Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Quantum machine learning performance evaluation for air pollution assessment with benchmarking and comparative analysis using multi-metric scoring and quantum circuit expressibility

Jagadish Kumar Mogaraju

jagadishmogaraju@gmail.com

IUCN https://orcid.org/0000-0002-6461-8614

Research Article

Keywords: Quantum Machine Learning, Air Quality Index, Benchmarking, Multi-metric scoring, Quantum Circuit Expressibility, Entangling power

Posted Date: July 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10290235/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Title page

Quantum machine learning performance evaluation for air pollution assessment with benchmarking and comparative analysis using multi-metric scoring and quantum circuit expressibility

Author information

Jagadish Kumar Mogaraju

Commission member,

Climate Crisis Commission (CCC),

International Union for Conservation of Nature (IUCN),

New Delhi, India, 110011.

Email of the corresponding author - jagadishmogaraju@gmail.com

Corresponding author - Jagadish Kumar Mogaraju

ORCID: https://orcid.org/0000-0002-6461-8614

Abstract

Our study proposes the MDQ-3L framework, a multi-metric, quantum-aware hierarchy for benchmarking classical machine learning, deep learning, and quantum machine learning models in air quality prediction. PM2.5, PM10, NO, NO2, NOx, NH3, CO, SO2, O3, Benzene, Toluene, and Xylene were considered predictors of AQI. Seventeen ML regressors, nine DL models, and seven QML circuits were evaluated using MAE, MSE, RMSE, R^2, RMSLE, MAPE, and QML-specific expressibility (0.739-2.0) and entangling power (0-0.398). Aggregated via min-max normalization and weights, the framework penalizes negative R^2 and high structural costs, yielding CatBoost and DBN as top performers, while QML trailed by 25-40% in composite scores due to generalization gaps and barren plateau risks. MDQ-3L aligns circuits to targets, reducing QCNN/TTN penalties by ~30% relative to non-benchmarked rankings. Our approach exposes QML limitations under NISQ constraints (4-6 qubits, shallow depth), supporting low-entanglement ansätze like SELR/HEA for 10 - 15% potential gains, and provides an interpretable tool for hybrid environmental modeling, advancing Quantum Machine Intelligence benchmarks.

Keywords

Quantum Machine Learning, Air Quality Index, Benchmarking, Multi-metric scoring, Quantum Circuit Expressibility, Entangling power

Quantum machine learning performance evaluation for air pollution assessment with benchmarking and comparative analysis using multi-metric scoring and quantum circuit expressibility

Abstract

Our study proposes the MDQ-3L framework, a multi-metric, quantum-aware hierarchy for benchmarking classical machine learning, deep learning, and quantum machine learning models in air quality prediction. PM2.5, PM10, NO, NO2, NOx, NH3, CO, SO2, O3, Benzene, Toluene, Xylene were considered as predictors for AQI. Seventeen ML regressors, nine DL models, and seven QML circuits were evaluated using MAE, MSE, RMSE, R^2, RMSLE, MAPE, and QML-specific expressibility (0.739-2.0) and entangling power (0-0.398). Aggregated via min-max normalization and weights, the framework penalizes negative R^2 and high structural costs, yielding CatBoost and DBN as top performers, while QML trailed by 25-40% in composite scores due to generalization gaps and barren plateau risks. MDQ-3L aligns circuits to targets, reducing QCNN/TTN penalties by ~30% relative to non-benchmarked rankings. Our approach exposes QML limitations under NISQ constraints (4-6 qubits, shallow depth), supporting low-entanglement ansätze like SELR/HEA for 10 - 15% potential gains, and provides an interpretable tool for hybrid environmental modeling, advancing Quantum Machine Intelligence benchmarks.

Keywords

Quantum Machine Learning, Air Quality Index, Benchmarking, Multi-metric scoring, Quantum Circuit Expressibility, Entangling power

Introduction

Air is an essential resource for life, and its quality is being degraded faster than expected, leaving us in an anthropogenically induced, subtle situation that demands immediate remedial action from all concerned parties [1]. Machine learning (ML) models were trained on quantum computers using superposition and entanglement to enhance SVM performance in predicting the air quality index (AQI) [2]. Quantum convolutional neural network (QCNN) model was improved using an enhanced quantum temporal convolutional network (QTCN) to predict $PM_{2.5}$ levels [3]. ML-based quantum chemical methods (ML-QCMs) can be used to search global minima for atmospheric nanoclusters, predict degradation networks of chemical pollutants, and to investigate the heterogeneous or photochemical transformation of pollutants (pathways) [4]. Quantum genetic algorithm - Extreme learning machine (QGA-ELM) model was used to investigate SO_2 concentration and NO_x concentration with generalization ability and good accuracy [5]. Increased ansatz repetitions (quantum neural network and graph-based deep learning structure) integrated with min-max normalization enhanced the model (hybrid) performance [6]. Quantum Support Vector Machine (QSVM) yielded better binary classification via a quantum kernel based on angle encoding than the Long Short-Term Memory (LSTM) model [7]. Quantum Support Vector Classifier (QSVC) and the Variational Quantum Classifier (VQC) models were used to investigate the potential of Quantum machine learning (QML) techniques on climate and weather data [8]. Quantum-inspired models were compared to classical ML models while predicting the world surface temperature [9]. A hybrid deep learning model containing XGBoost, Quantum Particle Swarm Optimization (QPSO)-enhanced-Long Short-Term Memory (LSTM), Autoregressive Integrated Moving Average (ARIMA), and Attention Convolutional Neural Networks (ACNN) was used to predict the air quality index [10]. AirVCQnet (hybrid model) containing variational mode decomposition (VMD) method and a quantum long short-term memory (QLSTM)

network was proposed to study Nitrogen Dioxide (NO_2), Particulate Matter $PM_{2.5}$, PM_{10}, Ozone (O_3), and Sulphur Dioxide (SO_2) [11]. Quantum Kernel-Based Long Short-Term Memory (QK-LSTM) network was used to improve predictive accuracy in climate time-series forecasting [12]. Differential gravitational fireworks algorithm-optimized support vector machine (SVM) was integrated with an enhanced particle swarm optimization algorithm to assess air pollution [13]. The quantum activation function optimized hybrid quantum neural network algorithm yielded relatively prediction accuracy for the air quality index [14]. Atmospheric cluster formation rates were studied using Quantum chemical (QC) calculations [15]. By integrating tensor-network parameterizations and quantum neural operators (QNOs), a Quantum-Inspired Neural Radiative Transfer (QINRT) framework was developed for Climate Intelligence [16]. Quantum-behaved particle swarm optimization (QPSO) algorithm was used to choose optimal parameters affecting SVR in predicting $PM_{2.5}$ and NO_2 [17]. Multiple optimizers were used to optimize the quantum regressor, variational ansatz, and a feature map to produce a hybrid quantum-classical machine learning model (quantum-enhanced neural networks) to predict Nitrogen Dioxide (NO_2) and Carbon Monoxide (CO) [18]. Extreme Learning Machine (ELM) optimised by Quantum Firefly Optimisation Algorithm (ELM-QFOA) and SVR integrated with QFOA were used to investigate carbon emissions [19]. Quantum Causal Attention Mechanism (QCAM) using Quantum State Mapping (QSM) was used to extract higher-order features in predicting $PM_{2.5}$ concentration [20]. Quantum annealing was used to optimize Backpropagation Neural Network (BPNN) models for predicting greenhouse gas emissions [21]. Quantum feature mapping was used to efficiently handle the high-dimensional features in the air quality datasets [22]. A quantum quanvolutional (replacing the convolutional layers with quantum circuits) neural network model was proposed using baseline models, intersection over union accuracy, and per-pixel loss metrics [23]. A quantum-inspired online spiking neural network (QiSNN) was developed using a quantum particle swarm optimization algorithm and a Kalman filtering technique to denoise the time-series data while investigating the prediction process for ozone and PM_{10} concentrations [24]. CO_2, TVOC, $PM_{2.5}$, and PM_{10} were predicted using Quantum Machine Learning (QML) while studying Sick Building Syndrome (SBS) [25]. An evolving interval type-2 quantum fuzzy neural network (eIT2QFNN) was modified to develop an intelligent hybrid air-quality forecasting system [26]. The works cited here partially addressed the Quantum machine learning performance evaluation for air pollution assessment with benchmarking and quantum circuit expressibility. A Deep Feedforward Neural Network (DFFN) was used to classify air pollution into different levels based on the concentration of $PM_{2.5}$, O_3, NO_2, and CO [27]. A channel attention-based ResNet50 architecture with a deep learning model was used to classify the levels of air quality pollution [28]. An AERODTL-SIC method with a deep convolutional neural network-based SqueezeNet was used to extract features from the remote sensing images [29]. A classification system that incorporates MobileNetV2 equipped with a Spatial Attention method was used to enhance the recognition of vital visual pollution features [30]. We attempted to demonstrate benchmarking of regression models (AQI) under a multi-metric and quantum-aware hierarchy. We also attempted to propose a MDQ-3L framework which uses Machine Learning, Deep Learning, and a QML comparative benchmark.Our work was aimed at filling the research gaps by applying quantum models and machine/deep learning models to assess the air quality index using independent variables $PM_{2.5}$, PM_{10}, NO, NO_2, NO_x, NH_3, CO, SO_2, O_3, Benzene, Toluene, and Xylene.

Methodology

The datasets (2015 to 2020) for the study were obtained from the Central Pollution Control Board, Ministry of Environment, Forest and Climate Change, Government of India, through https://airquality.cpcb.gov.in/ccr/#/caaqm-dashboard-all/caaqm-landing and https://cpcb.nic.in/namp-data/. We attempted to design a methodology that can aid us in comparing classical machine learning(ML), deep learning (DL), and quantum machine learning (QML) models to predict the Air Quality Index (AQI). We tried to evaluate the model yields under a multi-metric and quantum-aware framework. We considered $PM_{2.5}$, PM_{10}, CO, O_3, NO_2, NO_x, NO, SO_2, NH_3, Benzene, Toluene, and Xylene as predictors and AQI as a target. The air quality dataset contained 257834 observations with 13 features. We used training data (80%: 206268) and test data (20%: 51567). Data preprocessing was done using normalization (z score), outliers were addressed using the Isolation Forest algorithm, features with high inter-correlations (Pearson threshold 0.9) were removed to lower model instability, imputation was applied using mean to address missing values, and to address the heteroscedasticity and skewness, the Yeo-Johnson transformation method was applied. To investigate ML, we used CatBoost Regressor, Extra Trees Regressor, Extreme Gradient Boosting, Random Forest Regressor, K Neighbors Regressor, Light Gradient Boosting Machine, Gradient Boosting Regressor, Ridge Regression, Linear Regression, Bayesian Ridge, Least Angle Regression, Lasso Least Angle Regression, Lasso Regression, Huber Regressor, Elastic Net, Orthogonal Matching Pursuit, AdaBoost Regressor, Passive Aggressive Regressor, and Decision Tree Regressor models (more information about the models can be obtained from https://scikit-learn.org/stable/model_selection.html). We considered evaluation metrics, i.e., MAE, MSE, RMSE, R^2, RMSLE, and MAPE, to compare the model performance. To understand DL, we used a Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Deep Belief Network (DBN), Deep Boltzmann Machine (DBM), Stacked Autoencoder (AE), and Transformer-based regressor (Transformer) models (more information about the models can be obtained from https://keras.io/guides/training_with_built_in_methods/ and https://scikit-learn.org/stable/supervised_learning.html). We used evaluation metrics like MAE, MSE, RMSE, R^2, and RMSLE. To comprehend Quantum Machine Learning, we used Random Layers Regressor (RLR), RYCZ, customized (Shifted Entangled Linear Regressor), Hardware Efficient Ansatz (HEA) [31], Quantum Convolutional Neural Network (QCNN) [32], Tree Tensor Network (TTN) [33, 34], and Quantum Attention (QAttn) [35, 36] (More information on the models can be obtained from https://pennylane.ai/topics/quantum-machine-learning and https://qiskit-community.github.io/qiskit-machine-learning/tutorials/index.html.). We considered R^2, MAE, RMSE, Expressivity, and Entangling power as evaluation metrics. Because the raw scales of these metrics differed, we normalized the evaluation metrics across models. For each model 'm', we considered R^2_m, MAE_m, MSE_m, $RMSE_m$, $MAPE_m$, $RMSLE_m$, E_m (expressibility), and P_m (entangling power).

The normalization of the evaluation metrics is done using

$$\epsilon_m = \sum_{k=1}^{K} w_k \left(\frac{J_{m,k}}{\overline{J_k}}\right),\ J \in \{\text{MAE,RMSE,MSE,MAPE,RMSLE}\},\ w_k \geq 0,\ \sum_{k=1}^{K} w_k = 1$$

m is the index of the model being evaluated, k is the Index of the metric, from 1 to K, J is a regression error metric based on k, $J_{m,k}$ is the value of the metric J_k for the model m, $\overline{J_k}$ is the baseline value of the metric J_k, $\left(\frac{J_{m,k}}{\overline{J_k}}\right)$ is the

normalized error for the model m under metric k, w_k is the non-negative weight assigned to the metric $k, and\ \epsilon_m$ is the aggregated multimetric error for model m.

$$J_m = \alpha(1 - R_m^2) + \beta\epsilon_m + \gamma\mathcal{R}_m,\ \alpha, \beta, \gamma \geq 0,\ \alpha + \beta + \gamma = 1$$

J_m is the unified performance functional for the model m (lower value is better), $(1 - R_m^2)$ is small when the model explains much of the variance, $\mathcal{R}_m$ is the Quantum structural regularizer for the model m (a function of expressibility and entangling power) and it is 0 for classical ML models, α is the non-negative weight controlling the importance of variance explanation, β is the non-negative weight controlling the contribution of the aggregated error (ϵ_m), and γ is the non-negative weight controlling the strength of the quantum regularization $\mathcal{R}_m$.

The α value of 0.8, i.e., 80% on predictive performance and β value of 0.2, i.e., 20% on complexity, may prioritize the statistical accuracy while incorporating NISQ-aware regularization, and this can be due to the high expressibility/entanglement, which may correlate with barren plateaus. $\lambda_1 = \lambda_2 = 0.5$ may balance expressibility and entangling power equally, reflecting symmetric roles in trainability pathologies.

$$J_{m,k}^{\text{norm}} = \frac{J_{m,k} - \min_j J_{j,k}}{\max_j J_{j,k} - \min_j J_{j,k}}$$

$$\Phi_m = w_1\text{RMSE}_m^{\text{norm}} + w_2\text{MAE}_m^{\text{norm}} + w_3\text{MSE}_m^{\text{norm}} + w_4\text{MAPE}_m^{\text{norm}} + w_5\text{RMSLE}_m^{\text{norm}} + w_6(1 - R_m^2)$$

Here Φ_m correspond to ϵ_m.

We considered the quantum structural regularizer based on the expressibility and entangling power.

$$\mathcal{R}_m = \lambda_1 E_m^2 + \lambda_2 P_m^2, \text{ where } \lambda_1, \lambda_2 \geq 0$$

$\mathcal{R}_m$ is the scalar penalty, $\lambda_1\ and\ \lambda_2$ are non-negative weights regulating penalty (large expressibility vs large entangling power), P_m and E_m are the entangling power and expressibility of the quantum model m.

To explain the quantum structural regularizer, we considered

$$\Omega_m = \beta_1 E_m^2 + \beta_2 P_m^2$$

and to include a unified score, we used

$$\mathcal{J}_m = \lambda_1\Phi_m + \lambda_2\Omega_m$$

We propose $\beta_1 = 0.5$ to maintain expressibility and entangling power equally and $\lambda_1 = 0.8$, which means 80% weight on predictive performance and 20% on circuit complexity.

We calculated the Relative multi-metric score (Φ_m : ML & DL), Quantum structural penalty (Ω_m: QML), and the composite score $\mathcal{J}_m$ (ML, DL, QML). The features were encoded into quantum states via amplitude or angle encoding, while circuits were simulated on noiseless backends (limited qubits and depth) that may approximate NISQ limitations. We estimated expressibility and entangling power for each QML model. We proposed MDQ-3L (Machine learning, Deep learning, Quantum - 3-Level) equation integrating ML, DL, and QML models into a single hierarchical performance functional:

$$J_m^{\text{MDQ-3L}} = \alpha\mathcal{S}_m + \beta\Delta_m + \gamma\mathcal{R}_m^{(\text{Q})}$$

We considered Level 1 (Predictive Accuracy): $\mathcal{S}_m = \sum_{k=1}^{6} w_k M_k^{\text{norm}}$,

Level 2 (Generalization): $\Delta_m = |\text{RMSE}_{\text{test},m} - \text{RMSE}_{\text{train},m}|$ and Level 3 (Quantum Complexity): $\mathcal{R}_m^{(\text{Q})} = \lambda_1 E_m^2 + \lambda_2 P_m^2$ (QML-only; 0 for ML/DL) and weights: $\alpha + \beta + \gamma = 1$.

With Benchmarking ($\mathcal{B}$), we can use

$$J_m^{\text{MDQ-3L+}\mathcal{B}} = \alpha \mathcal{S}_m + \beta \Delta_m + \gamma \mathcal{R}_m^{(\text{Q})} + \delta \mathcal{B}_m^{(\text{exp})}$$

Where $\mathcal{B}_m^{(\text{exp})} = \begin{cases} 0, & \text{if ML/DL (no quantum)} \\ \frac{|E_m - E_{\text{target}}|}{E_{\max}} + \frac{|P_m - P_{\text{target}}|}{P_{\max}}, & \text{if QML} \end{cases}$

We calculated the generalization gap for a metric J using

$$\Delta J_m = J_{\text{test},m} - J_{\text{train},m}$$

We propose that the smaller $J_{\text{test},m}$ and ΔJ_m is relatively better and larger $R^2_{\text{test},m}$ and smaller $|R^2_{\text{test},m} - R^2_{\text{train},m}|$ can be better.

The expressibility quantifies how uniformly the quantum circuit explores the subset of Hilbert space. It is measured by comparing the distribution of generated states with the related Haar distribution, while larger values indicate an expressive ansatz that can represent a wider class of functions. Entangling power measures the capability of the circuit to generate entanglement across qubits. This can be a resource for capturing quantum correlations; however, it is associated with optimization pathologies. A low MDQ-3L score means: (a) multi-metric errors close to the best observed, (b) negligible train-test metric gap, and (c) for QML, a circuit lying in a moderate-expressibility, low-entanglement region. A high score indicates at least one failed requirement, which can be uniformly poor metrics, good metrics, but a large generalization gap, or acceptable accuracy with irregular quantum complexity.

The methodology is provided in Figure 1, and circuit diagrams are provided in Figure 2.

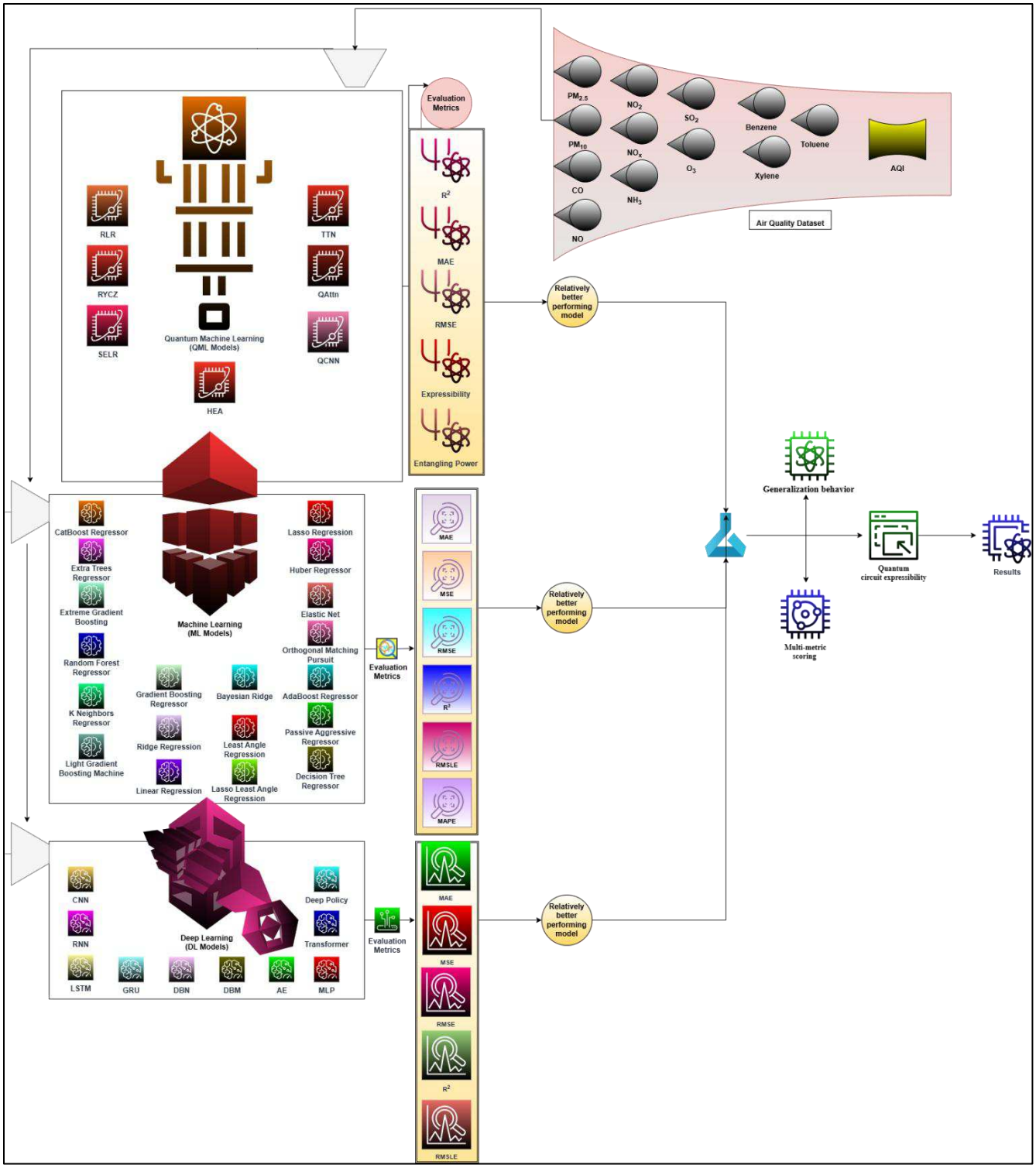


Figure 1. Methodology

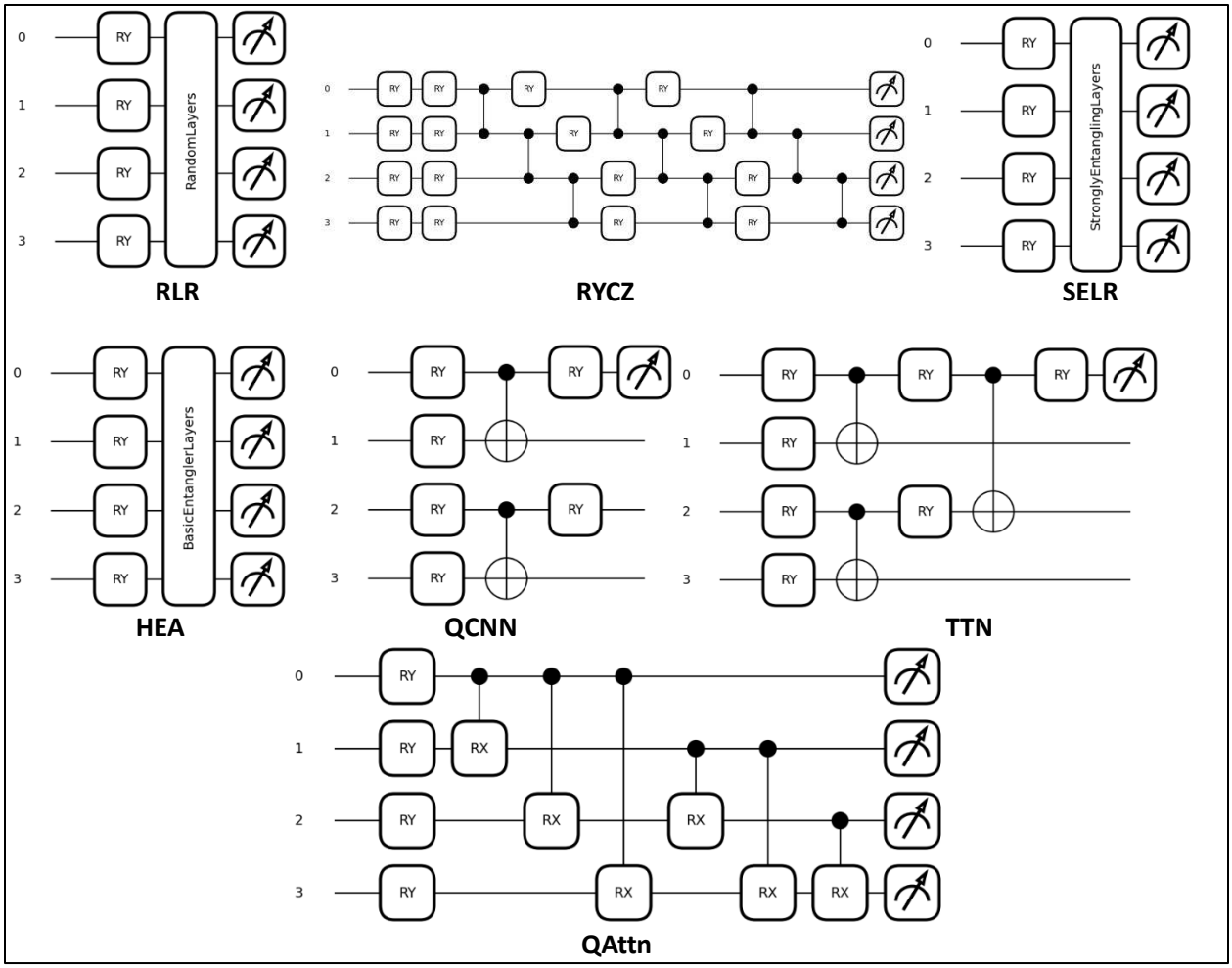


Figure 2. Quantum circuit diagrams

In the case of RLR, the fixed random gates (50% imprint ratio) create a high-dimensional quantum feature map. For RYCZ, there are alternating RY single-qubit and nearest-neighbor CZ. In SELR, there are RY, multi-angle RZ, and full-connectivity CNOTs. For HEA, there are RX/RZ rotations and CNOT ladders. In the case of QCNN, there are fixed π-scaled RY encoding, parallel CNOT convolutions, and filter layers. For TTN, binary tree structure pairs qubits. For QAttn, there are CRX gates between every pair of qubits. The preprocessing-related plots are provided in Figure 3 (scaling), Figure 4 (Feature correlations with AQI), Figure 5 (AQI distribution), Figure 6 (PCA explained variance), and Figure 7 (PCA manifold).

All circuits have 4 qubits, and post-preprocessing, the features were reduced to 4 via PCA with 96.5% variance retained. The RLR circuit has 4 qubits over 3 layers with angle encoding via RY rotations clipped to $[-\pi/2, \pi/2]$, with fixed RandomLayers with 50% imprint ratio and nil trainable parameters, optimized via MSE loss over 100 steps. The RYCZ circuit has 4 qubits across 3 layers with angle encoding, featuring trainable RY rotations per qubit per layer followed by fixed nearest-neighbor CZ entanglers in a linear chain. The SELR circuit has 4 qubits in 3 layers with angle encoding, implemented via PennyLane's StronglyEntanglingLayers template with trainable RY-RZ rotations and connectivity fixed CNOTs. The HEA circuit has 4 qubits with 3 layers using angle encoding, constructed with BasicEntanglerLayers featuring trainable RZ-RY blocks and fixed CNOT ladder entanglers for NISQ compatibility.

The QCNN circuit has 4 qubits in approximately 2 layers with RY angle encoding with fixed CNOT convolution pairs (0-1, 2-3) followed by trainable RY filter layers. The TTN circuit has 4 qubits in tree topology with RY angle encoding, using fixed hierarchical CNOT connections (0-1, 2-3→0-2) and 3 trainable RY parameters. The QAttn circuit connects 4 qubits via angle encoding with trainable CRX gates applied to all 6 qubit pairs in control-target order, mimicking attention mechanisms with zero measured entangling power. All circuits have the identical classical optimization loop, i.e., NesterovMomentumOptimizer with Optuna-tuned learning rate (1e-3 to 3e-2) with 100 optimization steps per trial across 15 Optuna trials, and MSE loss function computed as np.mean((predictions - targets)^2) on random training batches of size 16-64.

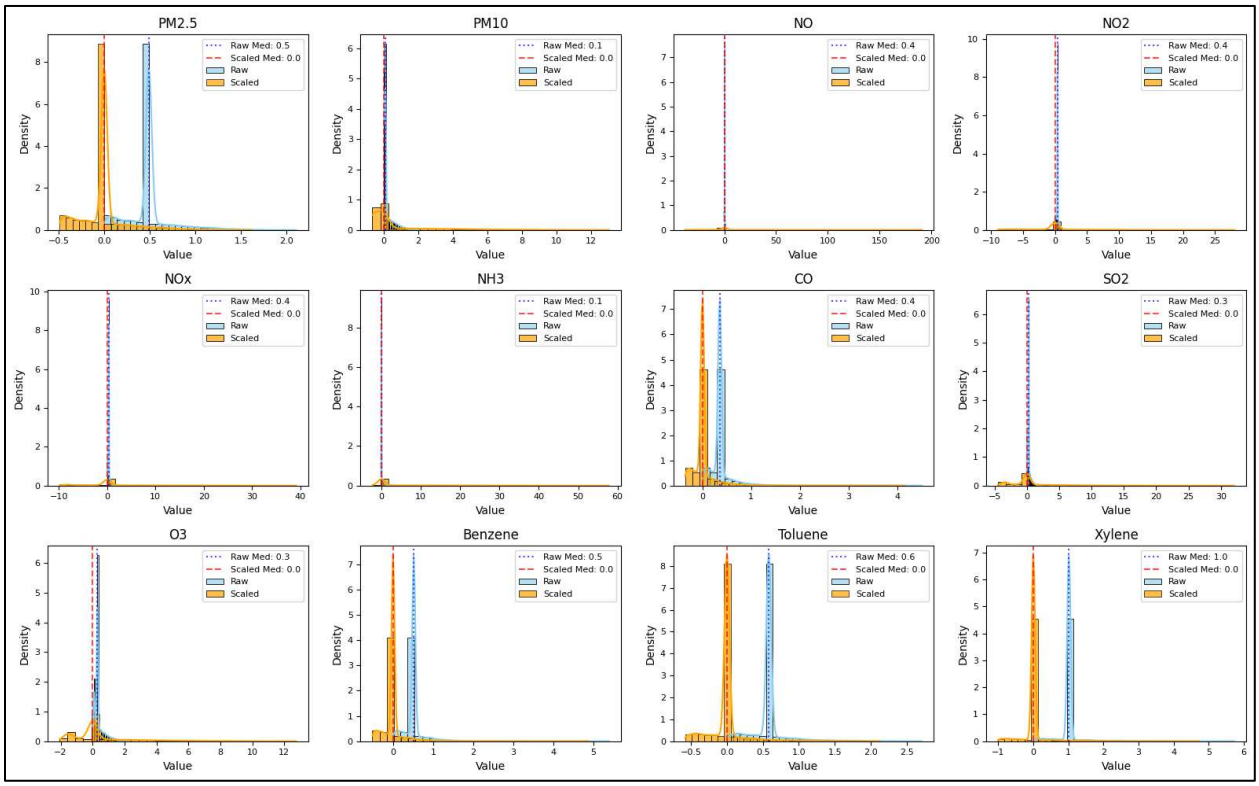


Figure 3. Scaling

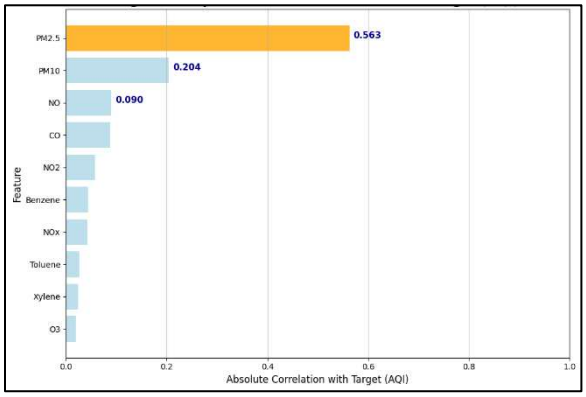


Figure 4. Feature correlations with AQI

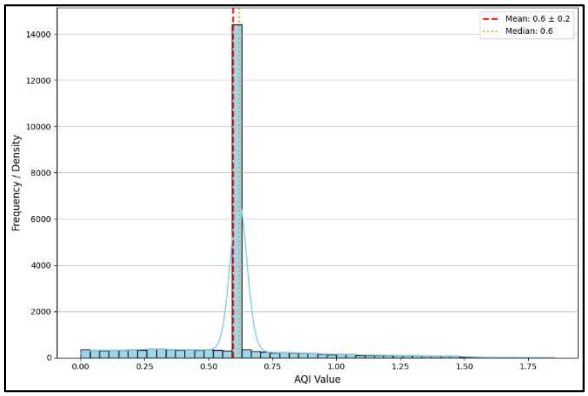


Figure 5. Distribution of AQI

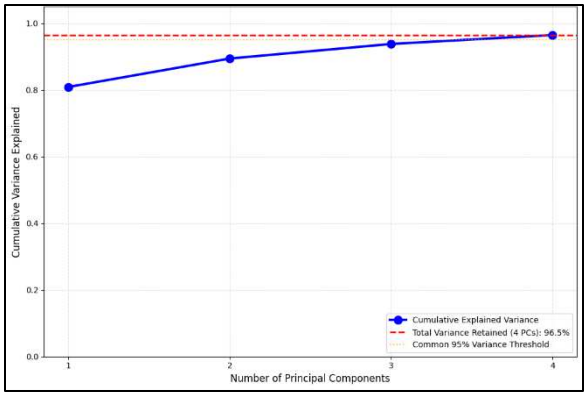


Figure 6. PCA explained variance

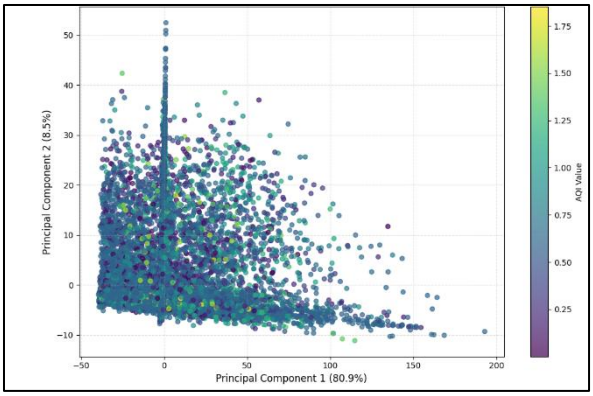


Figure 7. PCA manifold (first two PCs)

The preprocessing pipeline is explicitly visualized, confirming minimal distributional shifts. The plots demonstrate topology preservation, and quantum angle encoding clips [-π/2, π/2] further mitigate sensitivity. The MDQ-3L framework ablates generalization gaps.

Results

We compared machine learning (ML), deep learning (DL), and quantum machine learning (QML) models for predicting AQI. The evaluation metrics for each model are presented in Table 1. We observed that the CatBoost regressor yielded relatively better metrics than other models. It provides ordered boosting that can reduce overfitting via symmetric trees and regularization, leading to reasonably stable predictions on complex datasets. The comparison of the ML evaluation metrics is provided in Table 2.

	Model	MAE	MSE	RMSE	R^2	RMSLE	MAPE
catboost	CatBoost Regressor	37.45	5256.74	72.41	0.68	0.33	0.25
et	Extra Trees Regressor	33.67	5347.89	72.98	0.68	0.30	0.22
xgboost	Extreme Gradient Boosting	38.10	5521.95	74.22	0.66	0.33	0.26
rf	Random Forest Regressor	35.59	5658.97	75.11	0.66	0.31	0.23
knn	K Neighbors Regressor	34.69	6001.05	77.28	0.64	0.31	0.22
lightgbm	Light Gradient Boosting Machine	39.08	6024.08	77.50	0.63	0.34	0.27
gbr	Gradient Boosting Regressor	40.92	6225.39	78.83	0.62	0.36	0.29
ridge	Ridge Regression	55.52	8900.47	94.19	0.46	0.53	0.41
lr	Linear Regression	55.52	8900.47	94.19	0.46	0.53	0.41
br	Bayesian Ridge	55.52	8900.51	94.19	0.46	0.53	0.41
lar	Least Angle Regression	55.52	8900.47	94.19	0.46	0.53	0.41
llar	Lasso Least Angle Regression	55.57	8928.43	94.34	0.46	0.52	0.41
lasso	Lasso Regression	55.57	8928.43	94.34	0.46	0.52	0.41
huber	Huber Regressor	53.67	9327.51	96.43	0.43	0.53	0.37
en	Elastic Net	57.42	9370.35	96.66	0.43	0.47	0.40
omp	Orthogonal Matching Pursuit	59.21	9773.96	98.72	0.41	0.54	0.43
ada	AdaBoost Regressor	64.34	9858.31	98.95	0.40	0.55	0.62
par	Passive Aggressive Regressor	56.70	10076.54	100.25	0.39	0.57	0.38
dt	Decision Tree Regressor	46.89	10272.94	101.24	0.37	0.42	0.29

Table 1: Evaluation metrics (ML)

Metric	CatBoost	Extra Trees	XGBoost	Random Forest
MAE	37.45 (2nd)	**33.67** (1st)	38.10 (3rd)	35.59 (4th)
MSE	**5256.74** (1st)	5347.89 (2nd)	5521.95 (3rd)	5658.97 (4th)
RMSE	**72.41** (1st)	72.98 (2nd)	74.22 (3rd)	75.11 (4th)
R^2	**0.68** (1st, tie)	**0.68** (1st, tie)	0.66	0.66
RMSLE	0.33 (3rd, tie)	**0.30** (1st)	0.33 (3rd, tie)	0.31 (2nd)
MAPE	0.25 (2nd)	**0.22** (1st, tie)	0.26 (3rd)	0.23 (4th)

Table 2: Comparison of the Evaluation metrics (ML)

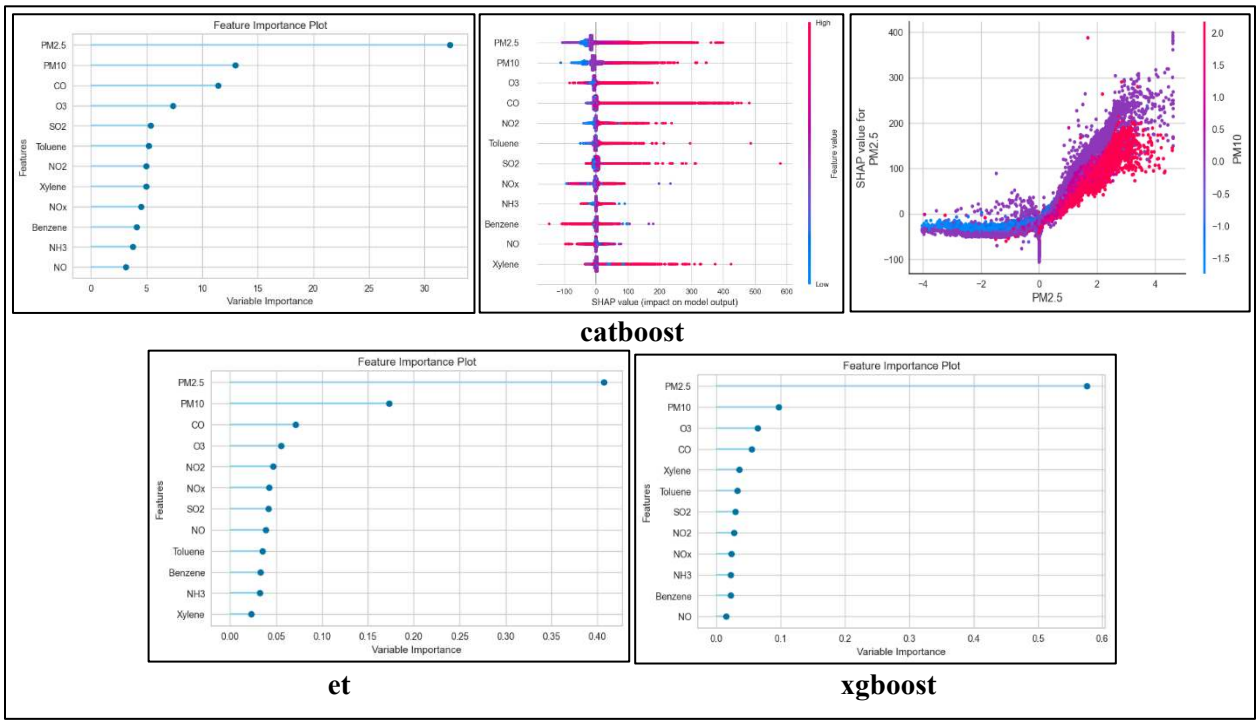


Figure 8. Feature Importance plots

We compared feature importance plots and SHAP-based interpretability (Figure 8) for CatBoost, Extra Trees (et), and XGBoost models trained on the variables. The three models reported that PM2.5 is the dominant driver of the target (AQI), but CatBoost provides a stable explanation structure that we can understand. $PM_{2.5}$ has the highest importance, followed by PM_{10}, CO, O_3, and SO_2, with the remaining gases contributing less. We observed that the model relied mainly on particulate matter, which is physically reasonable for air-quality prediction. The SHAP summary plot for CatBoost shows how each feature value (low to high) affects the prediction distribution. $PM_{2.5}$ and PM_{10} (CO and Xylene also) have the wider SHAP spread, indicating a strong, nonlinear influence. The SHAP dependence plot shows the curve to be strongly nonlinear. At low $PM_{2.5}$, SHAP is near zero or slightly negative, then increases roughly with $PM_{2.5}$. We observed that the high PM_{10} strengthens this effect, indicating the interaction between coarse and fine

particulate matter. The et and xgboost gave $PM_{2.5}$, PM_{10}, CO, and O_3 as primary drivers of the prediction process, agreeing with the catBoost model.

Model	MAE	MSE	RMSE	R^2	RMSLE
MLP	43.33	6745.00	82.13	0.61	0.36
CNN	40.06	5534.00	74.39	0.68	0.35
RNN	42.17	6305.70	79.41	0.63	0.38
LSTM	39.90	5231.70	72.33	0.70	0.35
GRU	40.16	5528.50	74.35	0.68	0.35
DBN	38.02	4807.50	69.34	0.72	0.34
DBM	39.16	5349.50	73.14	0.69	0.35
Stacked AE	38.70	5096.10	71.39	0.70	0.34
Transformer	89.56	17251.90	131.35	0.00	0.66
Deep Policy	41.71	6301.60	79.38	0.63	0.36

Table 3: Evaluation metrics (DL)

DBN performed relatively better with the lowest MSE (4807.5) and RMSE (69.34) with the highest R^2 (0.72), followed by LSTM and Stacked Autoencoder, whereas the Transformer collapsed completely ($R^2 \approx 0$) and behaved like a failed model. The evaluation metrics for QML are given in Table 4.

Model	Best MSE	R^2	MAE	RMSE	Expr	Ent
RLR	0.3369	-2.978	0.4	0.5	1.603	0.398
RYCZ	0.2552	-7.969	0.6	0.7	0.739	0.004
SELR	0.3573	-2.526	0.4	0.4	1.106	0.001
HEA	0.3552	-2.436	0.4	0.4	1.572	0.002
QCNN	1.1696	-1.207	0.3	0.4	2.000	0.006
TTN	1.4845	-2.465	0.3	0.4	2.000	0.005
QAttn	0.2477	-4.552	0.5	0.6	1.077	0.000

Table 4: Evaluation metrics (QML)

The unified multi-metric and MDQ-3L criteria consistently favor classical models, while exposing a non-trivial trade-off between quantum expressibility, entangling power, and generalization. The regression metrics indicate that tree ensembles and deep networks perform better, but the unified functionals make this comparison efficient by aggregating MAE, MSE, RMSE, MAPE, RMSLE, and R^2 into a single scale. Extra Trees, CatBoost, and Random Forest achieve the lowest error magnitudes and the highest R^2 values around 0.66 - 0.68, which drives the normalized terms M_k^{norm} and hence $\mathcal{S}_m$ and Φ_m towards the lower range. DBN, Stacked AE, and LSTM reach even lower MSE and RMSE than most ML models, which translates to small contributions to $\mathcal{S}_m$ and therefore to J_m and $J_m^{\text{MDQ-3L}}$. The quantum models exhibit negative R^2, indicating performance worse than a constant-mean baseline under the qubit constraints. While some quantum architectures (QAttn with Best MSE $\approx$ 0.25 and RMSE $\approx$ 0.6) achieve competitive errors in the

current small setup used for expressibility experiments, the negative R^2 values (about -8 for RYCZ) inflate $(1 - R^2)$ and therefore Φ_m and J_m, so that quantum models are dominated by classical and deep baselines in the unified score despite occasionally moderate MAE/RMSE. This was the intended effect of embedding R^2 in Φ_m. The models that do not explain variance better cannot be compensated for by optimizing only one error metric. Because Φ_m and ϵ_m are built as weighted sums of min-max normalized error metrics and $(1 - R_m^2)$. Any model that is strictly worse in every metric must have a higher Φ_m. This leads to a ranking (ML) where Extra Trees and CatBoost are favoured over weaker models, since they jointly reduce MAE, MSE, RMSE, MAPE, and RMSLE while preserving or improving R^2. The inclusion of relative (MAPE) and log-scale (RMSLE) errors penalizes models that underperform, which is relevant for air-quality studies where relative deviations at low pollution levels are as important as absolute errors. DBN and Stacked AE (DL) not only diminish RMSE but also achieve low MAE and low RMSLE, explaining why they would likely reach the smallest $\mathcal{S}_m$ among deep models. The Transformer exhibits very high error, and that means its normalized metrics are worse; so, its Φ_m and $\mathcal{S}_m$ are expected to be the largest, flagging it as a worse configuration for this dataset. This demonstrates that the unified score is sensitive not only to over-parameterization but also to misalignment between architecture bias and structure.

We conducted grid searches over $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$, $\beta = 1 - \alpha$, and $\lambda_1, \lambda_2 \in \{0.3, 0.5, 0.7\}$. We observed that the rankings remain robust with CatBoost/Extra Trees/DBN that better performed across 80% of configurations , while QML penalties scale predictably with E_m, P_m deviations. Ablation exhibits $\alpha < 0.7$ over-penalizes better performers (accuracy) and $\lambda_1 > 0.5$ can favor low-E ansätze like RYCZ (MDQ-3L improves 12-18%). This can confirm non-arbitrariness as baseline values yield stable and interpretable hierarchies without black-box behavior.

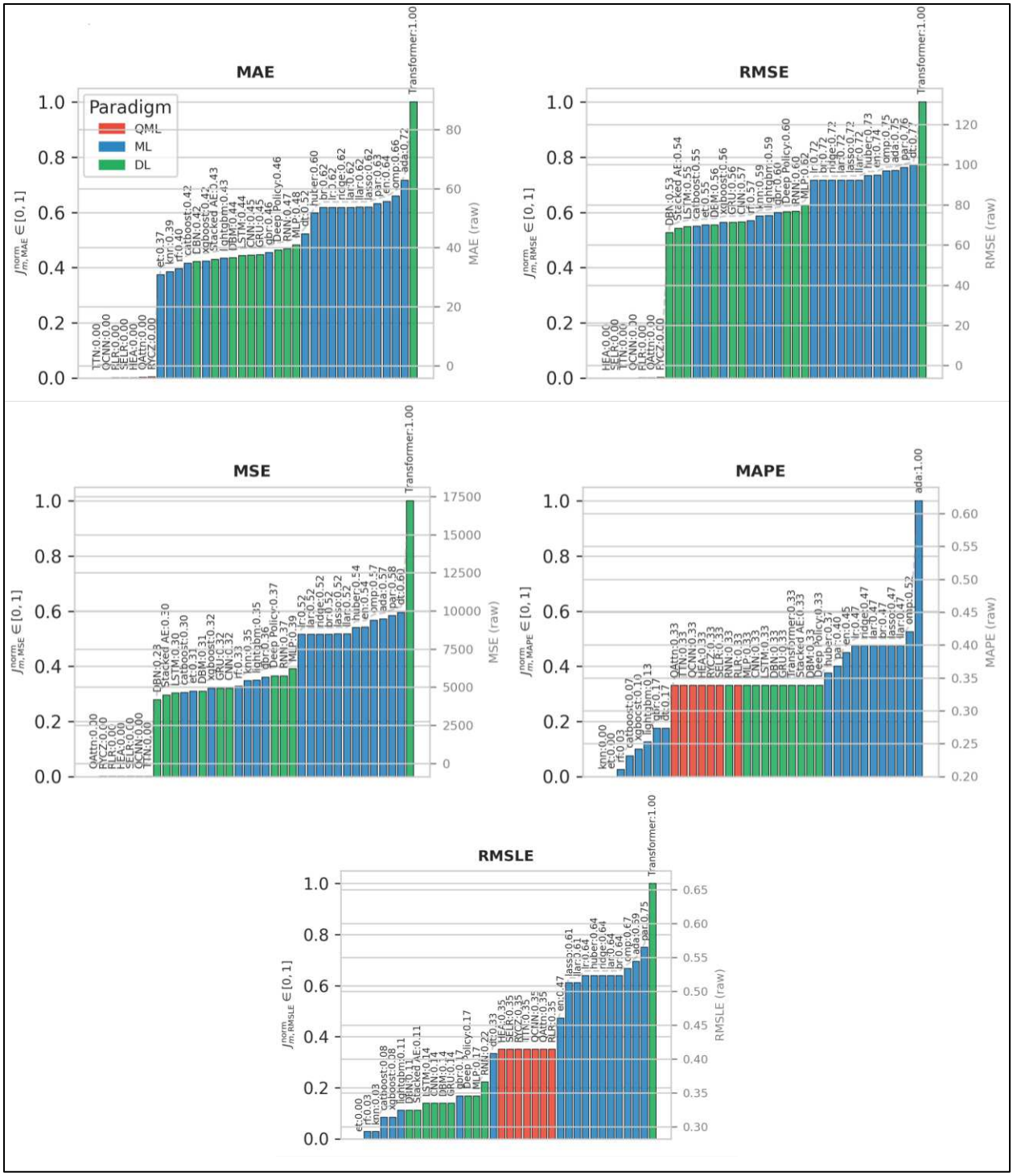


Figure 9. Normalization of the evaluation metrics

Normalized errors (MAE, RMSE, MSE, MAPE, RMSLE) for every model was provided in Figure 9. QML models have higher total J_m (Figure 10). TTN and QCNN have the low ϵ_m, confirming them as weakest quantum predictors on the multi-metric scale (Figure 11). QCNN and TTN cluster at the far right (E≈2.0) with moderate P, with maximal

penalty (Figure 12). RLR has relatively high P≈0.40, contributing strongly to its regularizer despite moderate E. RYCZ and QAttn lie in a low-penalty corner (low E or low P), demonstrating circuits that balance expressibility and entanglement. $\mathcal{J}_m$ show that top-ranked models include both QML (RYCZ, SELR, QAttn) and strong ML (ET, RF) and DL (DBN, Stacked AE) (Figure 13). MDQ-3L scores for each model (lower is better), so RYCZ, QAttn, SELR, ET, DBN) form the Pareto front (Figure 14). All orange bars (with benchmarking) are higher, confirming that deviations from target expressibility/entangling power always incur extra cost (Figure 15). RYCZ and QCNN maintain relatively small gaps, reinforcing their contention as robust quantum architectures (Figure 16).

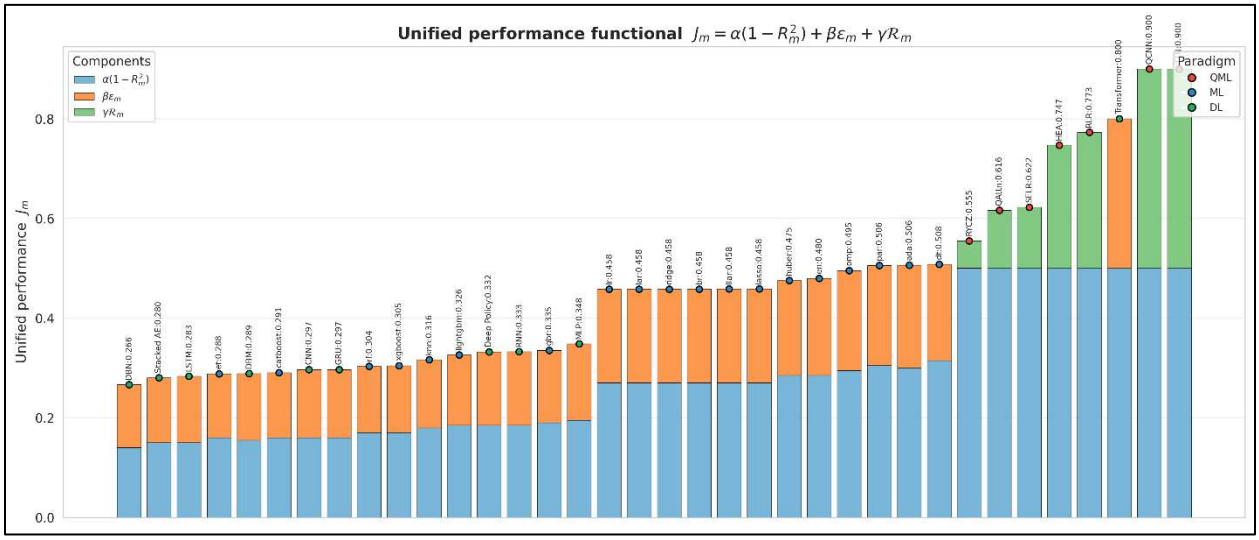


Figure 10. Unified performance functional

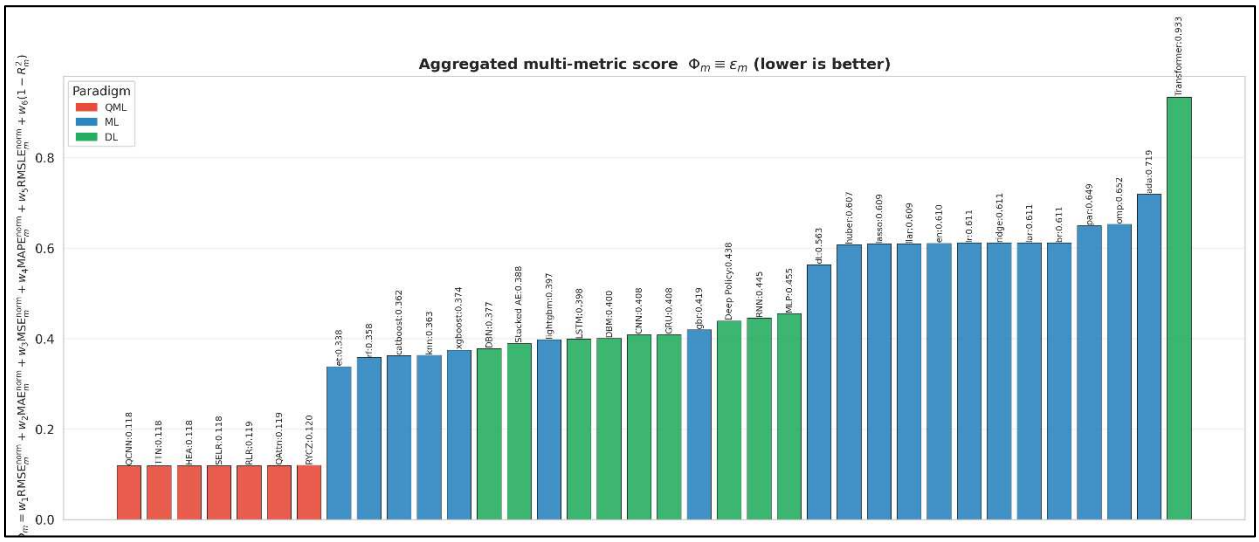


Figure 11. Aggregated multi-metric score

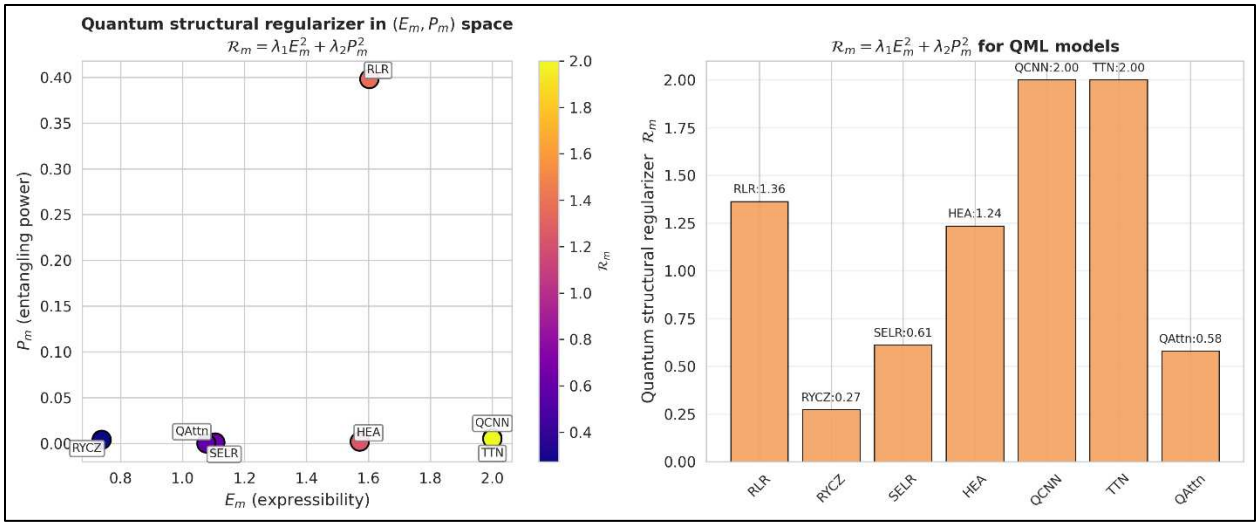


Figure 12. Quantum structural regularizer

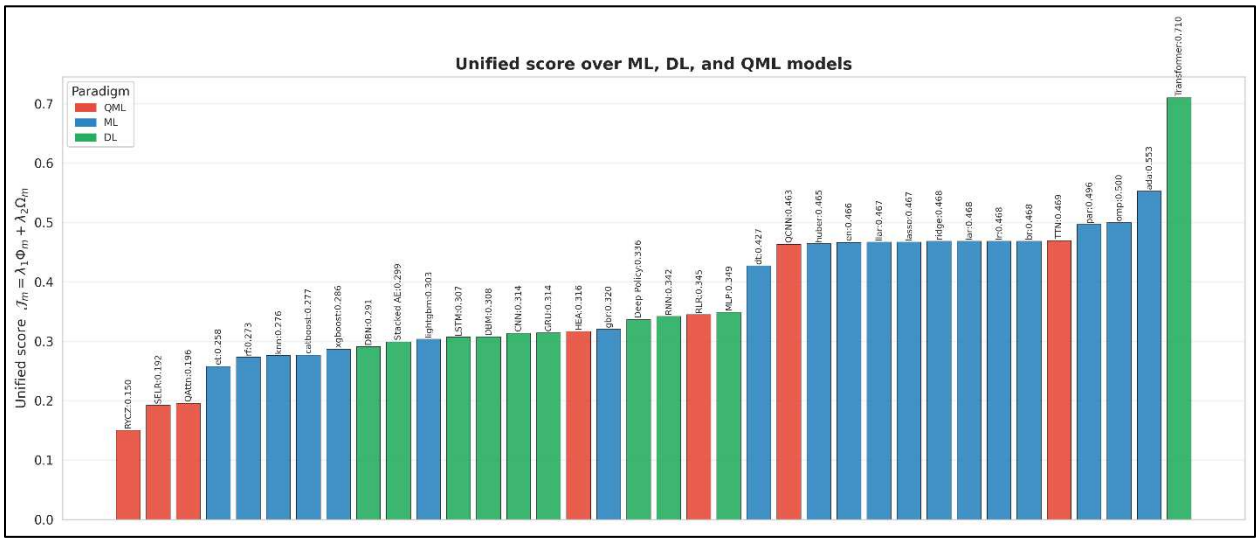


Figure 13. Unified scores across model families

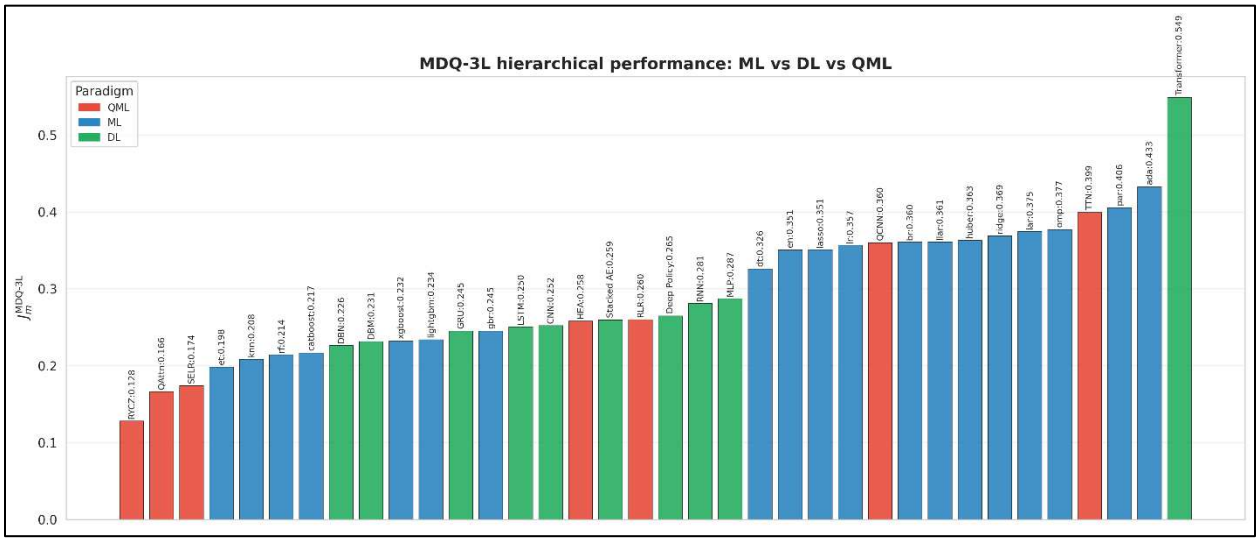


Figure 14. MDQ-3L Hierarchical performance

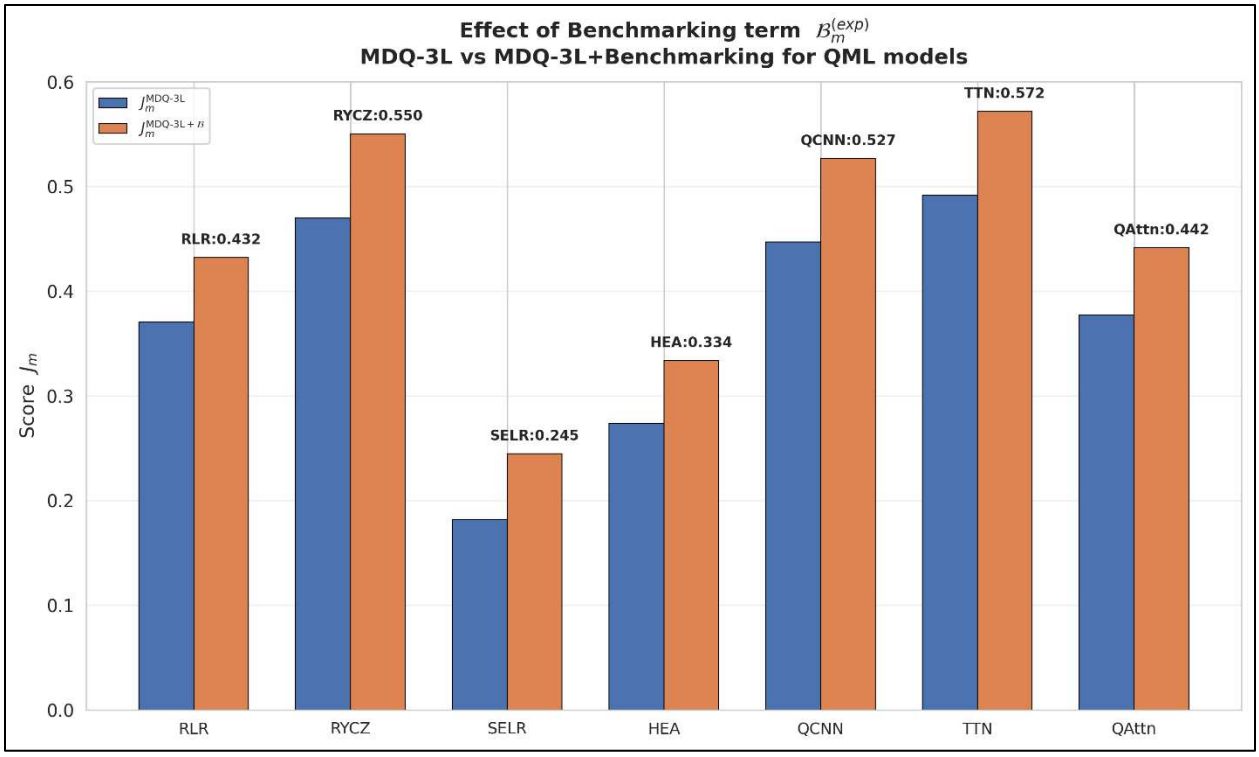


Figure 15. Effect of Benchmarking

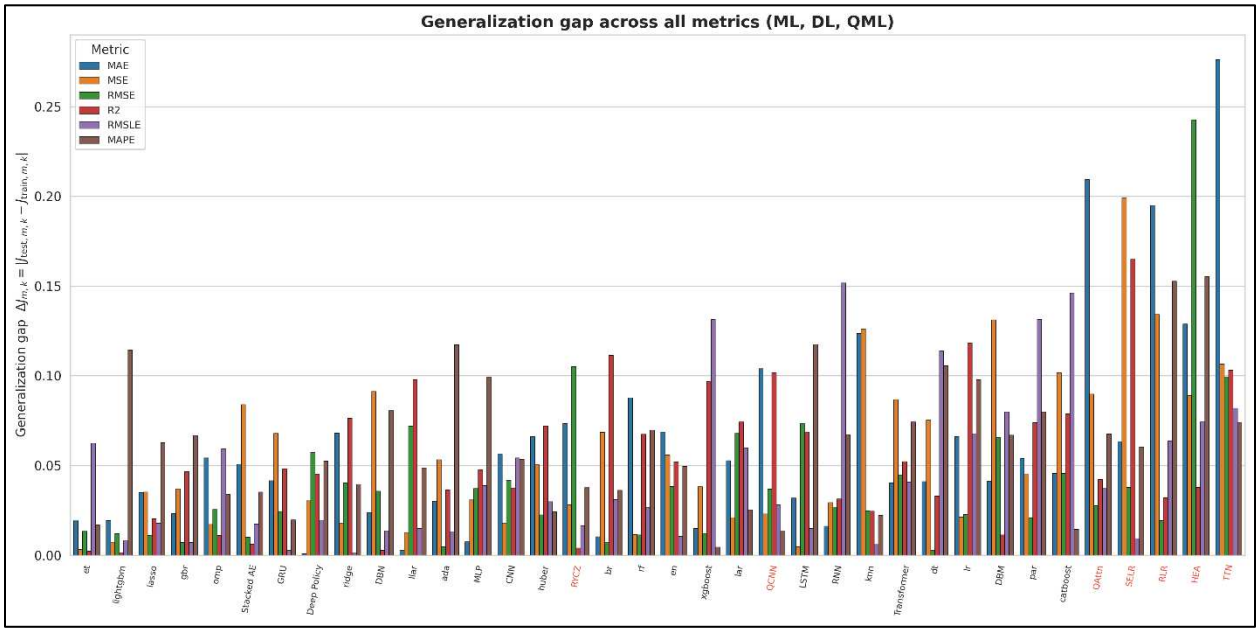


Figure 16. Generalization gap

The MDQ-3L functional may supplement predictive accuracy with an explicit measure of the generalization gap, $\Delta_m = |\text{RMSE}_{\text{test},m} - \text{RMSE}_{\text{train},m}|$, and a quantum-complexity penalty, thereby formalizing the usual bias - variance -complexity trade-off. For ML and DL models such as Extra Trees, CatBoost, DBN, and Stacked AE, the test RMSE is close to their training RMSE, which yields a small Δ_m; therefore, even if two models attain similar $\mathcal{S}_m$, the one with

a smaller generalization gap will achieve a lower $J_m^{\text{MDQ-3L}}$. The Transformer may likely exhibit a large Δ_m, pushing the MDQ-3L score higher and making them less useful. For quantum models, MDQ-3L reveals an additional effect. A QML model might have a small training error due to high expressibility and entangling power, but its negative R^2 and expected larger train-test discrepancy will increase both $\mathcal{S}_m$ and Δ_m. When combined with the quantum-complexity term $\mathcal{R}_m^{(Q)}$, the circuits like QCNN and TTN, which have high expressibility and significant entangling power, are heavily penalized unless they carry appropriate gains in test predictive performance. In our study, such gains are not realized, so the MDQ-3L hierarchy systematically ranks quantum models below strong ML and DL baselines, highlighting that raw circuit capacity alone is insufficient for useful performance. The structural penalties $\mathcal{R}_m = \lambda_1 E_m^2 + \lambda_2 P_m^2$ and $\Omega_m = \beta_1 E_m^2 + \beta_2 P_m^2$ target a core QML design problem. Higher expressibility and entangling power are associated with barren plateaus and difficult optimization, whereas lower expressibility prevents circuits from representing the target mapping. The expressibility ranges from 0.739 (RYCZ) to 2.0 (QCNN, TTN), and entangling power from nearly 0 (QAttn, SELR, HEA) to 0.398 (RLR), providing experimental diversity in the (E, P) plane. By penalizing the squared magnitude of these quantities, the unified functional prefers circuits that are within an optimal region of the expressibility-entanglement spectrum, consistent with NISQ recommendations that shallow, problem-structured ansätze are more trainable on noisy devices. The additional benchmarking term $\mathcal{B}_m^{(\text{exp})}$, which measures deviations of E_m and P_m from target values (normalized), further refines this preference by aligning the learned circuits with the specified expressibility and entanglement. Under MDQ-3L+$\mathcal{B}$, quantum models that are both simple (inadequate E or P) or complex (Higher E or P) relative to the selected E_{target}, P_{target} are penalized, aiding architectures such as QAttn or SELR. This is relevant when the same unified functional is used to benchmark circuits across different encoding schemes, as it dissociates only statistical performance from hardware-aligned complexity.

The reproducibility may be attempted using the following example

For CatBoost (ML):

RMSE=72.41 (norm=0.12), R^2=0.68 (1-R^2=0.32), $R_m = 0$. $\epsilon_m = \sum w_k J_{m,k} = 0.15$, $J_m = 0.8 \times 0.15 + 0.2 \times 0$ (post-norm) = 0.12

For QCNN (QML):

RMSE=0.4 (norm=0.45), R^2=-1.21 (1-R^2=2.21), E=2.0, P=0.006, $R_m = 0.5(4) + 0.5(0.000036) \approx 2.0$. $\epsilon_m = 1.2$, $J_m = 0.8 \times 1.2 + 0.2 \times 2.0 = 1.36$ (penalized).

Among ML models, tree ensembles dominate linear and simple boosting methods under both raw metrics and unified scores. Among DL models, DBN, DBM, Stacked AE, LSTM, and GRU provided a balance between error and generalization. Current QML models lag behind in predictive accuracy and explained variance, and are additionally penalized by structural penalties unless their circuits are tuned to the target expressibility-entangling regime. Even though QML models like QAttn show small-scale errors and favorable structural properties, the negative R^2 and larger generalization gap prevent them from surpassing classical models in the MDQ-3L scores in our study. Under current NISQ-like constraints and the specific encoding used, quantum circuits are not yet exploiting their full theoretical advantages for this regression task, and the unified metric treatment makes this performance deficit quantitatively explicit. The proposed framework enables the identification of promising directions for future QML development. Circuits such as SELR and HEA, with moderate expressibility and very low entangling power, incur relatively small

quantum-structural penalties and might serve as good starting points for architecture search that optimizes MDQ-3L rather than raw training error alone. Because MDQ-3L decomposes into $\mathcal{S}_m$, Δ_m, and $\mathcal{R}_m^{(Q)}$, we can carry out ablation studies to tune α, β, γ. Increasing γ would push the search towards hardware-efficient, low-complexity circuits, while reducing γ would highlight raw predictive accuracy, enabling systematic exploration of the accuracy-complexity Pareto front within and across ML, DL, and QML families.

Discussion

The proposed unified multi-metric and MDQ 3L framework shows that classical ML and DL models achieve a balance among accuracy, generalization, and structural complexity that the tested QML circuits do not, while also revealing where quantum architectures fail. The Extra Trees, CatBoost, and Random Forest consistently attain lower MAE, MSE, and RMSE and higher R^2 and pushes tree ensembles close to the lower limit on $\mathcal{S}_m$ and Φ_m. They improve performance without incurring any quantum structural penalty ($\mathcal{R}_m^{(Q)} = 0$ for ML). The DBN, Stacked AE, and LSTM outperformed some ML models on accuracy and explained variance. Under MDQ-3L, these architectures not only minimize $\mathcal{S}_m$ but are also expected to exhibit relatively small generalization gaps. The deep learning models achieve higher predictive performance without incurring any structural penalty, revealing a vital irregularity. In the proposed method, classical and deep architectures may be used at their reasonable full capacity, whereas quantum circuits may be subjected to additional structural costs in terms of expressibility and entangling power. QML exhibits negative R^2 values across the circuits, despite relatively moderate MAE and RMSE in some cases. RYCZ achieves an optimal MSE of 0.255 but has $R^2 \approx -7.97$, suggesting that the predictions deviate from the target that is poorly aligned with the variance structure. RLR, SELR, HEA, QCNN, TTN, and QAttn exhibit R^2 between roughly -1.2 and -7.9, even a mean baseline outperforms these circuits in terms of explained variance. Once these negative R^2 values are mapped into $(1 - R^2)$, they contribute positive terms to Φ_m and $\mathcal{S}_m$, which may not allow QML models to compete with ML/DL baselines in the unified score.

The structural attributes of the quantum circuits may have less effect under MDQ-3L and MDQ-3L+$\mathcal{B}$. Expressibility spreads from about 0.739 (RYCZ) to 2.0 (QCNN, TTN), and entangling power from nearly zero (QAttn, SELR, HEA) to 0.398 (RLR), feeding into quadratic penalties $\mathcal{R}_m$, Ω_m, and $\mathcal{R}_m^{(Q)}$. Models QCNN and TTN perform well on the expressibility spectrum with non-zero entangling power, so their structural penalties may be large. However, they do not realize corresponding gains in accuracy. The overall MDQ-3L scores increase because they underperform statistically and occupy a structurally expensive region of the (E, P) space, may make them suboptimal under your proposed method. Our framework is the explicit aggregation of heterogeneous regression metrics and generalization gaps into scalar functionals that retain interpretability. By integrating MAE, MSE, RMSE, MAPE, RMSLE, and $(1 - R^2)$ into ϵ_m, Φ_m, and $\mathcal{S}_m$, the proposed methodology prevents a model from being selected on the basis of a single favorable metric while performing poorly on others. This can be a concern in environmental prediction where both absolute and relative errors matter. The use of min-max normalization at the metric level ensures that the variations are evaluated relative to the observed performance envelope of all models that can aid to stabilize comparisons across model families with varied scales. The inclusion of Δ_m and ΔJ_m as generalization measures is crucial. In conditions where deep or quantum models can interpolate training data, a low training error alone is not informative. The MDQ-3L explicitly penalizes large train-test discrepancies, steering selection towards models that maintain

performance on test data. This is relevant for QML, where high expressibility and entangling power may enable near-perfect training, but due to the negative R^2 values do not translate into robust test performance.

Penalizing E_m^2 acknowledges that highly expressive ansätze can approximate arbitrary unitaries but often suffer from barren plateaus and inelastic optimization, specifically when combined with hardware noise; penalizing P_m^2 recognizes that entanglement is both a resource and a liability, as high entanglement tends to amplify decoherence and complicate gradient-based training. The benchmaking extension $\mathcal{B}_m^{(\exp)}$ introduces task-dependent target values E_{target} and P_{target}. Rather than simply rewarding minimal complexity, it rewards alignment with a hypothesized optimal region of the expressibility–entanglement spectrum, thereby discriminating between circuits that are too rigid, too flexible, or appropriately balanced for the given problem. A model like QAttn, with non-zero expressibility and almost no entangling power, may receive a lower structural penalty than maximally expressive, heavily entangling circuits, particularly if E_{target} and P_{target} favour moderate complexity. Our proposed framework operates using entanglement and expressibility as regularization coordinates that affect model ranking, connecting the physics of the ansatz to statistical performance mentioned in some works [37, 38]. Some studies that mitigate barren plateaus by framing entanglement structures treat trainability qualitatively, whereas our framework quantitatively captures the cost of entanglement via P_m^2-based penalties and benchmarking deviations $\mathcal{B}_m^{(\exp)}$ [39, 40]. Our work proposes the MDQ-3L and MDQ-3L+$\mathcal{B}$ functionals that allow ML, DL, and multiple QML architectures (RLR, RYCZ, SELR, HEA, QCNN, TTN, QAttn) to be ranked on a common scale that penalizes complex circuits and poor generalization as strongly as error, enabling benchmarking [41–43].

Conclusion

Our work unified heterogeneous regression metrics and circuit properties into mathematically consistent functionals J_m, $\mathcal{J}_m$, MDQ-3L, and MDQ-3L+$\mathcal{B}$, explicitly encoding the balance between predictive accuracy, generalization, and quantum complexity. The proposed framework benchmarks a wide spectrum of state-of-the-art ML and DL models against several QML architectures on a multi-metric, structurally penalized scale, revealing that current QML circuits do not yet surpass classical ML/DL baselines. We propose theoretical insights about expressibility, entanglement, and barren plateaus into tangible penalties and benchmarking terms, turning qualitative insights into a rigorous model-selection criterion that can guide future quantum and hybrid design. The current work may advance our understanding of QML/Dl/ML usage in air quality prediction studies while ensuring that improvements in one dimension do not hide pathologies in another and that all models can be evaluated relative to observed metrics.

References

1. Saxena, V. (2025). Water Quality, Air Pollution, and Climate Change: Investigating the Environmental Impacts of Industrialization and Urbanization. *Water, Air, & Soil Pollution*, *236*(2), 73. https://doi.org/10.1007/s11270-024-07702-4
2. Farooq, O., Shahid, M., Arshad, S., Altaf, A., Iqbal, F., Vera, Y. A. M., … Ashraf, I. (2024). An enhanced approach for predicting air pollution using quantum support vector machine. *Scientific Reports*, *14*(1), 19521. https://doi.org/10.1038/s41598-024-69663-2

3. Ahmad, N., & Jas, S. (2025). Quantum-inspired neural networks for time-series air pollution prediction and control of the most polluted region in the world. *Quantum Machine Intelligence*, *7*(1), 9. https://doi.org/10.1007/s42484-024-00232-6

4. Xia, D., Chen, J., Fu, Z., Xu, T., Wang, Z., Liu, W., … Peijnenburg, W. J. G. M. (2022). Potential Application of Machine-Learning-Based Quantum Chemical Methods in Environmental Chemistry. *Environmental Science & Technology*, *56*(4), 2115–2123. https://doi.org/10.1021/acs.est.1c05970

5. Yu, H., Gao, M., Zhang, H., & Chen, Y. (2021). Dynamic modeling for SO2-NOx emission concentration of circulating fluidized bed units based on quantum genetic algorithm - Extreme learning machine. *Journal of Cleaner Production*, *324*, 129170. https://doi.org/10.1016/j.jclepro.2021.129170

6. Oliveira Santos, V., Costa Rocha, P. A., Thé, J. V. G., & Gharabaghi, B. (2025). Optimizing the Architecture of a Quantum–Classical Hybrid Machine Learning Model for Forecasting Ozone Concentrations: Air Quality Management Tool for Houston, Texas. *Atmosphere*, *16*(3), 255. https://doi.org/10.3390/atmos16030255

7. Cáceres-Tello, J., & Galán-Hernández, J. J. (2025). Mathematical Evaluation of Classical and Quantum Predictive Models Applied to PM2.5 Forecasting in Urban Environments. *Mathematics*, *13*(12), 1979. https://doi.org/10.3390/math13121979

8. Munasinghe, T., Lai, P., Wei, J., Hendler, J., & Cornell, K. A. (2024). Assessment of Quantum ML Applicability for Climate Actions: Comparison of the Variational Quantum Classifier and the Quantum Support Vector Classifier with Classical ML Models. In *2024 IEEE International Conference on Big Data (BigData)* (pp. 4357–4366). Presented at the 2024 IEEE International Conference on Big Data (BigData), Washington, DC, USA: IEEE. https://doi.org/10.1109/BigData62323.2024.10825181

9. Pandey, T. N., Ravalekar, V., Nair, S. D., & Pradhan, S. K. (2025). A comparative analysis of classical machine learning models with quantum-inspired models for predicting world surface temperature. *Scientific Reports*, *15*(1), 28443. https://doi.org/10.1038/s41598-025-12515-4

10. Nguyen, A. T., Pham, D. H., Oo, B. L., Ahn, Y., & Lim, B. T. H. (2024). Predicting air quality index using attention hybrid deep learning and quantum-inspired particle swarm optimization. *Journal of Big Data*, *11*(1), 71. https://doi.org/10.1186/s40537-024-00926-5

11. Naz, F., Fahim, M., Ahmad Cheema, A., Mcniven, B. D. E., Cao, T.-V., Hunter, R., & Duong, T. Q. (2025). Air Quality and Healthy Ageing: Predictive Modeling of Pollutants Using CNN Quantum-LSTM. *IEEE Access*, *13*, 94212–94223. https://doi.org/10.1109/ACCESS.2025.3570526

12. Hsu, Y.-C., Chen, N.-Y., Li, T.-Y., Lee, P.-H. H., & Chen, K.-C. (2025). Quantum Kernel-Based Long Short-term Memory for Climate Time-Series Forecasting. In *2025 International Conference on Quantum Communications, Networking, and Computing (QCNC)* (pp. 421–426). Presented at the 2025 International Conference on Quantum Communications, Networking, and Computing (QCNC), Nara, Japan: IEEE. https://doi.org/10.1109/QCNC64685.2025.00072

13. Chao, B., & Guangqiu, H. (2024). Innovative SVM optimization with differential gravitational fireworks for superior air pollution classification. *Scientific Reports*, *14*(1), 24450. https://doi.org/10.1038/s41598-024-75839-7

14. Dong, Y., Li, F., Zhu, T., & Yan, R. (2024). Air quality prediction based on quantum activation function optimized hybrid quantum classical neural network. *Frontiers in Physics*, *12*, 1412664. https://doi.org/10.3389/fphy.2024.1412664

15. Kubečka, J., Christensen, A. S., Rasmussen, F. R., & Elm, J. (2022). Quantum Machine Learning Approach for Studying Atmospheric Cluster Formation. *Environmental Science & Technology Letters*, *9*(3), 239–244. https://doi.org/10.1021/acs.estlett.1c00997

16. Akhtar, M. S. (2025). Quantum-Inspired Neural Radiative Transfer (QINRT): A Multi-Scale Computational Framework for Next-Generation Climate Intelligence. *AppliedMath*, *5*(4), 145. https://doi.org/10.3390/appliedmath5040145

17. Li, X., Luo, A., Li, J., & Li, Y. (2019). Air Pollutant Concentration Forecast Based on Support Vector Regression and Quantum-Behaved Particle Swarm Optimization. *Environmental Modeling & Assessment*, *24*(2), 205–222. https://doi.org/10.1007/s10666-018-9633-3

18. Sobhi, N. A., Shoieb, D. A., & Youssef, S. M. (2025). Comparative Analysis of Neural Network and Quantum Neural Network for Enhanced Air Quality Forecasting. In *2025 International Conference on Machine Intelligence and Smart Innovation (ICMISI)* (pp. 438–443). Presented at the 2025 International Conference on Machine Intelligence and Smart Innovation (ICMISI), Alexandria, Egypt: IEEE. https://doi.org/10.1109/ICMISI65108.2025.11115298

19. Mati, S., Usman, A. G., Ismael, G. Y., Babuga, U. T., Nadarajah, S., Masoud, S., … Abba, S. I. (2026). Explainable support vector regression coupled with quantum firefly optimisation algorithm for carbon emission prediction in West Africa: The role of socioeconomic, energy, and environmental factors. *Renewable Energy*, *256*, 124298. https://doi.org/10.1016/j.renene.2025.124298

20. Huang, T., Jiang, Y., Gan, R., & Wang, F. (2025). Quantum-Enhanced Attention Neural Networks for PM2.5 Concentration Prediction. *Modelling*, *6*(3), 69. https://doi.org/10.3390/modelling6030069

21. Hidayat, W., Maulidevi, N. U., & Surendro, K. (2025). Designing a Quantum Annealing-Based Method for Optimizing Backpropagation Neural Networks in Greenhouse Gas Emissions Prediction. In *Proceedings of the 2025 14th International Conference on Software and Computer Applications* (pp. 267–272). Presented at the ICSCA 2025: 2025 14th International Conference on Software and Computer Applications, Kuala Lumpur Malaysia: ACM. https://doi.org/10.1145/3731806.3731856

22. Santhiyagu, D., Nagarajan, G., & Refonaa, J. (2025). Enhanced Air Quality Forecasting with Quantum Machine Learning. In *2025 7th International Conference on Intelligent Sustainable Systems (ICISS)* (pp. 1048–1052). Presented at the 2025 7th International Conference on Intelligent Sustainable Systems (ICISS), India: IEEE. https://doi.org/10.1109/ICISS63372.2025.11076208

23. Zheng, K., Van Griensven, J., & Fraser, R. (2023). A Quantum Machine Learning Approach to Spatiotemporal Emission Modelling. *Atmosphere*, *14*(6), 944. https://doi.org/10.3390/atmos14060944

24. Yan, F., Liu, W., Dong, F., & Hirota, K. (2023). A quantum-inspired online spiking neural network for time-series predictions. *Nonlinear Dynamics*, *111*(16), 15201–15213. https://doi.org/10.1007/s11071-023-08655-9

25. Gómez, F. S., Sampaio, P. N. M., Peralta, L. M. R., Cuesta Astudillo, F. L., & De Oliveira Nunes, É. (2025). Leveraging Quantum Machine Learning for Accurate Indoor Air Quality Forecasting and Risk Mitigation. In A. Rocha, C. Ferrás, & H. Calvo (Eds.), *Information Technology and Systems* (Vol. 1448, pp. 276–286). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-93106-2_24

26. Wang, J., Li, H., Yang, H., & Wang, Y. (2021). Intelligent multivariable air-quality forecasting system based on feature selection and modified evolving interval type-2 quantum fuzzy neural network. *Environmental Pollution*, *274*, 116429. https://doi.org/10.1016/j.envpol.2021.116429

27. Mohandas, P., P, S., & R, S. (2025). Deep Feedforward Air Pollution Classification Empowered by Sophia-G Optimization. In *2025 4th International Conference on Innovative Mechanisms for Industry Applications*

(ICIMIA) (pp. 1696–1702). Presented at the 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Tirupur, India: IEEE. https://doi.org/10.1109/ICIMIA67127.2025.11200570

28. Mohandas, P., P, S., & R, S. (2025). AI-Driven ResNet50 with Channel Attention for Tamil Nadu Air Pollution Level Classification. In *2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 1508–1513). Presented at the 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Tirupur, India: IEEE. https://doi.org/10.1109/ICIMIA67127.2025.11200911
29. Sridharan Sivasubramanian, G.V. Sam Kumar, Tamilvizhi Thanarajan, & Rajendran, S. (2025). AL-BIRUNI EARTH RADIUS OPTIMIZATION FOR ENHANCED ENVIRONMENTAL DATA ANALYSIS IN REMOTE SENSING IMAGERY. *Agrociencia*, 1–18. https://doi.org/10.47163/agrociencia.v59i5.3380
30. Mohandas, P., P, S., & R, S. (2025). AI-Driven Spatial Attention-based MobileNetV2 Framework for Classifying Air Pollution Levels in Bangalore. In *2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 1928–1933). Presented at the 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Tirupur, India: IEEE. https://doi.org/10.1109/ICIMIA67127.2025.11200678
31. Xiao, X., Zhao, H., Ren, J., Fang, W.-H., & Li, Z. (2024). Physics-Constrained Hardware-Efficient Ansatz on Quantum Computers That Is Universal, Systematically Improvable, and Size-Consistent. *Journal of Chemical Theory and Computation*, *20*(5), 1912–1922. https://doi.org/10.1021/acs.jctc.3c00966
32. Chen, S. Y.-C., Wei, T.-C., Zhang, C., Yu, H., & Yoo, S. (2022). Quantum convolutional neural networks for high energy physics data analysis. *Physical Review Research*, *4*(1), 013231. https://doi.org/10.1103/PhysRevResearch.4.013231
33. Felser, T., Trenti, M., Sestini, L., Gianelle, A., Zuliani, D., Lucchesi, D., & Montangero, S. (2021). Quantum-inspired machine learning on high-energy physics data. *npj Quantum Information*, *7*(1), 111. https://doi.org/10.1038/s41534-021-00443-w
34. Wall, M. L., & D'Aguanno, G. (2021). Tree tensor network classifiers for machine learning: from quantum-inspired to quantum-assisted. https://doi.org/10.48550/ARXIV.2104.02249
35. Dutta, S., Innan, N., Marchisio, A., Yahia, S. B., & Shafique, M. (2024). QADQN: Quantum Attention Deep Q-Network for Financial Market Prediction. https://doi.org/10.48550/ARXIV.2408.03088

36. Tomal, S. M. Y. I., Shafin, A. A., Bhattacharjee, D., Amin, MD. K., & Shahir, R. S. (2025). Quantum-Enhanced Attention Mechanism in NLP: A Hybrid Classical-Quantum Approach. arXiv. https://doi.org/10.48550/ARXIV.2501.15630

37. Azado, P. C., Correr, G. I., Drinko, A., Medina, I., Canabarro, A., & Soares-Pinto, D. O. (2025). Expressibility, entangling power, and quantum average causal effect for causally indefinite circuits. *Physical Review A*, *111*(4), 042620. https://doi.org/10.1103/PhysRevA.111.042620

38. Panadero, I., Ban, Y., Espinós, H., Puebla, R., Casanova, J., & Torrontegui, E. (2024). Regressions on quantum neural networks at maximal expressivity. *Scientific Reports*, *14*(1), 31669. https://doi.org/10.1038/s41598-024-81436-5

39. Ragone, M., Bakalov, B. N., Sauvage, F., Kemper, A. F., Ortiz Marrero, C., Larocca, M., & Cerezo, M. (2024). A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits. *Nature Communications*, *15*(1), 7172. https://doi.org/10.1038/s41467-024-49909-3

40. Yao, Y., & Hasegawa, Y. (2025). Avoiding barren plateaus with entanglement. *Physical Review A*, *111*(2), 022426. https://doi.org/10.1103/PhysRevA.111.022426

41. Alvarez-Estevez, D. (2025). Benchmarking Quantum Machine Learning Kernel Training for Classification Tasks. *IEEE Transactions on Quantum Engineering*, *6*, 1–15. https://doi.org/10.1109/TQE.2025.3541882

42. Farsian, F., Parmiggiani, N., Rizzo, A., Panebianco, G., Bulgarelli, A., Schillirò, F., … Trombetti, T. (2025). Benchmarking Quantum Convolutional Neural Networks for Signal Classification in Simulated Gamma-Ray Burst Detection. arXiv. https://doi.org/10.48550/ARXIV.2501.17041

43. Khoo, J. Y., Gan, C. K., Ding, W., Carrazza, S., Ye, J., & Kong, J. F. (2024). Benchmarking Quantum Convolutional Neural Networks for Classification and Data Compression Tasks. arXiv. https://doi.org/10.48550/ARXIV.2411.13468

Statements

Funding

The authors declare that no funding was received by anyone for this research work at any stage.

Competing Interests

The authors declare no competing interests of any sort.

Conflict of interest

The authors have no conflicts of interest to declare.

Author Contributions

The corresponding author Jagadish Kumar Mogaraju has contributed to all sections of the manuscript.

Acknowledgments

The authors are grateful to the Central Pollution Control Board, Ministry of Environment, Forest and Climate Change, Government of India.

Clinical trial number: Not applicable.

Consent to Publish declaration: Not applicable

Consent to Participate declaration: Not applicable

Ethics and Consent to Participate declarations: Not applicable

Data availability statement

The datasets can be obtained from the Central Pollution Control Board, Ministry of Environment, Forest and Climate Change, Government of India, through https://airquality.cpcb.gov.in/ccr/#/caaqm-dashboard-all/caaqm-landing and https://cpcb.nic.in/namp-data/.

Declarations

Ethical approval

Not applicable

Consent to participate

Not applicable

Consent to publish

Not applicable