



A machine learning approach to investigate the build-up of surface ozone in Mexico-City

M. Ahmad ^{a,*}, B. Rappenglück ^a, O.O. Osibanjo ^{a,c}, A. Retama ^b

^a University of Houston, Department of Earth and Atmospheric Sciences, Houston, TX, United States

^b Ciudad de México, Mexico

^c Now with FM Global Insurance, Boston, MA, USA

ARTICLE INFO

Keywords:

Air quality
Ozone predictors ranking
Boundary layer height
Deep neural network
Mexico City

ABSTRACT

Ground-level ozone is an important pollutant regarding air quality and climate. Mexico City frequently experiences severe ozone episodes due to a combination of strong ozone precursor emissions and its specific topographical environment which critically impacts meteorological conditions. High ozone levels during these episodes cause harmful effects to the public health and the environment. This necessitates ranking air quality and meteorological variables according to their contributions towards the build-up of ozone. In this study, three machine learning models are used to learn a prediction function with hourly data of eight predictors as input and hourly ground-level ozone mixing ratios as output. One-year hourly data of eight predictors collected in Mexico-City from March 2015 to February 2016 is employed to train and test the models. The best model, capturing ozone peak levels with 92% accuracy during 6–18 March 2016, is used to rank the predictors according to their importance in the build-up of ozone applying a shapley additive explanations approach based on the game theory shapley values. This 6–18 March 2016 period encompassed different meteorological and emission conditions and included a severe ozone smog episode from 12 to 17 March 2016. Such ranking of the air quality and meteorological variables is crucial for policy-making decisions regarding the prevention and mitigation of ozone detrimental effects during severe ozone episodes and provides insight into the functional dependency of ozone on its predictors. The proposed approach showcases Mexico City, but its principles can be applied for ozone episodes at any other location.

1. Introduction

In the atmosphere, ozone (O_3) exists both at ground level and in the stratosphere. Photochemical dissociation of oxygen molecules at wavelengths <290 nm into oxygen atoms and then recombining of oxygen atoms results in stratospheric ozone generation. Filtering Shortwave ultraviolet-B (UV-B) radiation and Shortwave ultraviolet-C (UV-C) radiation from the sun's ultraviolet (UV) radiation, the stratospheric ozone layer protects life on earth from harmful effects of UV radiation and is considered to be good ozone. At ground level, ozone is a harmful air pollutant with adverse effects on human and animal health, vegetation, and structures (Bell et al., 2006; Mills et al., 2007), as well as an important contributor to the radiative balance. Tropospheric ozone is formed in photochemical reactions between volatile organic compounds (VOCs) and nitrogen oxides (NO_x) (Jenkin and Clemishaw, 2000). Shortwave ultraviolet-A (UV-A) (315–400 nm) radiation reaches the

ground level without much attenuation in the upper atmosphere and plays a crucial role in ground-level ozone formation. Meteorological variables such as the wind direction (WD), wind speed (WS), relative humidity (RH), and temperature (TEMP) affect variations in surface ozone mixing ratios (Comrie and Yarnal, 1992; Davies et al., 1992; Rappenglück et al., 1993; Dueñas et al., 2002; Elminir, 2005; Solomon et al., 2000; Zhang et al., 1998). In addition, the planetary boundary layer height (PBLH) causes variations in ozone mixing ratios (Berman et al., 1999). In conclusion, the ambient ozone formation depends both on its precursors and meteorological variables. Furthermore, this dependence is highly complex and nonlinear.

Mexico City, the capital of Mexico, is a metropolitan area with 21.8 million inhabitants contained within a high-altitude basin (~2240 m above sea level). Due to its specific topographical environment in conjunction with meteorological and emission conditions Mexico City is frequently exposed to ozone episodes, which appear to have become

* Corresponding author. University of Houston, Department of Earth and Atmospheric Sciences, 4800 Calhoun Rd, Houston, TX, 77204-5007, United States.

E-mail address: mahmad18@uh.edu (M. Ahmad).

more severe again recently (Velasco and Retama, 2017). High O₃ levels during an ozone episode cause a great deal of concern for the public health and the environment. Thus, it necessitates simulating ozone levels, picking up peak O₃ values, and ranking the air quality and meteorological variables according to their contributions towards the build-up of ozone during an ozone episode. Such ranking of ozone predictors is of great significance for both policy-making decisions regarding the prevention and mitigation of ozone detrimental effects and understanding functional dependence of ozone on its predictors during an ozone episode. Generally, multiple linear regression (MLR), numerical, and machine learning (ML) models are used to predict ozone levels. Applying MLR and artificial neural network to predict concentration changes in ozone, Bekesiene et al. (2021) reported that artificial neural network model outperforms the MLR model. The MLR model is good at capturing linear relationship between a dependent variable and independent variables, but this model is unable to establish nonlinear dependences like the dependence of ambient ozone on its precursors and meteorological variables (Zielesny, 2011). MLR requires constructing model function with unknown parameters and then calculating the unknown parameters using the observed data. The model function guess may depart from the actual functional form and result in reduced accuracy. Furthermore, MLR is unable to rank air quality and meteorological variables with respect to their contributions to the build-up of ozone during an ozone episode. Numerical models such as the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) and the Community Multiscale Air Quality (CMAQ) model along with WRF are used to simulate ozone levels. But the following issues are inherent in the numerical models: (1) numerical models are computationally expensive and require extensive computing power and time (Kelly et al., 2021), (2) numerical models simulate ozone levels with reasonable accuracies at larger scales (Hu et al., 2016), but at the urban scale their performances are not satisfactory (Gradišar et al., 2016; Wang et al., 2020). ML algorithms have been increasingly employed to explore the dependence of surface ozone on environmental and meteorological variables, and to predict ozone concentrations (Table 1) due to their greater efficiency and less utilization of computing power and time. Gagliardi and Andenna, 2020 investigated the dependence of surface ozone on ozone precursors and meteorological variables in Agri Valley located in the south-west part of the Basilicata Region (Southern Italy) applying ML. The boosted regression trees (BRT) model was used to develop a functional relation between surface ozone and twelve predictors, namely methane (CH₄), non-methane volatile organic compounds (NMVOCs), carbon monoxide (CO), nitrogen monoxide (NO), nitrogen dioxide (NO₂), NO_x, relative humidity, wind speed, wind direction, temperature, atmospheric pressure, and solar radiation. The model predicts hourly ozone concentrations using hourly data of twelve predictors. The trained BRT model was used to calculate hourly ozone concentrations for the entire 2018 year. It was reported that the BRT model predicts ozone concentrations with R² equals 0.81; however, the BRT model does not estimate extreme ozone values properly (Gagliardi and Andenna, 2020). Eslami et al. (2020) applied a deep convolutional neural network (CNN) to predict hourly ozone concentrations 24 h in advance using current day hourly data of ozone and seven other predictors. The model predicts next day hourly ozone concentrations utilizing hourly data of wind speed, wind direction, surface temperature, dew point temperature, surface pressure, precipitation, NO_x, and ozone of the current day. The trained CNN model was used to calculate hourly, daily mean, and daily maximum ozone concentrations for all 25 continuous ambient monitoring stations across Seoul, South Korea, for the entire year 2017. The model predicts hourly ozone concentrations with sufficient accuracy (IOA = 0.84–0.89, r = 0.74–0.81), but the model underpredicts high ozone peaks during the hot summer days of 2017. Another study by Sayeed et al. (2020) deployed CNN to predict next day hourly ozone concentrations in Texas using the same set of predictors as used by Eslami et al., 2020 with solar radiation as one extra predictor. The model predicts next day hourly ozone concentrations

Table 1
Summary of the studies.

Author(s)	Study	Model(s) Applied	Study Area
Ibarra et al. (2000)	Real-time prediction of hourly O ₃ , NO ₂ and NO levels	MLR	Bilbao
Moustris et al. (2012)	O ₃ Forecast	MLR and ANNs	Athens, Greece
Jaioun et al. (2014)	Estimation of O ₃ concentrations	MLR	Chonburi, Thailand.
Gradišar et al., 2016	Local O ₃ forecast improvement	Multilayer perceptron neural network, WRF, and QualeAria air-quality models	Representative locations in Slovenia
Sousa et al., 2016	Next day hourly O ₃ concentrations prediction	MLR and ANNs based on principal components	Oporto, Portugal
Eslami et al., 2020	O ₃ concentrations prediction 24 h in advance	Deep CNN	Seoul, South Korea
Sayeed et al., 2020	O ₃ concentrations next day prediction	Deep CNN	Texas, USA
Verma et al. (2019)	24 h Advance Forecast of Surface O ₃	MLR, principal component regression, ANN	Site of Indo-Gangetic Plain
Allu et al. (2020)	Seasonal ground level O ₃ prediction	MLR	Hyderabad, India
Arsić et al. (2020)	Prediction of O ₃ concentration in ambient air	MLR and ANNs	City of Zrenjanin, Serbia
Gagliardi and Andenna (2020)	Dependence of surface O ₃ on its precursors and meteorological variables	BRT	Basilicata Region (Southern Italy)
Hafeez et al., 2020	Investigation of interacting effects of experimental parameters on O ₃ generation	ANN	-
Liang et al. (2020)	Prediction of air quality index using O ₃ and other predictors	RF, SVM, ANN	Zhongli, Changhua, Fengshan (Taiwan)
Wang et al. (2020)	Regional prediction of ground-level O ₃	Hybrid sequence-to-sequence deep learning approach	Beijing, China
Abirami and Chitra, 2021	Air quality forecast (O ₃ , PM ₁₀ , PM _{2.5} , SO ₂ , NO ₂ , CO, NH ₃ , pb)	Hierarchical deep learning model	Delhi, India
Bekesiene et al. (2021)	O ₃ concentration changes prediction	ANN and MLR	Eastern part of Lithuania
Juarez and Petersen, 2021	O ₃ concentrations prediction 1–24 h in the future	KNN, SVM, RF, AdaBoost, XGBoost, LSTM, MLR, DT	Delhi, India
Makarova et al. (2021)	Hourly tropospheric O ₃ concentrations prediction	ANNs	Southern parts of Russia
Yafouz et al. (2021)	Hourly O ₃ prediction	MLP, SVR, XGBoost, LSTM, CNN	Three sites in Malaysia
Balamurugan et al. (2022)	Modeling of urban surface O ₃ variability	XGBoost	Munich, southern German
Zhan et al. (2022)	O ₃ formation sensitivity	RF model	Beijing, China

Note: For an expanded list of the abbreviations see Table S1 in Appendix A Supplementary material.

utilizing current day hourly data of wind speed, wind direction, surface temperature, dew point temperature, solar radiation, surface pressure, precipitation, NO_x, and ozone. The trained model was applied to calculate hourly and daily maximum ozone concentrations for all 21 continuous ambient monitoring stations across Texas for the year 2017,

and average IOA (index of agreement) of 0.78 was reported for daily maximum ozone concentrations. Even though the model predicts daily maximum ozone concentrations average of IOA of 0.78, it underpredicts peak daily maximum ozone concentrations. Makarova et al. (2021) applied artificial neural network (ANN) with various configurations to predict hourly tropospheric ozone concentrations for the southern parts of the Russian Federation using hourly data of temperature, humidity, pressure, wind speed, and wind direction as predictors. The trained ANN model was used to predict hourly ozone concentrations for the entire 2017 year. It was reported that the ANN model with feed forward backpropagation (FFB) structure predicts ozone concentrations with IOA = 0.87; however, the model does not estimate properly high ozone concentrations occurring during June 2017. Sousa et al. (2007) in their paper described a new methodology based on feedforward artificial neural networks with principal components as inputs to predict next day hourly ozone concentrations. Comparing developed model with multiple linear regression, it was found that the use of principal components as inputs improved the model prediction by eliminating data collinearity; however, the model underestimated the highest ozone concentrations.

Even though machine learning models predict ambient ozone concentrations with higher accuracy for a longer period of time, the prediction of peak surface ozone values especially during an ozone episode, which are of most relevant interest for public health, continues to be an intricacy due to the non-linear dependence of tropospheric ozone formation on its precursors and the inherent challenge in capturing small-scale meteorological dynamics. Our proposed research fills in this gap, prediction of peak surface ozone values, applying suitable machine learning models with appropriate suite of predictors. In this study, we present and intercompare for the first-time machine learning (ML) models to simulate ozone during a severe ozone episode in Mexico City in an effort to assess the performance of these models to capture peak ozone events. Also, we incorporate continuous measurements of the PBLH as one of the predictors for ozone prediction. To our knowledge the variation of the PBLH has not yet been taken into account in ML models before. The models were trained to learn a functional relation between ozone and eight predictors. The model, with the best performance in picking up both low and peak ozone values, was used to analyze the importance of selected meteorological variables and surface chemistry in the build-up of surface ozone in Mexico-City during 6–18 March 2016. This period covered different meteorological and emission conditions and included a severe ozone smog episode during 12–17 March 2016. Strong advection and a deep upper tropospheric trough effected the first few days prior to the smog episode. On these days pollutants were flushed out, and maximum hourly O₃ mixing ratios did not exceed ~50 ppbv. After the departure of the storm, strong near surface temperature inversions were observed and overall stagnant, sunny conditions favored increased daytime O₃ mixing ratios up to ~200 ppbv, which resulted in one of the recent most severe smog episodes in Mexico City (Velasco and Retama, 2017; Osibanjo et al., 2021). Three ML models including random forest (RF), gradient tree boosting (GTB), and deep neural network (DNN) were used to learn a functional relation between eight predictors (six meteorological variables including UV-A, WD, WS, RH, TEMP, PBLH and two O₃ precursors nitric oxide (NO), nitrogen dioxide (NO₂)) and O₃ itself. The best model, the DNN model, was used to analyze the individual importance of the predictors in the build-up of ozone in Mexico City from 6 March 2016 to 18 March 2016. The Deep SHAP (Shapley Additive exPlanations) explainer method was used to interpret the O₃ predictions of the DNN model based on three scenarios: (1) considering all samples, (2) samples with UV-A values greater than 0.5 W/m², and (3) samples with UV-A values greater than 20 W/m². The first scenario was used to evaluate the predictions for the entire O₃ dataset, whereas the two latter approaches targeted at daytime O₃ data, the major time period for photochemical O₃ formation. The novelties of our proposed study are summed up in the following lines: (1) machine learning approach to simulate ozone levels

and rank ozone predictors during an ozone episode in Mexico City; the proposed approach can be applied for ozone episodes at any other location (2) prediction of peak surface ozone values applying suitable machine learning models with appropriate suite of predictors (3) the incorporation of the continuous planetary boundary layer height data retrieved by a microwave radiometer into the suite of machine learning models' predictors.

2. Data description

The hourly data of eight predictors include continuous hourly measurements of meteorological parameters, i.e. ambient surface temperature (TEMP [°C]), relative humidity (RH [%]), surface wind speed (WS [m/s]) and direction (WD [degree]), planetary boundary layer height (PBLH [m asl]), shortwave ultraviolet A radiation (UV-A [W/m²]) as well as ambient surface chemistry data including nitric oxide (NO [ppb]) and nitrogen dioxide (NO₂ [ppb]) mixing ratios (Table 2). All surface air quality and meteorological data was obtained at the RAMA (Red Automática de Monitoreo Atmosférico) air quality site Hospital General de México (HGM), apart from UV-A data which was obtained at the RAMA air quality site Merced (MER). Both sites are 3.7 km away from each other and are located in the downtown area of Mexico City. Continuous PBLH measurements were taken at a site about 5 km north of the center of Mexico City. Detailed information about these measurements can be found in Osibanjo et al. (2021, 2022). For the purpose of this paper, we only consider the nocturnal stable boundary layer height (SBLH) and the daytime convective boundary layer heights (CBLH) to indicate the evolution of the PBLH.

The models were trained and tested on one-year data from March 2015 to February 2016 using 5541 samples (66.35% of total); missing samples (2810 samples, 33.65% of total) were discarded. The trained models were used to estimate hourly ozone mixing ratios in Mexico City for the period 6–18 March 2016 (containing total 312 samples; 36 missing values were imputed applying k-nearest neighbors (KNN) imputer). The testing dataset gave us the opportunity to challenge the models with a complex scenario characterized by ample variations on the variables used as predictors.

3. Methodology

The research flow chart of the research conducted in this paper is depicted in Fig. 1. This study aims at utilizing machine learning (ML) approach to investigate the ozone build-up during an ozone episode in Mexico City. Random forest (RF), gradient tree boosting (GTB), and deep neural network (DNN) models are employed to predict hourly ground-level ozone mixing ratios. The predictors comprised of two O₃ precursors (nitric oxide (NO), nitrogen dioxide (NO₂)) and six meteorological variables (Shortwave ultraviolet-A radiation (UV-A), wind direction (WD), wind speed (WS), relative humidity (RH), ambient surface temperature (TEMP), planetary boundary layer height (PBLH). The

Table 2
Data description.

Variable	Abbreviation	units	Time resolution
Temperature	TEMP	Degree Celsius	Hourly
Relative humidity	RH	% Relative humidity	Hourly
surface wind speed	WS	m/s	Hourly
Wind direction	WD	Degree	Hourly
Planetary boundary layer height	PBLH	m	Hourly
Shortwave ultraviolet A radiation	UV-A	W/m ²	Hourly
Nitric oxide	NO	ppb	Hourly
Nitrogen dioxide	NO ₂	ppb	Hourly
Ozone	O ₃	ppb	Hourly

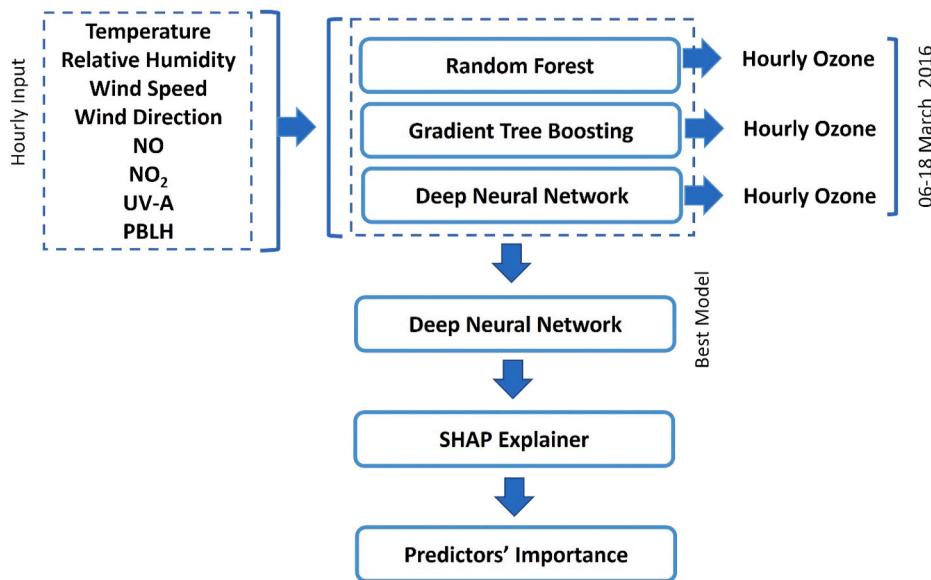


Fig. 1. Research flow chart of proposed research.

models were trained and tested on one-year hourly data collected in Mexico-City from March 2015 to February 2016. A comparison among the three trained ML models was performed based on two statistical indicators the coefficient of determination (R^2) and the index of agreement (IOA) to choose the best model in predicting hourly surface O₃ concentrations in Mexico-City from 6 to 18 March 2016. The DNN model, being the best model, was then used to assess the importance of predictors which were supposed to play an important role in the build-up of O₃. The Shapely Additive exPlanations (SHAP) approach, which is based on the game theory Shapley values, explains machine learning models results by calculating the contribution of each predictor towards models' outputs. The Deep SHAP explainer method was used to interpret the predictions of DNN model for the period 6–18 March 2016 and rank the predictors according to their importance with regard to their contribution in the build-up of ozone during this period.

3.1. Machine learning models

Three ML models including RF, GTB, and DNN were used to learn a prediction function with hourly data of eight predictors as input and hourly O₃ mixing ratios as output:

$$O_3 = f(TEMP, RH, WS, WD, NO, NO_2, UVA, PBLH) \quad (1)$$

In the ML model, external parameters that control the learning process are defined as hyperparameters. The choice of hyperparameters play a critical role, as they directly influence the learning process and have an important effect on the performance of the model. Hyperparameters are set manually prior to the training process, and they are not learned during the training process while the parameters of a model (like weights and biases in case of a neural network) are learned during the training process. Applying hyperparameters tuning, the optimal hyperparameters (hyperparameters that result in approximately 90% accuracy score on the test data) for these models were selected. In the following sections the models are arranged in ascending degree of complexity.

3.1.1. Random forest (RF)

A decision tree is a non-parametric supervised ML algorithm that divides the predictor space into a number of non-overlapping regions based on splitting rules and predicts output for a given point by locating the region to which it fits in and using mean output of all the training points belonging to that specific region. Decision trees require little data

preparation and are simple to understand and interpret, can be used for both regression and classification problems, but they generally overfit and have high variance.

Random Forest is a ML algorithm based on an ensemble (group) of uncorrelated decision trees (Breiman, 2001). The decision trees in the ensemble are built from the bootstrapping of data with replacement, and each split in a tree is carried out based upon a predictor chosen randomly from a subset of the predictors. Typical default values are \sqrt{m} (for classification) and $m/3$ (for regression) (Probst et al., 2019), where m is the total number of available predictors. At each split, such a condition is imposed to involve more randomness which leads to more uncorrelated trees and hence more accurate prediction at the end. The final prediction is the average of all the predictions from the decision trees, and it is more accurate than the individual predictions because the averaging process reduces the variance (error) inherited in the individual predictions (James et al., 2013).

A schematic diagram of a RF algorithm is shown in Fig. 2, where dataset D consists of input values and corresponding target values, and D_1, D_2, \dots, D_m are copies of the dataset D generated from bootstrapping

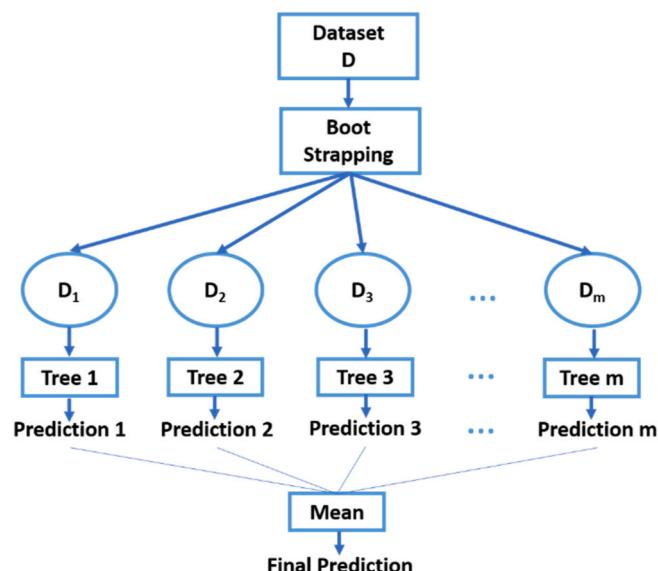


Fig. 2. Schematic diagram of a Random Forest algorithm.

with replacement. The number of trees and the maximum number of features for each split in a tree are the important tuning hyperparameters for the RF model. The RF model was implemented with optimal hyperparameters, 110 regression trees and maximum number of features equals 3 for each split in a tree. The RF model was set up using the Python Scikit-learn library (Pedregosa et al., 2011). The RF model was trained and tested on twelve months of hourly data (from March 2015 to February 2016); 80% of the data was used for training and 20% for testing. The trained RF was used to estimate hourly ozone mixing ratios in Mexico City for the period 6–18 March 2016 which contained the severe ozone smog episode during 12–17 March described in detail by Osibanjo et al. (2021).

3.1.2. Gradient tree boosting (GTB)

Gradient Tree Boosting is also an ensemble algorithm, but the ensemble is grown sequentially adding one tree, called a weak learner, at a time. Gradient Tree Boosting is not based on the bootstrapping of the data, and each tree in the ensemble is fitted to minimize the residuals (errors) calculated from its predecessor. In the end, a strong learner is generated by adding up all the weak learners (Friedman, 2001, 2002).

A schematic diagram of a GTB algorithm is shown in Fig. 3. In Fig. 3 dataset D consists of input values and corresponding target values, dataset R_1 contains input values and residuals calculated from Tree 1, R_2 is the dataset consisting of input values and residuals calculated from Tree 2, and R_{i-1} is the dataset with input values and residuals calculated from Tree $i-1$. The number of trees and the learning rate are the important tuning hyperparameters for the GTB model. The learning rate r controls the contribution of each tree by shrinking its contribution with a factor ($0 < r \leq 1$) which prevents overfitting, with a lower learning rate requiring a higher number of trees (Zhang and Haghani, 2015). The GTB model was implemented with optimal hyperparameters, 320 regression trees and 0.2 as learning rate. The GTB model was set up using the Python Scikit-learn library (Pedregosa et al., 2011). The GTB model was trained and tested on twelve months of hourly data (from March 2015 to February 2016); 80% of the data was used for training and 20% for testing. Finally, the trained GTB model was used to estimate hourly ozone mixing ratios in Mexico City for the same time period as for the RF algorithm.

3.1.3. Deep neural network (DNN)

An artificial neural network, commonly referred to as neural network, consists of simple processing units, called neurons, operating in parallel. Like a human brain, it can learn from experience (training) and generalize. During the training process the weights and biases of the neural network are adjusted so that specific inputs lead to the required output. The weights are the parameters representing strengths of links between input values and neurons, and between neurons and neurons; a bias is a value added to the weighted sum of input values to a neuron and its value may be zero, positive or negative, and it may be considered as a weight with corresponding constant input value 1. Generalization is the neural network's ability to generate satisfactory outcomes for inputs not seen during the training process (Haykin, 2009).

The block diagram of a neuron is shown in Fig. 4. It consists of input

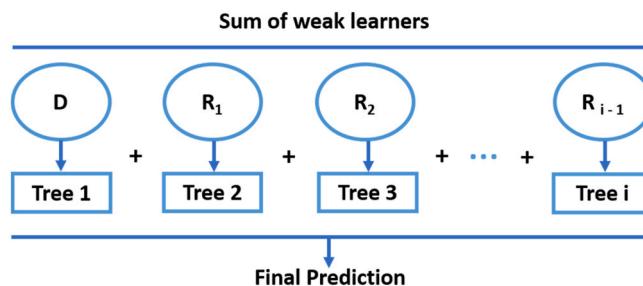


Fig. 3. Schematic diagram of a Gradient Tree Boosting algorithm.

values, weights, bias, summing junction, activation function, and an output value. The summing junction adds bias b_k to the weighted sum of input values to generate v_k , and the activation function f operates on v_k to produce the output y_k .

Mathematically, the output of the neuron depicted in Fig. 4 is described by the equations:

$$v_k = \sum_{i=1}^m w_i x_i + b_k \quad (2)$$

$$y_k = f(v_k) = f\left(\sum_{i=1}^m w_i x_i + b_k\right) \quad (3)$$

Where x_1, x_2, \dots, x_m are the input values; w_1, w_2, \dots, w_m are the weights; b_k is the bias; f is the activation function; and y_k is the output of the neuron.

An activation function is used to incorporate non-linearity into a neural network which allows the neural network to learn more complex functions. There are many activation functions available and the most commonly used are sigmoid, Rectified Linear Unit (ReLU), and hyperbolic tangent (Fig. 5). Generally, a sigmoid activation is used for a classification problem, and a Rectified Linear Unit is used for a regression problem.

In Fig. 6 an architectural graph of a neuron is shown. This graph is equivalent to the block diagram shown in Fig. 4, but it is more compact because the summing junction and activation function are not showing up explicitly and merged into the circular disk.

A multilayer feedforward neural network consists of layers of neurons and can be used for a classification problem, where the task of the network is to find class labels (discrete attributes used to distinguish one data item from another) corresponding to different inputs, and a regression problem, where the task of the network is to find required continuous values corresponding to different inputs. Multilayer feed-forward neural networks are universal approximators (Hornik et al., 1989).

A densely connected multilayer feedforward neural network for a regression problem, where every neuron in each layer of the network is linked to all the nodes in the immediate backward layer, is shown in Fig. 7. In case of a regression problem, the output layer of the network has no activation function so that the output of the network is not constrained to some range of values, and any value can come out as output. During training, the output of the network is compared to the required output, and weights of the network are adjusted applying backpropagation algorithm so that the loss function is minimum. The loss function is a function of the difference between the outputs of the network and the desired outputs and depends on weights and represents an error hypersurface graphically. There are many loss functions available, Mean Squared Error (MSE), Mean Absolute Error (MAE) etc., and we can choose a suitable loss function according to our problem. MSE is the average squared difference between the outputs of the network and the desired outputs, and MAE is the average absolute difference between the outputs of the network and the desired outputs:

$$MSE = \frac{\sum (y_p - y)^2}{n} \quad (4)$$

$$MAE = \frac{\sum |y_p - y|}{n} \quad (5)$$

where y_p is the predicted value, and y is the desired value.

To calculate weights corresponding to the minimum value of the loss function is the task of an optimization algorithm. Many optimization algorithms, like Gradient Descent, Stochastic Gradient Descent, Adam etc., are available and can be used according to the problem at hand. The number of hidden layers and the number of neurons in each hidden layer are the important tuning hyperparameters for the DNN model. The DNN

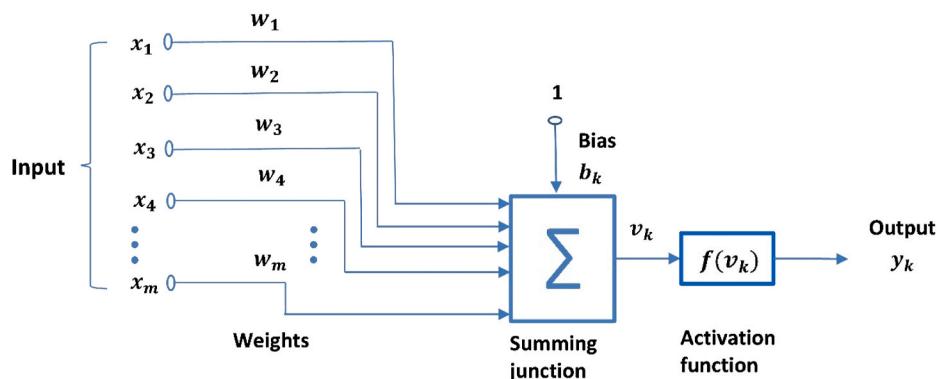


Fig. 4. Block diagram of a neuron (adapted from Haykin, 2009).

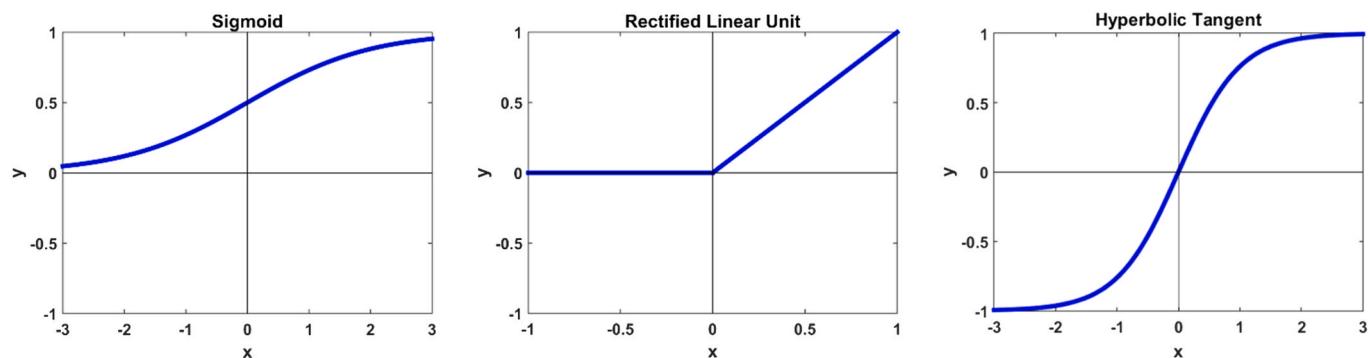


Fig. 5. Sigmoid, Rectified Linear Unit, and hyperbolic tangent activation functions.

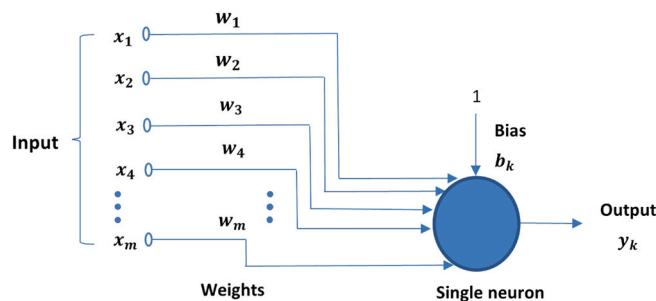


Fig. 6. Architectural graph of a neuron (adapted from Haykin, 2009).

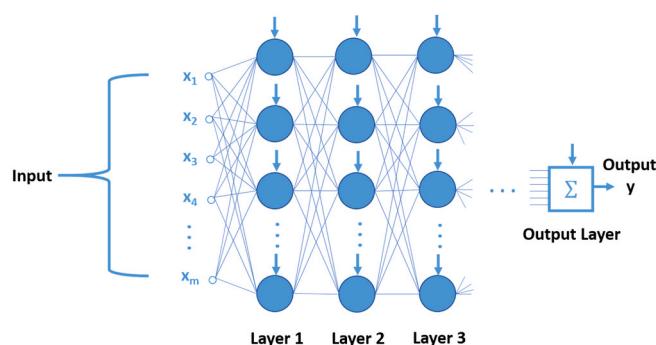


Fig. 7. Densely connected multilayer feedforward neural network for a continuous regression problem, vertical arrows represent biases (adapted from Haykin, 2009).

model was implemented with ReLU activation function, MSE loss function, Adam optimization algorithm, and optimal hyperparameters six fully connected hidden layers with 180 neurons in the first hidden layer, 128 neurons in the second hidden layer, 130 neuron in the third hidden layer, 130 neurons in the fourth hidden layer, 130 neurons in the fifth hidden layer, and 130 neurons in the sixth layer applying Python Keras deep learning library (Chollet, 2017). The criterion for the selection of the activation function, the loss function, and the optimization algorithm was also based on approximately 90% accuracy score on the test data. The DNN model was trained and tested on twelve months of hourly data (from March 2015 to February 2016); 80% of the data was used for training and 20% for testing. Finally, the trained DNN model was used to estimate hourly ozone mixing ratios in Mexico City for the same time period as for the previously mentioned RF and GTB algorithms.

3.2. Shapley values

The Shapley Additive exPlanations (SHAP) approach allows us to explain the outputs of machine learning models. It is based on the game theory Shapely values concept (Kuhn and Tucker, 1953), and it is used to estimate the reward that all the players receive corresponding to a certain outcome in a collaborative game. We can rank predictors locally (with respect to single output) or globally (with respect to more than one outputs) according to their importance in any machine learning model calculating their Shapely values. The SHAP method is applicable to both ensemble models (like RF and GTB models) and black-box models (like DNN and convolutional neural network models) (Lundberg and Lee, 2017). Python provides a SHAP library to calculate Shapely values for the predictors and visualize the results. The predictions of the DNN model for the period 6–18 March 2016 (containing 312 sample points) were interpreted applying Deep SHAP explainer, and the predictors were ranked according to their importance regarding their contribution

in the build-up of ozone during this period.

4. Results and discussion

We present two types of results. First, we describe the performance of RF, GTB, and DNN model in predicting hourly O₃ mixing ratios during 6–18 March 2016 period based on three scenarios: (i) considering all samples to evaluate the predictions for the entire O₃ data set, and two additional scenarios focusing on daytime O₃ data, i.e. (ii) samples with UV-A values greater than 0.5 W/m², and (iii) samples with UV-A values greater than 20 W/m². The latter two approaches are done in an effort to describe the performance of the models with respect to daytime peak O₃ values, which are of specific concern for defining O₃ exceedance days and public health alerts. The performance of the models is measured through two statistical indicators: the coefficient of determination (R^2) and the index of agreement (IOA). The models with R^2 and IOA close to 1 are highly accurate. The scatter plots between the observed and predicted hourly O₃ mixing ratios were also drawn to back up the analysis of the performance of the models during 6–18 March 2016 period. Second, we select the best model for each scenario and present the importance ranking of the input variables in predicting hourly O₃ mixing ratios during 6–18 March 2016 period applying the Deep SHAP explainer method which is based on SHAP. Shapely values for the predictors during 6–18 March 2016 period are calculated applying the Python SHAP library, and SHAP values plots (based on shapely values for the respective predictor) and mean absolute bar plots (based on mean of the absolute shapely values for the respective predictor) are drawn to visualize the importance of the predictors.

4.1. First scenario

The observed and model-estimated hourly ozone concentrations during March 2016 period based on first scenario (considering all samples) are shown in Fig. 8. All three models are picking up the low ozone mixing ratios nicely, but only the GTB and DNN models are picking up high ozone mixing ratios with the DNN model performing better than the GTB model.

The scatterplots between observed and model-estimated hourly ozone mixing ratios using eight predictors for March 2016 period along with performance indicators R^2 and IOA are shown in Fig. 9. Fig. 9a indicates the performance of the RF model; the model is only picking up

ozone mixing ratios up to 120 ppb, but starting with 80 ppb most values are being underestimated by the model. The scatterplot of observed versus estimated ozone mixing ratios obtained by the GTB model is depicted in Fig. 9b; the model is not only estimating low mixing ratios correctly, but also picking up better high mixing ratios up to 160 ppbv, albeit with some underprediction for O₃ values > 80 ppb similarly to the RF model. Also, the values of R^2 and IOA indicate that the performance of the GTB model is better than the RF model. Fig. 9c provides a graphical visualization of the DNN model performance; the model is also picking up peak ozone mixing ratios which are greater than 160 ppb. The overall data scatter, though, increase for O₃ > 100 ppb. It is worth noting that the DNN model outperforms both RF and GTB models.

Fig. 10 represents SHAP summary and standard bar plots with regard to the importance of each predictor during 6–18 March period. The predictors' names are arranged vertically in order of importance from top to bottom. In the SHAP summary plot the Shapely values are located horizontally with positive values on the right side of zero and negative values on the left side of zero. Each point represents a sample taken from the dataset, and the gradient color indicates the value of a predictor corresponding to a single sample.

In the SHAP summary plot (Fig. 10, left), three hundred and twelve samples for each predictor are showing up with their Shapely values. The low values of the predictor UV-A have negative Shapely values, and its high values correspond to high positive Shapely values. It means low values of UV-A push down O₃ mixing ratios prediction while high values of UV-A increase ozone mixing ratios prediction. The low values of the predictor NO have high positive Shapely values and increase O₃ mixing ratios prediction, and its high values with high negative Shapely values push down the ozone mixing ratios prediction. The low values of NO₂ have negative effect, while high values of NO₂ have a positive effect on the O₃ mixing ratios prediction. The high values of RH cause decrease in the O₃ mixing ratios prediction, and its low values enhance O₃ mixing ratios prediction. The predictor TEMP pushes down ozone concentrations prediction for its high values and pushes up for its low values. The sixth predictor WS decreases ozone concentrations prediction for its low values and increases for its high values. The high PBLH pushes down ozone concentrations while low PBLH pushes up. The behavior of the predictor WD is somewhat unclear because its high and low values are lumped together.

The mean absolute values bar plot for the eight predictors, shown in Fig. 10 (on the right), shows the global importance of the predictors.

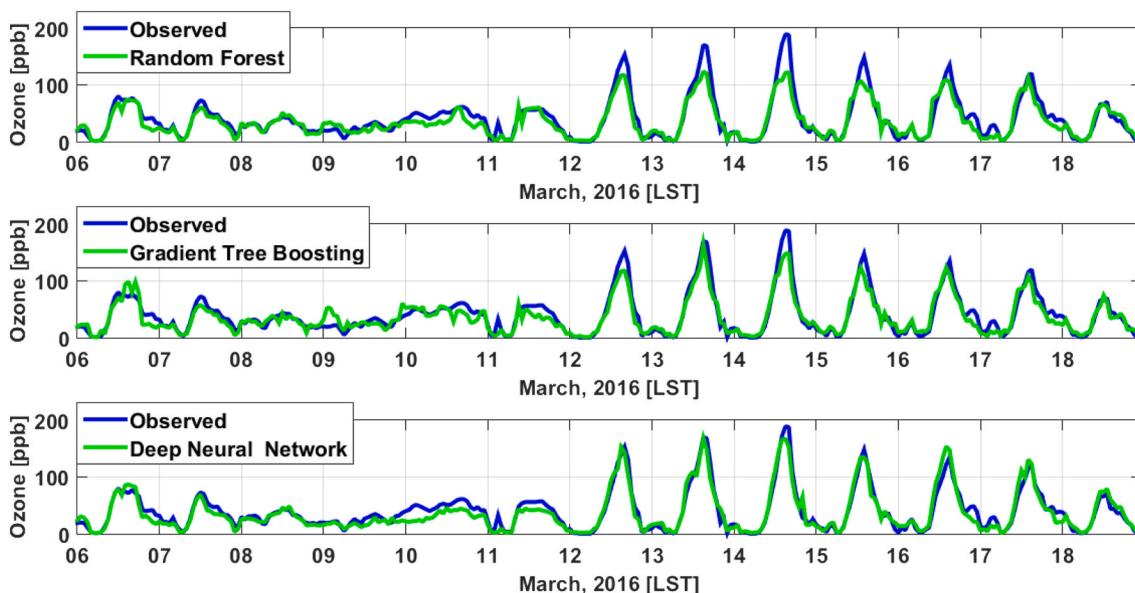


Fig. 8. Observed and model-estimated hourly ozone mixing ratios in Mexico-City during 06–18 March 2016 period based on first scenario (considering all samples). The estimations in the top, middle, and bottom figures are based on the RF, GTB, and DNN model respectively.

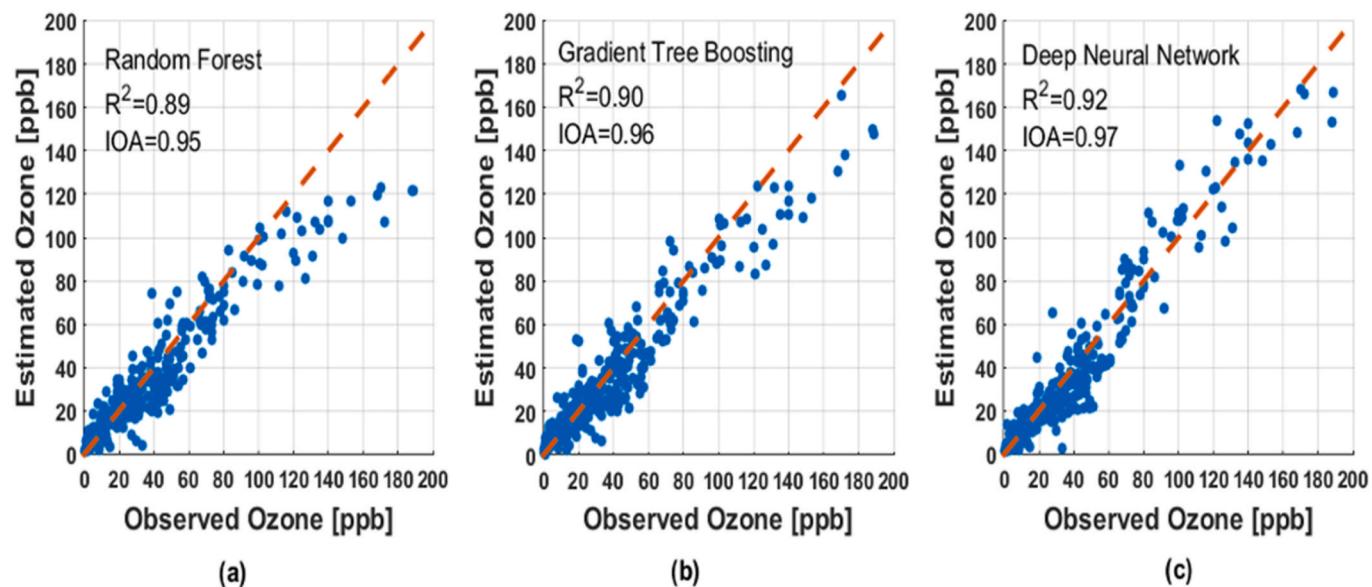


Fig. 9. Scatterplots depicting the performance of the RF, GTB, and DNN models in estimating hourly ozone mixing ratios using hourly data of eight predictors for the March 2016 period in Mexico-City period based on first scenario (considering all samples).

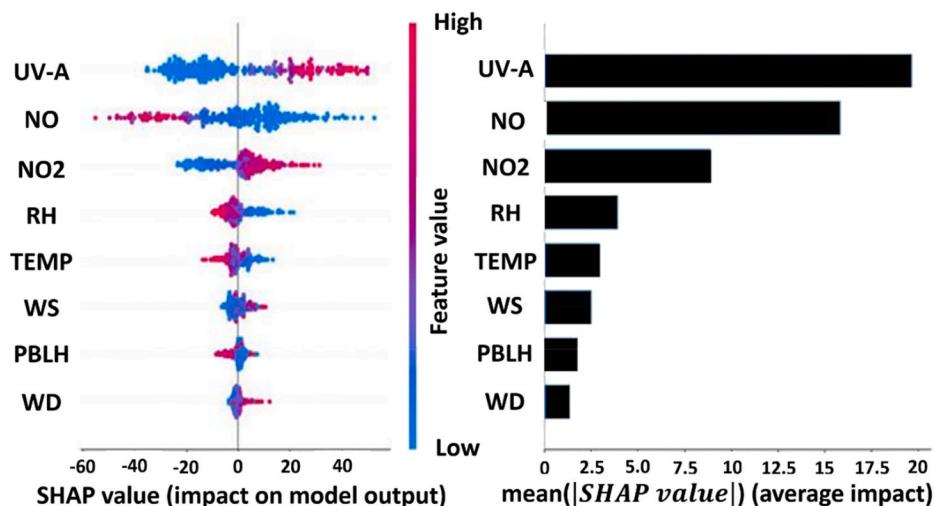


Fig. 10. SHAP plots for DNN based for the first scenario considering all samples to evaluate the predictions for the entire O₃ data set (left: SHAP summary plot; right: mean absolute values bar plot).

Each bar in the mean absolute values bar plot represents mean of the absolute shapely values for the respective predictor and provides an average impact of the predictor on the DNN model output. The top three predictors are UV-A, NO, and NO₂ with UV-A having highest average impact, NO having second highest average impact, and NO₂ having third highest impact on the build-up of O₃ during 6–18 March 2016 period. This ranking is in line with our scientific knowledge. While NO itself is an important primary pollutant emitted in combustion processes (e.g. fossil fuel combustion in vehicles, power plants) and the main precursor of NO₂, the photolysis of NO₂ through UV-A radiation will generate ground-level O₃ (Seinfeld and Pandis, 2016).

In an effort to confirm the most important predictors for high ozone mixing ratio we fitted the three models to different combinations of the predictors after removal of selected predictors. An example is shown in Fig. 11 with observed and model-estimated hourly ozone mixing ratios during March 2016 period, where UV-A was excluded from the set of eight predictors leaving TEMP, RH, WD, WS, NO, NO₂, and PBLH as the remaining seven predictors (Fig. 12 depicts corresponding scatter plots).

Fig. 11 reveals that in the absence of UV-A, all three models including GTB and DNN are not picking up high ozone mixing ratios (≥ 160 ppbv) and exclusion of UV-A has retarded their ability to predict high-peak ozone mixing ratios significantly. Similar results were obtained by excluding NO, then additionally NO₂, and then other combinations of NO, NO₂, UV-A (like (NO & NO₂, NO & UV-A, UV-A & NO₂, and UV-A, NO & NO₂) from the predictors' list and fitting the three models to the remaining predictors. All these results confirm that NO, NO₂, and UV-A together play very important role in the prediction of high ozone mixing ratios, and if any one of NO, NO₂, and UV-A is excluded from the eight predictors, the models will not be able to pick up high ozone mixing ratios and their performance decreases significantly. As a result, these predictors must be included in the list of predictors to obtain better prediction of ozone mixing ratios from the models.

Further analysis of the predictors revealed that in the presence of UV-A removal of TEMP or PBLH or both, TEMP and PBLH, from the predictors' list does not affect the models' predictions significantly, and in the presence of UV-A, both TEMP and PBLH are redundant and can be

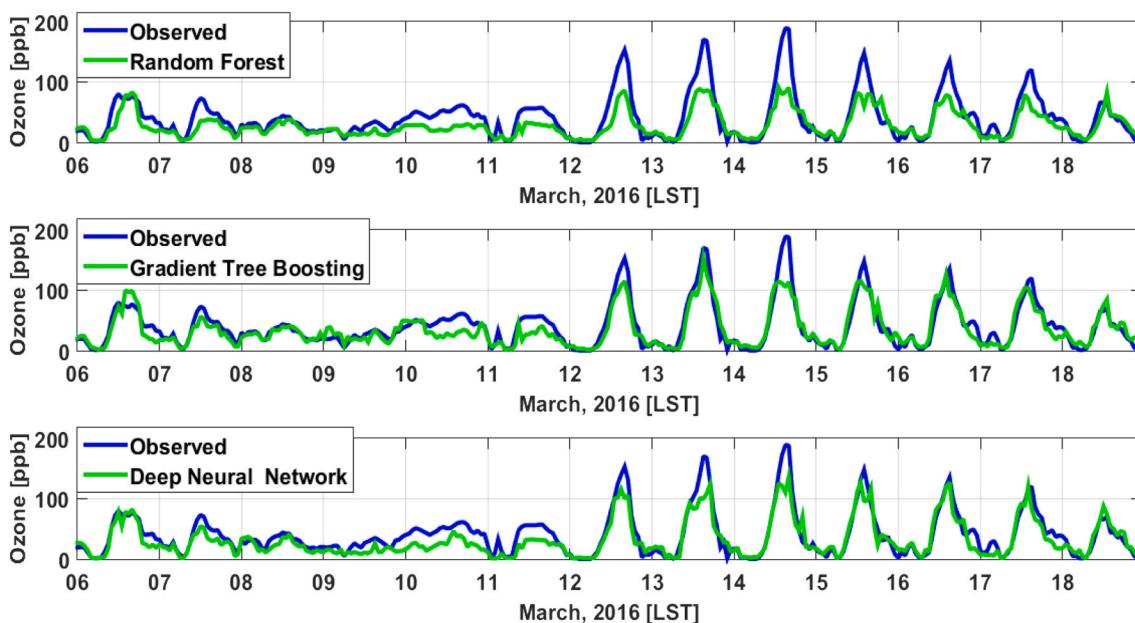


Fig. 11. Observed and model-estimated hourly ozone mixing ratios using TEMP, RH, WD, WS, NO, NO₂, and PBLH (UV-A excluded) during 06–18 March 2016.

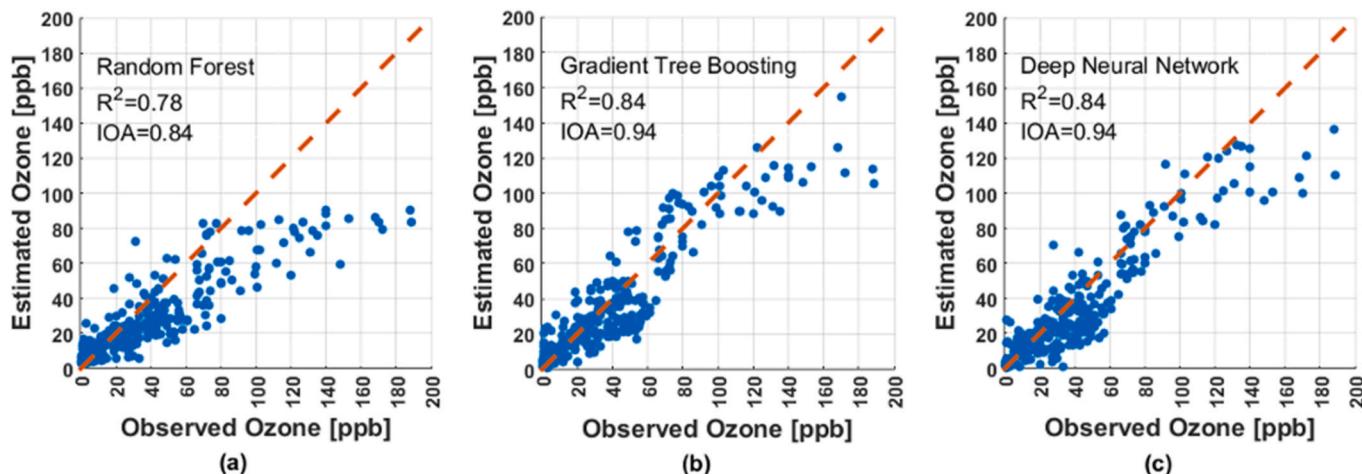


Fig. 12. Scatterplots depicting the performance of the models in estimating hourly ozone mixing ratios using TEMP, RH, WD, WS, NO, NO₂, and PBLH (UV-A excluded) during 06–18 March 2016.

removed safely from the predictors list. The observed and model-estimated hourly ozone mixing ratios during March 2016 period using TEMP, RH, WD, WS, NO, NO₂, UV-A (excluding PBLH) are shown in Fig. 13.

Fig. 13 reveals that in the absence of PBLH, there is no change in the prediction of the models, and GTB and DNN are still picking up high ozone mixing ratios (≥ 160 ppbv), and RF is still picking up ozone mixing ratios (≤ 120 ppbv). Similar results were obtained by excluding TEMP, and then both TEMP and PBLH from the predictors' list and fitting the models to the remaining predictors. All these results confirm that in the presence of UV-A, removal of TEMP or PBLH or both TEMP and PBLH from the predictors' list does not affect the models' predictions significantly.

4.2. Comparison of machine learning models to multiple linear regression

Multiple linear regression (MLR), a statistical model, is used to calculate functional dependence of a response variable on two or more explanatory variables. MLR entails constructing a linear relationship

involving a dependent variable and independent variables with unknown parameters and then calculating the unknown parameters using method of least squares. The fitted linear function represents a hyperplane geometrically.

To test the robustness of the machine learning models used in this paper, we compare the performance of the machine learning models to MLR model, fitted on the same training dataset, in estimating hourly ozone mixing ratios in Mexico City for the period 6–18 March 2016. Table 3 shows the accuracies of the machine learning models and MLR using two performance metrics namely R^2 and IOA. The MLR model heavily underestimates ozone peak values due to its inability to capture the nonlinear dependence of ozone on its predictors, and its performance is even worse than the RF model.

4.3. Second scenario

The scatterplots in Fig. 14 are based on the second scenario (considering all samples with UV-A $> 0.5 \text{ W/m}^2$), and the performance indicators R^2 and IOA indicate that the DNN model is still the best

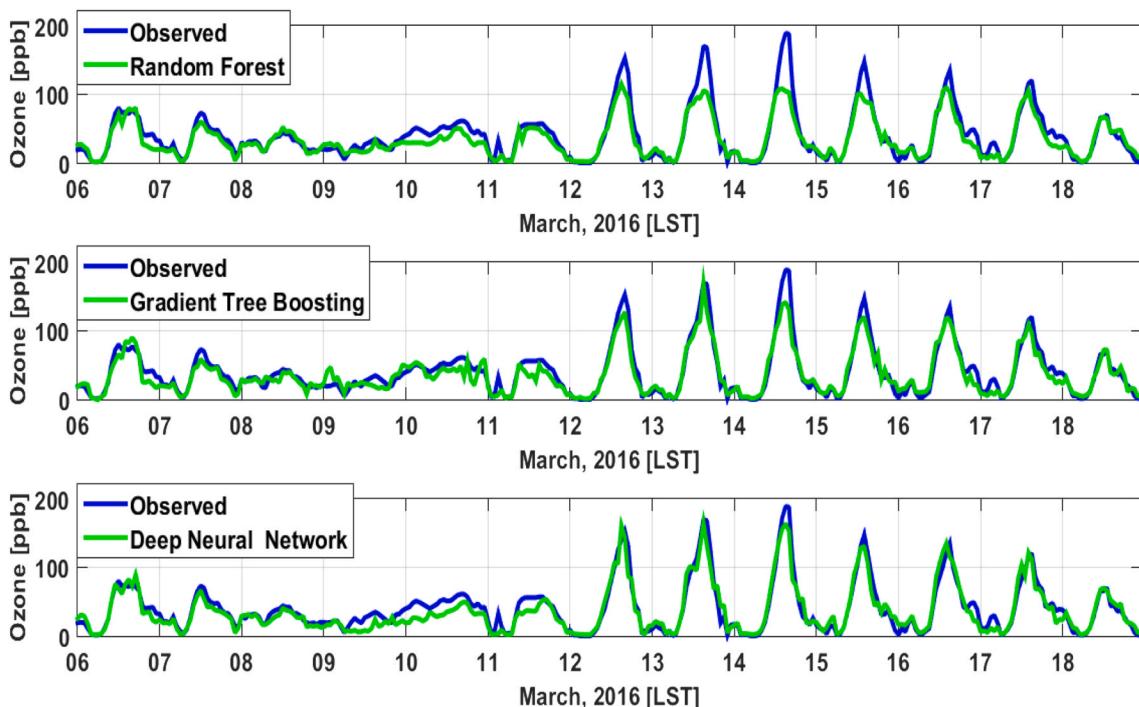


Fig. 13. Observed and model-estimated hourly ozone mixing ratios using TEMP, RH, WD, WS, NO, NO₂, and UV-A (PBLH excluded) during 06–18 March 2016.

Table 3

Performance metrics for ozone estimation.

Model	R Squared	IOA
DNN	0.92	0.97
GTB	0.90	0.96
RF	0.89	0.95
MLR	0.64	0.84

model. Fig. 15 represents SHAP summary plot and standard bar plot for the importance of each predictor during 6–18 March period based on second scenario ($UV\text{-}A > 0.5 \text{ W/m}^2$). In the SHAP summary plot, one hundred and sixty-one samples for each predictor are showing up with their Shapely values. In this case, the first three import predictors are NO, NO₂, and UV-A. And UV-A is standing at third position form the top. The least important predictor is PBLH.

4.4. Third scenario

The scatterplots in Fig. 16 are based on the third scenario (considering all samples with $UV\text{-}A > 20 \text{ W/m}^2$), and it shows that the DNN model is still the best model with highest R² and IOA.

Fig. 17 represents SHAP summary plot and standard bar plot for the importance of each predictor during 6–18 March period based on third scenario ($UV\text{-}A > 20 \text{ W/m}^2$). In the SHAP summary plot, eighty-four samples for each predictor are showing up with their Shapely values. In this case, the first three import predictors are NO, NO₂, and WS. And UV-A is showing up at the fifth position form the top. The least important predictor is PBLH.

5. Conclusions and recommendations

In this study, we propose a novel approach based on machine learning to simulate ozone (O₃) levels, pick up peak O₃ values, and rank the predictors according to their importance in the build-up of O₃ during Mexico City severe O₃ episode. Among various predictors which include

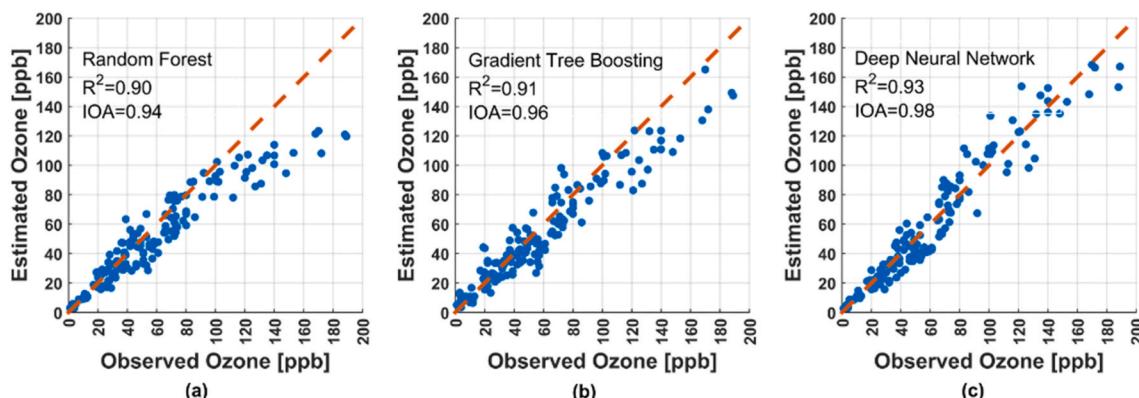


Fig. 14. Scatterplots depicting the performance of the RF, GTB, and DNN models in estimating hourly ozone mixing ratios using TEMP, RH, WD, WS, NO, NO₂, PBLH, and UV-A during 06–18 March 2016 based on the second scenario (considering all samples with $UV\text{-}A > 0.5 \text{ W/m}^2$).

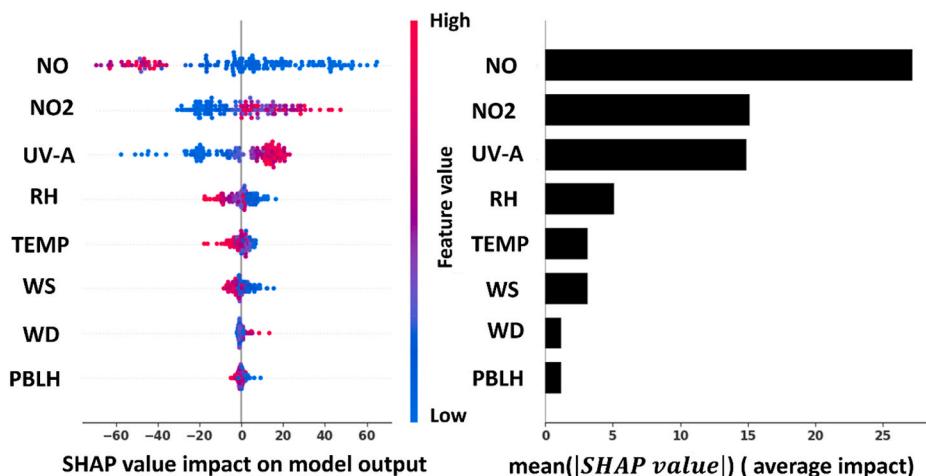


Fig. 15. SHAP summary plot (left) and mean absolute values bar plot (right) for the DNN model based on second scenario ($UV\text{-}A > 0.5 \text{ W/m}^2$) during 06–18 March 2016.

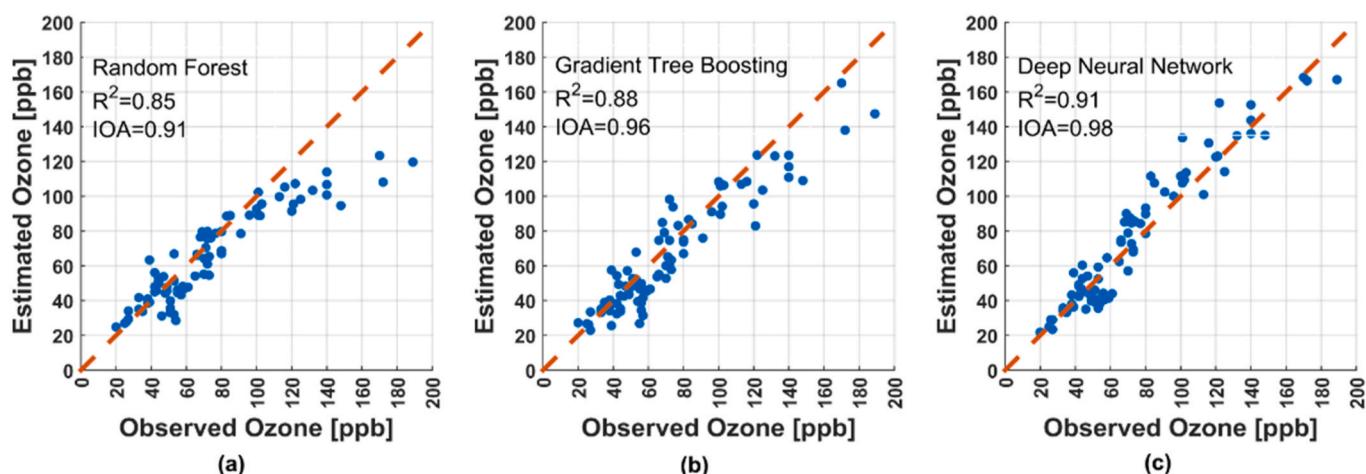


Fig. 16. Scatterplots depicting the performance of the RF, GTB, and DNN models in estimating hourly ozone mixing ratios using TEMP, RH, WD, WS, NO, NO₂, PBLH, and UV-A based on third scenario (considering all samples with $UV\text{-}A > 20\text{W/m}^2$) during 06–18 March 2016.

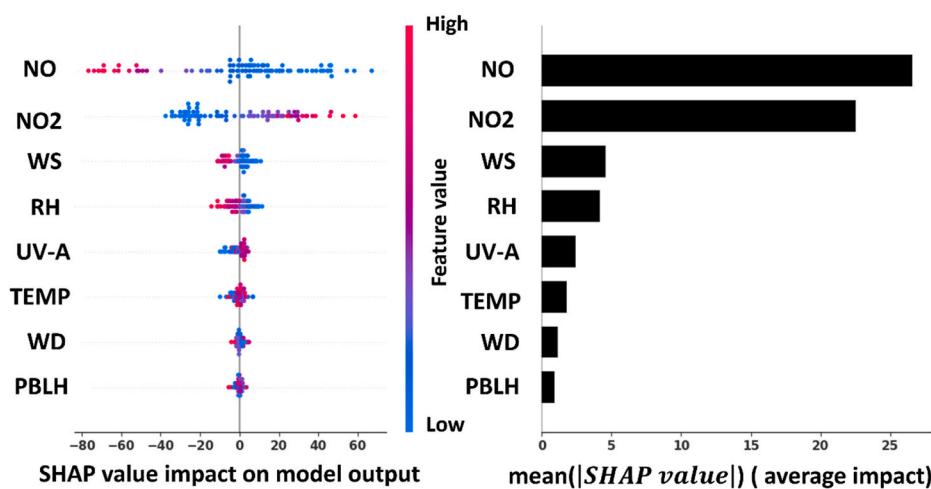


Fig. 17. SHAP summary plot (left) and mean absolute values bar plot (right) for DNN model based on third scenario ($UV\text{-}A > 20\text{W/m}^2$) during 06–18 March 2016.

chemical and surface meteorological data, this paper also incorporates for the first time continuous planetary boundary layer height data retrieved by a microwave radiometer into the suite of machine learning

models' predictors. The three machine learning models are depicting diurnal O₃ levels nicely (~90% accuracy), as long as O₃ mixing ratios are at moderate levels ($\leq 120 \text{ ppb}$), but only the Deep Neural Network and

Gradient Tree Boosting models are picking up high ozone mixing ratios (>160 ppb) with the Deep Neural Network model (92% accuracy) outperforming the Gradient Tree Boosting model (90% accuracy). The most important variables among all three scenarios are ultraviolet-A (UV-A) radiation, nitrogen monoxide (NO), nitrogen dioxide (NO_2) and wind speed (WS). While UV-A and WS are meteorological variables, NO and NO_2 are related to human activities and may be regulated. Such ranking of the air quality and meteorological variables is crucial for policy-making decisions regarding the prevention and mitigation of ozone detrimental effects during severe ozone episodes and provides insight into the functional dependency of ozone on its predictors. The proposed approach can be used for O_3 episodes at any other location to conduct importance analysis of predictors with regard to O_3 build-up. For future work, we intend to extend our study to other stations of the Mexico City atmospheric monitoring system, also incorporating more O_3 precursors such as volatile organic compounds which would result in a wider area coverage, a better understanding of O_3 dependence on its precursors, and better strategies for controlling ozone harmful effects. For such an effort the models presented in this paper will have to be tuned and trained again to accommodate larger datasets and more predictors.

CRediT authorship contribution statement

M. Ahmad: Methodology, Formal analysis, Machine learning models, Writing – original draft. **B. Rappenglück:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **O.O. Osibanjo:** Data curation. **A. Retama:** Data curation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The air quality and vertical profiles measurements were carried out by personnel from the Atmospheric Monitoring Directorate of the Secretariat of the Environment of Mexico City, we acknowledge all their effort and commitment. We thank O. Rivera-Hernández from the Mexico City Air Quality Monitoring Directorate for granting access to the profilers' data. We gratefully acknowledge the use of the University of Houston Research Computing Center (RCC). Mateen Ahmad acknowledges the financial support provided by the Higher Education Commission (HEC), Pakistan and University of Engineering and Technology (UET), Lahore, Pakistan.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jclepro.2022.134638>.

References

- Abirami, S., Chitra, P., 2021. Regional air quality forecasting using spatiotemporal deep learning. *J. Clean. Prod.* 283, 125341.
- Allu, S.K., Srinivasan, S., Maddala, R.K., Reddy, A., Anupoju, G.R., 2020. Seasonal ground level ozone prediction using multiple linear regression (MLR) model. *Model. Earth Syst. Environ.* 6 (4), 1981–1989.
- Arsić, M., Mihajlović, I., Nikolić, D., Živković, Ž., Panić, M., 2020. Prediction of ozone concentration in ambient air using multilinear regression and the artificial neural networks methods. *Ozone: Sci. Eng.* 42 (1), 79–88.
- Balamurugan, V., Balamurugan, V., Chen, J., 2022. Importance of ozone precursors information in modelling urban surface ozone variability using machine learning algorithm. *Sci. Rep.* 12 (1), 1–8.
- Bekesiene, S., Meidute-Kavalaiuskiene, I., Vasiliauskienė, V., 2021. Accurate prediction of concentration changes in ozone as an air pollutant by multiple linear regression and artificial neural networks. *Mathematics* 9 (4), 356.
- Bell, M.L., Peng, R.D., Dominici, F., 2006. The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations. *Environ. Health Perspect.* 114 (4), 532–536.
- Berman, S., Ku, J.Y., Rao, S.T., 1999. Spatial and temporal variation in the mixing depth over the northeastern United States during the summer of 1995. *J. Appl. Meteorol.* 38 (12), 1661–1673.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32.
- Chollet, F., 2017. Deep learning with python, Second Edition. Manning Publications CO. ISBN-10: 1617296864.
- Comrie, A.C., Yarnal, B., 1992. Relationships between synoptic-scale atmospheric circulation and ozone concentrations in metropolitan Pittsburgh, Pennsylvania. *Atmos. Environ. Part B - Urban Atmos.* 26 (3), 301–312.
- Davies, T.D., Kelly, P.M., Low, P.S., Pierce, C.E., 1992. Surface ozone concentrations in Europe: links with the regional-scale atmospheric circulation. *J. Geophys. Res. Atmos.* 97 (D9), 9819–9832.
- Dueñas, C., Fernández, M.C., Cañete, S., Carretero, J., Liger, E., 2002. Assessment of ozone variations and meteorological effects in an urban area in the Mediterranean Coast. *Sci. Total Environ.* 299 (1–3), 97–113.
- Elminir, H.K., 2005. Dependence of urban air pollutants on meteorology. *Sci. Total Environ.* 350 (1–3), 225–237.
- Eslami, E., Choi, Y., Lops, Y., Sayeed, A., 2020. A real-time hourly ozone prediction system using deep convolutional neural network. *Neural Comput. Appl.* 32 (13), 8783–8797.
- Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 1189–1232.
- Friedman, J.H., 2002. Stochastic gradient boosting. *Comput. Stat. Data Anal.* 38 (4), 367–378.
- Gagliardi, R.V., Andenna, C., 2020. A machine learning approach to investigate the surface ozone behavior. *Atmosphere* 11 (11), 1173. <https://doi.org/10.3390/atmos11111173>.
- Gradišar, D., Grašič, B., Božnar, M.Z., Mlakar, P., Kocijan, J., 2016. Improving of local ozone forecasting by integrated models. *Environ. Sci. Pollut. Control Ser.* 23 (18), 18439–18450.
- Hafeez, A., Taqvi, S.A.A., Fazal, T., Javed, F., Khan, Z., Amjad, U.S., et al., 2020. Optimization on cleaner intensification of ozone production using artificial neural network and response surface methodology: parametric and comparative study. *J. Clean. Prod.* 252, 119833.
- Haykin, S., 2009. *Neural Networks and Learning Machines*, 3/E. Pearson Education India.
- Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. *Neural Network.* 2 (5), 359–366.
- Hu, J., Chen, J., Ying, Q., Zhang, H., 2016. One-year simulation of ozone and particulate matter in China using WRF/CMAQ modeling system. *Atmos. Chem. Phys.* 16 (16), 10333–10350.
- Ibarra, G., Elias, A., Albizu, M.V., Aguirre, E., 2000. Multiple linear regression modelling for short-term real-time prediction of hourly ozone, NO_2 and NO levels in the area of Bilbao. *WIT Trans. Ecol. Environ.* 41.
- Jaioun, K., Saithanu, K., Mekparyup, J., 2014. Multiple linear regression model to estimate ozone concentration in chonburi, Thailand. *Int. J. Appl. Environ. Sci.* 9 (4), 1305–1308.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. *An Introduction to Statistical Learning*, vol. 112. Springer, New York, p. 18.
- Jenkin, M.E., Clemitshaw, K.C., 2000. Ozone and other secondary photochemical pollutants: chemical processes governing their formation in the planetary boundary layer. *Atmos. Environ.* 34 (16), 2499–2527.
- Juarez, E.K., Petersen, M.R., 2021. A comparison of machine learning methods to forecast tropospheric ozone levels in Delhi. *Atmosphere* 13 (1), 46.
- Kelly, J.T., Jang, C., Zhu, Y., Long, S., Xing, J., Wang, S., Pye, H.O., 2021. Predicting the nonlinear response of PM_{2.5} and ozone to precursor emission changes with a response surface model. *Atmosphere* 12 (8), 1044.
- Kuhn, H.W., Tucker, A.W. (Eds.), 1953. *Contributions to the Theory of Games*, vol. 2. Princeton University Press.
- Liang, Y.C., Maimury, Y., Chen, A.H.L., Juarez, J.R.C., 2020. Machine learning-based prediction of air quality. *Appl. Sci.* 10 (24), 9151.
- Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* 30.
- Makarova, A., Evstaf'eva, E., Lapchenco, V., Varbanov, P.S., 2021. Modelling tropospheric ozone variations using artificial neural networks: a case study on the Black Sea coast (Russian Federation). *Clean. Eng. Technol.* 5, 100293 <https://doi.org/10.1016/j.clet.2021.100293>.
- Mills, G., Buse, A., Gimeno, B., Bermejo, V., Holland, M., Emberson, L., Pleijel, H., 2007. A synthesis of AOT40-based response functions and critical levels of ozone for agricultural and horticultural crops. *Atmos. Environ.* 41 (12), 2630–2643.
- Moustris, K.P., Nastos, P.T., Larissi, I.K., Palaiatos, A.G., 2012. Application of multiple linear regression models and artificial neural networks on the surface ozone forecast in the greater Athens area, Greece. *Adv. Meteorol.* 2012, 894714 <https://doi.org/10.1155/2012/894714>, 8 pages.
- Osibanjo, O.O., Rappenglück, B., Retama, A., 2021. Anatomy of the March 2016 severe ozone smog episode in Mexico-City. *Atmos. Environ.* 244, 117945 <https://doi.org/10.1016/j.atmosenv.2020.117945>.

- Osibanjo, O.O., Rappenglück, B., Ahmad, M., Jaimes-Palomera, M., Rivera-Hernández, O., Prieto-González, R., Retama, A., 2022. Intercomparison of planetary boundary-layer height in Mexico City as retrieved by microwave radiometer, micro-pulse lidar and radiosondes. *Atmos. Res.* 271, 106088 <https://doi.org/10.1016/j.atmosres.2022.106088>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Probst, P., Wright, M.N., Boulesteix, A.L., 2019. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinay Rev.: Data Min. Knowl. Discov.* 9 (3), e1301 <https://doi.org/10.1002/widm.1301>.
- Rappenglück, B., Kourtidis, K., Fabian, P., 1993. Measurements of ozone and peroxyacetyl nitrate (PAN) in Munich. *Atmos. Environ. Part B - Urban Atmos.* 27 (3), 293–305.
- Sayeed, A., Choi, Y., Eslami, E., Lops, Y., Roy, A., Jung, J., 2020. Using a deep convolutional neural network to predict 2017 ozone concentrations, 24 hours in advance. *Neural Network.* 121, 396–408.
- Seinfeld, J.H., Pandis, S.N., 2016. Atmospheric Chemistry and Physics: from Air Pollution to Climate Change. Wiley, Hoboken, New Jersey, USA.
- Solomon, P., Cowling, E., Hidy, G., Furiness, C., 2000. Comparison of scientific findings from major ozone field studies in North America and Europe. *Atmos. Environ.* 34 (12–14), 1885–1920.
- Sousa, S.I.V., Martins, F.G., Alvim-Ferraz, M.C., Pereira, M.C., 2007. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environ. Model. Software* 22 (1), 97–103.
- Velasco, E., Retama, A., 2017. Ozone's threat hits back Mexico City. *Sustain. Cities Soc.* 31, 260–263.
- Verma, N., Kumari, S., Lakhani, A., Kumari, K.M., 2019. 24 hour advance forecast of surface ozone using linear and non-linear models at a semi-urban site of indo-gangetic plain. *Int.J.Environ. Sci. Nat. Res.* 18 (2), 46–55.
- Wang, H.W., Li, X.B., Wang, D., Zhao, J., Peng, Z.R., 2020. Regional prediction of ground-level ozone using a hybrid sequence-to-sequence deep learning approach. *J. Clean. Prod.* 253, 119841.
- Yafouz, A., Ahmed, A.N., Zaini, N.A., Sherif, M., Sefelnasr, A., El-Shafie, A., 2021. Hybrid deep learning model for ozone concentration prediction: comprehensive evaluation and comparison with various machine and deep learning algorithms. *Eng. Appl. Computat. Fluid Mech.* 15 (1), 902–933.
- Zhan, J., Liu, Y., Ma, W., Zhang, X., Wang, X., Bi, F., et al., 2022. Ozone formation sensitivity study using machine learning coupled with the reactivity of volatile organic compound species. *Atmos. Meas. Tech.* 15 (5), 1511–1520.
- Zhang, J., Rao, S.T., Daggupaty, S.M., 1998. Meteorological processes and ozone exceedances in the northeastern United States during the 12–16 July 1995 episode. *J. Appl. Meteorol.* 37 (8), 776–789.
- Zhang, Y., Haghani, A., 2015. A gradient boosting method to improve travel time prediction. *Transport. Res. C Emerg. Technol.* 58, 308–324.
- Zielesny, A., 2011. From Curve Fitting to Machine Learning, vol. 18. Springer, Berlin Heidelberg.