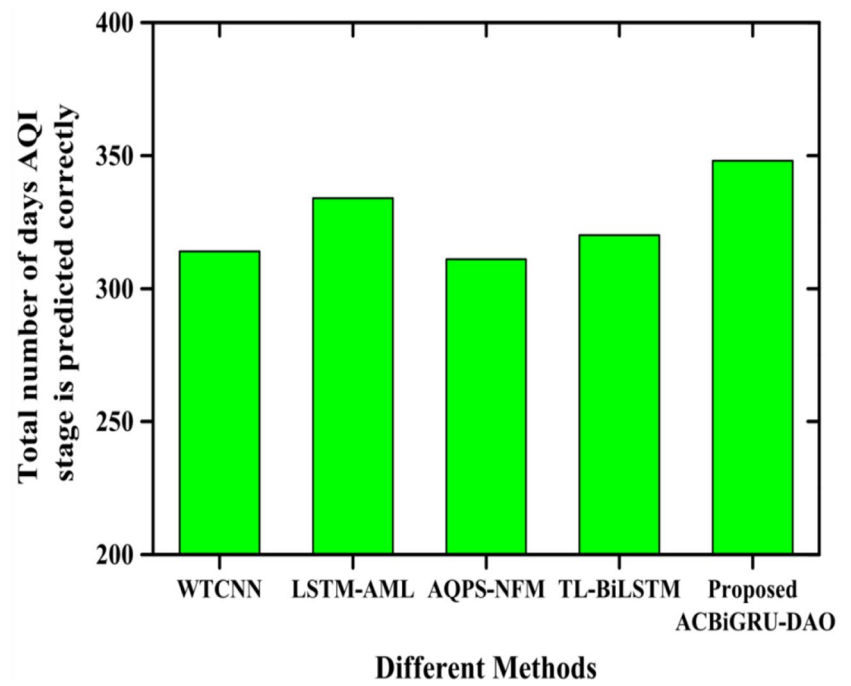


Fig. 9 MAE analysis

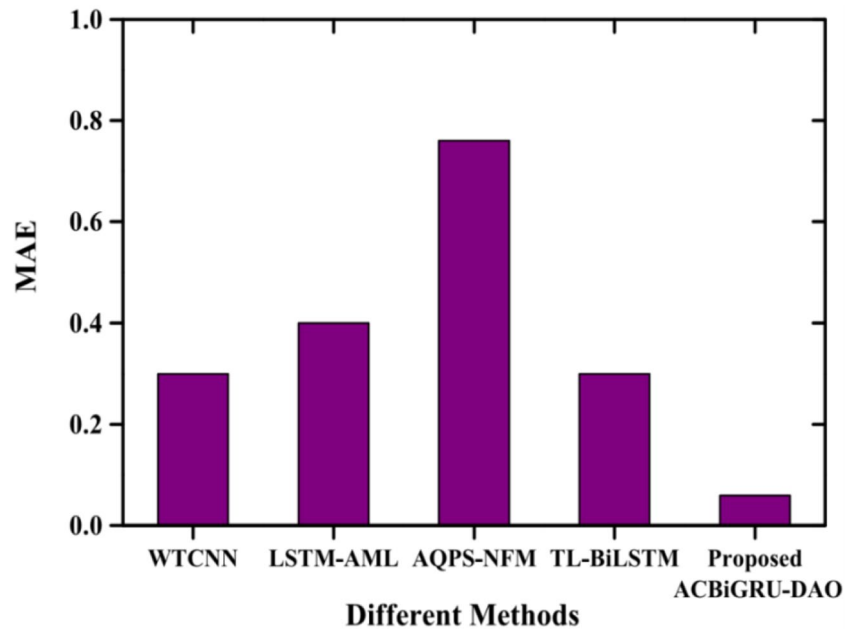
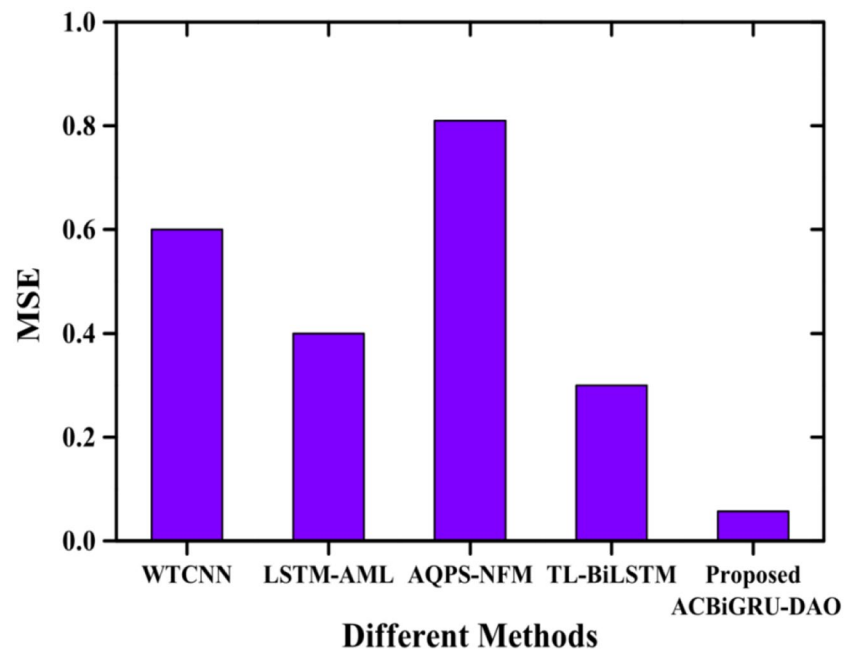


Fig. 10 MSE analysis



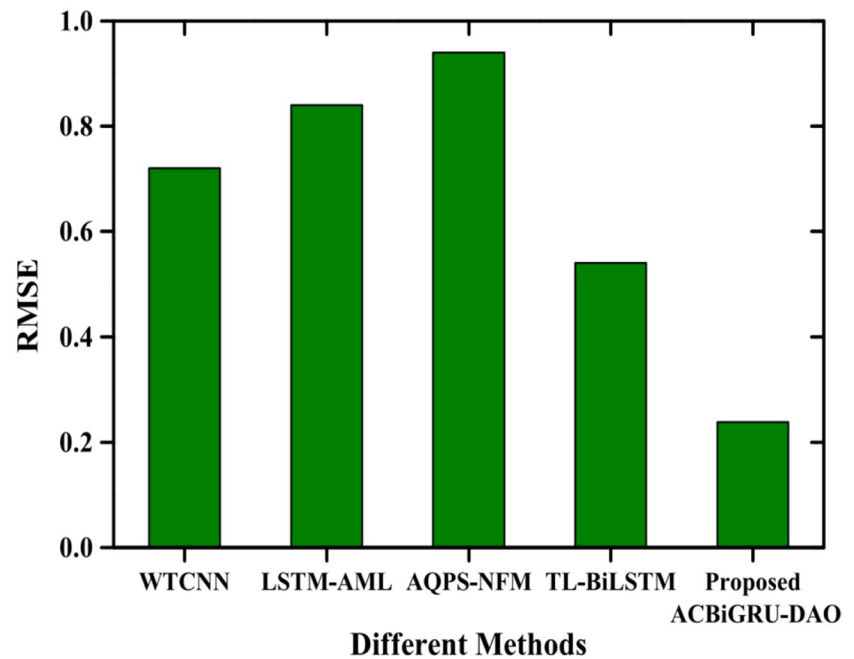
methods. The lower the error rate, the greater will be the prediction performance.

Figures 10 and 11 represent the diagrammatic representation of MSE analysis and RMSE analysis respectively with respect to different methods. The error rates (MSE and RMSE) achieved by the proposed ACBiGRU-DAO approach are very less when compared to existing methods such as WTCNN, LSTM-AML, AQPS-NFM, and TL-BiLSTM. The MSE and RMSE rate achieved by the proposed ACBiGRU-DAO approach is 0.057 and 0.238, respectively.

Conclusion

This paper presents a novel air quality prediction model “ACBiGRU-DAO approach” to accurately predict air quality and classify predicted results based on severity ranges into 6 air quality index (AQI) stages. The utilization of the DAO algorithm in the ACBiGRU model enhances prediction accuracy. The effectiveness of the proposed method is validated by air quality data of India that is acquired from the Kaggle website. The performance of the

Fig. 11 RMSE analysis



proposed ACBiGRU-DAO approach is examined using diverse evaluation indicators namely Accuracy, MPE, MAE, MSE, RMSE, and Correlation Coefficient. The accuracy rate achieved by the proposed ACBiGRU-DAO approach is greater than other compared methods with a percentage of 95.34%. In addition, the error rates validation namely MPE, MAE, MSE, and RMSE for the proposed method is diminished when compared to the existing WTCNN, LSTM-AML, AQPS-NFM, and TL-BiLSTM methods. This indicates the effectiveness of the proposed ACBiGRU-DAO approach over other existing methods. In future work, the prediction of accuracy in air quality will be further improved by exploring spatial domain features.

Author contribution All agreed on the content of the study. Vinoth Panneerselvam and Revathi Thiagarajan collected all the data for analysis. Vinoth Panneerselvam agreed on the methodology. Vinoth Panneerselvam and Revathi Thiagarajan completed the analysis based on agreed steps. Results and conclusions are discussed and written together. All authors read and approved the final manuscript.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval This article does not contain any studies with human or animal subjects performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

References

- Bekkar A, Hssina B, Douzi S, Douzi K (2021) Air-pollution prediction in smart city, deep learning approach. *J Big Data* 8(1):1–21
- Benhaddi M, Ouarzazi J (2021) Multivariate time series forecasting with dilated residual convolutional neural networks for urban air quality prediction. *Arab J Sci Eng* 46(4):3423–3442
- Dairi A, Harrou F, Sun Y, Khadraoui S (2020) Short-term forecasting of photovoltaic solar power production using variational auto-encoder driven deep learning approach. *Appl Sci* 10(23):8400
- Freeman BS, Taylor G, Gharabaghi B, Thé J (2018) Forecasting air quality time series using deep learning. *J Air Waste Manag Assoc* 68(8):866–886
- Ge L, Wu K, Zeng Y, Chang F, Wang Y, Li S (2021) Multi-scale spatiotemporal graph convolution network for air quality prediction. *Appl Intell* 51(6):3491–3505
- Guo Q, He Z (2021) Prediction of the confirmed cases and deaths of global COVID-19 using artificial intelligence. *Environ Sci Pollut Res* 28:11672–11682
- Guo Q, He Z, Li S, Li X, Meng J, Hou Z, Liu J, Chen Y (2020) Air pollution forecasting using artificial and wavelet neural networks with meteorological conditions. *Aerosol Air Qual Res* 20(6):1429–1439
- Guo Q, Wang Z, He Z, Li X, Meng J, Hou Z, Yang J (2021) Changes in air quality from the COVID to the post-COVID era in the beijing-tianjin-tangshan region in China. *Aerosol Air Qual Res* 21(12):210270
- Guo Q, He Z, Wang Z (2023a) Predicting of daily PM_{2.5} concentration employing wavelet artificial neural networks based on meteorological elements in Shanghai, China. *Toxics* 11(1):51

- Guo Q, He Z, Wang Z (2023b) Long-term projection of future climate change over the twenty-first century in the Sahara region in Africa under four Shared Socio-Economic Pathways scenarios. *Environ Sci Pollut Res* 30(9):22319–22329
- He Z, Zhang Y, Guo Q, Zhao X (2014) Comparative study of artificial neural networks and wavelet artificial neural networks for groundwater depth data forecasting with various curve fractal dimensions. *Water Resour Manag* 28:5297–5317
- He Z, Guo Q, Wang Z, Li X (2022) Prediction of monthly PM_{2.5} concentration in Liaocheng in China employing artificial neural network. *Atmosphere* 13(8):1221
- Huang G, Wang D, Du Y, Zhang Q, Bai Z, Wang C (2022) Deformation Feature Extraction for GNSS Landslide Monitoring Series Based on Robust Adaptive Sliding-Window Algorithm. *Frontiers in Earth Science* 487.
- Jin N, Zeng Y, Yan K, Ji Z (2021) Multivariate air quality forecasting with nested long short term memory neural network. *IEEE Trans Industr Inf* 17(12):8514–8522
- Khodadadi N, Snaes V, Mirjalili S (2022) Dynamic arithmetic optimization algorithm for truss optimization under natural frequency constraints. *IEEE Access* 10:16188–16208
- Kristiani E, Lin H, Lin JR, Chuang YH, Huang CY, Yang CT (2022) Short-term prediction of PM_{2.5} using LSTM deep learning methods. *Sustainability* 14(4):2068
- Li H, Wang J, Li R, Lu H (2019) Novel analysis–forecast system based on multi-objective optimization for air quality index. *J Clean Prod* 208:1365–1383
- Lin YC, Lee SJ, Ouyang CS, Wu CH (2020) Air quality prediction by neuro-fuzzy modeling approach. *Appl Soft Comput* 86:105898
- Liu H, Yan G, Duan Z, Chen C (2021) Intelligent modeling strategies for forecasting air quality time series: A review. *Appl Soft Comput* 102:106957
- Liu J, Yang Y, Lv S, Wang J, Chen H (2019) Attention-based BiGRU-CNN for Chinese question classification. *Journal of Ambient Intelligence and Humanized Computing* 1–12.
- Ma J, Cheng JC, Lin C, Tan Y, Zhang J (2019) Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques. *Atmos Environ* 214:116885
- Ma J, Li Z, Cheng JC, Ding Y, Lin C, Xu Z (2020) Air quality prediction at new stations using spatially transferred bi-directional long short-term memory network. *Sci Total Environ* 705:135771
- Maleki H, Sorooshian A, Goudarzi G, Baboli Z, Tahmasebi Birgani Y, Rahmati M (2019) Air pollution prediction by using an artificial neural network model. *Clean Technol Environ Policy* 21(6):1341–1352
- Mao W, Wang W, Jiao L, Zhao S, Liu A (2021) Modeling air quality prediction using a deep learning approach: Method optimization and evaluation. *Sustain Cities Soc* 65:102567
- Samal KKR, Babu KS, Acharya A, Das SK (2020) Long term forecasting of ambient air quality using deep learning approach. In 2020 IEEE 17th India Council International Conference (INDICON) (pp. 1–6). IEEE.
- Tiwari A, Gupta R, Chandra R (2021) Delhi air quality prediction using LSTM deep learning models with a focus on COVID-19 lockdown. *arXiv preprint arXiv:2102.10551*.
- Wang J, Song G (2018) A deep spatial-temporal ensemble model for air quality prediction. *Neurocomputing* 314:198–206
- Xu X, Yoneda M (2019) Multitask air-quality prediction based on LSTM-autoencoder model. *IEEE Trans Cybern* 51(5):2577–2586
- Zhang Y, Wang Y, Gao M, Ma Q, Zhao J, Zhang R, Wang Q, Huang L (2019) A predictive data feature exploration-based air quality prediction approach. *IEEE Access* 7:30732–30743
- Zhang K, Thé J, Xie G, Yu H (2020) Multi-step ahead forecasting of regional air quality using spatial-temporal deep neural networks: a case study of Huaihai Economic Zone. *J Clean Prod* 277:123231
- Zhang Z, Zeng Y, Yan K (2021) A hybrid deep learning technology for PM_{2.5} air quality forecasting. *Environ Sci Pollut Res* 28(29):39409–39422
- Zhao G, Huang G, He H, He H, Ren J (2019) Regional spatiotemporal collaborative prediction model for air quality. *IEEE Access* 7:134903–134919
- Zhu D, Cai C, Yang T, Zhou X (2018) A machine learning approach for air quality prediction: Model regularization and optimization. *Big Data Cogn Comput* 2(1):5
- Zou X, Zhao J, Zhao D, Sun B, He Y, Fuentes S (2021) Air quality prediction based on a spatiotemporal attention mechanism. *Mobile Information Systems* 2021:1–12.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.