Turkish Journal of Science & Technology
19(1), 1-12, 2024

Research Paper
https://doi.org/10.55525/tjst.1224661

Prediction of Air Pollution with Machine Learning Algorithms

Kahraman OĞUZ[1*], Muhammet Ali PEKİN[2]
[1] Turkish State Meteorological Service, 9th Directorate of Meteorology, Ankara
[2] Turkish State Meteorological Service, 12th Directorate of Meteorology, Erzurum
[*1] koguz@mgm.gov.tr, [2] mapekin@mgm.gov.tr

(Geliş/Received: 26/12/2022; Kabul/Accepted: 1/10/2023)

Abstract: Air pollution has become an important problem due to its threats. Air pollutants are in complex interaction with atmosphere and environment. For this reason, it is important to study air pollution issues. In recent years, studies on prediction of air pollutants with machine learning methods have gained momentum. In this study, some air pollutants are predicted with various machine learning algorithms considering meteorological factors. In machine learning phase, a separate study is conducted with various machine learning algorithms (multilayer perceptron neural network, stochastic gradient descent, ridge regression, cross decomposition) considering temperature, relative humidity, wind, pressure and air pollutant measurements of previous hour. Consistencies of these algorithms in estimating pollutant concentrations are compared. Various statistical metrics are used to analyze the consistencies. As a result, the coefficient of determination of all algorithms are found above 0.67, considering the test section. It is found that the coefficient of determination of the multilayer perceptron neural network algorithm provides better results than other algorithms.

Key words: Air pollution, machine learning, neural network, modeling, Çanakkale city.

Hava Kirliliğinin Makine Öğrenme Algoritmaları ile Tahmin Edilmesi

Öz: Hava kirliliği, canlı sağlığına yönelik tehditleri sebebiyle önemli bir problem haline gelmiştir. Hava kirliliği atmosfer ve çevre ile karmaşık ilişki içerisindedir. Bu nedenle hava kirliliği ile alakalı konuların çalışılması önemlidir. Son yıllarda hava kirleticilerinin makine öğrenmesi yöntemleriyle tahmin edilmesine yönelik çalışmalar hız kazanmıştır. Bu çalışmada, meteorolojik faktörler göz önüne alınarak çeşitli makine öğrenme algoritmaları ile bazı hava kirleticilerinin tahmini yapılmıştır. Makine öğrenmesi aşamasında, bir önceki saatin sıcaklık, bağıl nem, rüzgar, basınç ve hava kirletici ölçümleri dikkate alınarak çeşitli makine öğrenmesi algoritmaları (çok katmanlı algılayıcı sinir ağı, stokastik gradyan inişi, sırt regresyonu, çapraz ayrıştırma) ile ayrı ayrı çalışma yapılmıştır. Bu algoritmaların kirletici konsantrasyonlarını tahmin etmedeki tutarlılıkları karşılaştırılmıştır. Tutarlılıkları analiz etmek için çeşitli istatistiksel metrikler kullanılmıştır. Sonuç olarak, test bölümü dikkate alındığında tüm algoritmaların belirleme katsayısı 0.67'nin üzerinde bulunmuştur. Çok katmanlı algılayıcı sinir ağı algoritmasının belirleme katsayısının diğer algoritmalara göre daha iyi sonuçlar verdiği tespit edilmiştir.

Anahtar kelimeler: Hava kirliliği, makine öğrenmesi, sinir ağı, modelleme, Çanakkale ili.

1. Introduction

Air pollution has significant impact on public health and environment. Population, urbanization, industrial growth, energy consumption and usage of transportation vehicles have increased significantly in the last decade worldwide. This has resulted in rised emissions of air pollutants including greenhouse gases, ambient temperature and other atmospheric variables [1]. The main factors affecting air pollutant concentration are emission sources and meteorological factors. The existence of a significant relationship between meteorological factors and air pollutant concentration has been addressed by many studies [2-4]. Transportation, chemistry and deposition of particulate matter (PM) are mainly controlled by meteorological factors. Meteorological factors not only affect each other, but also form a closely linked system with PM. The effects of meteorological conditions on PM concentration are quite complex [5]. There are various methods for estimating PM concentration. Machine learning method is one of the most frequently used prediction methods in recent years.

Machine learning is a rapidly developing field that enables computers to learn based on data. Data can come from a variety of sources, including physical experiments, computer models, or a combination of both. Machine learning method has had successful applications in many areas [6]. Over the last three decades, it has also been increasingly applied in the field of air quality prediction due to the development of statistical models based on machine learning techniques and its ability to explore, analyze and make predictions on multiple and complex datasets. The main purpose of a machine learning algorithm is to provide a model that captures the general properties and interactions of the dataset (learned) to obtain information from data and make predictions [7-8].

* Corresponding author: koguz@mgm.gov.tr. ORCID Number of authors: [1] 0000-0001-5305-6145, [2] 0000-0002-6807-890X

There are various studies on the analysis of air pollution with machine learning methods. These studies are aimed at estimating air pollution in various regions with several machine learning algorithms and examining their consistency. Several studies about prediction of air pollution with machine learning algorithms are compared in Table 1. [9] studied PM2.5 prediction with several machine learning algorithms including linear regression, random forecast and ridge regression. [10] studied prediction of air quality index with several machine learning algorithms like k-Nearest Neighbor (KNN), Gaussian Naive Bayes, support vector regression (SVM). [11] studied prediction of PM10, CO, SO2, O2 and O3 with random forest, multiple linear regression algorithms. [12] studied PM10 prediction with artificial neural network. [13] studied PM2.5 prediction with several machine learning algorithms like long short-term memory (LSTM), gradient boosting model (GRU), convolutional neural network (CNN).

Table 1. Some studies in the literature.

Pollutant	**Proposed Method**	**Reference**
SO2, PM10	Random Forest, Decision Tree	[14]
SO2, NO2, O3, CO, PM10	Random Forest	[15]
PM10, PM2.5	Support Vector Regression, Autoregressive Integrated Moving Average, Long Short-Term Memory	[16]
PM2.5	Gradient Boosting Model	[17]
PM10, PM2.5	Linear Regression, Random Forest Regressor, Gradient Boosting Regressor, K Neighbors Regressor, MLP Regressor, and Decision Tree Regressor CART	[18]
AQI	Adaptive Boosting, Artificial Neural Network, Random Forest, Stacking Ensemble, and Support Vector Machine	[19]
PM2.5	LSTM, Bi-LSTM, GRU, Bi-GRU, CNN, and a hybrid CNN-LSTM	[13]
PM10	Artificial Neural Network	[12]
PM10, CO, SO2, O2, O3	Random Forest, Multiple Linear Regression	[11]
AQI	KNN, Gaussian Naive Bayes, SVM, RF, and XGBoost	[10]
PM2.5	Linear Regression, Random Forest, KNN, Ridge and Lasso, XGBoost, and AdaBoost	[9]

In this study, we performed modeling with various machine learning algorithms (multilayer perceptron neural network, stochastic gradient descent, ridge regression, cross decomposition) by considering temperature, relative humidity, wind, pressure and air pollutant measurements of the previous hour to predict pollutant concentrations. As a result, consistencies of pollutant concentrations of PM10, SO2, NO2, NOX and O3 were compared.

2. Material and Method

2.1. Study area and data

In this study, hourly meteorological data of temperature, relative humidity, wind and pressure parameters were obtained from Çanakkale Central Meteorology Station of Turkish State Meteorological Service (latitude: 40.1410, longitude: 26.3993). PM10, SO2, NO2, NOX and O3 pollutant data were obtained from Çanakkale Central air quality station (latitude: 40.1366, longitude: 26.4055) affiliated to Ministry of Environment, Urbanization and Climate Change. Locations of the stations are shown in Figure 1. The data covers the period from 1 January 2019 to 31 December 2021. While applying the machine learning method, 80% of the data set was randomly selected as training data and 20% was used as test data.

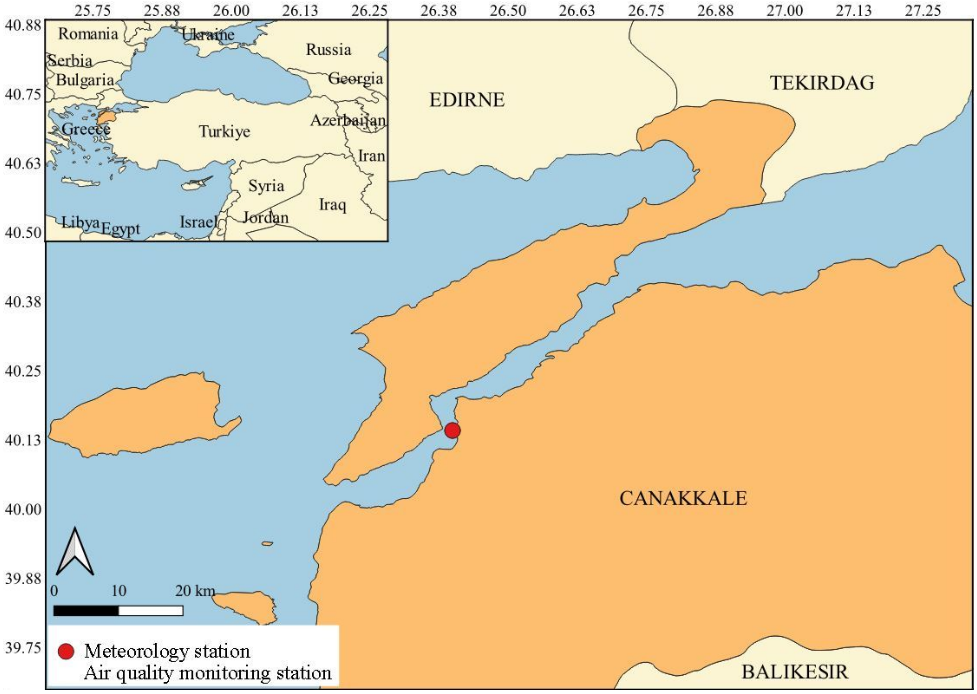

Figure 1. Locations of the meteorological and air quality.

Table 2 shows statistical information about meteorological parameters and air pollutants. PM10 average is 43.59 µg/m³, minimum and maximum values are 2.0 µg/m³ and 1003.0 µg/m³, respectively. SO2 average is 9.58 µg/m³, minimum and maximum values are 0 µg/m³ and 183.23 µg/m³, respectively. NO2 average is 20.62 µg/m³, minimum and maximum values are 0.26 µg/m³ and 145.86 µg/m³, respectively. NOX average is 40.02 µg/m³, minimum and maximum values are 1.53 µg/m³ and 392.95 µg/m³, respectively. The O3 average is 56.0 µg/m³, minimum and maximum values are 1.05 µg/m³ and 205.40 µg/m³, respectively.

Descriptive statistics of meteorological and air pollutant parameters are presented in Table 3. It was observed that the highest correlation coefficient (r=-0.35) of NO2 is with wind. It was observed that the highest correlation coefficient (r=0.558) of O3 is with temperature. This is followed by relative humidity with the value of r=-0.5. There is a weak negative relationship between NO2 and wind. There is a moderate positive relationship between O3 and temperature while moderate negative relationship between O3 and relative humidity. It is clear that the relationship between other air pollutants and meteorological parameters is very weak.

Table 2. Descriptive statistics of meteorological and air pollutant parameters.

	T	RH	W	PS	PM10	SO2	NO2	NOX	O3
Valid	26145	25815	26117	25979	25254	24714	24712	24521	25514
Missing	157	487	185	323	1048	1588	1590	1781	788
Mean	17.40	66.96	3.35	1014.75	43.59	9.58	20.62	40.02	56.00
Std. Dev.	7.94	16.40	2.24	6.31	35.40	9.99	14.57	28.68	26.07
Skew.	0.05	-0.34	1.41	0.29	16.09	4.35	1.73	2.54	0.22
Range	43.70	84.00	18.30	48.90	1001.00	183.23	145.60	391.42	204.35
Min.	-4.20	16.00	0.00	992.20	2.00	0.00	0.26	1.53	1.05
Max.	39.50	100.00	18.30	1041.10	1003.00	183.23	145.86	392.95	205.40
T: Temperature (°C), RH: Relative humidity (%),W: Wind (m sñ¹), PS: Pressure (hpa).									

Table 3. Correlations of the parameters used in the study.

Variable		T	RH	W	PS	PM10	SO2	NO2	NOX	O3
T	Pearson's r	—								
	p-value	—								
RH	Pearson's r	-0.6	—							
	p-value	< .001	—							
W	Pearson's r	0.048	-0.14	—						
	p-value	< .001	< .001	—						
PS	Pearson's r	-0.45	0.074	-0.12	—					
	p-value	< .001	< .001	< .001	—					
PM10	Pearson's r	0.116	-0.06	0.012	-0.06	—				
	p-value	< .001	< .001	0.063	< .001	—				
SO2	Pearson's r	-0.02	-0.05	-0.14	0.028	0.064	—			
	p-value	0.001	< .001	< .001	< .001	< .001	—			
NO2	Pearson's r	-0.13	0.171	-0.35	0.053	0.12	0.31	—		
	p-value	< .001	< .001	< .001	< .001	< .001	< .001	—		
NOX	Pearson's r	-0.14	0.16	-0.24	0.049	0.074	0.28	0.849	—	
	p-value	< .001	< .001	< .001	< .001	< .001	< .001	< .001	—	
O3	Pearson's r	0.558	-0.5	0.23	-0.3	-0.02	-0.09	-0.45	-0.47	—
	p-value	< .001	< .001	< .001	< .001	0.002	< .001	< .001	< .001	—

2.2. Multilayer Perceptron Neural Network (MLPNN)

Artificial neural network is a data processing method that mimics the neural structure of the human brain. It establishes relationships between inputs and outputs. It has parallel data processing architecture like the human nervous system [20]. Artificial neural network is one of the most used machine learning algorithms.

MLPNN is a type of feed forward artificial neural network, which uses Boolean function. MLPNN includes layers of nodes, which allow unidirectional forward connections of inputs and outputs. MLPNN consist of 3 layers including input layer, output layer and a hidden layer (between input and output layers). Data are transferred from input layer to output layer [21]. A schematic diagram of multilayer perceptron neural network presented in Figure 2.

The main aim of MLPNN is to predict future trends in a dataset given current and previous conditions. Function logic is related to modeling the connections between variables. The objective of MLPNN application is to find unknown function f with input vectors in X and output vectors in Y:

$$Y = f\ (X) \tag{1}$$

where X = [n x k], Y = [n x j], n is number of training patterns, k the number of input nodes/variables and j the number of output nodes/variables. The training data is represented with the matrices of X and Y. The f function is defined with regulable network weights. During training the function f is optimised, such that the network output for the input vectors in X is as close as possible to the target values in Y [22].

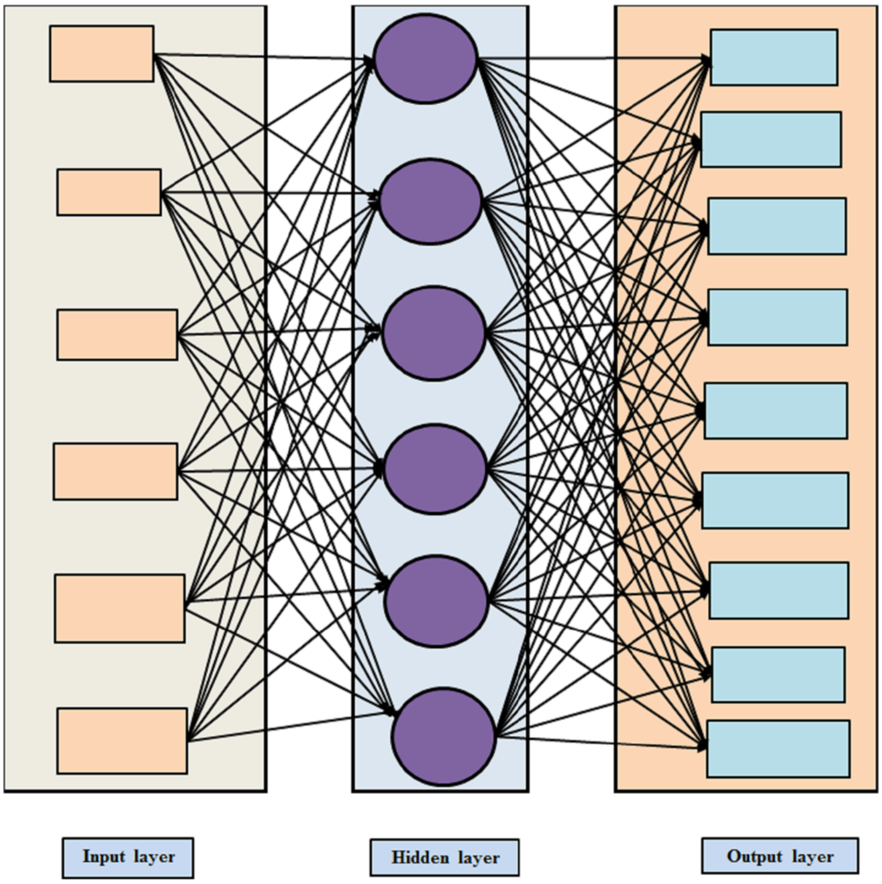

Figure 2. A schematic diagram of multilayer perceptron neural network [21].

2.3. Stochastic Gradient Descent (SGD)

SGD is another most used algorithms in machine learning. SGD is a type of optimization combining classical gradient descent with random subsampling within the target functional [23]. SGD shows good performance in convex and non-convex optimization. SGD minimizes a loss function selected through a linear function. The algorithm approximates a true gradient, considering one sample at a time, and simultaneously updates the model according to the gradient of the loss function [24].

SGD is calculated as below:

$$\theta = \theta - \alpha \nabla_{\theta} J(\theta; x^{(i)}, y^{(i)}) \tag{2}$$

where $x^{(i)}$ is training example, $y^{(i)}$ is label, θ is vector parameters of J(θ), α is learning rate [25-26].

2.4. Ridge Rigression (RR)

RR is an algorithm to analyze multivariate regression data and it is the biased estimation methods. The aim is to find the factors that minimize the error sum of squares by implementing a penalty to these factors. It is durable to over-fitting and it tenders an answer to multidimensionality.

RR is calculated as below:

$$\beta^* = \left(X^{(T)}X + kI\right)^{-1} X^{(T)}Y \tag{3}$$

where β*= (p-1)x1 dimensional vector of ridge regression coefficients, I= (p-1)x(p-1) dimensional unit matrix and k constant value, which is almost 0≤k≤1. [27].

2.5. Cross Decomposition (CD)

CD algorithms are useful for finding relations between two matrices (datasets). Examples of cross decomposition are partial least squares regression and canonical correlation analysis. Main aim of the algorithms are to determine multidimensional aspect in the X space that clarifies the maximum multidimensional variance aspect in the Y space. The mathematical model of cross decomposition is below [28-29]:

Assuming the separations of X and Y are done in a way that maximizes the covariance between T and U:

$$X = TPT + E \tag{4}$$
$$Y = UQT + F \tag{5}$$

where, X is nxm matrix of predictors, Y is nxp matrix of responses, T nx1 projections of X, U is nx1 projections of Y, P is mx1 orthogonal loading matrix, Q is px1 orthogonal loading matrix, E is error term and F is error term.

2.6. Data Preprocessing

Data preprocessing is one of the most crucial part for machine learning. It is a data mining technique which converts raw data into a more intelligible, advantageous and fertile format. The main task for a model to be precise in estimations is that the algorithm should be able to smoothly commentate the data's features [30-31].

The data set used in the study is not complete. In time series estimates, it is required that the data set to be completed. Therefore, it is important to complete the missing data. However, this study is not a time series estimation study. Therefore, it is a convenient method to remove rows with missing columns. In our study, we removed the rows with missing columns from the data set and did not use them in the train/test stages.

In the study, the input data are scaled between 0-1. In this way, it was ensured that the different digits of the parameters were on the same scale. This ensures that the parameters have the same effect when calculating the output variable (PM10, NO2, etc.).

In this study, we used max-min normalization methodology. The max-min normalization is calculated as below:

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{6}$$

2.7. Evaluation metrics

The R-Squared (Coefficient of determination), RMSE (Root Mean Square Error) and MAE (Mean absolute error) metrics are mainly used to evaluate model prediction performance in regression analysis. R-squared shows the coefficient of how well the values fit compared to the real values. Root Mean Square Error is very common and the standard deviation of the residuals (prediction errors). MAE shows the variation between the real and predicted values extracted by averaged the absolute difference over the data set [32].

The matrics are calculated as below:

$$r^2 = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}} \tag{7}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum(y - x)^2} \tag{8}$$

$$\text{MAE} = \frac{1}{n}\sum |y - x| \tag{9}$$

where, y is donated predicted, x is observed and n is number of samle.

3. Results and Discussions

The performance results of machine learning models are presented in Table 4. It may be seen from the table that MLPNN model performed better than others. MLPNN model estimated O3 parameter best with the scores of r=0.887, RMSE=8.81 and MAE=5.81. MLPNN then successfully estimated SO2 (r=0.783), PM10 (r=0.752), NOX (r=0.716) and NO2 (0.694), respectively. The results are close to each other when train and test scores are evaluated together.

Table 3. Performance results of machine learning models.

		r2		RMSE		MAE	
Parameter	**Model**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**
NO2	**SGD**	0.697	0.691	8.05	7.99	5.2	5.22
	RR	0.697	0.692	8.04	7.99	5.17	5.19
	MLPNN	0.701	0.694	7.99	7.95	5.03	5.06
	CD	0.697	0.691	8.04	7.99	5.16	5.19
NOX	**SGD**	0.674	0.711	16.54	14.99	9.44	8.86
	RR	0.675	0.711	16.53	14.96	9.39	8.82
	MLPNN	0.687	0.716	16.22	14.85	9.03	8.52
	CD	0.675	0.712	16.53	14.96	9.35	8.79
O3	**SGD**	0.881	0.880	9.01	9.06	5.88	5.96
	RR	0.881	0.880	9.01	9.05	5.86	5.94
	MLPNN	0.886	0.887	8.82	8.81	5.73	5.81
	CD	0.881	0.880	9.01	9.06	5.86	5.94
PM10	**SGD**	0.707	0.748	18.72	16.87	6.99	6.74
	RR	0.713	0.752	18.54	16.74	6.83	6.56
	MLPNN	0.714	0.752	18.51	16.76	6.69	6.38
	CD	0.714	0.752	18.5	16.76	6.73	6.45
SO2	**SGD**	0.746	0.779	5.13	4.63	2.53	2.47
	RR	0.747	0.780	5.11	4.62	2.49	2.43
	MLPNN	0.751	0.783	5.07	4.59	2.45	2.4
	CD	0.748	0.780	5.11	4.63	2.46	2.41

Scatter plots of air pollutants according to the models are shown in Figure 3. In the scatterplots, blue graphs on the left show train results and red graphs on the right show test results. The results are mostly clustered around the regression line. The scatter plot of O3 variable shows a more successful result than others. Variables of O3, NOX and NO2 are more harmonious with regression line. It is seen that PM10 and SO2 variables give the least successful results than others. It is also seen that especially large values of these two parameters deviate considerably from the regression line.

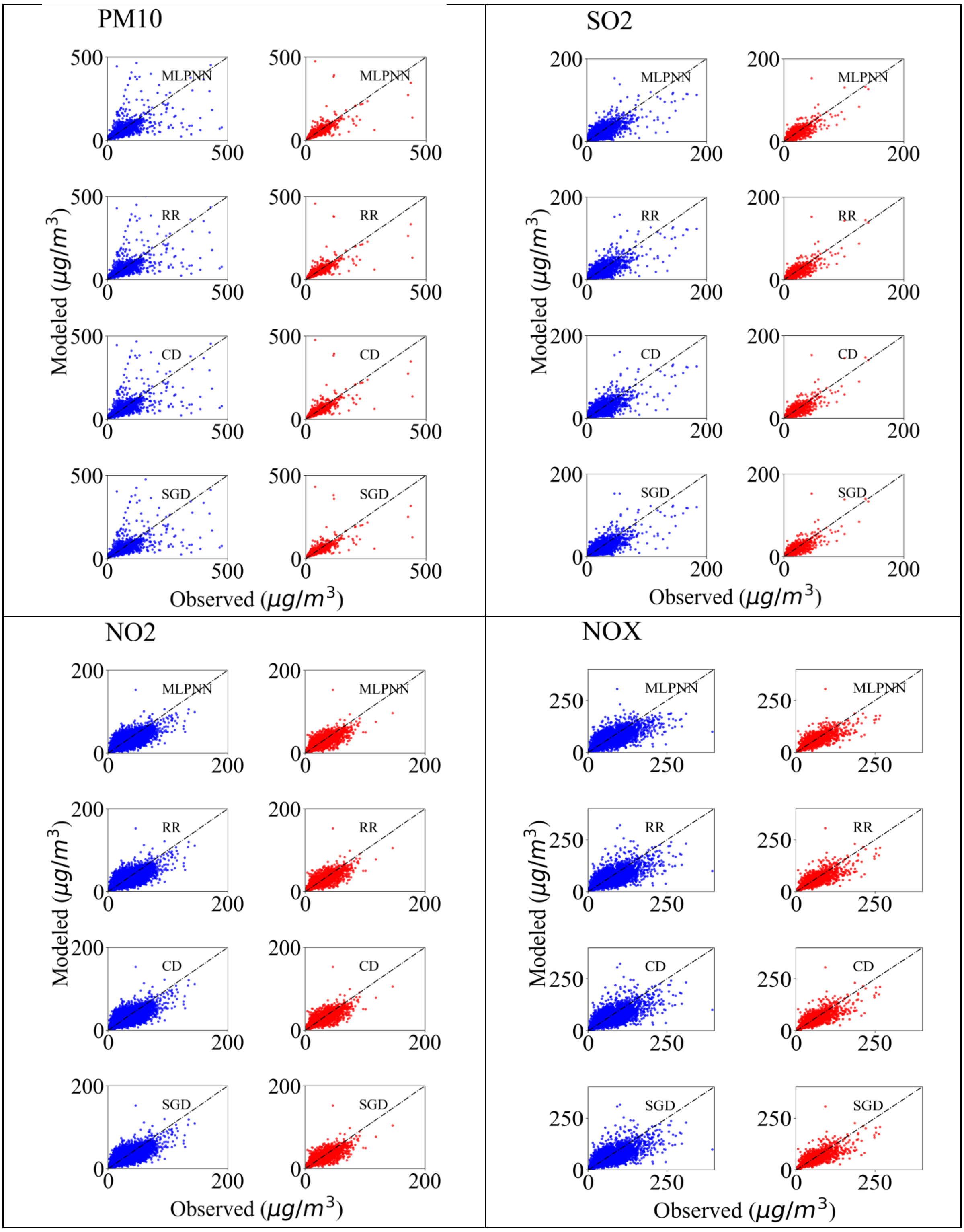

Figure 3. Scatter plots of some air pollutants according to the models.

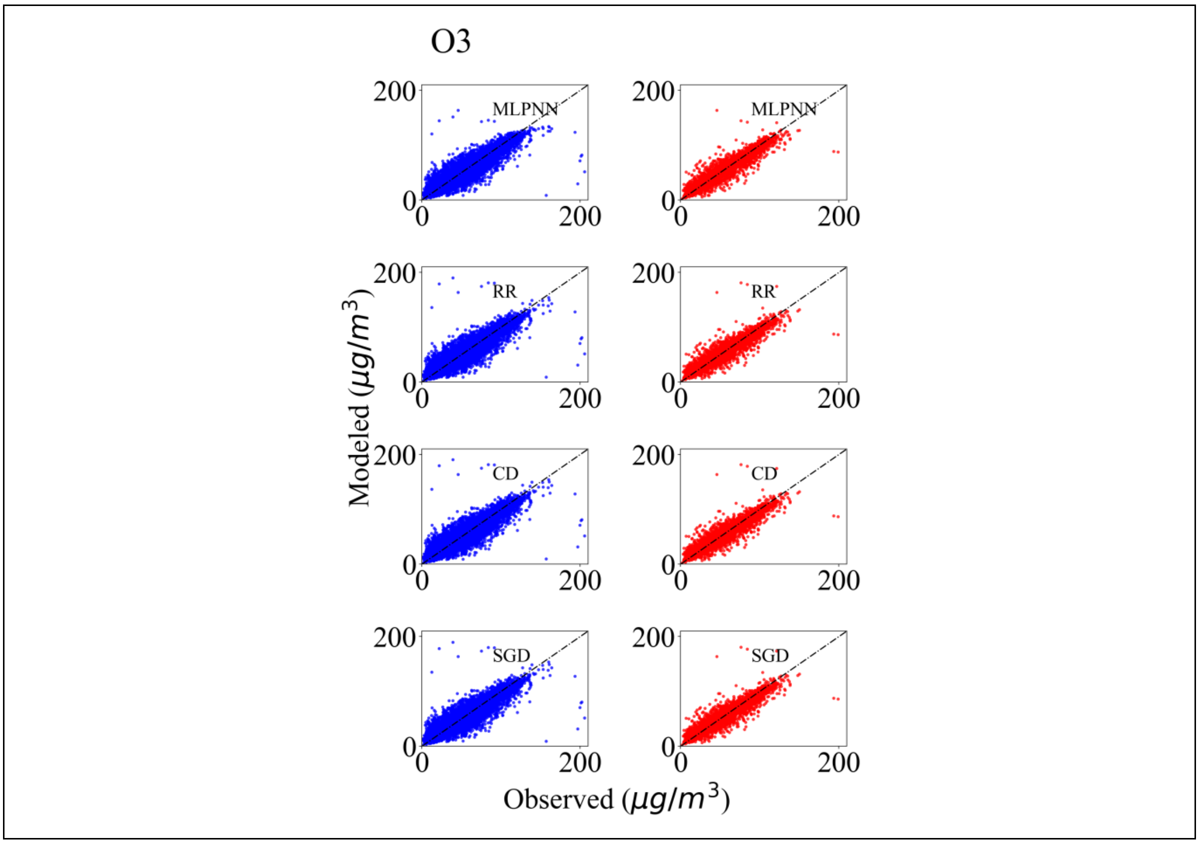

Figure 3 (cont.). Scatter plots of some air pollutants according to the models.

Error histogram graphics of the test section are shown in Figure 4. Histogram plots show the frequency distribution of the datasets. The clustering of data around zero in the error histogram indicates that the model is successful. According to the graphs, it is seen that the error amounts of O3, NOX and NO2 variables are clustered closest to 0 and its surroundings. It is also seen that PM10 and SO2 parameters are clustered a little further away from 0 compared to other parameters.

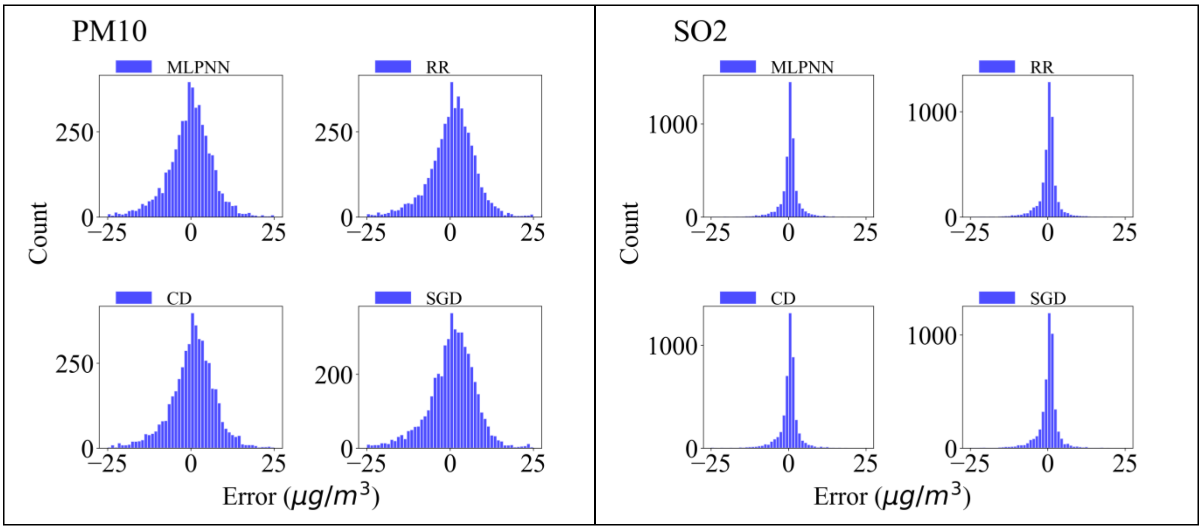

Figure 4. Error histograms of the test sections of the machine learning models.

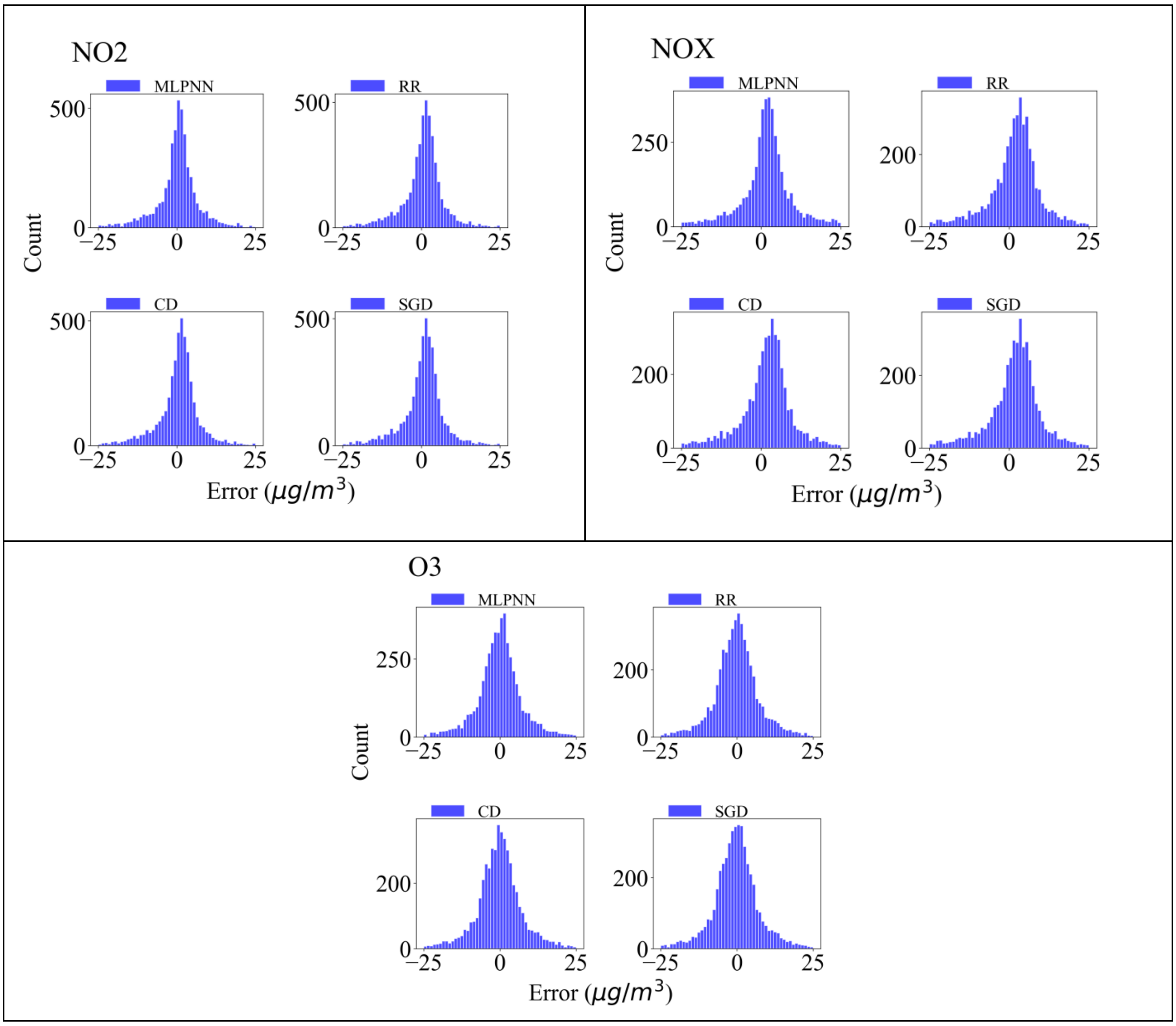

Figure 4 (cont.). Error histograms of the test sections of the machine learning models.

4. Conclusions

Air pollution is one of the most important environmental problems. Uncontrolled industrialization causes dramatic levels of air pollution and trigger environmental problems all over the world. Machine learning, which has been successfully applied in many fields nowadays, should be studied in the field of air pollution more. In this study, we performed modeling with various machine learning algorithms (multilayer perceptron neural network, stochastic gradient descent, ridge regression, cross decomposition) by considering temperature, relative humidity, wind, pressure and air pollutant measurements of the previous hour to estimate PM10, SO2, NO2, NOX pollutant concentrations.

According to the findings, MLPNN model is more successful than other methods for both train and test phases. Many machine learning studies stated the success of MLPNN method. Our study also confirms its success. MLPNN method is a nonlinear method and this structure gives it a clear advantage compared to other methods. Besides, O3 variable was better predicted. The correlation of O3 variable with the input parameters (0.558 with T, -0.5 with RH, 0.23 with W, -0.3 with PS) is higher than the correlations of other variables. It has been understood that the high correlation of the estimating variable with the input parameters increases the model performance in machine learning. From this finding, it was also understood that the O3 variable is more correlated with temperature, humidity, wind and pressure parameters than other variables (PM10, SO2, NO2, NOX). The error amounts of the methods are spread homogeneously over 0 and its surroundings in most of the variables. This finding indicates that the learning process shows efficient results and the models are not biased.

Only meteorological variables and the air quality variable of the previous hour were used as input parameters in this study. Anthropogenic pollutants have also a huge impact on air quality. More successful results can be obtained if anthropogenic pollutant sources are used as input parameters. However, all the models used in this study have yielded successful results. We especially recommend the MLPNN method to be used in air quality modeling studies due to its successful performance and fast predictions. In addition, we recommend that the input variables to be used while creating the model should be selected from the parameters that show high correlation with the output variable.

Acknowledgment

The authors would like to thank Turkish State Meteorological Service and Ministry of Environment, Urbanization and Climate Change to provide data used in this study.

References

[1] Abed Al Ahad M, Sullivan F, Demšar U, Melhem M, Kulu H. The effect of air-pollution and weather exposure on mortality and hospital admission and implications for further research: A systematic scoping review. PloS one 2020; 15(10): e0241415.
[2] Hrdlickova Z, Michalek J, Kolar M, Vesely M. Identification of factors affecting air pollution by dust aerosol PM10 in Brno City, Czech Republic. Atmos Environ 2008; 42(37): 8661–8673.
[3] Ei-Sharkawy MF, Zaki GR. Effect of meteorological factors on the daily average levels of particulate matter in the Eastern Province of Saudi Arabia: a cross-sectional study. J Sci Technol 2015; 5(1): 18–29.
[4] Oğuz, K. Nevşehir İlinde Hava Kalitesinin ve Meteorolojik Faktörlerin Hava Kirliliği Üzerine Etkilerinin İncelenmesi. Doğal Afetler ve Çevre Dergisi 2020; 6(2): 391-404.
[5] Qin YG, Yi C, Dong GL, Min JZ. Investigating the influence of meteorological factors on particulate matters: A case study based on path analysis. Energy & Environment 2019; 31(3): 1-13.
[6] Panda N, Osthus D, Srinivasan G, O'Malley D, Chau V, Oyen D, Godinez H. Mesoscale informed parameter estimation through machine learning: A case-study in fracture modeling. Journal of Computational Physics 2020; 420: 1-15.
[7] Alpaydin E. Introduction to Machine Learning. The MIT Press: Cambridge, MA, USA, 2010.
[8] Gagliardi RV, Andenna C. A Machine Learning Approach to Investigate the Surface Ozone Behavior. Atmosphere 2020; 11(11): 1173.
[9] Kothandaraman D, Praveena N et al. Intelligent Forecasting of Air Quality and Pollution Prediction Using Machine Learning, Adsorption Science & Technology 2022; (5086622), 15.
[10] Kumar K, Pande BP. Air pollution prediction with machine learning: a case study of Indian cities. Int. J. Environ. Sci. Technol. 2022; 1-16.
[11] Ünaldı S, Yalçın N. Hava Kirliliğinin Makine Öğrenmesi Tabanlı Tahmini: Başakşehir Örneği. Mühendislik Bilimleri ve Araştırmaları Dergisi 2022; 4(1): 35-44.
[12] Oğuz K, Pekin MA. Makine Öğrenme Algoritmaları ile PM10 Konsantrasyon Tahmini. Journal of Advanced Research in Natural and Applied Sciences 2022; 8(2): 201-213.
[13] Bekkar A, Hssina B, Douzi S, Douzi K. Air-pollution prediction in smart city, deep learning approach. J Big Data 2021; 8(1): 161.
[14] Gültepe YA Comparative Assessment on Air Pollution Estimation by Machine Learning Algorithms. European Journal of Science and Technology 2019; (16): 8-15.
[15] Irmak ME, Aydilek İB. Using Ensemble Regression Algorithms for Improving the Prediction Success of Air Quality Index. Academic Platform Journal of Engineering and Science 2019; 7(3): 507-514.
[16] Dobrea M et al. Machine Learning algorithms for air pollutants forecasting. 2020 IEEE 26th International Symposium for Design and Technology in Electronic Packaging (SIITME) 2020; 109-113.
[17] Lee M, Lin L, Chen CY et al. Forecasting Air Quality in Taiwan by Using Machine Learning. Sci Rep 10, 2020; 4153.
[18] Doreswamy KSH, Yogesh KM, Gad I. Forecasting air pollution particulate matter (PM2.5) using machine learning regression models. Procedia Computer Science 2020; 171(2020): 2057-2066.
[19] Liang YC, Maimury Y, Chen AHL, Juarez JRC. Machine learning-based prediction of air quality. Appl. Sci. 10, 2020; 9151.
[20] Haykin S. Neural Networks: A Comprehensive Foundation. Prentice Hall: USA, 1999.
[21] Al-Saif AM, Abdel-Sattar M, Aboukarima AM, Eshra DH. Application of a multilayer perceptron artificial neural network for identification of peach cultivars based on physical characteristics. PeerJ 2021; 9(e11529).
[22] Gardner MW, Dorling SR, Artificial neural networks (the multilayer perceptron) —a review of applications in the atmospheric sciences. Atmospheric environment 1998; 32(14-15): 2627-2636.
[23] Latz J. Analysis of stochastic gradient descent in continuous time. Stat Comput. 2021; 31(39): 1-25.
[24] Orange Data Mining, 2022, Stochastic Gradient Descent, available at: https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/widgets/model/stochasticgradient.html (accessed: 1 June 2022).
[25] UFLDL Tutorial, 2022, Optimization: Stochastic Gradient Descent, available at: http://deeplearning.stanford.edu/tutorial/supervised/OptimizationStochasticGradientDescent/ (accessed: 15 June 2022).

[26] Ruder, 2016, An overview of gradient descent optimization algorithms, available at: https://ruder.io/optimizing-gradient-descent/#:~:text=Stochastic%20gradient%20descent,-Stochastic%20gradient%20descent&text=%CE%B8%3D%CE%B8%E2%88%92%CE%B7%E2%8B%85%E2%88%87%CE%B8J(%CE%B8%3B,%3B%20y%20(%20i%20)%20)%20 (accessed: 30 June 2022).
[27] Topal M, Eyduran E, Yağanoğlu AM, Sönmez A, Keskin S. Çoklu Doğrusal Bağlantı Durumunda Ridge ve Temel Bileşenler Regresyon Analiz Yöntemlerinin Kullanımı. Atatürk Üniversitesi Ziraat Fakültesi Dergisi 2013; 41(1): 53-57.
[28] Scikitlearn, 2022, Cross decomposition, available at: https://scikit-learn.org/stable/modules/cross_decomposition.html (accessed: 5 July 2022).
[29] ML-science, 2022, Cross decomposition, available at: https://www.ml-science.com/cross-decomposition (accessed: 5 July 2022).
[30] Singh, 2020, Understanding Data Preprocessing, available at: https://towardsdatascience.com/data-preprocessing-e2b0bed4c7fb (accessed: 28 July 2022).
[31] Baheti, 2022, A Simple Guide to Data Preprocessing in Machine Learning, available at: https://www.v7labs.com/blog/data-preprocessing-guide (accessed: 3 August 2022).
[32] DataTechNotes, 2019, Regression Model Accuracy (MAE, MSE, RMSE, R-squared) Check in R, available at: https://www.datatechnotes.com/2019/02/regression-model-accuracy-mae-mse-rmse.html (accessed: 10 August 2022).