

Article

Evaluation of Machine Learning Models in Air Pollution Prediction for a Case Study of Macau as an Effort to Comply with UN Sustainable Development Goals

Thomas M. T. Lei ^{1,*} , Jianxiu Cai ², Altaf Hossain Molla ³ , Tonni Agustiono Kurniawan ⁴ 
and Steven Soon-Kai Kong ⁵ 

¹ Institute of Science and Environment, University of Saint Joseph, Macau, China

² Faculty of Applied Sciences, Macau Polytechnic University, Macau, China; jianxiu.cai@mpu.edu.mo

³ Department of Mechanical and Manufacturing Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), Bangi 43600, Selangor, Malaysia; altafhossain1994@gmail.com

⁴ College of Environment and Ecology, Xiamen University, Xiamen 361102, China; tonni@xmu.edu.cn

⁵ Department of Atmospheric Sciences, National Central University, Taoyuan 32001, Taiwan; kongsk@gmail.com

* Correspondence: thomas.lei@usj.edu.mo

Abstract: To comply with the United Nations Sustainable Development Goals (UN SDGs), in particular with SDG 3, SDG 11, and SDG 13, a reliable air pollution prediction model must be developed to construct a sustainable, safe, and resilient city and mitigate climate change for a double win. Machine learning (ML) and deep learning (DL) models have been applied to datasets in Macau to predict the daily levels of roadside air pollution in the Macau peninsula, situated near the historical sites of Macau. Macau welcomed over 28 million tourists in 2023 as a popular tourism destination. Still, an accurate air quality forecast has not been in place for many years due to the lack of a reliable emission inventory. This work will develop a dependable air pollution prediction model for Macau, which is also the novelty of this study. The methods, including random forest (RF), support vector regression (SVR), artificial neural network (ANN), recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU), were applied and successful in the prediction of daily air pollution levels in Macau. The prediction model was trained using the air quality and meteorological data from 2013 to 2019 and validated using the data from 2020 to 2021. The model performance was evaluated based on the root mean square error (RMSE), mean absolute error (MAE), Pearson's correlation coefficient (PCC), and Kendall's tau coefficient (KTC). The RF model best predicted PM₁₀, PM_{2.5}, NO₂, and CO concentrations with the highest PCC and KTC in a daily air pollution prediction. In addition, the SVR model had the best stability and repeatability compared to other models, with the lowest SD in RMSE, MAE, PCC, and KTC after five model runs. Therefore, the results of this study show that the RF model is more efficient and performs better than other models in the prediction of air pollution for the dataset of Macau.

Keywords: air pollution prediction; machine learning; deep learning; air quality; sustainability; sustainable development



Citation: Lei, T.M.T.; Cai, J.; Molla, A.H.; Kurniawan, T.A.; Kong, S.S.-K. Evaluation of Machine Learning Models in Air Pollution Prediction for a Case Study of Macau as an Effort to Comply with UN Sustainable Development Goals. *Sustainability* **2024**, *16*, 7477. <https://doi.org/10.3390/su16177477>

Academic Editor: Marc A. Rosen

Received: 6 August 2024

Revised: 25 August 2024

Accepted: 27 August 2024

Published: 29 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Air pollution is a trending topic due to repeated disturbances and its adverse effects on human health. Air pollution from human activities includes emissions from industrial factories, cars and motorcycles, airplanes, and the burning of biomass, coal, kerosene, and aerosol cans [1]. Dangerous air pollutants, including CO, CO₂, PM, NO₂, SO₂, O₃, NH₃, and Pb, are emitted into the atmospheric environment daily [1]. The World Health Organization (WHO) showed that air pollution is responsible for around 7 million premature deaths annually [2]. It also leads to other devastating outcomes, such as acidic rain, climate change,

and smog [3,4]. One of the most critical parameters directly correlated with air pollution is NO_2 , released primarily from diesel and gasoline engines during road transportation [5,6]. NO_2 was the most reactive pollutant emitted during the Industrial Revolution and is closely related to anthropogenic activities [5]. In addition, the air pollution level is significantly affected by various weather variables, such as atmospheric pressure, temperature, humidity, solar radiation, wind direction, and wind speed [7].

Air pollution has led to different forms of mortality, such as pneumonia [8,9], heart stroke [8,9], ischemic heart diseases [8,9], chronic obstructive pulmonary disease (COPD) [9,10], and lung cancer [9,10]. Nevertheless, air is a fundamental element for sustaining life and is polluted by both anthropogenic and natural factors; approximately 80% of cities worldwide and 90% of cities in middle-income countries exceed the recommended WHO air quality guidelines [11,12]. Accurately predicting air pollutants and identifying pollution trends can help scientists investigate effective emission control policies [11]. Likewise, air pollution negatively impacts the economy and human health; many countries encounter air pollution along their development path [13]. Epidemiological and experimental evidence shows $\text{PM}_{2.5}$ to be closely related to respiratory [13–15] and cardiovascular morbidity rates [16–18], lowering life expectancy [19,20] and increasing threats to public health even at a low concentration level [21–23]. Therefore, a highly accurate prediction model is much needed; traditional statistical models are frequently applied in the forecasting of air pollutant concentrations in different countries, including China and the USA [24–29], but these models have an issue of over-simplification and have challenges in capturing the nonlinear interaction relationship between multiple variate factors and pollutant concentrations [30].

A previous systematic review by Essamalali et al. [31] focused on supervised learning algorithms such as LSTM, RF, ANN, and SVR as instruments for a cleaner and healthier environment and on accurately forecasting key pollutants, such as PM, NO_x , CO, and O_3 , which allows urban planners and policymakers to make decisions to counter urban air pollution problems. This study showed that LSTM and RF are among the most popular DL and ML methods in air pollution prediction. A previous study by Zareba et al. [32] employed ML techniques for air pollution prediction in Krakow, Poland, and the best prediction result was achieved using linear ML methods. The best performance was observed when sudden changes occurred within the data due to the stability and immunity to overfitting. Another systematic review by Bellinger et al. [33] showed that early implementations preferred using ANNs. Still, more recent work has applied decision trees, support vector machines, and k-means clustering for air pollution prediction in Europe, China, and the USA. In addition, studies from Krichen et al. [34] and Raman et al. [35] verified that ML and AI are formal and reliable methods in prediction models.

As a popular tourism destination, Macau welcomed over 28 million tourists in 2023 and over 39 million tourists before the COVID-19 pandemic. Nevertheless, previous studies have focused on using one or more DL or ML methods to predict the same air quality dataset, and most studies have focused on predicting the air quality index (AQI). The novelty of this study is in applying four DL and two ML methods in the same work and performing the daily predictions of air pollution, which is more difficult to predict due to higher variability compared to AQI prediction. Due to the lack of a reliable emission inventory in Macau, developing an accurate air pollution prediction model in recent years has been very difficult. With the advancements in machine learning algorithms, a reliable and precise air pollution prediction method was developed in this work, which is also the novelty of this study.

The air pollution and weather variables were considered in this work because it is known that weather conditions significantly contribute to the levels of air pollutants. Also, this is the first study in the Macau region combining DL and ML methods in air pollution predictions using Macau's dataset. As an effort to ensure the excellent health and well-being of the residents (in compliance with SDG 3: Good Health and Well-Being), construct a sustainable, safe, and resilient city (in compliance with SDG 11: Sustainable Cities and Communities), and take urgent action to combat climate change (in compliance with SDG

13: Climate Action), a reliable and accurate air pollution prediction model must be built to warn the residents ahead of an air pollution episode, which ensures that vulnerable groups of people with pre-existing health conditions such as heart or lung diseases avoid outdoor activities during poor air quality days and is also an effort to create safe and sustainable cities for Macau. By accurately identifying and reducing air pollution problems, the effects of climate change may also be mitigated for a double win. Hence, using computational ML methods may be a solution for accurate air pollution predictions in Macau, and this study will investigate the efficiency and accuracy of the proposed models.

2. Materials and Methods

2.1. Data Collection

This study obtained the daily observations of air pollution and weather data from the Direcção dos Serviços Meteorológicos e Geofísicos (SMG) and Hong Kong Observatory (HKO) from 2013 to 2021 to build and train the DL and ML models. As shown in Figure 1, the data collection sites are located in the city center of the Macau peninsula (represented by a red circle on the map). Over 190,000 data entries were applied in this work. The air pollution data collected by SMG were obtained using EPA-equivalent methods, which is on par with the international standards of air quality monitoring stations (AQMSs) in the other regions. The data selected for the model execution included air pollution and weather data, which are crucial for studies of this nature. The air pollutants present in the ambient air during the previous day and hour are known to have a significant effect on the current air quality, while the weather parameters, including temperature, wind speed, and wind direction, are also known to have a direct impact on the current air quality. Therefore, these categories of data were identified and selected for this study.



Figure 1. Map of AQMS locations in Macao (adapted from SMG).

Table 1 shows the air quality, meteorological, and other variables applied in the prediction models in the training and validation datasets. A previous study by Zareba

et al. [36] shows the importance of including climate and weather data in air pollution predictions.

Table 1. Different variables applied in the air pollution prediction models.

Types of Data	Variables	Description of Variables
Air Quality	PM ₁₀ , PM _{2.5} , NO ₂ , CO	Mean hourly concentration values (micrograms per cubic meter)
	16D1, 23D0, 23D1, 23D2, 23D3	16D1: 24 h concentration averaging period between 16:00 of D1 and 15:00 of D0 23D0: 24 h concentration averaging period between 0:00 and 23:59 of D0 23D1: 24 h concentration averaging period between 0:00 and 23:59 of D1 23D2: 24 h concentration averaging period between 0:00 and 23:59 of D2 23D3: 24 h concentration averaging period between 0:00 and 23:59 of D3
	D0, D1, D2, D3	D0: Prediction Day; D1: Prediction Day Minus One; D2: Prediction Day Minus Two; D3: Prediction Day Minus Three
Meteorological Variables	H1000, H850, H700, H500	Geopotential height of one thousand hectopascal, eight hundred and fifty hectopascal, seven hundred hectopascal, and five hundred hectopascal (meters)
	TAR925, TAR850, TAR700	Air temperature of nine hundred and twenty-five hectopascal, eight hundred and fifty hectopascal, seven hundred hectopascal (degree Celsius)
	HR925, HR850, HR700	Relative humidity of nine hundred and twenty-five hectopascal, eight hundred and fifty hectopascal, seven hundred hectopascal (percentage)
	TD925, TD850, TD700	Dew point temperature of nine hundred and twenty-five hectopascal, eight hundred and fifty hectopascal, seven hundred hectopascal (degree Celsius)
	THI850, THI700, THI500	Thickness of eight hundred and fifty hectopascal, seven hundred hectopascal, and five hundred hectopascal (meters)
	STB925, STB850, STB700	Stability of nine hundred and twenty-five hectopascal, eight hundred and fifty hectopascal, seven hundred hectopascal (degree Celsius)
	T_AIR_MX, T_AIR_MD, T_AIR_MN	Highest, mean, and lowest air temperature (degree Celsius)
	HRMX, HRMD, HRMN	Highest, mean, and lowest relative humidity (percentage)
	TD_MD	Mean dew point temperature (degree Celsius)
	RRTT	Rainfall (millimeters)
	VMED	Mean wind speed (meters per second)
	PREV_WDIR	Predominant wind direction (degree)
Other Variables	DD	Number of sunshine hours per day (hour)
	FF	Weekday indicator: non-weekend = 0 and weekend = 1

2.2. Experimental Phase

The air pollution forecasting models were built using one air quality variable (e.g., PM₁₀) and all meteorological variables and other variables for each type of pollutant. This process was repeated for PM_{2.5}, NO₂, and CO. The entire dataset was first split into the training set and testing set in a ratio of 7:2. The training set contains data from 2013 to 2019, while the testing set contains data from 2020 to 2021. In the training set, the data for the year 2019 were allocated as the validation set to optimize the hyperparameters of models. After the determination of hyperparameters, the models were trained on the entire training dataset from 2013 to 2019, and the evaluation of model performance on the testing dataset from 2020 to 2021 was performed. Each of the six models (ANN, GRU, LSTM, RF, RNN, and SVR) was run five times in Python using TensorFlow 2.6.2, and the result shown is the average of the five model runs. The parameters used to evaluate the models are known as the model performance indicators, which include root mean square error (RMSE), mean absolute error (MAE), Pearson's correlation coefficient (PCC), and Kendall's tau coefficient

(KTC). Figure 2 shows the workflow for the air pollution prediction in this study. The study began with data collection, followed by data curation (converting the units and removing invalid data), and then entered the experimental phase, with both ML (RF and SVR) and DL (ANN, GRU, LSTM, RNN) methods being applied. Each of the models was run five times to obtain an average result. The model performance indicator (RMSE, MAE, PCC, KTC) was generated to evaluate the performance of each model.

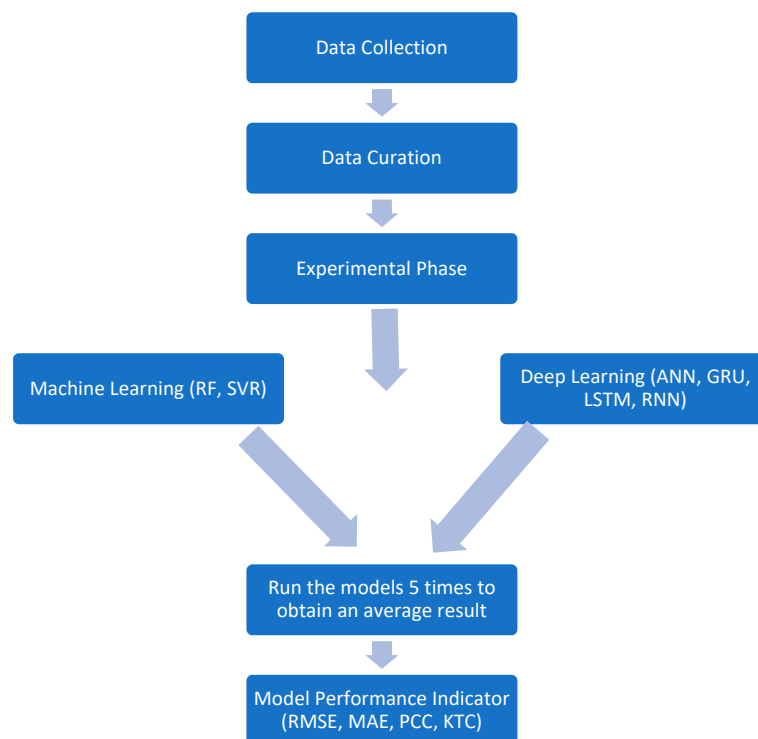


Figure 2. Study workflow for air pollution prediction.

2.3. Learning Algorithms

2.3.1. Artificial Neural Network (ANN)

ANNs are a very versatile computational method that can be used for classification, forecasting, and pattern recognition. The inherent mathematical structure of ANNs allows them to capture complex relationships within the dataset [37]. The ANN model has a neurological network that includes the input, the hidden, and the output layer. With the correct calibration and enough datasets, it can tackle any linear or nonlinear problems [7,38]. Also, previous studies found that ANNs are preferred when predicting PM [13]. The concept of ANNs covers many areas of expertise, including biology, mathematics, programming, engineering, statistics, and informatics, combined to simulate the function of a human brain and be able to model nonlinear data, scalability, and rational and contextualized outcomes [5]. For instance, humans learn new things by training the neurons within the brain using examples and storing the latest information in the memory system. At the same time, an ANN is fed input data into its artificial neurons for training, so the network is calibrated to achieve the best outcome in a prediction task [4]. The advantages of ANNs include the fact that neural networks can implement tasks that a linear model cannot perform. In contrast, the disadvantages of ANNs include the required training to operate and the fact that they often require a high processing time for extensive neural networks.

2.3.2. Gated Recurrent Unit (GRU)

GRU is a simple version of LSTM that integrates the forget and the input gates [39]. The RNN also inspires the architecture of GRU, and the info within the GRU is regulated by a two-gate system, which has a quicker training time due to less computational load [5,40].

Also, GRU is appropriate for processing text, speech, and time series data, which utilize a gate structure to facilitate information flow with the reset and the update gate [37,41]. GRU's high efficiency and excellent performance make it a good choice for many sequential data tasks [42,43]. The advantages of GRU include fewer parameters, making it computationally less expensive and faster to train than LSTM. In contrast, the disadvantages of GRU included being more prone to overfitting on smaller datasets and requiring careful tuning of hyperparameters.

2.3.3. Long Short-Term Memory (LSTM)

LSTM is an advanced version of RNN which possesses extensive memory capable of dealing with long-term dependencies and memorizing information over an extensive period, with main components such as cell state holding information throughout the data processing [39]. LSTM has three logic gates that control how information flows through the system and can also overcome the weakness of RNNs [5]. LSTM is an extraordinary deep learning model with improved capabilities to overcome the challenging conditions faced by RNNs with three main functional areas: the forget gate, the input gate, and the output gate [37,44,45]. The advantages of LSTM included handling long sequences and being appropriate for processing data with long-range dependencies. The disadvantages of LSTM included computational complexity and being more intensive compared to other neural networks, being prone to overfitting when there are insufficient training data, and requiring several hyperparameter tunings.

2.3.4. Random Forest (RF)

RF has evolved from several variations of decision trees with forecasting based on the average prediction provided by a series of trees [39]. RF is a supervised learning ensemble algorithm, in combination with decision trees, to form a forest with the bagging concept, which adds randomness to the trees, making it capable of solving classification and regression cases [4]. The decision tree comprises three parts, the root, the leaf, and the internal node, with the root node storing all the data, the internal node classifying the features, and the leaf node representing the output results [7,46]. In addition, RF is known to be a simple yet efficient and interpretable method, making it one of the most favorable techniques for air pollution forecasting [8]. The advantages of RF include the bias remaining the same with multiple decision trees and it being straightforward and user-friendly to make predictions. In contrast, the disadvantages of RF include not training well on smaller datasets and the time required to train multiple big decision trees.

2.3.5. Recurrent Neural Network (RNN)

RNN can work with both time series and sequential data, with an internal memory that can implement neuron feedback itself. This allows the model to have short-term memory, making it possible for time series forecasting [39]. RNNs represent a favorable choice for time series-based forecasting, and a standard RNN has simple functions for input data and forecasting with minimal errors and simple mechanisms. Some drawbacks of RNNs include gradient exploding/disappearing and data morphing [5,47–49]. The advantages of RNNs include their effectiveness in sequence modeling tasks, their explicit design for processing sequential data, and their ability to process sequences of variable length. In contrast, the disadvantages of RNNs include limited memory, being problematic when dealing with long sequences, and being biased toward recent data in the sequence.

2.3.6. Support Vector Regression (SVR)

SVR is a regression derived from the support vector machine (SVM), which can be used in classification problems [39]. SVR is established by optimizing SVM based on deep learning functions and prevailing overfitting, underfitting, and local optimization using mathematical methods and techniques [8,13,50]. The advantages of SVR include good robustness against outliers, good forecasting accuracy, ease of implementation, resistance to

overfitting, and prevailing against nonlinear limitations and uncertainties. In contrast, the disadvantages of SVR include needing to be more suitable for large datasets and performing better when there is more noise in the dataset.

2.3.7. Shapley Additive Explanation (SHAP)

SHAP values were applied to interpret the feature score in the different forecasting models, which determines the impact of each feature on the forecasting model [7,8,51,52]. SHAP values offer a clear explanation for the contribution of each feature, which is essential to trace the source that leads to severe air pollution problems.

2.4. Model Evaluation Metrics

Four evaluating metrics were applied to validate the model's prediction accuracy: RMSE, MAE, PCC, and KTC. Specifically, RMSE measures the average difference between values predicted by a model and the actual values, MAE measures the average absolute differences between the predicted values and the actual target values, PCC measures the linear correlation with a number between -1 and 1 to determine the strength and direction of the relationship between two variables, and KTC measures the ordinal association between two variables and a nonparametric measure of strength and direction of association.

The equations are shown in the following:

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(\hat{y}_i - y_i)^2}{N}} \quad (1)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (2)$$

$$PCC = 1 - \frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sum_{i=1}^N (y_i - \bar{y})^2 \sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2} \quad (3)$$

where y_i represents the ground truth value of regressive models; $\hat{y}_i$ denotes the i th sample of prediction values of the regression model; and N is the sample number.

$$KTC = \frac{P - Q}{\sqrt{(P + Q + T)(P + Q + U)}} \quad (4)$$

where P is the number of concordant pairs, Q is the number of discordant pairs, T is the number of ties only in x , and U is the number of ties only in y .

2.5. Features and Testing Scenario

Table 2 shows the model parameters and hyperparameters used in this study that are closely associated with the model's performance in air pollution prediction. The model parameters and hyperparameters were determined after a rigorous review of the previous literature for the best performance in air pollution forecasting [53].

Table 2. Different model parameters and hyperparameters were applied to this study.

Models	Model Parameters and Hyperparameters	
ANN	learning rate	0.0005
	epochs	100
	batch_size	32
	validation split	0.3
GRU	optimizer	adam
	layers	100
	epochs	500
	batch size	32

Table 2. Cont.

Models	Model Parameters and Hyperparameters	
LSTM	optimizer	adam
	epochs	20
	batch size	64
RF	n_estimators	100
	criterion	squared error
	min_sample_split	2
	max_depth	none
RNN	optimizer	adam
	layers	100
	epochs	500
	batch size	32
SVR	kernal	linear
	degree	3
	gamma	scaler

3. Results and Discussion

The models applied in this work can be grouped into two categories, namely DL models and ML models. A total of six models were applied, specifically four DL models (ANN, GRU, LSTM, and RNN) and two ML models (RF and SVR). DL models are generally more complicated and require more effort to develop. In contrast, ML models are more straightforward and designed for user-friendly applications due to having fewer parameters and hyperparameters. The air pollutants predicted in this work included PM₁₀, PM_{2.5}, NO₂, and CO, which caused the most health concerns and received the most public attention. Table 3 shows the model performance indicator of DL and ML models (average RMSE, MAE, PCC, and KTC) after five runs for each model.

The ANN, RF, and SVR models showed the best performance in predicting PM_{2.5} with the highest PCC and KTC, while the GRU model showed the best performance in forecasting NO₂. Figure 3 shows the measured and predicted PM_{2.5} concentrations using ANNs in 2020 and 2021, which offer a high PCC of 0.87 and a KTC of 0.71. Figure 4 shows the measured and predicted PM_{2.5} concentration using RF in 2020 and 2021, which provides a high PCC of 0.99 and a KTC of 0.92. Figure 5 shows the SHAP value of the RF model in PM_{2.5} prediction, demonstrating that the past days' air pollutants and weather data play a vital role in the prediction. Figure 6 shows the measured and predicted CO concentrations using SVR in 2020 and 2021, which offers a high PCC of 0.72 and a KTC of 0.51. Although the SVR model for CO did not receive the highest PCC and KTC, it provides the lowest RMSE and MAE (0.28 and 0.24, respectively). Figure 7 shows the measured and predicted NO₂ concentrations using GRU in 2020 and 2021, which offers a high PCC of 0.80 and a KTC of 0.62.

Table 3. Model performance indicator of DL and ML models in forecasting air pollutants (PM₁₀, PM_{2.5}, NO₂, CO), trained using 2013 to 2019 data and validated with 2020 and 2021 data.

Categories	Model	Pollutant	Model Performance Indicators			
			RMSE	MAE	PCC	KTC
Deep Learning Model	ANN	PM ₁₀	16.15	12.63	0.84	0.67
		PM _{2.5}	6.94	5.13	0.87	0.71
		NO ₂	13.34	10.16	0.82	0.64
		CO	0.58	0.51	0.37	0.25
Deep Learning Model	GRU	PM ₁₀	24.26	19.19	0.74	0.55
		PM _{2.5}	12.18	9.19	0.73	0.56
		NO ₂	15.07	12.03	0.80	0.62
		CO	0.38	0.30	0.61	0.42

Table 3. Cont.

Categories	Model	Pollutant	Model Performance Indicators			
			RMSE	MAE	PCC	KTC
Deep Learning Model	LSTM	PM ₁₀	16.69	14.20	0.88	0.70
		PM _{2.5}	9.58	8.16	0.86	0.69
		NO ₂	11.64	9.32	0.86	0.71
		CO	0.29	0.24	0.69	0.49
Machine Learning Model	RF	PM ₁₀	6.81	4.86	0.96	0.84
		PM_{2.5}	2.04	1.20	0.99	0.92
		NO ₂	6.93	4.84	0.94	0.81
		CO	0.12	0.09	0.82	0.58
Deep Learning Model	RNN	PM ₁₀	21.16	18.06	0.78	0.58
		PM _{2.5}	12.76	10.35	0.67	0.48
		NO ₂	19.08	15.86	0.71	0.52
		CO	0.39	0.31	0.44	0.29
Machine Learning Model	SVR	PM ₁₀	7.45	5.63	0.95	0.82
		PM_{2.5}	3.52	2.46	0.96	0.84
		NO ₂	12.19	10.12	0.92	0.77
		CO	0.28	0.24	0.72	0.51

The best performance in each category is shown in bold.

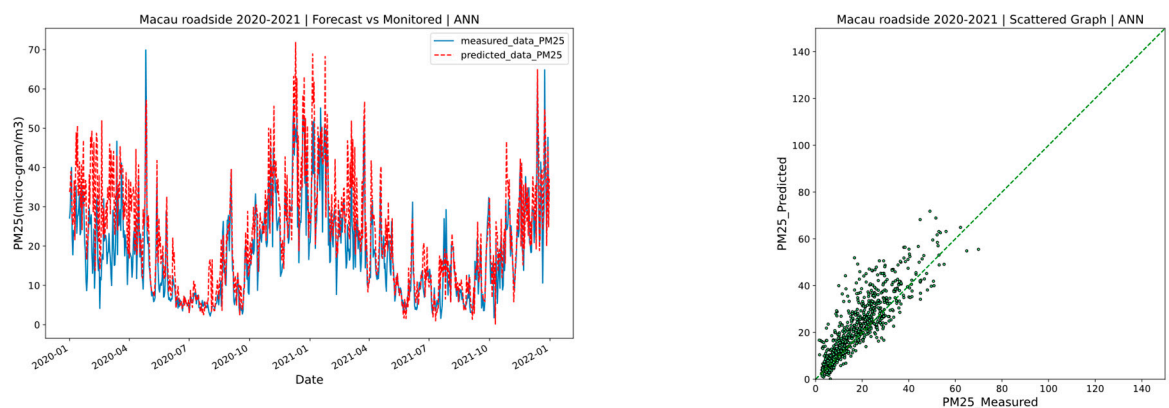


Figure 3. Measured and predicted PM_{2.5} concentrations using ANN for 2020 and 2021.

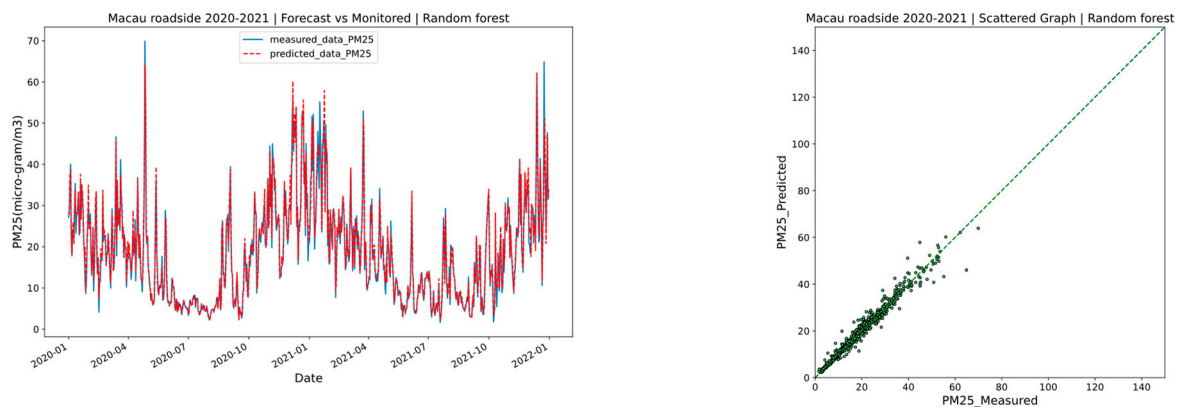


Figure 4. Measured and predicted PM_{2.5} concentrations using RF for 2020 and 2021.

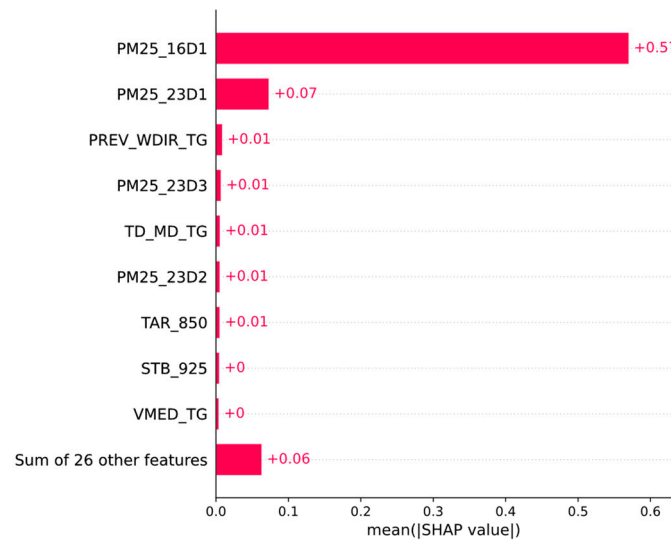


Figure 5. SHAP values for the RF model of PM_{2.5} prediction.

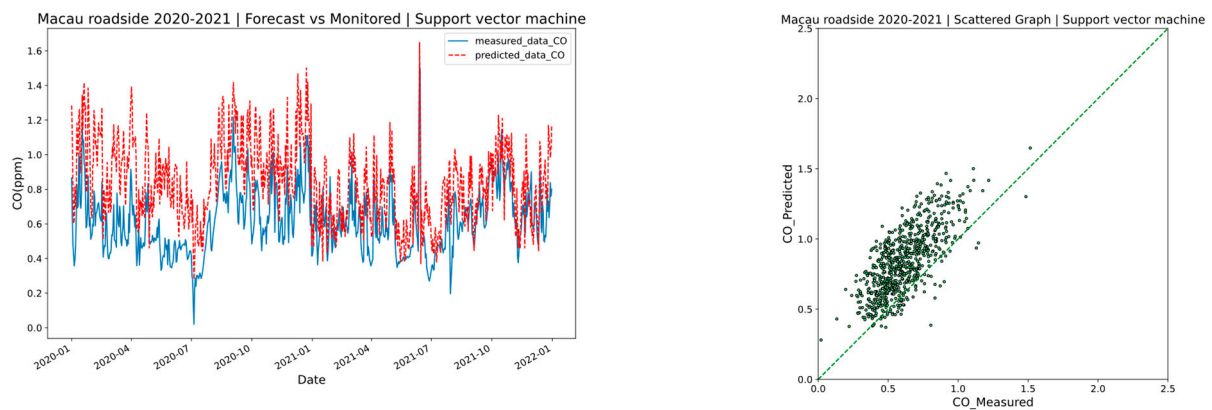


Figure 6. Measured and predicted CO concentrations using SVR for 2020 and 2021.

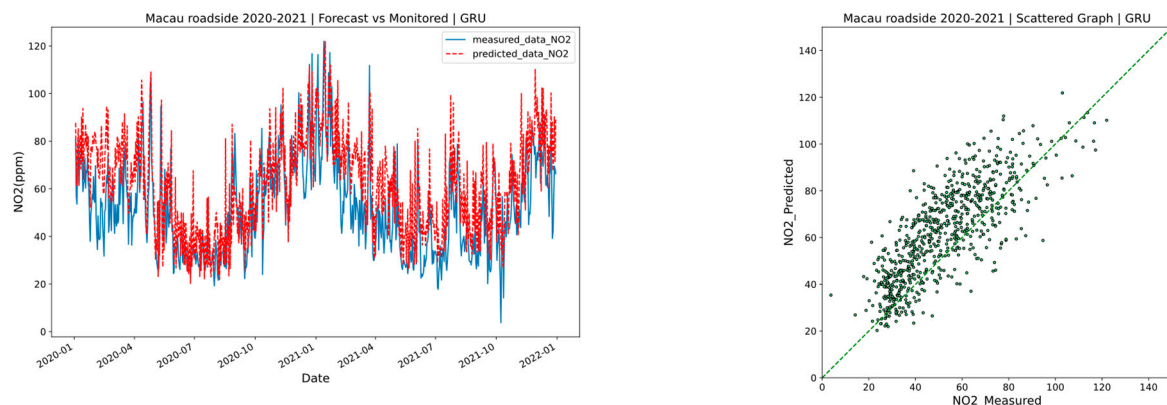


Figure 7. Measured and predicted NO₂ concentrations using GRU for 2020 and 2021.

In addition, the RNN model showed the best performance in predicting PM₁₀ with the highest PCC and KTC. Also, the LSTM model showed the best performance in predicting PM₁₀ with the highest PCC. Figure 8 shows the measured and predicted PM₁₀ concentrations using RNN in 2020 and 2021, which offers a high PCC of 0.78 and a KTC of 0.58. Figure 9 shows the measured and predicted PM₁₀ concentrations using LSTM in 2020 and 2021, which provides a high PCC of 0.88 and a KTC of 0.70.

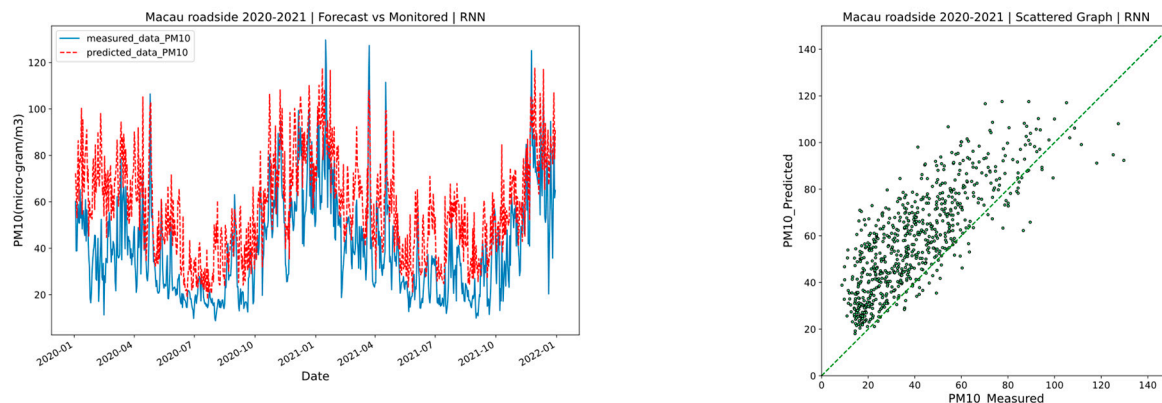


Figure 8. Measured and predicted PM₁₀ concentrations using RNN for 2020 and 2021.

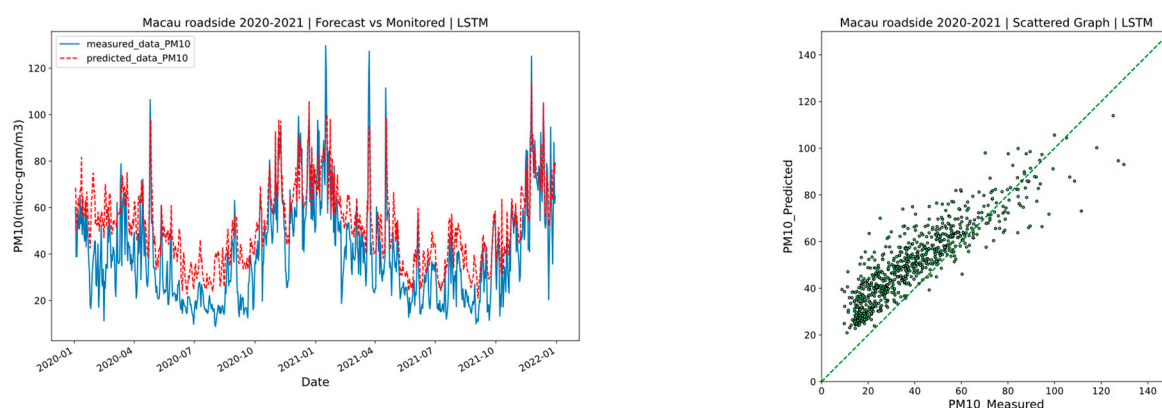


Figure 9. Measured and predicted PM₁₀ concentrations using LSTM for 2020 and 2021.

Overall, the RF model provided the best prediction for all pollutants, specifically for PM₁₀ (PCC = 0.96 and KTC = 0.84), PM_{2.5} (PCC = 0.99 and KTC = 0.92), NO₂ (PCC = 0.94 and KTC = 0.81), and CO (PCC = 0.82 and KTC = 0.58), respectively, while the SVR model performed the second best in terms of prediction for all pollutants, specifically for PM₁₀ (PCC = 0.95 and KTC = 0.82), PM_{2.5} (PCC = 0.96 and KTC = 0.84), NO₂ (PCC = 0.92 and KTC = 0.77), and CO (PCC = 0.72 and KTC = 0.51), respectively.

Table 4 shows the standard deviation (SD) of the average five model runs for the different models. It is essential to assess the SD in terms of the PCC and KTC of the model, especially after five model runs, to ensure the stability and repeatability of the model. The ANN and GRU models have the lowest SD for PM_{2.5}; the LSTM model has the lowest SD for PM₁₀; the RF model has the lowest SD for PM₁₀, PM_{2.5}, and NO₂; the RNN model has the lowest SD for CO; and the SVR model has the lowest SD for PM₁₀, PM_{2.5}, NO₂, and CO.

Overall, the SVR model has a 0.00 value with the lowest SD for RMSE, MAE, PCC, and KTC for all pollutants after five model runs, which shows that the SVR models performed the best in terms of stability and repeatability amongst the prediction models applied in this study.

The obtained results of this study show that the RF model has the highest PCC and KTC values for the air pollution prediction for all pollutants, including PM₁₀, PM_{2.5}, NO₂, and CO, while the SVR model has the lowest SD for all model evaluation metrics, including RMSE, MAE, PCC, and KTC. A previous study by Liang et al. [4] showed that traditional machine learning methods, including AdaBoost and stacking ensemble, outperform deep learning methods, including ANN, in Taiwan's air quality index (AQI) forecast. Also, Mampitiya et al. [37] show that the traditional ensemble model best predicted PM₁₀ levels in Sri Lanka amongst deep learning models, including ANN, LSTM, and GRU.

Table 4. SD of the average five model runs for the different DL and ML models.

Categories	Model	Pollutant	SD of the Five Model Runs			
			RMSE	MAE	PCC	KTC
Deep Learning Model	ANN	PM ₁₀	4.06	3.78	0.06	0.06
		PM_{2.5}	1.02	0.76	0.03	0.03
		NO ₂	2.03	1.71	0.05	0.05
		CO	0.21	0.19	0.09	0.07
Deep Learning Model	GRU	PM ₁₀	6.21	5.41	0.08	0.07
		PM_{2.5}	1.28	1.03	0.03	0.02
		NO ₂	2.31	2.00	0.04	0.05
		CO	0.06	0.06	0.04	0.03
Deep Learning Model	LSTM	PM₁₀	1.68	1.47	0.02	0.02
		PM _{2.5}	2.54	2.29	0.03	0.04
		NO ₂	2.48	2.36	0.02	0.03
		CO	0.04	0.04	0.07	0.06
Machine Learning Model	RF	PM ₁₀	0.04	0.05	0.00	0.00
		PM_{2.5}	0.03	0.01	0.00	0.00
		NO ₂	0.02	0.02	0.00	0.00
		CO	0.00	0.00	0.00	0.01
Deep Learning Model	RNN	PM ₁₀	2.93	2.63	0.06	0.06
		PM _{2.5}	4.35	3.62	0.09	0.09
		NO ₂	2.08	1.96	0.04	0.04
		CO	0.10	0.08	0.04	0.02
Machine Learning Model	SVR	PM ₁₀	0.00	0.00	0.00	0.00
		PM_{2.5}	0.00	0.00	0.00	0.00
		NO ₂	0.00	0.00	0.00	0.00
		CO	0.00	0.00	0.00	0.00

The best performance in each category is shown in bold.

4. Conclusions

Air pollution is primarily caused by gaseous pollutants such as CO₂, NO_x, SO_x, and PM_{2.5}. SO_x is known to form acid rain, while CO₂ is known to be a primary contributor of greenhouse gasses (GHGs). Developing tools to forecast air pollutants accurately is essential to build a more sustainable and environmentally friendly future [54]. In addition, a good and reliable forecast is significant as it could partially fulfill the requirements of SDG 3, SDG 11, and SDG 13, which include protecting the life and well-being of vulnerable groups of people, making the region a safe and sustainable place to live, and mitigating the effect of climate change through immediate action. The application of DL and ML models to forecast the daily mean concentration of PM₁₀, PM_{2.5}, NO₂, and CO for tomorrow, from 2020 to 2021, was successful in the coastal city of Macau, located in Southern China. The results show that traditional ML models performed better with higher stability in the air pollution prediction for all pollutants, which aligned with the findings from previous studies [1,4,7,8,13,37]. Overall, the RF model performed best in air pollution prediction for PM₁₀, PM_{2.5}, NO₂, and CO, with the highest PCC and KTC. In addition, SVR has the best stability and repeatability compared to other models, with the lowest SD in RMSE, MAE, PCC, and KTC after five model runs. The SVR models also come second best regarding air pollution prediction for all pollutants, closely behind the RF models. The result of this study confirmed that the RF and SVR models outperform the other models in air pollution prediction during a case study in Macau. Despite the similarity between PM₁₀ and PM_{2.5}, these particles are often derived from different emission sources with different chemical compositions. Therefore, the results may not be directly comparable. Also, previous studies [53,55] showed a marginal accuracy advantage in the prediction of PM_{2.5} over PM₁₀, which also aligned with the results of this study.

The novelty of this work is that an air pollution prediction model has been successfully developed and may become operational at the discretion of the local authorities. This is a significant advancement for a small region such as Macau, where an air pollution prediction was not in place for many years. In addition, by identifying and reducing the air pollution problems, the effect of climate change may also be mitigated for a double win. The results obtained in this study may be due to DL models requiring a more extensive dataset to reach optimal performance. In contrast, the ML models do fine with a relatively small dataset, with over 190,000 entries of data being applied in this study. A comparison study between ML and DL models in the prediction of air pollution in major Chinese cities with a sample size of 7 years also confirms that ML models outperform DL models [56], with the reason being the consideration of smaller sample size, which is similar to the study period of this work, a sample size of 9 years. The RF and SVR models may be considered the priority choice for air pollution predictions in future datasets. Nevertheless, the application of hybrid DL and ML methods to more datasets globally may be considered for future studies to validate that ML models outperform DL models in air pollution prediction.

Author Contributions: Conceptualization, T.M.T.L. and J.C.; methodology, T.M.T.L. and J.C.; software, J.C.; validation, T.M.T.L. and S.S.-K.K.; formal analysis, J.C.; investigation, T.M.T.L. and T.A.K.; resources, T.M.T.L.; data curation, J.C.; writing—original draft preparation, T.M.T.L., A.H.M., T.A.K. and S.S.-K.K.; writing—review and editing, T.M.T.L., A.H.M., T.A.K. and S.S.-K.K.; visualization, J.C.; supervision, T.M.T.L. and T.A.K.; project administration, T.M.T.L.; funding acquisition, T.M.T.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study used third party data. Restrictions apply to the availability of these data.

Acknowledgments: The developed work was supported by The Macao Meteorological and Geophysical Bureau (SMG).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kumar, K.; Pande, B.P. Air Pollution Prediction with Machine Learning: A case study of Indian cities. *Int. J. Environ. Sci. Technol.* **2022**, *20*, 5333–5348. [CrossRef] [PubMed]
2. World Health Organization. Air Pollution. Available online: https://www.who.int/health-topics/air-pollution#tab=tab_1 (accessed on 6 May 2024).
3. Ghorani-Azam, A.; Riahi-Zanjani, B.; Balali-Mood, M. Effects of Air Pollution on Human Health and Practical Measures for Prevention in Iran. *J. Res. Med. Sci.* **2016**, *21*, 65. [CrossRef]
4. Liang, Y.-C.; Maimury, Y.; Chen, A.H.-L.; Juarez, J.R.C. Machine Learning-Based Prediction of Air Quality. *Appl. Sci.* **2020**, *10*, 9151. [CrossRef]
5. Cican, G.; Buturache, A.-N.; Mirea, R. Applying Machine Learning Techniques in Air Quality Prediction—A Bucharest City Case Study. *Sustainability* **2023**, *15*, 8445. [CrossRef]
6. Jonson, J.E.; Borken-Kleefeld, J.; Simpson, D.; Nyiri, A.; Posch, M.; Heyes, C. Impact of excess NO_x emissions from diesel cars on air quality, public health and eutrophication in Europe. *Environ. Res. Lett.* **2017**, *12*, 094017. [CrossRef]
7. Lei, T.M.T.; Ng, S.C.W.; Siu, S.W.I. Application of ANN, XGBoost, and Other ML Methods to Forecast Air Quality in Macau. *Sustainability* **2023**, *15*, 5341. [CrossRef]
8. Lei, T.M.T.; Siu, S.W.I.; Monjardino, J.; Mendes, L.; Ferreira, F. Using Machine Learning Methods to Forecast Air Quality: A Case Study in Macao. *Atmosphere* **2022**, *13*, 1412. [CrossRef]
9. WHO. *World Health Statistics 2021: Monitoring Health for the SDGs, Sustainable Development Goals*; WHO: Geneva, Switzerland, 2021. Available online: <https://iris.who.int/bitstream/handle/10665/342703/9789240027053-eng.pdf?sequence=1> (accessed on 1 May 2024).
10. Zaheer, J.; Jeon, J.; Lee, S.-B.; Kim, J.S. Effect of Particulate Matter on Human Health, Prevention, and Imaging Using PET or SPECT. *Prog. Med. Phys.* **2018**, *29*, 81. [CrossRef]

11. Mathew, A.; Gokul, P.R.; Raja Shekar, P.; Arunab, K.S.; Ghassan Abdo, H.; Almohamad, H.; Abdullah Al Dughairi, A. Air Quality Analysis and PM_{2.5} modelling using machine learning techniques: A study of Hyderabad City in India. *Cogent Eng.* **2023**, *10*, 2243743. [\[CrossRef\]](#)
12. Raju, L.; Gandhimathi, R.; Mathew, A.; Ramesh, S.T. Spatio-temporal modelling of particulate matter concentrations using satellite derived aerosol optical depth over coastal region of Chennai in India. *Ecol. Inform.* **2022**, *69*, 101681. [\[CrossRef\]](#)
13. Ma, X.; Chen, T.; Ge, R.; Cui, C.; Xu, F.; Lv, Q. Time Series-based PM_{2.5} concentration prediction in Jing-jin-ji Area using machine learning algorithm models. *Heliyon* **2022**, *8*, e10691. [\[CrossRef\]](#)
14. Al-Hemoud, A.; Gasana, J.; Al-Dabbous, A.; Alajeel, A.; Al-Shatti, A.; Behbehani, W.; Malak, M. Exposure levels of air pollution (PM_{2.5}) and Associated Health Risk in Kuwait. *Environ. Res.* **2019**, *179*, 108730. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Apte, J.S.; Brauer, M.; Cohen, A.J.; Ezzati, M.; Pope, C.A. Ambient PM_{2.5} reduces global and regional life expectancy. *Environ. Sci. Technol. Lett.* **2018**, *5*, 546–551. [\[CrossRef\]](#)
16. Bu, X.; Xie, Z.; Liu, J.; Wei, L.; Wang, X.; Chen, M.; Ren, H. Global PM_{2.5}-attributable health burden from 1990 to 2017: Estimates from the global burden of disease study 2017. *Environ. Res.* **2021**, *197*, 111123. [\[CrossRef\]](#)
17. Burnett, R.T.; Pope, C.A.; Ezzati, M.; Olives, C.; Lim, S.S.; Mehta, S.; Shin, H.H.; Singh, G.; Hubbell, B.; Brauer, M.; et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ. Health Perspect.* **2014**, *122*, 397–403. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Diao, B.; Ding, L.; Zhang, Q.; Na, J.; Cheng, J. Impact of Urbanization on PM_{2.5}-Related Health and Economic Loss in China 338 Cities. *Int. J. Environ. Res. Public Health* **2020**, *17*, 990. [\[CrossRef\]](#)
19. Geng, G.; Zheng, Y.; Zhang, Q.; Xue, T.; Zhao, H.; Tong, D.; Zheng, B.; Li, M.; Liu, F.; Hong, C.; et al. Drivers of PM_{2.5} air pollution deaths in China 2002–2017. *Nat. Geosci.* **2021**, *14*, 645–650. [\[CrossRef\]](#)
20. Xing, Y.; Xu, Y.; Shi, M.; Lian, Y. The impact of PM_{2.5} on the human respiratory system. *J. Thorac. Dis.* **2016**, *8*, 69–74. [\[CrossRef\]](#)
21. Feng, S.; Gao, D.; Liao, F.; Zhou, F.; Wang, X. The health effects of ambient PM_{2.5} and potential mechanisms. *Ecotoxicol. Environ. Saf.* **2016**, *128*, 67–74. [\[CrossRef\]](#)
22. Ouyang, R.; Yang, S.; Xu, L. Analysis and Risk Assessment of PM_{2.5}-Bound PAHs in a Comparison of Indoor and Outdoor Environments in a Middle School: A Case Study in Beijing, China. *Atmosphere* **2020**, *11*, 904. [\[CrossRef\]](#)
23. Yu, W.; Guo, Y.; Shi, L.; Li, S. The association between long-term exposure to low-level PM_{2.5} and mortality in the State of Queensland, Australia: A modelling study with the difference-in-differences approach. *PLoS Med.* **2020**, *17*, e1003141. [\[CrossRef\]](#)
24. Alyousifi, Y.; Ibrahim, K.; Kang, W.; Zin, W.Z. Markov chain modeling for Air Pollution Index based on maximum a posteriori method. *Air Qual. Atmos. Health* **2019**, *12*, 1521–1531. [\[CrossRef\]](#)
25. Faganeli Pucer, J.; Pirš, G.; Štrumbelj, E. A bayesian approach to forecasting daily air-pollutant levels. *Knowl. Inf. Syst.* **2018**, *57*, 635–654. [\[CrossRef\]](#)
26. Liu, Y.; Guo, H.; Mao, G.; Yang, P. A bayesian hierarchical model for urban air quality prediction under uncertainty. *Atmos. Environ.* **2008**, *42*, 8464–8469. [\[CrossRef\]](#)
27. Polat, E.; Gunay, S. The comparison of partial least squares regression, principal component regression and ridge regression with multiple linear regression for predicting PM₁₀ concentration level based on meteorological parameters. *J. Data Sci.* **2021**, *13*, 663–692. [\[CrossRef\]](#)
28. Riccio, A.; Barone, G.; Chianese, E.; Giunta, G. A hierarchical bayesian approach to the spatio-temporal modeling of air quality data. *Atmos. Environ.* **2006**, *40*, 554–566. [\[CrossRef\]](#)
29. Sun, W.; Zhang, H.; Palazoglu, A.; Singh, A.; Zhang, W.; Liu, S. Prediction of 24-hour-average PM_{2.5} concentrations using a hidden Markov model with different emission distributions in Northern California. *Sci. Total Environ.* **2013**, *443*, 93–103. [\[CrossRef\]](#)
30. Ni, X.Y.; Huang, H.; Du, W.P. Relevance analysis and short-term prediction of PM 2.5 concentrations in Beijing based on multi-source Data. *Atmos. Environ.* **2017**, *150*, 146–161. [\[CrossRef\]](#)
31. Essamlali, I.; Nhaila, H.; El Khaili, M. Supervised Machine Learning Approaches for Predicting Key Pollutants and for the Sustainable Enhancement of Urban Air Quality: A Systematic Review. *Sustainability* **2024**, *16*, 976. [\[CrossRef\]](#)
32. Zareba, M.; Cogiel, S.; Danek, T.; Weglinska, E. Machine Learning Techniques for Spatio-Temporal Air Pollution Prediction to Drive Sustainable Urban Development in the Era of Energy and Data Transformation. *Energies* **2024**, *17*, 2738. [\[CrossRef\]](#)
33. Bellinger, C.; Mohamed Jabbar, M.S.; Zaiane, O.; Osornio-Vargas, A. A systematic review of data mining and machine learning for Air Pollution Epidemiology. *BMC Public Health* **2017**, *17*, 907. [\[CrossRef\]](#)
34. Krichen, M.; Mihoub, A.; Alzahrani, M.Y.; Adoni, W.Y.; Nahhal, T. Are formal methods applicable to machine learning and Artificial Intelligence? In Proceedings of the 2022 2nd International Conference of Smart Systems and Emerging Technologies (SMARTTECH), Riyadh, Saudi Arabia, 9–11 May 2022. [\[CrossRef\]](#)
35. Raman, R.; Gupta, N.; Jeppu, Y. Framework for formal verification of machine learning based complex system-of-Systems. *Insight* **2023**, *26*, 91–102. [\[CrossRef\]](#)
36. Zareba, M.; Weglinska, E.; Danek, T. Air pollution seasons in urban moderate climate areas through Big Data Analytics. *Sci. Rep.* **2024**, *14*, 3058. [\[CrossRef\]](#)
37. Mampitiya, L.; Rathnayake, N.; Hoshino, Y.; Rathnayake, U. Forecasting PM₁₀ levels in Sri Lanka: A comparative analysis of machine learning models PM₁₀. *J. Hazard. Mater. Adv.* **2024**, *13*, 100395. [\[CrossRef\]](#)

38. Elangasinghe, M.A.; Singhal, N.; Dirks, K.N.; Salmond, J.A. Development of an ANN-based air pollution forecasting system with explicit knowledge through sensitivity analysis. *Atmos. Pollut. Res.* **2014**, *5*, 696–708. [\[CrossRef\]](#)
39. Méndez, M.; Merayo, M.G.; Núñez, M. Machine learning algorithms to Forecast Air Quality: A survey. *Artif. Intell. Rev.* **2023**, *56*, 10031–10066. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 26–28 October 2014. Available online: <https://arxiv.org/pdf/1406.1078> (accessed on 1 May 2024).
41. Abdallah, M.S.; Samaan, G.H.; Wadie, A.R.; Makhmudov, F.; Cho, Y.-I. Light-Weight Deep Learning Techniques with Advanced Processing for Real-Time Hand Gesture Recognition. *Sensors* **2023**, *23*, 2. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Liu, Z.; Li, W.; Feng, J.; Zhang, J. Research on Satellite Network Traffic Prediction Based on Improved GRU Neural Network. *Sensors* **2022**, *22*, 8678. [\[CrossRef\]](#)
43. Thanthawy Sukanda, A.J.; Adytia, D. Wave forecast using bidirectional GRU and GRU method case study in Pangandaran, Indonesia. In Proceedings of the 2022 International Conference on Data Science and Its Applications (ICoDSA), Bandung, Indonesia, 6–7 July 2022. [\[CrossRef\]](#)
44. Moharm, K.; Eltahan, M.; Elsaadany, E. Wind speed forecast using LSTM and Bi-LSTM algorithms over Gabal El-Zayt Wind Farm. In Proceedings of the 2020 International Conference on Smart Grids and Energy Systems (SGES), Perth, Australia, 23–26 November 2020. [\[CrossRef\]](#)
45. Mampitiya, L.; Rathnayake, N.; Leon, L.P.; Mandala, V.; Azamathulla, H.M.; Shelton, S.; Hoshino, Y.; Rathnayake, U. Machine Learning Techniques to Predict the Air Quality Using Meteorological Data in Two Urban Areas in Sri Lanka. *Environments* **2023**, *10*, 141. [\[CrossRef\]](#)
46. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
47. Wu, L.; Noels, L. Recurrent neural networks (RNNs) with dimensionality reduction and break down in computational mechanics; application to multi-scale localization step. *Comput. Methods Appl. Mech. Eng.* **2022**, *390*, 114476. [\[CrossRef\]](#)
48. Wei, X.; Zhang, L.; Yang, H.; Zhang, L.; Yao, Y. Machine learning for pore-water pressure time-series prediction: Application of recurrent neural networks. *Geosci. Front.* **2021**, *12*, 453–467. [\[CrossRef\]](#)
49. Wang, J.; Li, X.; Li, J.; Sun, Q.; Wang, H. NGCU: A new RNN model for time-series data prediction. *Big Data Res.* **2022**, *27*, 100296. [\[CrossRef\]](#)
50. Rybarczyk, Y.; Zalakeviciute, R. Machine learning approaches for outdoor air quality modelling: A systematic review. *Appl. Sci.* **2018**, *8*, 2570. [\[CrossRef\]](#)
51. Zhu, W.; Wang, J.; Zhang, W.; Sun, D. Short-term effects of air pollution on lower respiratory diseases and forecasting by the group method of data handling. *Atmos. Environ.* **2012**, *51*, 29–38. [\[CrossRef\]](#)
52. Zhang, Y.; Bocquet, M.; Mallet, V.; Seigneur, C.; Baklanov, A. Real-time air quality forecasting, part I: History, techniques, and current status. *Atmos. Environ.* **2012**, *60*, 632–655. [\[CrossRef\]](#)
53. Guo, Q.; He, Z.; Wang, Z. Prediction of hourly PM_{2.5} and PM₁₀ concentrations in Chongqing City in China based on Artificial Neural Network. *Aerosol Air Qual. Res.* **2023**, *23*, 220448. [\[CrossRef\]](#)
54. Kurniawan, T.A.; Haider, A.; Khan, S.; Mohyuddin, A.; Lei, T.; Goh, H.H.; Othman, M.H.D.; Anouzia, A.; Aziz, F.; Mahmoud, M. Technological Solutions for air pollution to mitigate climate change: A strategy to facilitate glbal transition towards blue sky and net-zero emissions. *Chem. Pap.* **2024**. [\[CrossRef\]](#)
55. Mamić, L.; Gašparović, M.; Kaplan, G. Developing PM_{2.5} and PM₁₀ prediction models on a national and regional scale using open-source remote sensing data. *Environ. Monit. Assess.* **2023**, *195*, 644. [\[CrossRef\]](#)
56. Ayus, I.; Natarajan, N.; Gupta, D. Comparison of machine learning and deep learning techniques for the prediction of Air Pollution: A Case Study from China. *Asian J. Atmos. Environ.* **2023**, *17*, 4. [\[CrossRef\]](#)

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