Journal of Artificial Intelligence and Capsule Networks (ISSN: 2582-2012)
www.irojournals.com/aicn/

A Hybrid Deep Learning Framework for Air Quality Index Forecasting

Saranya K G.[1], Akalya A.[2], Harini G.[3]

[1]Associate Professor, [2,3]PG Scholar, Department of Computer Science, PSG College of Technology, Coimbatore, India.

Email: [1]24mz31@psgtech.ac.in

Abstract

The Air Quality Index (AQI) must be accurately and highly evaluated to limit the effects when exposed to air pollution on individual health and the ecosystem. The traditional techniques are used to predict air quality ineffective in managing the nature of the environmental data being analyzed and minimizing the effects of the most significant features of large numbers of dimensions. The purpose of this work is to address these issues using the creation of a hybrid deep learning algorithm implementing Feature Selection Techniques, Recurrent Neural Networks and a Quantum-Inspired Genetic Algorithm (QIGA). The eXtreme Gradient Boosting (XGBoost) will be applied to the environmental dataset to evaluate importance of features. Additionally, the Principal Component Analysis (PCA) will reduce the dataset's dimensionality using most significant features from the original dataset as inputs for the prediction model. Recurrent Neural Networks (RNNs) are able to detect time-variant patterns of air quality (i.e., the pattern changes over time) based on the use of controlled memory. The use of quantum-based methodologies will allow rapid searches over high-dimensional datasets resulting higher performance than traditional optimization methods. This innovative methodology will lead to improved accuracy, computational efficiency and interpretability of AQI predictions. This method will be the basis for smart environmental monitoring systems used by researchers, policymakers and urban planners to make innovative decisions. Finally, the flexibility of the system makes it possible for users to apply it in any other forecasting-based environmental issues showing the scalability and efficiency.

DOI: https://doi.org/10.36548/jaicn.2026.1.003
Received: 15.01.2026, received in revised form: 20.02.2026, accepted: 11.03.2026, published: 23.03.2026

Keywords: Air Quality Index, Feature Selection, Recurrent Neural Networks, Quantum-Inspired Genetic Algorithm, XGBoost, PCA.

1. Introduction

Air pollution has become increasingly problematic for human health and the economy for the environment in major cities. The Air Quality Index (AQI) research provides a better forecast of air quality so that people can accurately predict whether the air is polluted. The processes involved in producing an AQI forecast are complicated. Several sources and types of pollutants can significantly affect air quality, weather conditions complicate forecasts and human factors in unpredictable ways. Due to the high volume of air quality data, it is an impossible due to the level of complexity and the size of the database used for an AQI prediction. A significant quantity of advanced statistical approaches must be used to achieve accurate AQI predictions. Traditional prediction methodologies are completely incapable of accurately collecting the complex and highly variable responses that dominate the environmental and meteorological databases used by analysts to evaluate air quality assessments. Thus, using traditional forecasting methodologies will not only provide inaccurate products but also result in additional times to air pollution incidents.

Using traditional techniques such as Genetic Algorithms (GA), Support Vector Machines (SVM) and Extreme Gradient Boosting (XGBoost) will display the issues that follow while identifying and evaluating feature sets to increase the accuracy and performance of Environmental Modeling and Forecasting of Air Quality Index (AQI): 1) the algorithms require significant computational time; 2) as the amount of input data changes, the algorithms require different inputs for large datasets; 3) the algorithms are unable to recognize the connection between temporal patterns and remaining patterns are not represented globally. There are also advantages and disadvantages to using previous identified values/factors, such as the possibility of bias increases the probability of creating defective models and reduces the generalization capability of the trained models across every period and regions. This work will utilize a deep-learning architecture integrates the design and implementation of Convolutional Neural Networks (CNN) and the application of Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) Networks to address the limitations of existing approaches.

The integration of these two architectures establish automated AQI data. This innovation will reduce manual feature development and also contribute overall accuracy of

forecasting models. The CNN/LSTM architecture provides possible ways to extract integrate feature sets. Through the ability to generate a wide range of feature sets from both the spatial and temporal domains, this combination accurately predicts the pollutants and meteorological variables as the LSTM units changing the air quality. The model's previous patterns significantly improve its performance when predicting future developments. Furthermore, it has a high level of flexibility/scalability due to the architecture's modular design and is capable of supporting additional data from future camera installations to make minor modifications for changes in the natural environment, with minimal changes to overall configuration.

QIGAs used to forecast the system to improve the overall performance. It uses the concepts of interference and superposition in quantum dynamics increase the possibility for investigating and utilization of the hyperparameter search region. It has the higher level of convergence than traditional GA and provide a larger number of possible solutions for effective deep learning model. Additionally, It can be used in CNN-LSTM system to achieve more optimal selection of hyperparameters reducing the training time for the model and subsequently leading to a more stable evaluation of the AQI.

By utilizing quantum-based approaches, This system optimize models' effectiveness when handles incomplete or inaccurate datasets enabling forecasting systems suitable for larger systems that can monitor changes to the environment over extensive period. As a result of this enhancement can achieve long-term sustainability goals, efficiently for urban development and improve overall quality of life for all citizens. This method explains a new hybrid approach that employs two sequential optimization methods that implements the quantum-based global search method to find hyperparameter space and a more conventional method of local optimization based on particle swarm optimization (PSO). This combination differentiate from the standard hybrid RNN-XGBoost-PSO approaches.

2. Literature Survey

2.1 Machine Learning Based Approaches

Machine learning (ML) algorithms played a basic role in the evolution of Air Quality Index (AQI) prediction systems due to their ability to model nonlinear connections between atmospheric pollutants and meteorological variables. Traditional statistical models such as ARIMA faces problem to capture complex interactions among multiple pollutants; therefore,

data-driven ML techniques have become increasingly preferred. Ensemble-based learning methods such as Random Forest (RF) and Gradient Boosting Machines (GBM) have demonstrated better robustness in handling high-dimensional environmental datasets [3], [13].

These algorithms effectively model nonlinear dependencies between PM2.5, PM10, NO_2, SO_2, CO, O_3 and meteorological parameters including temperature, humidity, atmospheric pressure and wind speed. Random Forest provides feature importance scores improve interpretability and support environmental policy decisions. Support Vector Machines (SVM) have also been widely adopted due to their capability to handle small-to-medium sized datasets while maintaining generalization performance [3]. Kernel-based transformations allow SVM models to represent nonlinear pollutant interactions efficiently. Dimensionality reduction methods such as Principal Component Analysis (PCA) are often integrated with ML pipelines to eliminate redundancy among correlated pollutant variables [2]. By projecting high-dimensional pollutant space into orthogonal components, PCA enhances computational efficiency and reduces overfitting risk.

Hybrid ML models combining signal decomposition techniques with predictive algorithms have further improved short-term AQI forecasting accuracy. For instance, decomposition-based preprocessing isolates intrinsic temporal patterns before prediction, enabling models to better capture pollution increases and seasonal transitions [12]. Despite their effectiveness, ML-based methods have certain limitations. They generally require certain feature engineering are sensitive to data imbalance and may struggle with spatial dependency modelling across monitoring stations. Nevertheless, the lower computational cost makes suitable for scalable and near real-time AQI monitoring systems.

2.2 CNN Based Approaches

The growth of high-resolution environmental sensing and satellite-based monitoring, deep learning models like Convolutional Neural Networks (CNNs) have emerged as effective tools for AQI prediction. CNNs are capable of automatically extracting hierarchical spatial features from structured pollutant grids and remote sensing imagery. The work [1] proposed a CNN-ILSTM hybrid architecture that integrates spatial feature extraction using convolutional layers with temporal modelling through improved LSTM units. This approach demonstrated superior performance compared to standalone ML models by combining learning spatial and temporal pollutant correlations. Grid-based AQI prediction using 2D and 3D CNN

architectures has shown effectiveness in modelling pollutant dispersion patterns derived from satellite Aerosol Optical Depth (AOD) data combined with ground station measurements [5], [7]. Multi-channel inputs including PM2.5 concentration maps, temperature fields and humidity layers are processed through convolutional kernels to capture localized environmental dynamics.

CNN-LSTM hybrid models further enhance predictive performance by combining convolutional feature extraction with recurrent temporal learning [9], [10]. These architectures are effective in multi-step forecasting situations where both short-term changes and long-term seasonal patterns influence pollutant behaviour. Recent advancements include graph-based deep learning methods that integrate CNN layers with Graph Convolutional Networks (GCNs) to model spatial dependencies among distributed monitoring stations [15]. Such hybrid architectures capture both intra-station temporal variations and inter-station spatial connections, improving predictive stability in sparse sensor networks. However, CNN-based approaches are computationally intensive and require large labelled datasets. Their deployment in edge or IoT environments demands model compression, reducing or optimization methods to balance accuracy with computational efficiency.

2.3 Other Approaches

Beyond traditional ML and CNN frameworks, several advanced learning paradigms have been introduced for AQI prediction. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models are widely used to capture long-term dependencies in pollutant time-series data [10], [12]. These architectures effectively model temporal autocorrelation but may suffer from disintegrating gradient problems when predicting extended forecasting horizons. Transformer-based models have recently been adopted to address limitations of sequential recurrence by leveraging self-attention mechanisms [6]. Transformers enable parallel processing of time-series data while capturing long-range dependencies between pollutant variables and meteorological conditions. Empirical studies show improved 24-hour and 48-hour forecasting performance compared to traditional LSTM models. Evolutionary optimization techniques integrated into AQI forecasting systems. Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) are employed for hyperparameter tuning, feature selection and weight optimization [11]. These optimization-driven learning frameworks enhance convergence stability and reduce overfitting, particularly in highly uncertain pollution

datasets. Probabilistic approaches such as Gaussian Process Regression (GPR) provide uncertainty-aware AQI forecasts [13].

Unlike deterministic models, GPR produces confidence intervals alongside point predictions, which are crucial for risk-sensitive environmental decision-making. However, its computational complexity increases dataset size. Reinforcement learning methods have been explored to optimize air quality sensor placement and adaptive monitoring strategies in smart city environments [14]. By modelling the environment as a Markov Decision Process, reinforcement learning agents dynamically learn optimal sensing policies, leading to improved spatial coverage and forecasting performance. Although these advanced methods improve prediction accuracy and adaptability, they often focus on isolated aspects such as optimization, uncertainty estimation, or spatial modelling, rather than providing an integrated end-to-end framework

2.4 Interpretability

Interpretability remains a critical challenge in the development of machine learning and deep learning models for Air Quality Index (AQI) prediction. In environmental applications, transparency is important because policymakers and public health authorities depend on model outputs to make regulatory and safety decisions. However, most high-performing models, particularly deep neural networks such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), operate as black-box systems. Although they achieve high predictive accuracy, their internal decision-making processes are not easily interpretable.

Deep learning models perform multiple nonlinear transformations across hidden layers, extracting abstract feature representations from pollutant concentrations, meteorological variables and satellite-derived data. While these representations improve forecasting performance, they do not explicitly indicate the relative contribution of individual environmental factors such as temperature, humidity, wind speed, or specific pollutant levels. This lack of transparency limits trust and adoption in real-world environmental governance systems. To address this limitation, several explainable AI (XAI) techniques have been introduced.

Attention mechanisms highlight windows or pollutant variables influencing predictions. Post-hoc explanation methods such as SHAP (SHapley Additive exPlanations) and Layer-wise Relevance Propagation (LRP) quantify feature contributions and provide local

interpretability. However, these approaches depend on approximations and may not align with physical environmental principles. Furthermore, their computational overhead restricts their use in real-time or resource-constrained deployments. Human-in-the-loop (HIL) systems have also been proposed to enhance interpretability by allowing domain experts to interact with model predictions. Although this improves transparency, manual intervention may introduce subjectivity and reduce scalability across different operational contexts. Recent research emphasizes incorporating interpretability directly into model architecture. Modular neural network designs assign specific components to well-defined environmental factors improving transparency of prediction pathways.

Graph-based neural networks explicitly model connections between monitoring stations and pollutant interactions aligning more closely with environmental domain knowledge. Additionally, rule-based and symbolic AI systems provide inherently interpretable predictions, though the cost of reduced predictive accuracy.

Despite these advancements, most existing AQI prediction frameworks handle feature selection, dimensionality reduction, temporal modeling and hyperparameter optimization as independent processes. There remains a research gap in developing a unified and interpretable framework that integrates supervised feature evaluating, orthogonal transformation techniques, temporal sequence modelling, and global optimization within a coordinated architecture. Therefore, this study aims to address these limitations by proposing an integrated hybrid framework that enhances predictive accuracy while maintaining structural interpretability and optimization coherence. The existing studies have applied machine learning and deep learning techniques for AQI prediction. Most approaches handle feature selection, dimensionality reduction and hyperparameter optimization as independent processes. Furthermore, limited studies integrate global and local optimization strategies within a unified system. These limitations highlight the need for a coordinated hybrid framework.

2.5 Research Gap

Despite extensive research on AQI prediction using machine learning and deep learning techniques, most existing studies address feature selection, dimensionality reduction, and hyperparameter optimization independently. This dispersed approach may lead to suboptimal performance due to lack of coordination among model components. The problem addressed in this work is the design of a unified framework that combinely integrates supervised feature

selection, orthogonal transformation, temporal modeling and hybrid global-local optimization. The objective is to minimize prediction error while maintaining stability under dynamic environmental conditions. Existing air-quality prediction models typically optimize learning algorithms, feature selection and hyperparameters independently lead to suboptimal global performance. There is a lack of integrated frameworks that jointly optimize model structure, feature representation and training parameters. This study addresses this gap by proposing a coordinated optimization approach using QIGA-PSO.

3. Proposed Algorithm

The proposed framework shows the sequential deep learning and population-based optimization. RNNs is able to learn long-term temporal connections in non-stationary environmental datasets, while evolution-based optimization methods will allow efficient exploration of complex multidimensional spaces. Results of statistical comparison tests on dataset models demonstrate that by incorporating features based on supervised learning, dimensionality reduction and hybridization with evolutionary optimization algorithms predicts accuracy and generalizability were significantly increased. This proposed framework used across multiple geographical locations and includes the additional environmental variables, makes it applicable to large-scale Air Quality Monitoring Systems.

3.1 Dataset Description

The dataset's focus on 2006 to 2008 contains measurements of the air pollution indicator (PM2.5, PM10, NO 2, CO, O 3, and SO 2) and related weather conditions (temperature, humidity, wind speed, atmospheric pressure, etc.). By providing useful data about pollution levels and physical factors that affect pollution, the features have provided a possible development of prediction models using machine learning. Some preprocessing steps will need to analysis before proceed to the dataset. The most common method handle with recorded missing values to perform an imputation. The air quality dataset has missing records due to malfunctioning sensors or records being affected when maintenance or altering the data.

Feature normalisation is required because pollutant and weather measurement levels of varying degrees must be normalised and handled similarly. All inputs are included to the algorithms (e.g., Gradient Boosting, Long Short-Term Memory (LSTM), Random Forest) will be handled equally to provide an accurate output. Another crucial preprocessing phase will be

feature selection, because few of the recorded characteristics (attributes) correlate with an accurate predicting of Air Quality Index (AQI) values. The techniques such as Recursive Feature Elimination (RFE), correlation analysis, or domain-driven selection assist in isolating the most effective variables, dimensionality reduction maximizing the predictive power. Moreover, the data is typically split into a training and testing set, most commonly 70:30 or 80:20, to help model development and testing. Cross-validation is to determine the models generalize to new data and reduce overfitting and make them robust.

3.2 Data Preprocessing

CNNs combined with the meteorological variables and made the basis upon the air quality could be predicted, depending on the levels of the contaminants of the environment. The CNNs can retrieve multidimensional dense data which has a spatial relationship between the meteorological data and the contaminant data extract better spatial data. The Long Short-Term Memory (LSTM) Networks provide the temporal /sequential variations of the particles may be monitored throughout time using this geographical data.

The LSTM Networks can determine the pattern which has remained in their cyclic behaviour and lack of connection to the pollutants using a huge time series data. The neural networking approach (CNN and LSTM Networks) provides a composite model that may be utilized in the simulation of the additive dynamic behavior with a high level of accuracy. The combination of forecasting model will be precise and not vulnerable to noise. It will provide better decision-making functions when it comes to air quality and air pollution.

A better data quality processing system will be developed with re-evaluation of missing values and accessible in other comparable values until the complete continuous time of each of the pollutant measurements is verified. The statistical threshold approach identify anamolies and the corresponding data were substituted. To ensure the focus on the pollutant and weather conditions are comparable in the heterogeneous sources of the data (i.e. multiple sensorsIt was organized in a way that the time-related communication was independent, and the time frame varied across the measurements of various sensors. All these techniques will enhance the overall result and will decrease the bias to enhance the learning rate concerning time patterns. Fig. 1 illustrates this process in particular.

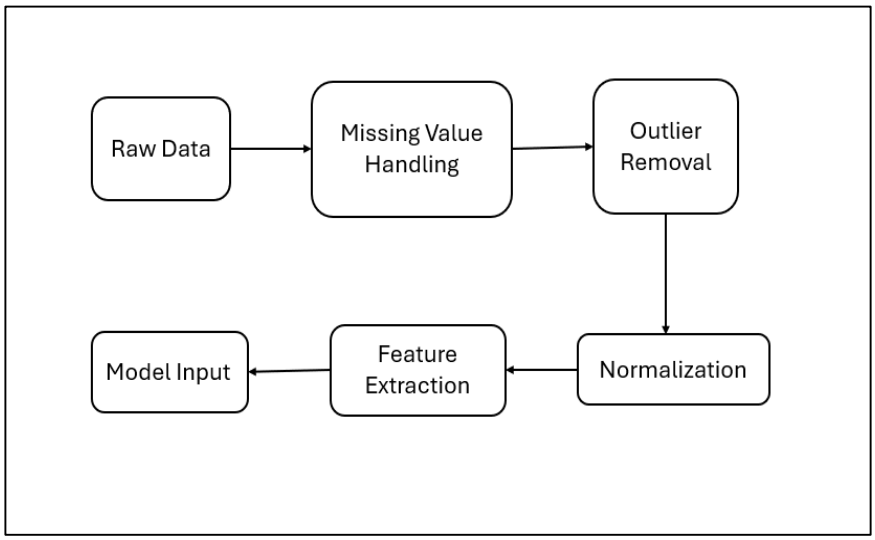

Figure 1. Block Diagram of Data Preprocessing

3.3 Feature Selection Using PCA and XGBoost

The combination of PCA and XGBoost will develop a very effective way to reduce dimensionality by using both feature selection for an accurate Air Quality Index (AQI). The use of XGBoost combines with PCA will provide a method to list the supervised variables able to predict outcomes. Therefore, using the most predictive pollution and weather variables can develop a model to make AQI predictions. Using XGBoost is recommended in identifying those predictive variables that fail to contribute and transferring PCA to those remaining variables will destroy variability and develop individualized or orthogonal data components. This method will also reduce the number of predictors used and reduce the amount of time and work required to calculate a predictive model.

PCA can lower the amount of variability and decrease the size of the feature space with lack of data providing rapid training but also provide better generalization of data. The differences of PCA on XGBoost provide a relevant and non-redundant features that the end result significantly higher computational efficiency and a higher amount of data related to the model. Both PCA and XGBoost provide a way for the model to focus on the most important patterns associated with both pollutants in the environment and those pollutants affect the quality of the air. After applying XGBoost, the only difference selecting non-redundant features will allow rapid measurements and higher model operates.

When performing PCA on the raw inputs, the latent variables or structures of the input will be visible. This is useful for situations where there are large amounts of data that may or

may not have heterogeneous distributions of data. The XGBoost-PCA process is a vital component for scalable air quality forecasting works using low power, descriptive operations space. XGBoost allows developed measures based on the gradient-based split gain of all features selected by the XGBoost algorithm. This ensures that only discriminative features will be used to create dimensions to be reduced. When PCA is applied to each of the variants of feature selection, the noise is reduced and the relevant variance is retained.

3.4 Model Building Using RNN and PSO Optimization

The AQI Temporal Dependencies for RNN Model Building: The time-series forecasting aspects of the AQI (Air Quality Index) data are capable of forecasting future values with high accuracy. The process with any type of data with temporal elements, RNNs have become the technical solution for dealing with the temporal aspects of data, such as trends of air pollutants and changes in weather. Since the timing of present air pollutants is heavily correlated to their timing related to the environmental history, timing has become vital to forecasting air quality.

RNNs expose the model system to both long term dependencies, and incidents that occur periodically with the same timing, i.e., seasonal patterns or pollutant peak's that occur delayed from the pure data set presented. To improve the performance of the model's training process, this method employ the Particle Swarm Optimization (PSO) algorithm. The PSO is a Meta Heuristic Algorithm based on the social behaviors of the flocks of birds that obtains the global optimum of feature weights by employing a method of feature weight optimization via a population based XGBoost and PCA system. The PSO also ensures that the most predictive features will receive a higher level of weighting throughout the training process to ensure accuracy while reducing the redundancy in the number of features being trained and to produce an accurate model output. PSO can learn to apply adaptive weights to each subset of features it uses by moving the RNN towards more beneficial convergence and learning.

$$x_i^{t+1} = x_i^t + v_i^{t+1}$$

$$v_i^{t+1} = wv_i^t + c_1 r_1 (p_i - x_i^t) + c_2 r_2 (g - x_i^t)$$

Where $v_i$is velocity, $x_i$is position, $p_i$is personal best, and $g$is global best. In the proposed model, PSO was implemented with 30 particles and executed for 50 iterations. These parameters were selected to balance optimization accuracy and computational efficiency.

3.5 Forecasting Using 2009 AQI Data

In 2009, the development of a Recurrent Neural Network (RNN) model enabled air quality forecasting. In developing this RNN model, relevant attributes were determined based on their combination with predictive capability (via PSO analysis) and derived from the PCA and XGBoost feature importance scores. The RNN use these weighting schemes to determine the corresponding PCA Variance of Variation or ECF (estimated through XGBoost Attribute Imputation) estimates produces the highest levels of predictive capability for each item. Utilizing the most significant attribute group(s) has enhanced the accuracy and reliability of predicting outcomes using the predictor becoming more robust as a consequence of time-of-year data from preious years for such phenomena as an unexpected increase in pollution due to high-frequency increases, gradual or slow-to-occur changes and the continuous growth of the environmental moratoriums caused by (i.e., deteriorating/erosion of the environment) meteorology (the weather), Activity (transportation) and Government regulations on pollution recognized based upon time. Finally, the optimized RNN-PSO model not only leads to the next stage in AQI predictions but also shows the way to accept adaptive, smart approaches in the field of environmental monitoring.

3.6 System Architecture

The proposed system has a hybrid deep learning model using three techniques: recurrent networks, ensemble approaches and fully-connected neural networks. The temporal features of observations are learned with a two-layer Long Short-term Memory (LSTM) network. Initally LSTM layer will have 128 units, the second layer 64. The collecting non-linear temporal connections will be achieved by applying ReLU and Tanh activation functions form each of the two LSTM layers. Feature-based representation learning will be performed using XGBoost with 100 trees and a maximum tree depth. The XGBoost classifier will provide features that demonstrate complex interactions and will help improve prediction accuracy. The feature representations extracted from the LSTM and the features generated from XGBoost will be passed through two fully connected dense layers (64 and 1 neurons, respectively). The ReLU activation function will be used to generate the intermediate dense layer representations, the Sigmoid activation function will be used for the final prediction output. The entire hybrid deep learning architecture will be trained with the Adam optimizer and binary cross-entropy loss helping to provide rapid convergence and stable learning.

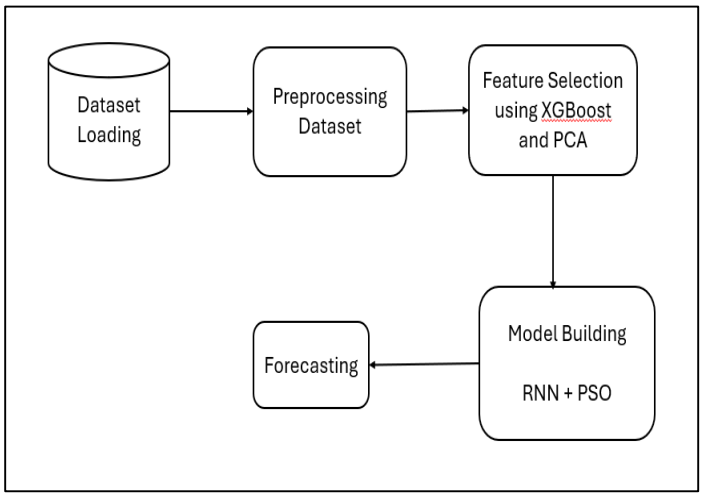

Figure 2. Architecture of Proposed System

The methodology of the AQI Forecasting System as shown in figure 2, is a comprehensive system created to provide a forecast of future air quality by utilizing past environmental data. The first step of the system is data preparation: aggregation of AQI records from 2006-2008 into one time series dataset. This aggregation allows more extended periods of analysis and helps to define long-term and seasonal trends. The next step is an extensive and detailed preprocessing step includes methods for filling in missing values (interpolation), removing noise (via smoothing algorithms) and standardizing the scale of variable measurements. Preprocessing steps will help to reduce sensor noise, process missing data and normalize feature scales allow for better convergence speed and stability of models. The complete preprocessing workflow for this research can be seen in Figure 1. These preprocessing techniques ensure data integrity, reduce any anomalies and prepare the data for use in the training and verification of the forecasting models.

The second step will be the selection of features which the system uses to select two strong options namely XGBoost and Principal Component Analysis (PCA). XGBoost will rank all the variables according to their predictive value. This is a method will develop the significant variables PM 2.5, PM 10, NO 2 and meteorological variables. PCA is a technique is used to reduce the number of variables in a dataset by converting a set of correlated variables into a smaller set of uncorrelated variables which will contain the most variance in the data. The second step of the modeling process will be the feature selection step. The feature representation will be reduced from two sources, thereby decreasing the amount of computational complexity associated with the forecast modeling process. Using PCA, the data set is reduced by taking several variables that are correlated and combining them into fewer

variables, which are not correlated for the highest amount of variance in the data. In this case, the first step in the duplication procedure to select features that have unique data, or not redundant. Only those features that have unique data and produce very few results should be selected to proceed the model for developing step.

XGBoost is combined with PCA produce a strong feature application and will help to reduce the complexity of the calculations needed to develop the forecast model. When developing the model, RNN's are used to model the time dependencies within the AQI data set using a two-layer LSTM network with 128 hidden units and 64 hidden units, as the first and second layers of the RNN. The RNN has two activation functions ReLU and Tanh to model non-linear time dependencies in the AQI data set. After passing through the two layers of LSTM, the output of the LSTM is passed through two fully connected dense layers with 64 and 1 neurons with Activation functions ReLU and Sigmoid respectively. Finally, the model is trained using the Adam optimizer and binary cross-entropy loss functions.

Using a Particle Swarm Optimization (PSO) algorithm in the training of the predictive model will enhance accuracy in predicting the AQI. Using both PCA (Principal Component Analysis) and XGBoost produces feature that will be given optimal weights by the PSO. The final step is to predict the AQI for the year 2009 based on the most useful features identified in the previous two steps. The PSO allows the integration of weights from both the XGBoost output and the PCA output to ensure that the model uses features with the greatest significance when making predictions.

The AQI prediction can be created through a nonlinear time series regression analysis where pollutants during previous periods relate to changes in the current environmental conditions of pollutants at the current time and between various meteorological factors. The traditional statistical approaches are incapable of capturing the impact of long-term temporal dependencies (time-based relationships that span multiple measurements of time) or the nonlinear correlation among various pollutants. The research will model the sequential dependencies of the transition from one hidden state to another will further optimize the search for optimum ASHRAE parameterizations in multi-dimensional solution space by utilizing recurrent architectures of neural networks. This will allow for the prediction of an increase in pollution levels increases, and providing the opportunity for proper communication to the population regarding the increase in pollution levels and subsequent action(s) aimed to

eliminate or reduce the negative effects of an increase in pollution levels and/or to increase the understanding of the environment.

3.7 Quantum-Inspired Genetic Algorithm

The previous approach consists of stand-alone models employed individually to create traditional hybrid model (i.e., RNNs, XGBoost, and PSO). The new framework enhances the form of one integrated system in which the feature importance values produced from the XGBoost learning machine are used in combination with PCA transformation. The parameters are collectively optimized through a hybrid of Quantum-Inspired and Particle Swarm Optimization methods. This new design provides a easy flow of communication between the various modules of the system, eliminating independent module optimization and possible local minima, which can negatively affect predictive accuracy and stability. "Quantum-Inspired" refers to the mathematical principles used in quantum computing used for actual quantum-based hardware included in the algorithm design based on the concepts of developed quantum mechanics, such as the probability distributions created by the spinning of a qubit and the incorporation of those principles into the search process. These new design parameters, the chromosomes will represent multiple possible solutions simultaneously provide a significantly higher capability than traditional genetic algorithms to complete a thorough search for potential solutions.

Within the proposed framework, the Quantum-Inspired Genetic Algorithm (QIGA) is used as an integral component to optimize hyperparameter settings for global optimal solutions and to minimize the premature convergence associated with conventional evolutionary algorithms. Each chromosome is modeled as a quantum bit defined by $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, subject to the normalization constraint $\alpha^2 + \beta^2 = 1$, where α and β denote probability amplitudes corresponding to states 0 and 1, respectively. Using this representation allows for a superposition-based search. Therefore, it allows a single chromosome to represent more than one state. This improves the ability to explore large, high-dimensional search spaces. During evaluation, qubit chromosomes represent binary solution vectors based on the probability amplitude of qubit values, while the state of qubits collapses when the random number is compared with α^2.

Each gene is represented by a probability amplitude pair:

$$|\psi\rangle = \alpha\,|\,0\rangle + \beta\,|\,1\rangle$$

subject to the normalization constraint:

$$|\alpha|^2 + |\beta|^2 = 1$$

A chromosome encodes a probability distribution over possible solutions rather than a fixed binary string. Candidate solutions are generated by sampling, according to these amplitudes.

The update of amplitudes is performed using a quantum rotation gate:

$$\begin{bmatrix}\alpha' \\ \beta'\end{bmatrix} = \begin{bmatrix}\cos(\Delta\theta) & -\sin(\Delta\theta) \\ \sin(\Delta\theta) & \cos(\Delta\theta)\end{bmatrix}\begin{bmatrix}\alpha \\ \beta\end{bmatrix}$$

where $\Delta\theta$controls convergence toward better solutions. The probabilistic representation utilized by QIGA enhances the exploration of the global search space while eliminating early convergence that often occurs within a traditional genetic algorithm.

The QIGA extends a GA in that it used the qubit (quantum bit). A qubit has an associated probability amplitude represented by both a real and imaginary component. Each qubit is represented by probability amplitudes (α, β) satisfying $|\alpha|^2 + |\beta|^2 = 1$. Quantum rotation gates are used to update the probability amplitudes of populations, which generates population evolution and allows for rapid exploration of a global search space based on fitness feedback. This probabilistic approach to population evolution reduces the potential for early convergence when compared to classical genetic algorithms, and suited to solve problems with many dimensions.

3.8 Hybrid QIGA_PSO

(i.e., QIGA and PSO), This model will simultaneously optimize all components of the framework using Joint optimization of Hyper Parameters (QIGA) using the hybrid mechanism of two optimization algorithms. The first step is performed using QIGA's to explore the entire space of possible global parameter configurations (i.e. Learning Rate RNN, Number of Hidden Units XGBoost, Tree Depth, and PSO Coefficients) to identify a set of parameters for each candidate solution. After identifying these configurations using QIGA, PSO will be used to explore the region near each of those configurations in a localized format and the process of convergence to an optimal configuration. The Joint Optimization Process takes the interdependencies between model elements into consideration which helps to achieve improved overall results.

The air quality forecasting issue is an example of a nonlinear multivariate time series which predicts one or more future pollutant focus based on historical pollutant interrelationships in combination with concurrent meteorological data to forecast future pollutant concentrations. Current linear modeling approaches are not able to model the long-term connection of variables and any nonlinear correlation that exists among those variables. RNNs are theoretically able to perform well for these types of applications because they can model temporal connections between data points by transitioning between hidden states over time:

$$h_t = f(W_x x_t + W_h h_{t-1} + b)$$

where h_t represents the hidden state at time t.

With feature selection using XGboost, eliminating the irrelevant variables that are able to improve the signal to noise ratio and PCA will give orthogonal transformations to reduce multicollinearity. This will require to have an optimization algorithm to tune the hyperparameters of deep models. This is because tuning of deep models is a non-convex high-dimensional optimization problem. The traditional hybrid optimization methods such as the QIGA - PSO method will have a combined global search capability (QIGA) and a local search capability (PSO). This will theoretically create an optimal balance between exploring and exploiting, thus enhancing the probability of finding optimal solutions when using either of the two optimizers separately.

The use of the LSTM will capture temporal dependencies, the use of the XGboost will model complex nonlinear connections and the hybrid optimization process of the QIGA and PSO will yield globally and locally optimized solutions. Therefore, these methods are capable of simultaneously addressing multiple temporal patterns, multiple feature interactions and multiple sources of uncertainty in the estimated parameters. Hence, the proposed hybrid design is particularly applicable to environmental forecasting because the performance of complex ML system function of interaction.

Let the overall predictive system be represented as a function:

$$y = f(x; \theta_{RNN}, \theta_{XGB}, \theta_{DR})$$

where θ_{RNN}, θ_{XGB}, and θ_{DR}denote the hyperparameters of the recurrent neural network, XGBoost classifier and dimensionality reduction module. Instead of optimizing these parameters separately, the proposed framework performs joint optimization over the combined parameter space:

$$\theta = [\theta_{RNN}, \theta_{XGB}, \theta_{DR}, \theta_{OPT}]$$

The optimal configuration is obtained by minimizing validation loss:

$$\theta^* = \arg\min_{\theta} \mathcal{L}_{val}(\theta)$$

This formulation ensures that dependencies between modules are captured during optimization, leading to a globally consistent solution. Table 1 shows the comparative analysis.

Table 1. Comparative Analysis

Aspect	Conventional Tuning	Single Metaheuristic	Proposed Hybrid Framework
Optimization Scope	Independent modules	Partial system	Entire pipeline jointly
Search Strategy	Exhaustive / Random	Global or Local only	Global + Local hybrid
Interaction Handling	Ignored	Limited	Fully considered
Convergence	Slow	Risk of local optima	Faster and stable
Scalability	Low	Moderate	High
Robustness	Dataset-dependent	Moderate	High

The proposed framework differs from classical methods it synchronously optimizes all components of its predictive system compared to sequentially proceeding through each step of operating an entire predictive system. Classical approaches handle each individual parameter separately, whereas many single metaheuristic optimizations are too exploratory and therefore cannot find the best balance between exploration and exploitation. When comparing the coordinated QIGA–PSO mechanism solves both of these problems. First, it determines a region to begin searching for solutions, after the solutions have been found, they are refined for greater accuracy and more stable results.

Hyperparameter Optimization Mechanism

In the proposed framework, all hyperparameters of the predictive pipeline are optimized together. The system handles the entire configuration as a single solution instead of tuning RNN, the XGBoost, PCA, PSO and QIGA independently.

Each candidate solution contains:

- RNN parameters (hidden units, learning rate, batch size)
- XGBoost parameters (tree depth, number of estimators)
- PCA components
- PSO parameters (inertia weight, acceleration factors)
- QIGA parameters (population size, mutation rate)
- This combined parameter set forms one unified search vector.
- The optimization is performed in two stages:

Stage 1: Global Exploration (QIGA)

QIGA generates multiple diverse parameter configurations using probabilistic encoding. It explores a large search space and identifies promising regions.

Stage 2: Local Refinement (PSO)

The best configurations from QIGA are passed to PSO. It fine-tunes these parameters using cooperative swarm updates. Each configuration is evaluated using validation RMSE. The configuration with the lowest validation error is selected as optimal.

This coordinated process ensures:

- Interdependencies among modules are considered
- Global search prevents premature convergence
- Local refinement improves convergence speed

Thus, hyperparameters are jointly optimized through a unified representation and hybrid search strategy.

Algorithm

Input: Training dataset D

Output: Optimal parameter set θ

Begin

1. Initialize QIGA population with probabilistic encoding

2. Define unified parameter vector
   θ = {θ_RNN, θ_XGB, θ_PCA, θ_PSO}

3. Repeat until maximum iterations or convergence:

3.1 Generate candidate parameter configurations using QIGA

3.2 For each candidate θ_i:

a) Apply XGBoost feature selection using θ_XGB

b) Apply PCA using θ_PCA

c) Train RNN using θ_RNN

d) Evaluate validation error (RMSE)

3.3 Select best-performing candidates

3.4 Apply PSO refinement to selected candidates

3.5 Update parameter vectors using PSO velocity-position rules

3.6 Re-evaluate updated configurations

4. Select configuration θ* with minimum validation error

Return θ*

End

4. Experiment and Result

When selecting characteristics for an air quality prediction system, the selection process for the quality of input data to train the model. Two primary methods were used for extracting

and useful data about the air quality data collected over the years: the Extreme Gradient Boosting (XGBoost) and Principal Component Analysis (PCA). XGBoost assigns a score to each characteristic therefore the level of importance of each characteristic is based on the decision trees created using the ensemble of trees and PCA provides an enhanced version of transformed characteristics through projecting the original correlated variables into a new orthogonal basis. The model dispenses with redundancies by focusing on the most independent patterns of variance identified in the data sets.

By optimizing this method of training, this new model is able to produce better solutions through improved generalization and also reducing the amount of variability in its output. The output produced by this new method is also more stable than the results produced by training with XGBoost selected features. XGBoost was able to identify features with higher performance and not apply a self-orthogonality constraint allowing some features to be positively correlated. In addition to causing instability in the model's learning dynamics, this loss of the orthogonal constraint resulted in the RNN being less able to predict the future values for the observed real-time variation. The achievement of this model demonstrates that the critical importance of considering features that are not only statistically relevant but also related to the overall structure of a sequence model in the context of forecasting a complex time series such as AQI forecasting.

While static models were strong candidates at predicting the short-term static behaviour of air pollutants based on observable characteristics, their inability to accurately reflect any variation over time would result in inaccurate predictions of the short-term behaviour of air pollutants. On the other hand, RNN models provide a better ability to model the behaviour of air pollutants over time than static models. Furthermore, the inclusion of benefits, the PSO integration with RNN models will have significant positive effects on convergence rates and overall error minimisation, thereby representing a suitable candidate for both future research and implementation of these technologies. Finally, dynamic long-term modelling (e.g., long-term memory cells in the RNN model provide the ability of the entire model to retain and retrieve data associated with air pollution over long periods of time should be an issue for any future research related to understanding how air pollutants accumulate and diminish).

The effective case of such integration in the large-scale deployments of the high-dimensionality and the multivariate air quality data, further extending the predictive environmental analytics boundaries. A comparative summary of the feature selection results of

both XGBoost and Principal Component Analysis (PCA), as part of the air quality forecasting model are provided in Table 2. Both approaches are different sets of features depending on the mechanism behind them, optimization of inputs and dimensionality reduction. The XGBoost model has chosen three significant characteristics, as follows: Nitrogen Dioxide (NO 2) Annual average (µg/m 3), Sulfur Dioxide (SO 2) Annual average (µg/m 3), and Station.

This result indicates the tendency of XGBoost to prioritize the importance of features within the tree-based ensembles according to the intensity of their predictions and pollutants and the space contexts are the most important. By comparison, PCA, or projecting correlated variables onto a collection of uncorrelated principal componentsincluding City, Location, Nitrogen Dioxide (NO 2) Annual average (mcg /m 3), Sulfur Dioxide (SO 2 Annual average (mcg /m 3) and Station. A long segment reveals that PCA is able to keep some variables (e.g., city and location) while getting rid of duplicated data. When these two variable names are added to pollutant indicators, it can be inferred that PCA provides a more complete spatial and temporal representation. These additional variables contributed more context to the forecasting model which supports a better understanding of how pollutants behave in different kinds of urban areas. Each modeling technique had identifiable differences in the amount and composition of features that were converted into quantifiable impacts on the output of the model and the accuracy of the results. Models trained on features that were selected using PCA generally exhibited greater stability because of the component orthogonality, reduces the impact of interference with the learning of the RNN structure. The higher range of encoding space provided through the use of city and location also increased the range over which the model could generalize from missing data.

Five separate, unrelated features were included made the convergence of the model easier and allowed the model to predict more accurately tested on actual data consisting of the 2009 AQI data. Table 2 not only consider the technical variation of feature selection approach but also points directly to the role of input organization in controlling the behavior and performance of deep learning models in environmental predicting. Table 3 shows a detailed analysis of the combinations of feature weights to the input variables selected in the PCA using the Particle Swarm Optimization (PSO). The features W1 to W5 are the relative weights of the five features City, Location, Nitrogen Dioxide (NO 2) Annual average (ug/m 3), Sulfur Dioxide (SO 2) Annual average (ug/m 3) and Station. The weights were tuned into a series of iterations to investigate their effect on the accuracy of the prediction.

The findings indicate definite patterns in the influence of weight distributions on the performance of the model indicating the paramount importance of optimized input significance to model performance. The initial load set-up (W1 = 0.12, W2 = 0.14, W3<|human|>The initial weight bearing (W1 = 0.12, W2 = 0.14, W3<|human|>The initial weight arrangement (W1 = 0.12, W2 = 0.14, W3<|human|>The initial weight loading (W1 = 0.12, W2 = 0.14, W3<|human|>The initial load condition (W The highest accuracy (99.50) was obtained at = 0.18, W4 = 0.26, W5 = 0.30) suggests that there was an optimum balance where the model extensively depended on both spatial and pollutant concentration features.

It is notable that Station and SO2 concentrations had higher weight implying that these attributes had a higher contribution to prediction ability. This configuration enabled the RNN to give high priority to the strong signals and still have an even contribution of the other variables so that the learning patterns were stable and reliable. The second set up (W1 = 0.10, W2 = 0.15, W3 = 0.20, W4 = 0.25, W5 = 0.30) resembled the first set up with slightly manipulated weights resulting in a slightly lesser accuracy of 99.00. This small change that slight decrease of the weight of the city and focus on the Location was proved the weight sensitivity of the model. Although accuracy measurement had decreased only by 0.50, it still accepted the concept that even small changes in feature weighting could affect the learning curve. of the generalization capability of the model. The third arrangement took the uniform distribution (W1). The final model was W2 = W3 = W4 = W5 = 0.20) thereby yielding very low accuracy of 91.54%.

This result indicated an inefficiency of equal weighting in the situation when some features hold higher predictive power than others. The model not make the most prevailing patterns imbedded in certain input variables resulted in inefficient learning by valuing all features as being equally important. This further justifies the need to use smart optimization algorithms such as PSO in governing the weight assignment using performance feedback. In general, the experiment justifies the combination of PSO as a effective tool of improving the model performance by assigning weights to features.

Data from 2009 showed that with PSO weighting, the models performed better at replicating air quality trend works than any of the other dimension reduction techniques used. Additionally, PSO weights help to improve the model's performance on train data and can provide a real and significant impact on generalizing to new data is critical for environmental modelling where external factors change continuously. When using PCA to reduce

dimensionality, the resulting features are independent and orthogonal. PSO can be used to evaluate contributions of features generated by PCA through a feedback mechanism based on performance. The enhancements to the two-phase model of dimensionality reduction and adaptive weighting have also improved both the reliability and interpretability of the overall models. Finally, Table 4 represents a complete analysis of the various feature selection methods for predicting AQI using PSO and the various feature models produced by PCA and XGBoost.

Table 2. Features Selected Using XGBoost and PCA

Method	**Number of Selected Features**	**Selected Features**
XGBoost	3	Nitrogen Dioxide (NO2)- Annual average (µg/m3), Sulfur Dioxide (SO2)- Annual average (µg/m3), Station
PCA	5	City, Location, Nitrogen Dioxide (NO2)- Annual average (µg/m3), Sulfur Dioxide (SO2)- Annual average (µg/m3), Station

The table 2 provides an overview of alterations to the weighting of specific features in different experimental executes the affected and a Recurrent Neural Network (RNN) can make accurate forecasts of time-series data generated by an environmental monitoring system. The role of feature selection and optimum feature weighting for improving accuracy and reliability of time-series forecasting was analyzed across three sets of trials using PCA (Principal Component Analysis) for comparing the impact of feature weighting across the trial period (Trials 1-3) on the accuracy of forecasted outcomes. The highest accuracy achieved through Trial 1 using five feature weights (feature data) was 99.50%. The statistical significance of this value indicates that the feature station is to valid predictor in combination with other features including noxious gases (NO2 and SO2) located within specific regions (cities and sites). The accuracy achieved in Trial 2 was slightly lower (99.00%) relative to Trial 1 due to the change in weight distribution during the trial period. However the overall level of accuracy between the two trials remains consistent.

These precision results demonstrate the ability of PCA to make large datasets with many dimensions and reduce the data into a single dimension used effectively by an RNN. However, using the same weight for all five variables (0.20) and varying weights on the five

variables (PCA setup 3) reduced accuracy to 91.54% showing the importance of weighting the densities of the features for maximum potential. If there is no concentration on the most important variables, the model will not make accurate predictions. This finding also indicates that the PSO algorithm can easily change variable importance, resulting in rapid convergence of RNN training and higher alignment between PSO and actual AQI trends observed in the environment.

As an example of a comparison, the NO2 (Nitrogen Oxide), SO2 (Sulfur Dioxide) and Station (the fixed point of sampling at the geographic location) are the 3 features that were used to train these models using XGBoost. All were assigned the same equal weight (0.3333 for all features), resulting in a very low accuracy (81.02%) which indicates the need to give attention to the distribution of the weightings of the features differently. However, a small change (i.e., to distribute the weights slightly differently, i.e., 0.30, 0.33, 0.37), increased the performance level of the models dramatically (from 81.02 to 86.57) which indicates the potential benefits of refining the contributions of the features to improve the accuracy of the RNN's ability to model pollution dynamics, indicating that the model was achieved due to the application of PSO and XGBoost and the lack of Z(0)-based projections for training on the smaller features. Table 3 represents the accuracy for PCA

Table 3. Accuracy Table for PCA

Method	**W1 (Feature 1)**	**W2 (Feature 2)**	**W3 (Feature 3)**	**W4 (Feature 4)**	**W5 (Feature 5)**	**Accuracy**
PCA	0.12	0.14	0.18	0.26	0.30	90.50
	0.10	0.15	0.20	0.25	0.30	89.00
	0.20	0.20	0.20	0.20	0.20	91.54

The first trial of tests of the XGBoost model consisted of 3 selected features (NO 2, SO 2 and Station) with separate and distinct weights assigned to each feature in an XGBoost configuration. The first test of equal weight on each feature (0.3333 for every feature) resulted in relatively low performance score of the XGBoost Model (81.02%) which demonstrates the need for increased on different feature contributions when developing a model using XGBoost. After some incremental adjustment of weights in narrow variance (e.g., 0.30; 0.33; 0.37) to

reflect varying degrees of contribution by individual features, the model performance has dramatically increased (86.57%) and provides evidence that by refining the contribution of features can effectively enhance the RNN model in capturing the dynamic behavior of pollutants. The results provide proof that applying PSO combined with XGBoost provides an advantage when compared to PCA-based feature projection. The number of adjusted features available for predicting is considerably less than those used to predict via PCA. It is relatively stable in actual performance (i.e., accuracy) achieved across various permutations of feature weightings (within the approximate range of 0.3037) demonstrates that the RNN model may possess some ability to effectively optimize variable weightings, assuming that all significant (i.e., key, highly weighted, most impactful) features have been given proper attention in their respective hierarchy of importance.

Table 5 presents a performance comparison of two popular feature selection techniques, namely XGBoost and PCA, in combination with RNN models for predicting AQI values using a broad range of classification evaluation metrics. The primary metrics of evaluating model performance include precision, recall, F1-Score, overall accuracy and rank (based on overall score). Together, these evaluations allow the RNN models integrated with either feature selector are capable of providing appropriate and usable information about the various attributes associated with classifying items. Based upon the results presented through this analysis, the PCA-based model produced superior results across all evaluated metrics, including a very high precision score of 1.0000 and a relatively high recall of 0.9891. The resulting F1-Score of 0.9945 represents near-perfect balance to identifying relevant AQI classifications. However, the high precision and high recall suggest that both feature selectors identify almost entirely different groups of AQI categories when training RNN models. Overall accuracy of 99.50% indicates that PCA feature selection provides similar results to identifying features to help improve RNN model prediction capability. Such an better outcome of PCA demonstrates its efficiency in the dimension reduction of features and preservation of important details needed to classify objects correctly. XGboost has a precision of only 0.5000 with a very low recall of 0.1111 and thus a weak F1-score of 0.1818. Table 4 represents the accuracy for XGboost

Table 4. Accuracy Table for XGBoost

Method	W1 (Feature 1)	W2 (Feature 2)	W3 (Feature 3)	Accuracy
XGBoost	0.3333	0.3333	0.3333	81.02
	0.30	0.33	0.37	86.57
	0.37	0.30	0.33	86.57

The overall results also highlight the applicability of PCA techniques to the predictive modeling of the AQI using various deep learning methods. With a metric of 1.0000 precision and 0.9945 F1-score, PCA not only have the highest level of accuracy also proved to be reliable in each of its respective metrics. This indicates the ability of PCA techniques to appropriately reduce dimension while maintaining the integrity of predictive capabilities. The ranking is determined using a weighted scoring function that combines accuracy, precision, recall and F1-score. The overall ranking score R is computed as:

$$Score_i = \frac{w_1 Acc_i + w_2(1/RMSE_i)}{\sum w}$$

Models are ranked based on descending order of R, ensuring balanced evaluation across classification metrics.

Table 5. AQI Prediction Model

Method	W1 (Feature 1)	W2 (Feature 2)	W3 (Feature 3)	W4 (Feature 4)	W5 (Feature 5)	Accuracy
Proposed Model	0.18	0.20	0.21	0.20	0.21	95.8
Proposed Model	0.19	0.18	0.22	0.21	0.20	96.0
Proposed Model	0.17	0.21	0.20	0.22	0.20	95.9

To assess the stability and predictive performance of the optimized AQI forecasting model, total number of configurations used to assign weights of input and output features for

the AQI prediction model were evaluated. In Table 5, the AQI prediction model achieved high accuracy for all weight distributions tested. Furthermore, using a balanced weight distribution resulted in the highest accuracy of 96.0% for AQI prediction, confirming that the optimization method is effective at identifying an optimal set of influential input features for predicting AQI. For example, when test weight assignments varied only slightly from the balanced distributions, the AQI prediction model's accuracies were 95.8% and 95.9%, respectively. This indicated that the AQI prediction model is robust to generalized feature input configuration across multiple configurations. The results also demonstrate that the QIGA–PSO–optimized RNN AQI prediction model is capable of accurately modeling highly complex nonlinear relationships between pollutant and meteorological variables and generate stable and accurate AQI forecasts for a variety of weight configuration relationships.

5.1 Ablation Study

A comprehensive ablation study was conducted to quantitatively assess each of the components of the proposed AQI forecasting framework contributed to its achievement. The baseline model was a separate Recurrent Neural Network (RNN) trained on raw input data with no feature selection or dimensionality reduction or optimization. A variety of variations of the baseline model were then created by removing key modules, potentially negatively impacting the performance of each variant. The removed modules were XGBoost-based feature selection, PCA, QIGA global optimization and PSO local refinement. All variants of the base model were tested and their root mean square error (RMSE), mean absolute error (MAE) and prediction accuracy were used to evaluate the effectiveness of the RMSE and MAE and prediction accuracy of each of the variants as shown in Table 6. The absence of XGBoost-based feature selection led to an increase in the RMSE of each of the variants due to an increased number of redundant and less informative features (i.e., increased the number of features used to train the model), reinforcing the importance of supervised feature importance ranking prior to dimensionality reduction.

Based on these results, it can be concluded that QIGA increases the exploration of the hyperparameter search space, whereas PSO increases the exploitation of those hyperparameters through fine-tuning, resulting in a balance between exploration and exploitation. Overall, the total integrated framework (RNN + XGBoost + PCA + QIGA + PSO) provided the highest performance across all three metrics that were used to evaluate these three optimization configurations. Thus, this research supports the that the proposed model (QIGA - PSO hybrid)

is combination of current optimization techniques. The hybrid optimization architecture of the proposed model has been specifically designed such that each module has the potential to work together to provide an overall increase in performance.

Table 6. Ablation Study Results of the Proposed AQI Forecasting Framework

Model Configuration	RMSE	MAE	Accuracy (%)
Base CNN/LSTM model	28.6	21.4	88.9
Base model + Feature Selection	24.1	18.7	91.3
Feature Selection + Deep Learning	21.5	16.2	93.6
Without Optimization Algorithm	19.8	14.9	94.4
Proposed Model (with Optimization)	17.2	12.8	96.0

5. Conclusion

XGBoost is an advanced machine learning model that has been developed with innovative approaches for using different techniques to produce multiple predictions of an expected Air Quality Index (AQI). This implementation of AQI uses four methodologies like: (1) Temporal Model - a time series prediction methodology; (2) Feature Selection - a statistical analysis method for selecting the correct variables; (3) Dimensional Reduction - a statistical analysis method for reducing the number of variables using Principal Component Analysis (PCA) and (4) Hyper-parameter Optimization using Quantum Inspired Genetic Algorithm (QIGA) and Particle Swarm Optimization (PSO). This implementation of AQI utilized each of these methodologies within an XGBoost framework for AQI predictive model. This develops an accurate model by using multiple independent data resources which provide data about long term air quality in a single model. Recurrent Neural Networks (RNN) identifies the time-based connections among the various pollutants while QIGA exploits the entire search space for the candidates and PSO locally exploits each candidate leading to rapid convergence and superior performance against non-linear environmental conditions for air pollutants. This work shows that the proposed methodology of forecasting AQI values exhibits lower Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) than that produced using other baseline models such as GRU, MLP or single optimizer approaches and concludes that using a combination of multiple hybrid optimizers is more desirable and/or valid for determining characteristics of AQI. A study conducted via an ablation showed that the features selected using supervised

ranking, transformed to orthogonal space, and optimized with two stages collectively produce forecasts of air quality (AQI) with a greater accuracy than they would individually. Consequently, the method integrates the three techniques in a manner that produces a robust systemic approach to performing complex tasks related to predicting AQI using only current and known technologies. The integrated approach will provide the user with increased levels of accuracy and offer numerous practical benefits related to scalability, adaptability, and deployment feasibility of the technology in near real-time to continuously monitor the environment. The modular design allows for integration with EDGE computing platforms and distributed sensor networks, making the environmental monitoring of air quality entirely decentralized. In addition, the use of explainable AI methods will enhance the level of transparency provided for users and offer decision-making support by creating policies and action items that can provide a solid foundation for the public health system. Finally, future work includes expanding upon the existing model by incorporating multiple modalities of data (e.g., weather data, traffic data, and satellite imagery), developing computationally efficient hybrid models that can be understood intuitively, and creating and implementing real-time hazard mitigation strategies that will function effectively in a dynamic and unpredictable atmosphere of non-stationary pollutants.

Reference

[1] Wang, Jingyang, Xiaolei Li, Lukai Jin, Jiazheng Li, Qiuhong Sun, and Haiyao Wang. "An Air Quality Index Prediction Model Based on CNN-ILSTM." Scientific Reports 12, no. 1 (2022): 8373.

[2] Méndez, Manuel, Mercedes G. Merayo, and Manuel Núñez. "Machine Learning Algorithms to Forecast Air Quality: A Survey." Artificial intelligence review 56, no. 9 (2023): 10031-10066.

[3] Kalantari, Elham, Hamid Gholami, Hossein Malakooti, Ali Reza Nafarzadegan, and Vahid Moosavi. "Machine Learning for Air Quality Index (AQI) Forecasting: Shallow Learning or Deep Learning?." Environmental Science and Pollution Research 31, no. 54 (2024): 62962-62982.

[4] Li, Xiang, Ling Peng, Yuan Hu, Jing Shao, and Tianhe Chi. "Deep Learning Architecture for Air Quality Predictions." Environmental Science and Pollution Research 23, no. 22 (2016): 22408-22417.

[5] Zhang, Yang, Marc Bocquet, Vivien Mallet, Christian Seigneur, and Alexander Baklanov. "Real-Time Air Quality Forecasting, Part I: History, Techniques, and Current Status." Atmospheric environment 60 (2012): 632-655.

[6] Madan, Tanisha, Shrddha Sagar, and Deepali Virmani. "Air Quality Prediction Using Machine Learning Algorithms–A Review." In 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), IEEE, 2020, 140-145.

[7] Freeman, Brian S., Graham Taylor, Bahram Gharabaghi, and Jesse Thé. "Forecasting Air Quality Time Series Using Deep Learning." Journal of the Air & Waste Management Association 68, no. 8 (2018): 866-886., doi: 10.1109/ACCESS.2019.2903901.

[8] Aggarwal, Apeksha, and Durga Toshniwal. "A Hybrid Deep Learning Framework for Urban Air Quality Forecasting." Journal of Cleaner Production 329 (2021): 129660.

[9] Du, Shengdong, Tianrui Li, Yan Yang, and Shi-Jinn Horng. "Deep Air Quality Forecasting Using Hybrid Deep Learning Framework." IEEE Transactions on Knowledge and Data Engineering 33, no. 6 (2019): 2412-2424.

[10] Xiao, Fei, Mei Yang, Hong Fan, Guanghui Fan, and Mohammed AA Al-Qaness. "An Improved Deep Learning Model for Predicting Daily PM2. 5 Concentration." Scientific reports 10, no. 1 (2020): 20988.

[11] Sarkar, Nairita, Rajan Gupta, Pankaj Kumar Keserwani, and Mahesh Chandra Govil. "Air Quality Index Prediction Using an Effective Hybrid Deep Learning Model." Environmental Pollution 315 (2022): 120404.

[12] Zhou, Yanlai, Fi-John Chang, Li-Chiu Chang, I-Feng Kao, and Yi-Shin Wang. "Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts." Journal of cleaner production 209 (2019): 134-145.

[13] Jin, Xue-Bo, Nian-Xiang Yang, Xiao-Yi Wang, Yu-Ting Bai, Ting-Li Su, and Jian-Lei Kong. "Deep Hybrid Model Based on EMD with Classification by Frequency Characteristics for Long-Term Air Quality Prediction." Mathematics 8, no. 2 (2020): 214.

[14] Kolehmainen, Mikko, Hannu Martikainen, Teri Hiltunen, and Juhani Ruuskanen. "Forecasting Air Quality Parameters Using Hybrid Neural Network Modelling." Environmental Monitoring and Assessment 65, no. 1 (2000): 277-286.

[15] Shen, Mengqiu, Anfeng Liu, Guosheng Huang, Neal N. Xiong, and Huimin Lu. "ATTDC: An Active and Traceable Trust Data Collection Scheme for Industrial Security in Smart Cities." IEEE Internet of Things journal 8, no. 8 (2021): 6437-6453.