# Time series modeling of air pollution and its association with season and climate variables in istanbul, turkey

**hazem al-najjar** ( hazem_najjar@yahoo.com )
gelisim unversitesi https://orcid.org/0000-0002-6143-2734

**Nadia Al-Rousan**
Gelisim Universitesi

**Ismail A. Elhaty**
Gelisim Universitesi

**Research**

**Keywords:** Air quality forecast, feature selection, artificial neural network, NARX, season effect

**Posted Date:** June 30th, 2020

**DOI:** https://doi.org/10.21203/rs.3.rs-27820/v4

# Time series modeling of air pollution and its association with season and climate variables in istanbul, turkey

Hazem Al-Najjar1*, Nadia AL-Rousan1, Ismail A. Elhaty2

[1]Department of Computer Engineering, Architecture and computer engineering college, Istanbul Gelishm universitesi, Istanbul, Turkey

[2]Department of Nutrition and Dietetics, School of Health Sciences, Istanbul Gelisim University, Istanbul, Turkey

* Correspondence: hazem_najjar@yahoo.com

**Abstract**

Air pollution depends on seasons, wind speed, temperature, wind direction and air pressure. The effect of different seasons on air pollution is not fully addressed in the reported works. The current study investigated the impact of season on air pollutants including $SO_2$, $PM_{10}$, NO, $NO_X$, and $O_3$ using Nonlinear AutoregRessive network with eXogenous inputs (NARX) method. In the applied methodology, a feature selection was used with each pollutant to find the most important season(s). Afterward, six models are designed based on the feature selection to show the impact of seasons in finding the concentration of pollutants. A case study is conducted on Esenyurt which is one of the most populated and industrialized places in Istanbul to validate the proposed framework. The performance of using all of the designed models with different pollutants showed that using season effect led to improving the performance of predictor and generating high R2 and low error functions.

**Keywords:** Air quality forecast, feature selection, artificial neural network, NARX, season effect.

## 1. Introduction

Air pollution increases with the increasing of industrialization and urbanization.

Public health and economic development in metropolitan cities are affected by air pollution. Air pollutants' levels and types vary from place to another based on the sources of air pollutants such as cars, power plants, oil refineries, industrial facilities, and factories [1-3]. Air quality is monitored using the most common air pollutants (indicators) including $SO_2$, $NO_2$, $CO_2$, $O_3$, NO, NOx, $PM_{2.5}$, and $PM_{10}$ [4]. These indicators can be found in different levels in the ambient air, unfortunately, exceeding the concentration levels of these pollutants will threaten the human health and may cause many serious problems namely long term and short-term problems. One of the most hazardous events produced by air pollutants is the great smog of 1952 in London, which continued for five days and killed 4000. Monitoring and detecting of the concentration of air pollutants can help decision markers to take the right decisions for the current and future plans. An enormous number of researches have been conducted to forecast the pollutants' concentrations and to understand the most suitable way to evaluate the air quality.

Cogliani [5] studied the relationship between metrological variables and daily pollution index in three Italian cities using linear multiple partial correlation analysis. The results of forecasting the concentration of pollutants showed high evaluation error

and the methods can be used in the surrounded areas of the observation stations while forecasting the pollutants in places that far away from the station is inapplicable.

Zhu et al., [6] investigated the relationship between low respiratory diseases and monthly average concentration of $SO_2$, $NO_2$ and $PM_{10}$ by considering the effect of seasons especially winter season. Besides, the study tried to estimate the dataset covered period from January 2001 to December 2005 and was collected from Xigu District's hospitals. The results found a relationship between short-term pollution and low respiratory diseases and found strong relation with winter season on low respiratory diseases. Feng et al.[7] proposed a novel daily $PM_{2.5}$ forecasting to improve the performance of artificial neural network by using air mass trajectory analysis and wavelet transformation. The dataset is collected from 13 stations from different locations in China including Beijing, Tianjin, and Hebei provinces. The results showed that the new hybrid method can reduce the root mean square error up to 40%. Furthermore, the results indicated that the proposed model is efficient to be applied in different countries.

Fortelli et al., (2016) investigated the relationship between local metrological variables and PM10 in Naple, IItaly. Afterward, metrological variables are used to

forecast $PM_{10}$ for couple of days. The results found a relationship between air pollution crises and geopotential heights. The prediction model showed a high correlation between $PM_{10}$ observations and the predicted values with 0.8 as a correlation coefficient. Alimissis et al. [9] evaluated two interpolation prediction models including Artificial Neural Networks and Multiple Linear Regression, to predict the quality of air in Athens, Greece. The quality of air is majored using five pollutants including Nitrogen dioxide, Nitrogen monoxide, Ozone, Carbon monoxide and Sulphur dioxide. The results showed that artificial neural networks are found in most cases to be significantly superior, especially where the air quality network density is limited.

Yu et al. [10] proposed a fast forecasting method to estimate $PM_{2.5}$ concentrations in six cities including Baoding, Beijing, Dezhou, Shijiazhuang, Tianjin, and Tangshan, which located in the north of China. The forecasting method is based on source–receptor relationship modeling with backward Lagrangian stochastic particle dispersion model and emission inventory inversion. The forecasting method is built using a dataset collected in 2015, where another dataset collected in 2016 is used for forecasting purposes. The results showed that applying the new techniques can achieve better results and high correlation coefficients compared with non-optimized model. Wang et

al. [11] developed forecasting model to predict an interval PM2.5 concentration using meteorological factors based on multilayer perceptron. To select the most important input variables from the list of variables an interval grey incidence analysis is adopted. In addition, the dataset is collected from three different locations in Beijing, China. The results indicated that the developed model is accurate and stabile than other models in the literatures. Liu et al. [12] developed three-stage hybrid algorithm based on neural network to forecast $PM_{2.5}$. The dataset is collected from four cities in China, including Beijing, Tianjin, Shijiazhuang, and Tangshan. The results indicated that the accuracy of the proposed model is efficient compared than conventional methods in the field.

Lu [13] investigated the relationship between $PM_{2.5}$ and $PM_{10}$ in different locations within Hong Kong province. Based on a relationship between two pollutants, a predictive model is built to estimate the concentration of $PM_{2.5}$ using $PM_{10}$ by using three prediction strategies including local, remote and mixed strategies. The results showed that the three used strategies are able to estimate the missing or unmonitored values using the surrounded stations. Zhu et al. [14] proposed two-step-hybrid prediction model to estimate concentrations of $NO_2$ and $SO_2$ pollutants in four cities in Central China region. The model is divided into three steps starting with finding high-

frequency and low-frequency sequences followed by applying Support Vector Regression based on combining the Cuckoo Search algorithm and Grey Wolf Optimizer algorithm and finally, forecasting data of low frequency and high frequency. The results indicated that different hybrid models should be used to efficiently predict high and low frequencies.

Catalano et al.[15] studied the relationship between the hourly mean concentration of $NO_2$ and the factors that reflecting the $NO_2$ level (i.e. traffic and weather conditions). Both neural network and Autoregressive Integrated Moving Average with Explanatory (ARIMAX) forecasting methods were used to predict the pollution peaks along with using a combination of these models to forecast the air quality. The results revealed that ARIMAX outperformed neural network in pollution peak forecasting, while neural network could better represent the realistic pollution's concentration association with wind attribute. Integrating both forecasting models could efficiently predict extreme pollution concentrations than using both models separately.

Durao et al. [16] forecasted the concentration level of $O_3$ using a combination of metrological and air quality and industrial emissions data for Sines Portuguese region, Portugal. Two forecasting models including Multi-Layer Perceptron (MLP) and

Classification and Regression Trees (CART) were used to predict $O_3$ concentration. The results revealed that MLP successfully predicted $O_3$ concentration within 24 hours ahead.

Corani et al. [17] designed multi-label classifier to predict multiple air pollution variables. Bayesian networks were used as a learning technique to predict the level of $PM_{2.5}$ and ozone in different three studies, and to compare the results with other classifiers. It is found that using multi-label classifier performed better than other independent classifiers.

Shi et al. [18] investigated the most metrological variables that have direct or indirect responsibility in serving $PM_{2.5}$ in Central East China. The results showed increasing in $PM_{2.5}$ concentration comes as a normal changing in wind direction from south to north, besides increasing in important meteorological factors (i.e., large-scale subsidence, and radiative cooling). Mo et al. [19] investigated the growth of surface ozone concentration and its effect on health. The study tried to develop a new model based on combining different machine learning techniques together as one model. The dataset is collected from four stations in China, then the dataset is divided into training dataset (i.e., 1 May 2014 to 31 May 2017), and testing dataset (i.e., 1 June 2017 to 30

May 2018). The results showed that the proposed model is accurate and stable, besides the model can be used in different locations. Researchers have investigated the capability of predicting the air quality using different prediction models including Artificial intelligent [20-23], Autoregression models [24], and other hybrid models as discussed in [25-27].

However, pollutants can make acid rain that has a harmful effect on plants, buildings, monuments, groundwater, soil composition and living organisms inside seas, ponds and rivers. Thus, this study is important, hence, it gives an opportunity for decision-makers to take into account the level of air pollution when developing future plans, especially Turkey is working to increase exports, attract investments and expand construction of factories. Besides, few researchers have focused on finding a relationship between increasing or decreasing pollutants including $SO_2$, NO, $NO_2$, NOX, $O_3$ and $PM_{10}$ in air and seasons. Besides, researchers have tried to test different combinations between input variables without considering any scientific or mathematical reasons. Therefore, this article comes as important research for both international and national researchers. Our contributions in this research are explained as follow:

- lack of studies considers the impact of season in understanding the trend of pollutants in air.
- lack of studies uses a feature selection with seasons and pollutants to determine the most effective season(s) on each pollutant.
- The study presents a new methodology that can be followed to improve pollutants forecasting including $SO_2$, NO, $NO_2$, NOX, $O_3$ and $PM_{10}$. Besides, it explains the main behavior between each independent variable and each pollutant to understand the trend of each independent variables on each pollutant.
- The study presents the most effective season(s) on the most important pollutants in Turkey. This can help different researchers to follow the same steps to explain the movement of pollutants in different countries.
- The study proves the capability of Nonlinear AutoregRessive network with eXogenous inputs (NARX) model in forecasting different pollutants in Turkey, which gives a hint to many researchers in the field to consider the model without considering other models. This can save researchers' time and effort.

In order to improve the capability of NARX model in forecasting pollutants, it is

therefore beneficial to develop a new methodology based on previous studies and feature selection of the most effective season(s) on each pollutant. Afterward, new models will be designed to cover different scenarios. Hence, dataset from one of the most populated sites in Istanbul (Esenyurt) is considered between 2015 and 2019, to validate the used methodology. Furthermore, to compare between different models a determination coefficient with different error functions are used to find the best and most suitable model on each pollutant.

## 2. Research Methodology

To build a forecasting model for gases in air, a feature selection method between seasons' variables and each pollutant is used as shown in the following subsections.

### *2.1 Data collection and analysis*

The case study was conducted in Esenyurt, one of the most polluted sites in Istanbul, Turkey. The dataset covers 5 years and the total collected hours for all the collected data are 37645. To build a forecaster model, the dataset is divided into two datasets including training and testing. The training dataset is used to train the forecast model about the historical information of all the gases in ambient air, where the testing dataset is used to check the capability of the trained forecaster to forecast

the future data. The dataset is divided based on the year of the collected information. In which, from 2015 to 2018 is used for training and 2019 is used for testing. To implement the effect of seasons in the collected dataset, four dummy variables for each season are considered. The dummy variable assigns one for starting to ending date and zero otherwise. The statistical description of the training and testing datasets are shown in Tables 1 and 2, respectively. In which, Time, Time, Day, Month, Year, Temperature (°C), Wind direction(Degree), Wind speed ($ms^{-1}$), Relative humidity(%), Air pressure(mbar) and seasons (based on the results of feature selection) are used as input variables and one of the pollutants including PM10 ($\mu gm^{-3}$), SO2 ($\mu gm^{-3}$), NO ($\mu gm^{-3}$), NO2 ($\mu gm^{-3}$), NOX ($\mu g\ m^{-3}$), O3 ($\mu g\ m^{-3}$) as output variable.

*2.2 Design seasons' models based on feature selection*

After collecting the data from the source, data processing should be conducted. Since, the collected data has no missing value, no outlier values, therefore, sample modelling and design should be the next step to validate the relationship between independent and dependent variables, besides, to find the capability of forecasting air pollution. The study uses six type of air pollutants including PM10, SO2, NO, NO2, NOX and O3, besides Turkey has four seasons, so the total number of models that can

be generated is equal to 144 models (24 “seasons’ combinations” X 6 “models”). Trying all the combinations is time consuming and a huge number of results will be generated.

To minimize the total number of models, a subset attributed evaluator with greedy stepwise search method is used as a feature selection method. The target of using a feature selection is to find the most important seasons(s) that connected with each studied pollutant. The process starts by selecting one of the pollutant gases as a target and all the seasons as input where the rest of metrological variables are used without using feature selection based on the previous studies. The most important season(s) for each pollutant is /are considered. The results of all the models are used to create models. The created models are used to design forecasting models for all the pollutants as shown in Figure 1.

In our dataset, a long time hourly (i.e., 30840 hourly readings) data is used to find the most related seasons that connected with each pollutant, then there will be 6806 time series samples for forecasting purposes. The most important seasons and the designed models will be discussed in the results section.

*2.3 Air pollution gases forecasting based on narx*

After building models using a feature selection and metrological variables. The NARX forecasting model will be used to forecast different pollutants. The first step in building a forecasting model is to train the NARX model using a training dataset between 2015 and 2018. The NARX will be ran many times until the best model is achieved. The best model that has highest determination coefficient and minimum error are considered for each pollutant. The testing dataset that is not used in training phase between Jan 2019 to Dec 2019 is used to forecast the performance of the trained NARX. The previous two steps are repeated to find the most appropriate weights for NARX models. The best model for each pollutant is used to calculate the performance metrics. For each pollutant, all the generated models are considered and the best model that achieved the best performance metrics are denoted as the best model based on testing dataset (not training dataset). The generated results from NARX and feature selection method are compared together to draw conclusion about the pollutant type and season(s).

Figure 2 shows the NARX network with one output (pollutant) denoted as y, u inputs and b as bias, the process of NARX starts by creating serial parallel architecture

(opened loop network), then parallel architecture (closed loop network). The target of creating opened and closed network is to improve the forecasting process and to increase the efficiency of the network by using the previous direct data. To forecast the trained NARX model, the first two inputs with training data are used by forecaster to adjust the predicted values.

To sum up all the used methodology in this study a flow chart in Figure 3 is considered.

*2.4 Performance analysis*

After building pollutant time series model for each pollutant using a training dataset, a testing dataset is used to determine the most effective and suitable model for each pollutant. The results of training and testing dataset using different generated models are evaluated using four metrics including determination coefficient (R2), mean square error (MSE), mean absolute error (MAE) and root mean square error (RMSE) .The calculations of all the performance metrics are presented as follows:

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

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N}(y_i - \hat{y}_i)^2} \quad (2)$$

$$MSE = \frac{1}{N}\sum_{i=0}^{N}(y_i - \hat{y}_i)^2 \quad (3)$$

$$MAE = \frac{1}{N}\sum_{i=0}^{N}|y_i - \hat{y}_i| \quad (4)$$

The higher $R^2$ and the low error function are considered as a best model, robust and accurate, besides the seasons of the best results are pollutants considered the most effected results that can explain the concentrations of the pollutant.

## 3. Results, Analysis and Discussion

### *3.1.Study area*

This study was conducted in Esenyurt, a district of Istanbul Province that belongs to a metropolitan municipality of the city. Esenyurt is situated on the European part of the city and km far from Marmara Sea. it's surrounded by Avgelar (Avcılar) district and Lake Kochokmaje (Küçükçekmece four) on the eastern part; Buyukchakmaje (Büyükçekmece) on the western region; Bashakshaheer (Başakşehir), Arna'outkoy (Arnavutköy) and Trans-European Motorway (TEM) on the northern region; and Bailekdazou (Beylikdüzü) and E-5 motorway on the southern portion. Its municipality exists since 1989 and its total area is 43.0 km2. The constructions and population are growing rapidly during the last decade. The population estimate is 954579 in 2019 with population density 22.20/km2 [28].

*3.2.Relationship between input variables and air pollutants*

To show the impact of each independent variable (input) on the concentration of air pollutants including $PM_{10}$, $SO_2$, NO, $NO_X$, $NO_2$ and $O_3$, correlation coefficients are shown in Table 3. The most effective date variable on the pollutants is year variable, which shows a decrease in a concentration of all pollutants per year (except $SO_2$). Decreasing the concentration of air pollutants (except $SO_2$) with time may be attributed to the high percentage of modern cars in Istanbul which emit less pollutants comparing with the old ones. In addition, such decreasing may be attributed to the rules issued by Turkish government to control the emitted pollutants from cars and factories. The growing industry in Turkey, in particular Istanbul, may interpret the increasing concentration of $SO_2$ with time. In general, sulfur dioxide releases from fossil fuel burning power stations, industrial processes such as extracting metal from ore and the burning of fuels with a high sulfur content by locomotives, large ships and non-road equipment [29].

In Esenyut, the temperature ranged between -5 °C in winter and 38 °C in summer during the study time period. Temperature has a negative correlation with all pollutants except ozone which showed a positive correlation. These results are almost consistent

with correlation between air pollutants and winter and summer. In this study, the negative correlation could be attributed to the decrease in usage of the domestic heating in summer. In case of ozone, it's mainly formed by a photochemical reaction consequently the more intense the solar radiation (temperature), the more $O_3$ concentrations [30]. Wind speed and direction vary widely in Esenyurt. The average wind speed is 2.225 m/s during the study while wind direction at the most is NNE. Wind direction has a positive low correlation with $PM_{10}$, $SO_2$, NO, $NO_2$ and $NO_X$, where $O_3$ has negative intermediate relationship with wind direction. Air pollutants in Esenyurt release mainly from local sources such as local industries, traffic and domestic heating. Wind transport air pollutants from Esenyurt to the surrounding places or the inverse. Positive correlation indicates that wind may transport air pollutants towards the monitoring sits. Ozone is unstable molecule and it may be affected by the air movement. In general, it has been found that ozone concentration is higher in the places surrounding the city than inside the city center [31].

In general, as the humidity increases, the concentration of air pollutants decreases because of the washing effect [32]. The interference between the relative humidity and other parameters such as temperature explained the variety of the correlation between

relative humidity and the concentration of air pollutants indicating that it's difficult to analyze meteorological variables in isolation [33]. Air pressure showed a positive correlation with all gases except with $O_3$. The positive results agree with the reported works [20]. The negative effect of pressure on ozone may attributed to the depletion of ozone under pressure.

### *3.3.Designing forecasting models based on season's variables*

Before building a forecasting model and to avoid a lot of combinations between seasons, a feature selection between the season variables and each pollutant gas is considered. Table 4 represents the best effected season on each gas. Based on Table 4 and after combining the most effective seasons, 6 models could be generated as represented in Table 5. Models 1 and 6 represent no season effect and all the season effect, respectively. Model 3 represents spring and Autumn, respectively, where Models 4 and 5 represent summer-winter and Autumn-spring seasons respectively. Therefore, instead of testing 144 models only 6 models are considered in this research. Each gas from the pollution list will use all the models and the best model for each gas are recorded separately as shown in the next section.

### *3.4.Air pollution gases forecasting results*

Before building a forecasting model and to avoid a lot of combinations between seasons, a feature selection between the season variables and each pollutant gas is considered. Table 4 represents the best effected season on each gas. Based on Table 4 and after combining the most effective seasons, 6 models could be generated as represented in Table 5. Models 1 and 6 represent no season effect and all the season effect, respectively. Model 3 represents spring and Autumn, respectively, where Models 4 and 5 represent summer-winter and Autumn-spring seasons respectively. Therefore, instead of testing 144 models only 6 models are considered in this research. Each gas from the pollution list will use all the models and the best model for each gas are recorded separately as shown in the next section.

*3.5.Air Pollution gases forecasting results*

The correlation results in Table 3 showed that autumn and spring seasons have a weak correlation with NO, which means that these two seasons may have a strong nonlinear relationship. The results of training and testing NO forecasting are shown in Table 6. The results showed that the best model for training is Model 5 with $R^2$, MSE, MAE and RMSE equal to 0.95, 508, 10. 23, respectively. While Model 6 showed that the best results for testing are achieved with $R^2$, MSE, MAE and RMSE equal to 0.901,

431, 12, 21, respectively. The results indicate that Autumn and Spring seasons are efficient for training the forecasting model, where for testing, Model 6 showed better performance.

To show the fluctuation of errors in the designed models Figure 4 showed that forecasting errors of NO gas in 2019. All models showed higher errors on February and March, where the results of months showed lower errors. The results indicated that using all the seasons variables with metrological variables and time to forecast NO gas is efficient. This indicate that NO gas has a relationship with the seasons, and it change every season.

The results of forecasting $NO_2$ gas are shown in Figure 5 and Table 7. NO forecasting results showed that Model 5 and Model 3 are the best models for training and testing respectively. The best $R^2$, MSE, MAE and RMSE are 0.959, 24, 3 and 5, respectively for training, where the best $R^2$, MSE, MAE and RMSE for testing are 0.973,37, 4 and 6, respectively. The forecasting error of $NO_2$ for 2019 is shown in Figure 7, the results of errors showed that February and March have the highest errors using all the models, which indicated that using the metrological variables with Autumn season could achieve the best results for training and testing datasets. The results

revealed that using the season effect can improve the $NO_2$ forecasting.

For $NO_X$ training results, Model 5 showed the best performance with $R^2$, MSE, MAE and RMSE are 0.951,1332, 17 and 37, respectively, where for testing results, Model 1 showed the best performance in $R^2$, MSE, MAE and RMSE with values equal to 0.951, 1142, 16 and 34, respectively as shown in Table 8. Figure 6 showed the forecasting errors of $NO_X$ gas using Models 1 to 6. The forecasting results showed that February and March have the highest errors. The results indicated that no season effect is the best model to describe the behavior of $NO_X$.

The forecasting of $O_3$ is shown in Table 9 and Figure 7, the results showed that Model 5 is the best model for training and testing datasets. The results came in line with a correlation analysis in Table 3. The analysis shows that spring and Autumn are the most effective variables on $O_3$ gas. Model 5 shows the lowest errors compared with other models as presented in Figure 9. The results revealed that using a season effect has a great benefit in training all the models, as well as testing dataset.

$PM_{10}$ forecasting results presented in Table 10 and Figure 8, Models 4 and 2 showed best performance for training and testing, respectively. The $PM_{10}$ results showed that

spring season has the best results compared with other models. The results indicate that using metrological variables with spring could retrieve the best performance metrics. $SO_2$ forecasting has very bad results compared with NO, $NO_2$, $NO_X$, $O_3$ and $PM_{10}$, if Models 2 to 5 are used as shown in Table 11 and Figure 9. Training and testing datasets showed that Model 6 is the best model in improving the performance metrics. The testing performance metrics are 0.888, 16, 3, 4 for $R^2$, MSE, MAE and RMSE, respectively. The results indicate that using all the seasons could improve the performance of $SO_2$ forecasting.

*3.6.Discussion*

Autumn, spring, summer and winter seasons have strong connections in increasing and decreasing the concentrations of $PM_{10}$, $SO_2$, NO, $NO_X$, $NO_2$ and $O_3$ in air as shown in Table 12. Both training and testing datasets have different behavior in forecasting. 50% of overall cases showed that Autumn and spring seasons are the dominant model in improving the performance metrics, while, without season effect model showed improvement for 20% of overall cases. In addition, summer-winter model, spring and autumn models showed 30% improvement in both datasets. Combining the results of training and testing datasets and eliminating the best results from training dataset. The

best performance of $PM_{10}$ is shown with metrological variable and spring season. $SO_2$ and NO behaved very good with metrological variables with all seasons. $NO_2$ and $O_3$ showed a good concentration estimation with spring season and Autumn-spring seasons, respectively, where metrological variables are used as input with the seasons. $NO_X$ has no improvement with any season and only using metrological variables are affected in $NO_X$ forecasting. The results of $PM_{10}$ and $NO_2$ come in line with the feature selection method, where $SO_2$, NO, $NO_X$ and $O_3$ are not in line with feature selection. This indicated that feature selection is appropriate for $PM_{10}$ and $NO_2$ only.

## 4. Conclusion

This work proposes a methodology of using feature selection to find the most effective season(s) on each air pollutant including $PM_{10}$, $SO_2$, NO, $NO_2$, $NO_X$ and $O_3$. Based on the results of feature selection, 6 models are proposed, NARX method is used to build a forecasting model using hourly training dataset between 2015 and 2018, where hourly testing dataset is used for validating the developed models. The dataset is adopted from Esenyurt, Istanbul. The main finding of this study can be summarized as follow:

- This paper is one of the few studies that considered the effect of season through feature selection method.
- It found that Autumn and spring season are the most effective seasons on the concentration of $NO_2$ and O3 gases, where spring season and autumn season are effective on $PM_{10}$ and $NO_X$, respectively. $SO_2$ and NO gases, on the other hand, have impact with all seasons.
- NARX model is the most effective and accurate model to build prediction models for different gases' concentration.
- All the prediction models that used to construct different pollutants showed very good results except for $SO_2$, the results were bad compared with other models.

This paper considered only one of the most polluted sites in Istanbul, Esenyurt and further sites needed to be conducted to validate different sites and different gases. As a next step in this study, different locations with extra metrological variables will be used to build more accurate and robust models for different gases in air.

## Declarations

### Availability of data and materials

All the used data is available publicly in State Institute of Statistics, Republic of

Turkey …

## Competing interests

The authors declare they have no competing interests.


## Funding

Not applicable.


## Authors' contributions

H. Al-Najjar and N AL-Rousan designed the research model and analysis. The first version is written by H. Al-Najjar and N AL-Rousan. I. Elhaty modified the article, checked and modified the scientific information in the article. All the authors read and approved the final manuscript.

## References


1. Koop G, McKitrick R, Tole L. Air pollution, economic activity and respiratory illness: evidence from Canadian cities, 1974–1994. Environmental Modelling & Software. 2019;25(7): 873-85.
2. Wyrwa A. An optimization platform for Poland's power sector considering air pollution and health effects. Environmental Modelling & Software. 2015;74:227-37.

3. Sacks D, Lloyd M, Zhu Y, Anderton J, Jang, Hubbell B, Fann N. The Environmental Benefits Mapping and Analysis Program–Community Edition (BenMAP–CE): A tool to estimate the health and economic benefits of reducing air pollution. Environmental Modelling & Software. 2018;104: 118-29.
4. Araujo N, Belotti T, Alves A, de Souza Tadano Y, Siqueira H. Ensemble method based on Artificial Neural Networks to estimate air pollution health risks. Environmental Modelling & Software. 2020;123, 104567.
5. Cogliani E. Air pollution forecast in cities by an air pollution index highly correlated with meteorological variables. Atmospheric Environment. 2001;35: 2871-7.
6. Zhu W, Wang J, Zhang, W, Sun D. Short-term effects of air pollution on lower respiratory diseases and forecasting by the group method of data handling. Atmospheric environment 2012;51:29-38.
7. Feng X, Li Q, Zhu Y, Hou J, Jin L, Wang J. Artificial neural networks forecasting of PM2. 5 pollution using air mass trajectory based geographic model and wavelet transformation. Atmospheric Environment. 2015;107:118-28.

8. Fortelli A, Scafetta N, Mazzarella A. Influence of synoptic and local atmospheric patterns on PM10 air pollution levels: a model application to Naples (Italy). Atmospheric environment. 2016;143:218-28.

9. Alimissis A, Philippopoulos K, Tzanis G, Deligiorgi D. Spatial estimation of urban air pollution with the use of artificial neural network models. Atmospheric Environment. 2018;191:205-13.

10. Yu M, Cai X, Song Y, Wang X. A fast forecasting method for PM2. 5 concentrations based on footprint modeling and emission optimization. Atmospheric Environment. 2019;219:117013.

11. Wang Z, Chen L, Ding Z, Chen H. An enhanced interval PM2. 5 concentration forecasting model based on BEMD and MLPI with influencing factors. Atmospheric Environment. 2019;117200.

12. Liu H, Chen C. Prediction of outdoor PM2. 5 concentrations based on a three-stage hybrid neural network model. Atmospheric Pollution Research, 2020;11:469-481.

13. Lu W. Z. Comparison of three prediction strategies within PM2. 5 and PM10 monitoring networks. Atmospheric Pollution Research, 2020;11:590-7.

14. Zhu S, Qiu X, Yin Y, Fang M, Liu X, Zhao X, Shi Y. Two-step-hybrid model based on data preprocessing and intelligent optimization algorithms (CS and GWO) for NO2 and SO2 forecasting. Atmospheric Pollution Research. 2019;10:1326-35.

15. Catalano M, Galatioto F, Bell M, Namdeo A, Bergantino S. Improving the prediction of air pollution peak episodes generated by urban transport networks. Environmental science & policy. 2016;60:69-83.

16. Durao M, Mendes T, Pereira J. Forecasting O3 levels in industrial area surroundings up to 24 h in advance, combining classification trees and MLP models. Atmospheric Pollution Research. 2016;7:961-970.

17. Corani G, Scanagatta M. Air pollution prediction via multi-label classification. Environmental modelling & software. 2016;80:259-264.

18. Shi C, Nduka C, Yang Y, Huang Y, Yao R, Zhang H, Yim L. Characteristics and meteorological mechanisms of transboundary air pollution in a persistent heavy PM2. 5 pollution episode in Central-East China. Atmospheric Environment. 2020; 223: 117239.

19. Mo Y, Li Q, Karimian H, Fang S, Tang B, Chen G, Sachdeva S. A novel framework for daily forecasting of ozone mass concentrations based on cycle reservoir with regular jumps neural networks. Atmospheric Environment. 2020, 220, 117072.

20. Dursun S, Kunt F, Taylan O. Modelling sulphur dioxide levels of Konya city using artificial intelligent related to ozone, nitrogen dioxide and meteorological factors. International journal of environmental science and technology. 2015;12: 3915-28.

21. Bai Y, Li Y, Wang X, Xie J, Li C. Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions. Atmospheric pollution research. 2016;7:557-66.

22. Biancofiore F, Busilacchio M, Verdecchia M, Tomassetti B, Aruffo E, Bianco S, ... Di Carlo P. Recursive neural network model for analysis and forecast of PM10 and PM2.5. Atmospheric Pollution Research 2017;8: 652-9.

23. Dotse Q, Petra I, Dagar L, De Silva C. Application of computational intelligence techniques to forecast daily PM10 exceedances in Brunei Darussalam. Atmospheric Pollution Research. 2018;9: 358-368.

24. Wang P, Zhang H, Qin Z, Zhang G. A novel hybrid-Garch model based on ARIMA and SVM for PM2. 5 concentrations forecasting. Atmospheric Pollution Research. 2017;8:850-60.

25. Carbajal-Hernández J, Sánchez-Fernández P, Carrasco-Ochoa A, Martínez-Trinidad F. Assessment and prediction of air quality using fuzzy logic and autoregressive models. Atmospheric Environment. 2012;60:37-50.

26. Franceschi F, Cobo M, Figueredo M. Discovering relationships and forecasting PM10 and PM2. 5 concentrations in Bogotá, Colombia, using artificial neural networks, principal component analysis, and k-means clustering. Atmospheric Pollution Research. 2018;9:912-22.

27. Yang H, Tao W, Liu Y, Qiu M, Liu J, Jiang K, ... Tao S. The contribution of the Beijing, Tianjin and Hebei region's iron and steel industry to local air pollution in winter. Environmental pollution. 2019;245:1095-1106.

28. State Institute of Statistics, Republic of Turkey Retrieved from http://www.turkstat.gov.tr/Start.do in March 9, 2020

29. Perez P. Prediction of sulfur dioxide concentrations at a site near downtown Santiago, Chile. Atmospheric Environment. 2001;35:4929-35.

30. Grosjean E, Grosjean D. Formation of ozone in urban air by photochemical oxidation of hydrocarbons: Captive air experiments in Porto Alegre, RS. Journal of the Brazilian Chemical Society 1998;9:131-143.

31. Salem A, Soliman A, El-Haty A. Determination of nitrogen dioxide, sulfur dioxide, ozone, and ammonia in ambient air using the passive sampling method associated with ion chromatographic and potentiometric analyses. Air Quality, Atmosphere & Health. 2009;2:133-145.

32. Kwak Y, Ko J, Lee S, Joh H. Identifying the correlation between rainfall, traffic flow performance and air pollution concentration in Seoul using a path analysis. Transportation research procedia. 2017;25:3552-63.

33. Dursun S, Kunt F, Taylan O. Modelling sulphur dioxide levels of Konya city using artificial intelligent related to ozone, nitrogen dioxide and meteorological factors. International journal of environmental science and technology. 2015;12: 3915-28.

**Table 1:** Statistical description of training data

| Variable | Mean | Err | Stan Dev. | Variance | Min | Max |
|---|---|---|---|---|---|---|
| Time | 0.482 | .002 | 0.286 | 0.082 | .000 | .958 |
| Day | 15.734 | .050 | 8.827 | 78 | 1.00 | 31 |
| Month | 6.476 | .020 | 3.436 | 12 | 1 | 12 |
| Year | 2016 | 0 | 1.099 | 1 | 2015 | 2018 |
| Temperature | 16.244 | .045 | 7.886 | 62 | -5 | 38 |
| Wind direction | 195.34 | .551 | 96.682 | 9347 | 0 | 360 |
| Wind speed ($ms^{-1}$) | 2.225 | .006 | 1.038 | 1 | 0 | 7 |
| Relative humidity(%) | 73.645 | .093 | 16.336 | 267 | 0 | 100 |
| Air pressure (mbar) | 1011.94 | 0.06 | 10.64 | 113 | 0 | 1036 |
| Summer | 0.257 | 0.002 | 0.437 | 0 | 0 | 1 |
| Winter | 0.243 | 0.002 | 0.429 | 0 | 0 | 1 |
| Spring | 0.259 | 0.002 | 0.438 | 0 | 0 | 1 |
| Autumn | 0.241 | 0.002 | 0.428 | 0 | 0 | 1 |
| PM10 ($\mu gm^{-3}$) | 81 | 0 | 78 | 6013 | 0 | 985 |
| SO2 ($\mu gm^{-3}$) | 6 | 0 | 7 | 46 | 0 | 220 |
| NO ($\mu gm^{-3}$) | 43 | 0 | 71 | 5047 | 0 | 880 |
| NO2 ($\mu gm^{-3}$) | 24 | 0 | 17 | 281 | 0 | 130 |
| NOX ($\mu gm^{-3}$) | 90 | 1 | 116 | 13390 | 0 | 1338 |
| O3 ($\mu gm^{-3}$) | 39 | 0 | 26 | 692 | 0 | 866 |

**Table 2:** Statistical description of testing data

| Variable | Mean | Error | S.Dev. | Variance | Min | Max |
|---|---|---|---|---|---|---|
| Time | 0.486 | 0.004 | 0.289 | 0.083 | 0 | 1 |
| Day | 15.634 | 0.104 | 8.602 | 73.999 | 1 | 31 |
| Month | 5.834 | 0.038 | 3.148 | 9.911 | 1 | 12 |
| Year | 2019 | 0 | 0.000 | 0.000 | 2019 | 2019 |
| Temperature (°C) | 17.509 | 0.094 | 7.756 | 60.152 | 0.12 | 34 |
| Wind direction (Degree) | 192.874 | 0.921 | 75.970 | 5771.462 | 34.86 | 350 |
| Wind speed ($ms^{-1}$) | 2.110 | 0.012 | 1.021 | 1.043 | 0.06 | 6 |
| Relative humidity(%) | 73.382 | 0.198 | 16.304 | 265.829 | 16.08 | 100 |
| Air pressure (mbar) | 1010.477 | 0.068 | 5.573 | 31.060 | 993.39 | 1028 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Summer | 0.273 | 0.005 | 0.446 | 0.199 | 0 | 1 |
| Winter | 0.201 | 0.005 | 0.401 | 0.161 | 0 | 1 |
| Spring | 0.305 | 0.006 | 0.460 | 0.212 | 0 | 1 |
| Autumn | 0.221 | 0.005 | 0.415 | 0.172 | 0 | 1 |
| $PM_{10}$ ($\mu gm^{-3}$) | 60 | 1 | 42 | 1801 | 6 | 509 |
| $SO_2$ ($\mu gm^{-3}$) | 10 | 0 | 7 | 51 | 0 | 100 |
| NO ($\mu gm^{-3}$) | 20 | 1 | 45 | 2033 | 0 | 696 |
| $NO_2$ ($\mu gm^{-3}$) | 29 | 0 | 18 | 332 | 2 | 152 |
| $NO_X$ ($\mu gm^{-3}$) | 61 | 1 | 82 | 6760 | 2 | 1219 |
| $O_3$ ($\mu gm^{-3}$) | 24 | 0 | 17 | 292 | 0 | 97 |

**Table 3:** Correlation coefficient between input and output variables

| Variable | PM10 | SO2 | NO | NO2 | NOX | O3 |
|---|---|---|---|---|---|---|
| Time | 0.110 | 0.082 | -0.003 | 0.159 | 0.021 | 0.092 |
| Day | -0.028 | -0.021 | -0.007 | -0.004 | -0.005 | 0.004 |
| Month | -0.011 | -0.096 | 0.048 | -0.037 | 0.049 | -0.043 |
| Year | -0.226 | 0.190 | -0.075 | -0.020 | -0.067 | -0.050 |
| Temperature | -0.055 | -0.136 | -0.254 | -0.109 | -0.245 | 0.380 |
| Wind direction | 0.057 | 0.045 | 0.197 | 0.094 | 0.198 | -0.212 |
| Wind speed | -0.225 | -0.061 | -0.355 | -0.416 | -0.395 | 0.441 |
| Relative humidity | -0.071 | -0.115 | 0.149 | 0.018 | 0.142 | -0.395 |
| Air pressure | 0.074 | 0.025 | 0.130 | 0.051 | 0.130 | -0.136 |
| Summer | -0.102 | -0.117 | -0.205 | -0.134 | -0.207 | 0.255 |
| Winter | -0.005 | 0.245 | 0.143 | 0.041 | 0.136 | -0.187 |
| Spring | 0.116 | -0.039 | -0.024 | 0.047 | -0.027 | 0.036 |
| Autumn | -0.011 | -0.083 | 0.093 | 0.048 | 0.106 | -0.114 |

**Table 4:** The most effect season(s) on each gas using subsets attribute evaluator with greed stepwise search method.

| Gas | Most Effected Season |
|---|---|
| $PM_{10}$ | Spring |
| $SO_2$ | Summer, winter |
| NO | Summer, winter |
| $NO_2$ | Autumn |
| NOX | Autumn, Spring |
| $O_3$ | Spring, Autumn, Spring, Winter |

**Table 5:** Model design and description of each model.

| Model No. | Description |
|---|---|
| 1 | Without Season effect |
| 2 | Spring |
| 3 | Autumn |
| 4 | Summer, Winter |
| 5 | Autumn, Spring |
| 6 | Spring, Autumn, Spring, Winter |

**Table 6:** The performance of NO gas prediction using training and testing datasets

| Dataset | Model | Gas | R2 | MSE | MAE | RMSE |
|---|---|---|---|---|---|---|
| Training | 1 | NO | 0.944 | 559.89 | 10.86 | 23.66 |
| | 2 | NO | 0.946 | 540.50 | 10.99 | 23.25 |
| | 3 | NO | 0.943 | 571.06 | 10.82 | 23.90 |
| | 4 | NO | 0.944 | 551.49 | 10.81 | 23.48 |
| | 5 | NO | 0.950 | 508.47 | 9.84 | 22.55 |
| | 6 | NO | 0.949 | 507.04 | 10.62 | 22.52 |
| Testing | 1 | NO | 0.894 | 430.62 | 10.18 | 20.75 |
| | 2 | NO | 0.864 | 533.95 | 11.18 | 23.11 |
| | 3 | NO | 0.877 | 565.95 | 14.14 | 23.79 |
| | 4 | NO | 0.890 | 456.10 | 11.22 | 21.36 |
| | 5 | NO | 0.867 | 540.75 | 10.83 | 23.25 |
| | 6 | NO | 0.901 | 431.54 | 11.59 | 20.77 |

**Table 7:** The performance of $NO_2$ gas prediction using training and testing datasets

| Dataset | Model | Gas | $R^2$ | MSE | MAE | RMSE |
|---|---|---|---|---|---|---|
| Training | 1 | $NO_2$ | 0.945 | 30.81 | 3.73 | 5.55 |
| | 2 | $NO_2$ | 0.910 | 48.93 | 4.79 | 7.00 |
| | 3 | $NO_2$ | 0.958 | 24.29 | 3.47 | 4.93 |
| | 4 | $NO_2$ | 0.949 | 29.15 | 3.74 | 5.40 |
| | 5 | $NO_2$ | 0.959 | 24.20 | 3.48 | 4.92 |
| | 6 | $NO_2$ | 0.938 | 34.57 | 4.01 | 5.88 |
| Testing | 1 | $NO_2$ | 0.961 | 56.87 | 5.35 | 7.54 |
| | 2 | $NO_2$ | 0.960 | 40.44 | 4.14 | 6.36 |
| | 3 | $NO_2$ | 0.973 | 36.81 | 4.28 | 6.07 |
| | 4 | $NO_2$ | 0.953 | 44.20 | 4.32 | 6.65 |
| | 5 | $NO_2$ | 0.971 | 49.07 | 4.97 | 7.01 |
| | 6 | $NO_2$ | 0.950 | 49.06 | 4.48 | 7.00 |

**Table 8:** The performance of $NO_X$ gas prediction using training and testing datasets

| Dataset | Model | Gas | $R^2$ | MSE | MAE | RMSE |
|---|---|---|---|---|---|---|
| Training | 1 | $NO_X$ | 0.946 | 1419.45 | 17.64 | 37.68 |
| | 2 | $NO_X$ | 0.944 | 1478.56 | 18.86 | 38.45 |
| | 3 | $NO_X$ | 0.942 | 1526.45 | 18.14 | 39.07 |
| | 4 | $NO_X$ | 0.944 | 1474.48 | 18.21 | 38.40 |
| | 5 | $NO_X$ | 0.951 | 1332.15 | 16.68 | 36.50 |
| | 6 | $NO_X$ | 0.949 | 1357.84 | 17.99 | 36.85 |
| Testing | 1 | $NO_X$ | 0.915 | 1141.55 | 15.53 | 33.79 |
| | 2 | $NO_X$ | 0.892 | 1463.17 | 18.60 | 38.25 |
| | 3 | $NO_X$ | 0.902 | 1403.59 | 20.39 | 37.46 |
| | 4 | $NO_X$ | 0.907 | 1277.21 | 18.33 | 35.74 |
| | 5 | $NO_X$ | 0.892 | 1471.76 | 17.43 | 38.36 |
| | 6 | $NO_X$ | 0.913 | 1241.92 | 19.37 | 35.24 |

**Table 9:** The performance of $O_3$ gas prediction using training and testing datasets

| Dataset | Model | Gas | R2 | MSE | MAE | RMSE |
|---|---|---|---|---|---|---|
| Training | 1 | $O_3$ | 0.947 | 72.72 | 4.74 | 8.53 |
| | 2 | $O_3$ | 0.954 | 64.54 | 4.91 | 8.03 |
| | 3 | $O_3$ | 0.959 | 57.85 | 4.91 | 7.61 |
| | 4 | $O_3$ | 0.959 | 56.64 | 4.62 | 7.53 |
| | 5 | $O_3$ | 0.965 | 50.86 | 4.50 | 7.13 |
| | 6 | $O_3$ | 0.949 | 70.56 | 5.20 | 8.40 |
| Testing | 1 | $O_3$ | 0.946 | 65.04 | 6.32 | 8.06 |
| | 2 | $O_3$ | 0.928 | 69.67 | 6.25 | 8.35 |
| | 3 | $O_3$ | 0.948 | 57.66 | 6.30 | 7.59 |
| | 4 | $O_3$ | 0.959 | 53.79 | 6.15 | 7.33 |
| | 5 | $O_3$ | 0.963 | 47.69 | 5.75 | 6.91 |
| | 6 | $O_3$ | 0.941 | 42.82 | 4.99 | 6.54 |

**Table 10:** The performance of $PM_{10}$ gas prediction using training and testing datasets

| Dataset | Model | Gas | R2 | MSE | MAE | RMSE |
|---|---|---|---|---|---|---|
| Training | 1 | $PM_{10}$ | 0.943 | 669.14 | 13.09 | 25.87 |
| | 2 | $PM_{10}$ | 0.950 | 590.93 | 12.66 | 24.31 |
| | 3 | $PM_{10}$ | 0.947 | 631.69 | 12.81 | 25.13 |
| | 4 | $PM_{10}$ | 0.952 | 579.37 | 14.06 | 24.07 |
| | 5 | $PM_{10}$ | 0.944 | 664.06 | 13.17 | 25.77 |
| | 6 | $PM_{10}$ | 0.947 | 640.62 | 12.91 | 25.31 |
| Testing | 1 | $PM_{10}$ | 0.925 | 314.89 | 9.84 | 17.75 |
| | 2 | $PM_{10}$ | 0.927 | 256.98 | 8.79 | 16.03 |
| | 3 | $PM_{10}$ | 0.911 | 309.40 | 9.48 | 17.59 |
| | 4 | $PM_{10}$ | 0.919 | 291.38 | 10.18 | 17.07 |
| | 5 | $PM_{10}$ | 0.906 | 342.02 | 9.63 | 18.49 |
| | 6 | $PM_{10}$ | 0.915 | 303.95 | 9.05 | 17.43 |

**Table 11:** The performance of $SO_2$ gas prediction using training and testing datasets

| Dataset | Model | Gas | R2 | MSE | MAE | RMSE |
|---|---|---|---|---|---|---|
| Training | 1 | $SO_2$ | 0.811 | 16.01 | 2.57 | 4.00 |
| | 2 | $SO_2$ | 0.683 | 24.80 | 2.95 | 4.98 |
| | 3 | $SO_2$ | 0.330 | 41.07 | 3.91 | 6.41 |
| | 4 | $SO_2$ | 0.550 | 32.20 | 3.44 | 5.67 |
| | 5 | $SO_2$ | 0.727 | 22.08 | 2.73 | 4.70 |
| | 6 | $SO_2$ | 0.803 | 16.59 | 2.52 | 4.07 |
| Testing | 1 | $SO_2$ | 0.824 | 34.60 | 4.37 | 5.88 |
| | 2 | $SO_2$ | 0.039 | 69.57 | 5.52 | 8.34 |
| | 3 | $SO_2$ | 0.338 | 48.53 | 4.62 | 6.97 |
| | 4 | $SO_2$ | 0.462 | 44.23 | 4.36 | 6.65 |
| | 5 | $SO_2$ | 0.756 | 28.08 | 3.46 | 5.30 |
| | 6 | $SO_2$ | 0.888 | 16.47 | 2.74 | 4.06 |

**Table 12:** The best model for each gas using training and testing datasets.

| **Dataset** | **Gas** | **$R^2$** | **Best Model** | **Description** |
|---|---|---|---|---|
| Training | $PM_{10}$ | 0.952 | 4 | Autumn Spring |
| | $SO_2$ | 0.811 | 1 | Without Season effect |
| | NO | 0.950 | 5 | Autumn Spring |
| | $NO_2$ | 0.959 | 5 | Autumn Spring |
| | $NO_X$ | 0.951 | 5 | Autumn Spring |
| | $O_3$ | 0.965 | 5 | Autumn Spring |
| Testing | $PM_{10}$ | 0.927 | 2 | Spring |
| | $SO_2$ | 0.888 | 6 | All seasons |
| | NO | 0.901 | 6 | All seasons |
| | $NO_2$ | 0.973 | 3 | Autumn |
| | $NO_X$ | 0.915 | 1 | Without Season effect |
| | $O_3$ | 0.963 | 5 | Autumn Spring |

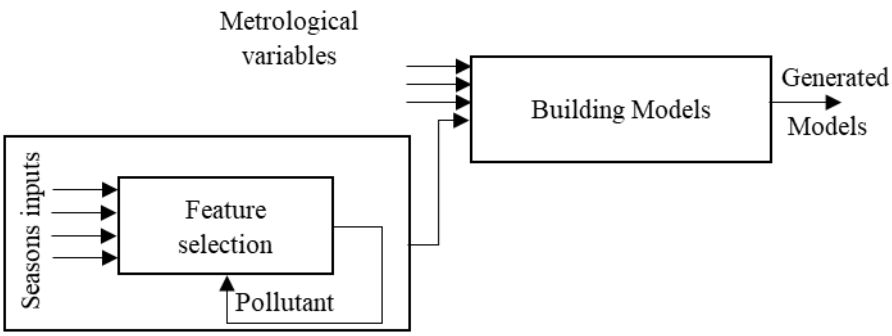



**Figure 1:** The created models' methodology.

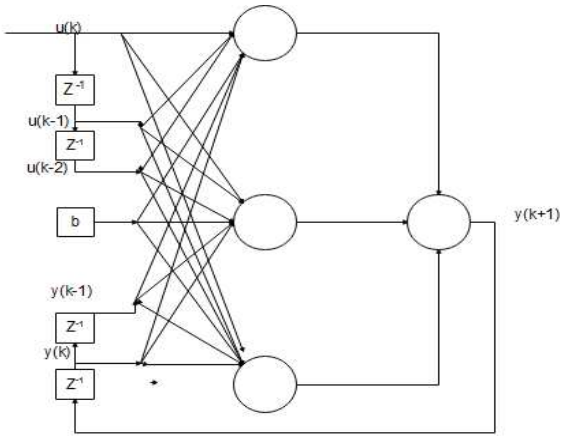



**Figure 2:** NARX model with tapped delay line at input.

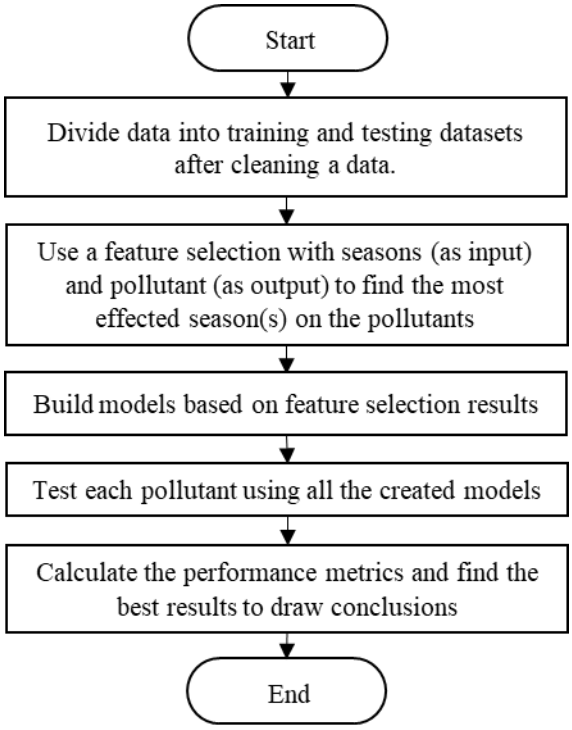



**Figure 3:** The used methodology to build models.

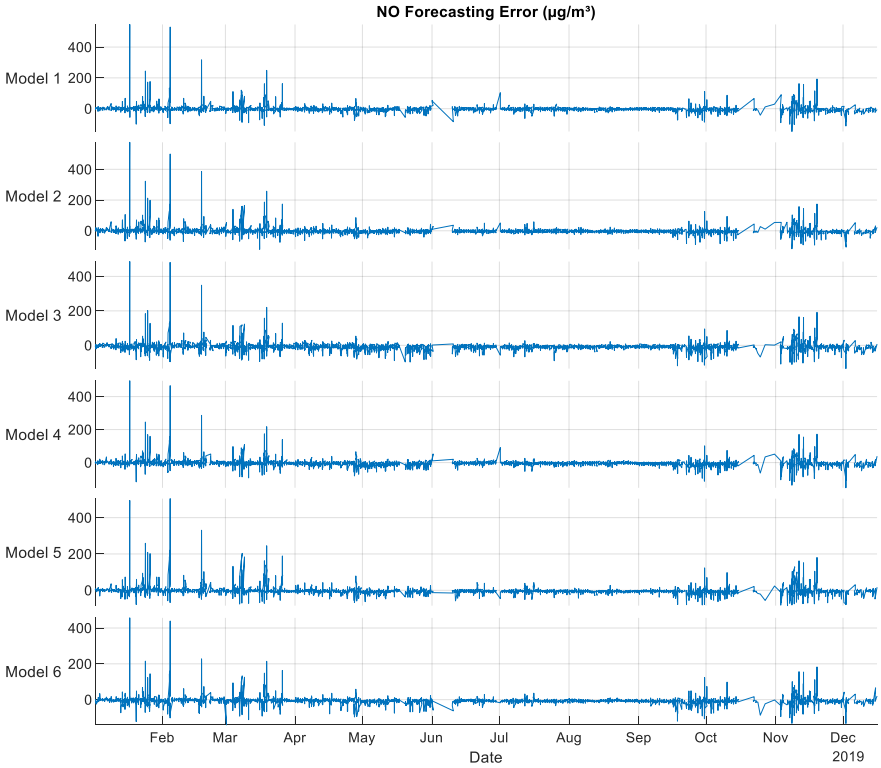



**Figure 4:** NO error (µgm$^{-3}$) forecasting for 2019 using NARX.

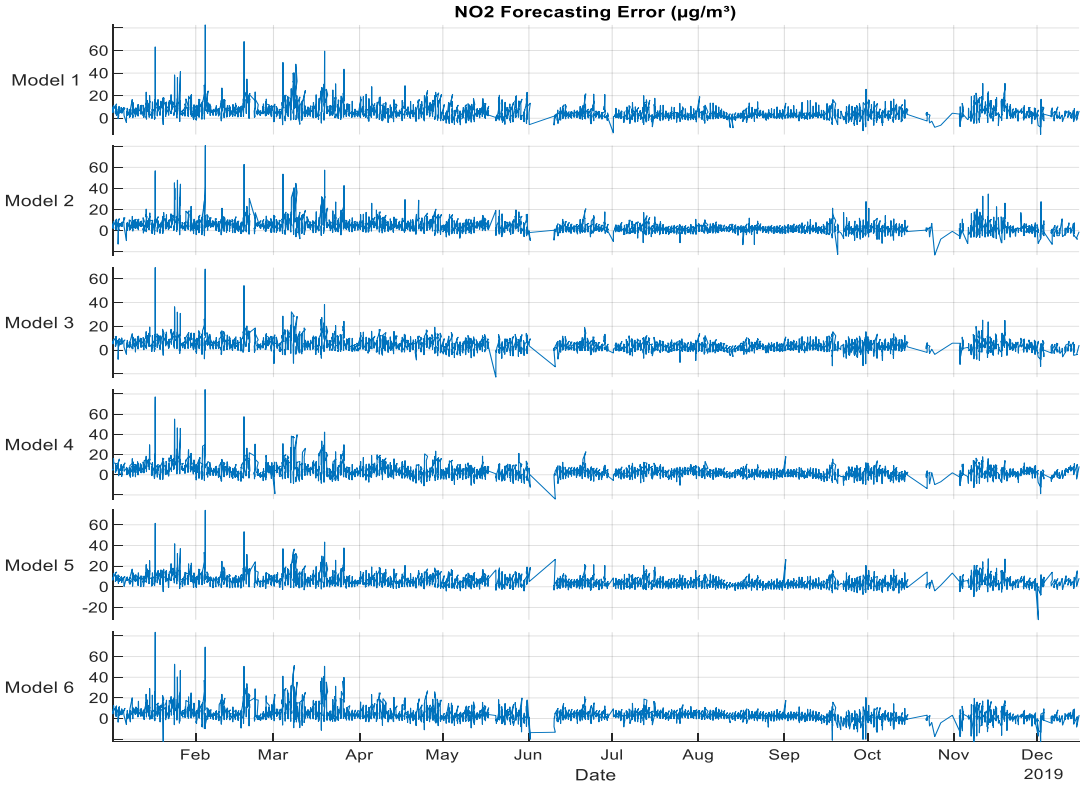



**Figure 5:** $NO_2$ error (µg/m³) forecasting for 2019 using NARX.

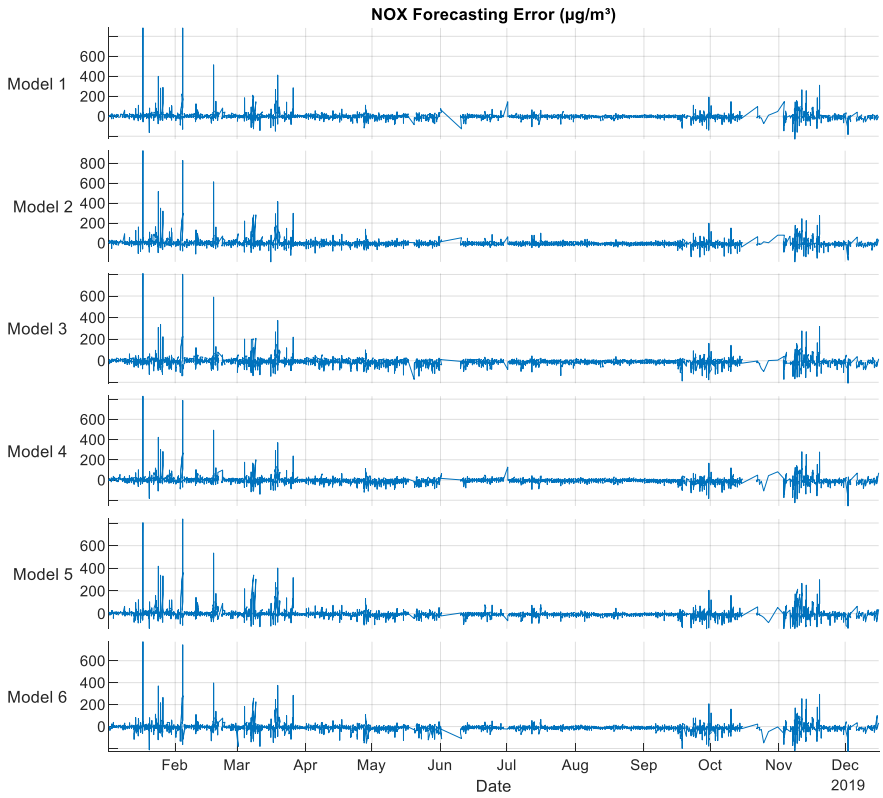



**Figure 6:** $NO_X$ error (µg/m³) forecasting errors for 2019 using NARX.

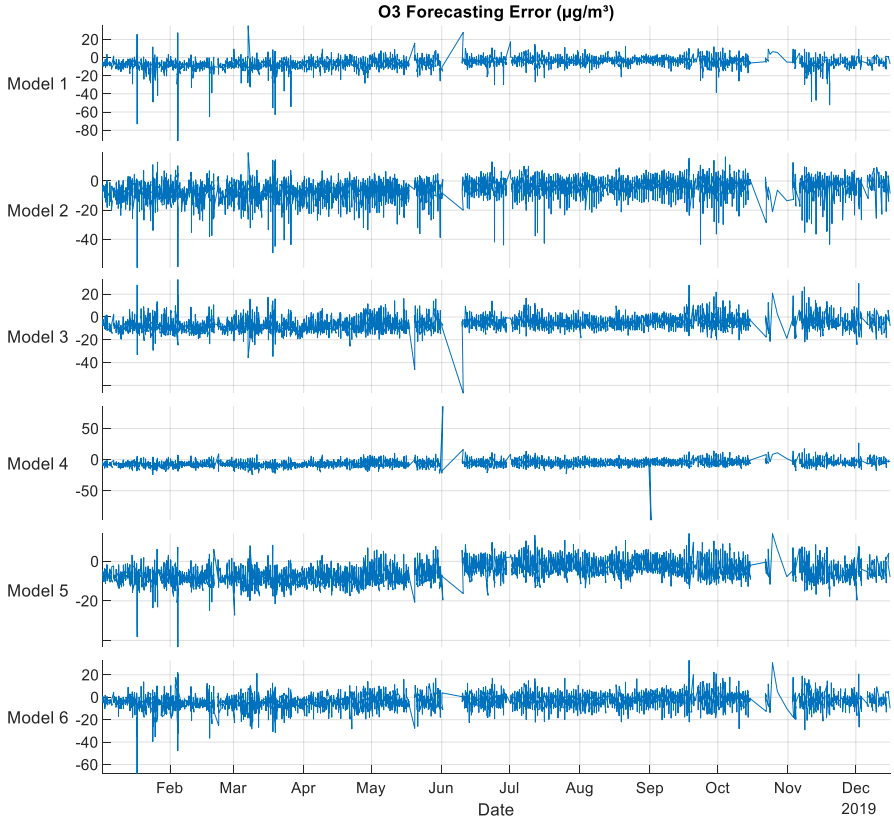



**Figure 7:** $O_3$ error (μg/m³) forecasting errors for 2019 using NARX.

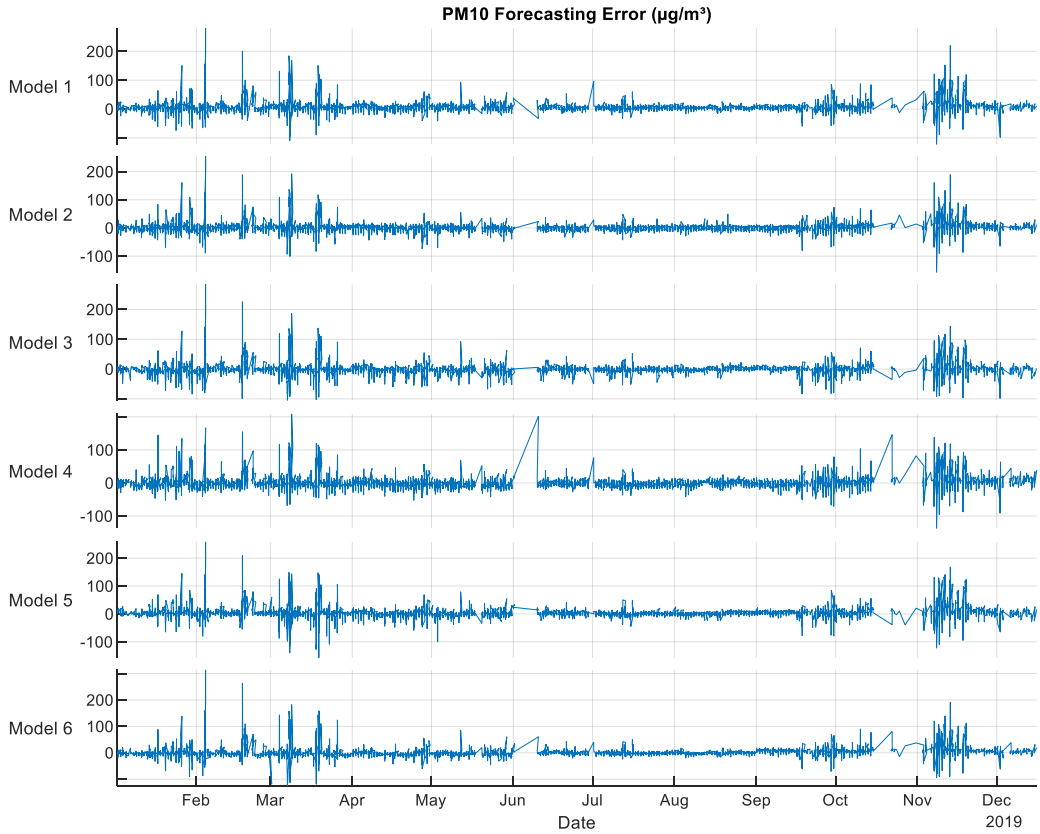



**Figure 8:** $PM_{10}$ error (µg/m³) forecasting errors for 2019 using NARX.

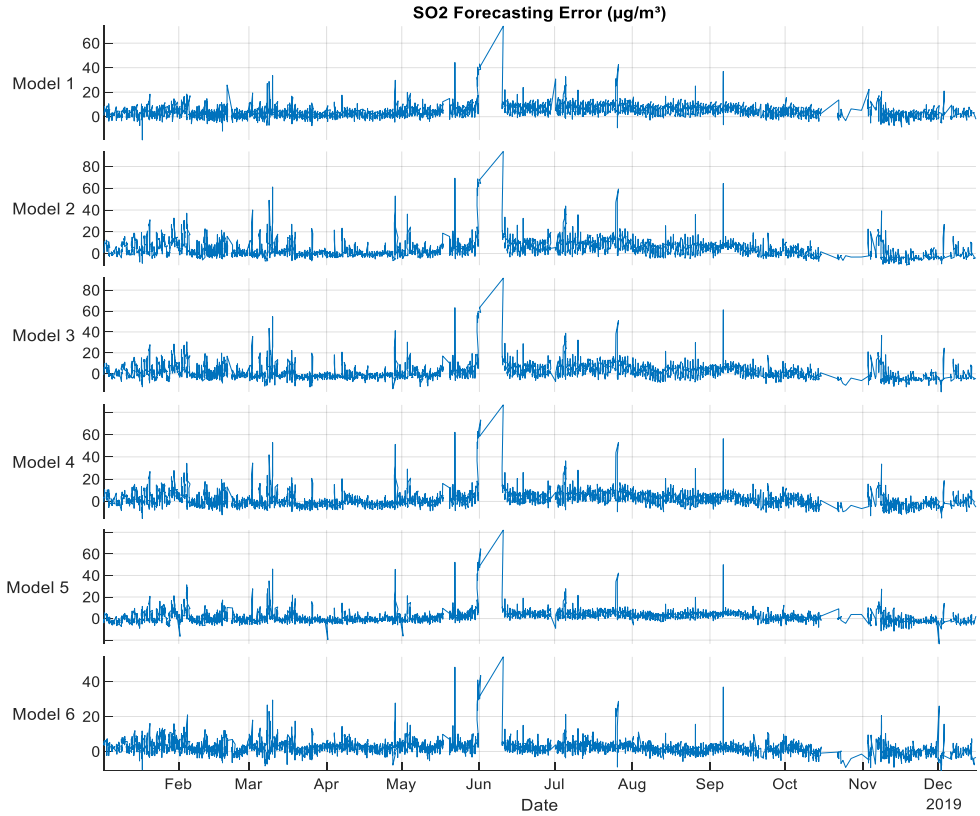



**Figure 9:** $SO_2$ error (µg/m³) forecasting errors for 2019 using NARX.

# Figures

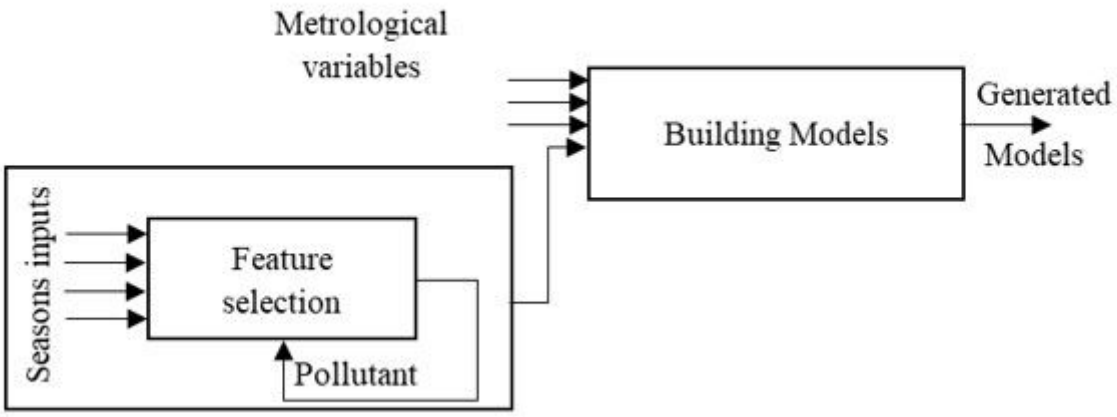



**Figure 1**

The created models' methodology

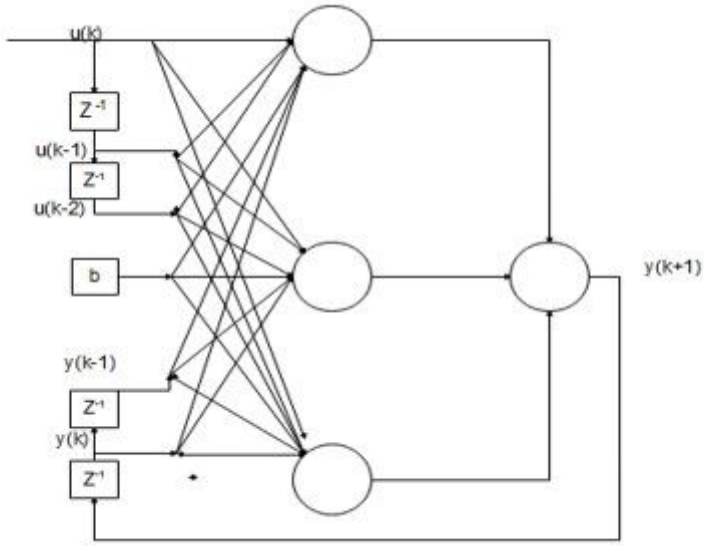



**Figure 2**

NARX model with tapped delay line at input. To sum up all the used methodology in this study a flow chart in Figure 3 is considered.

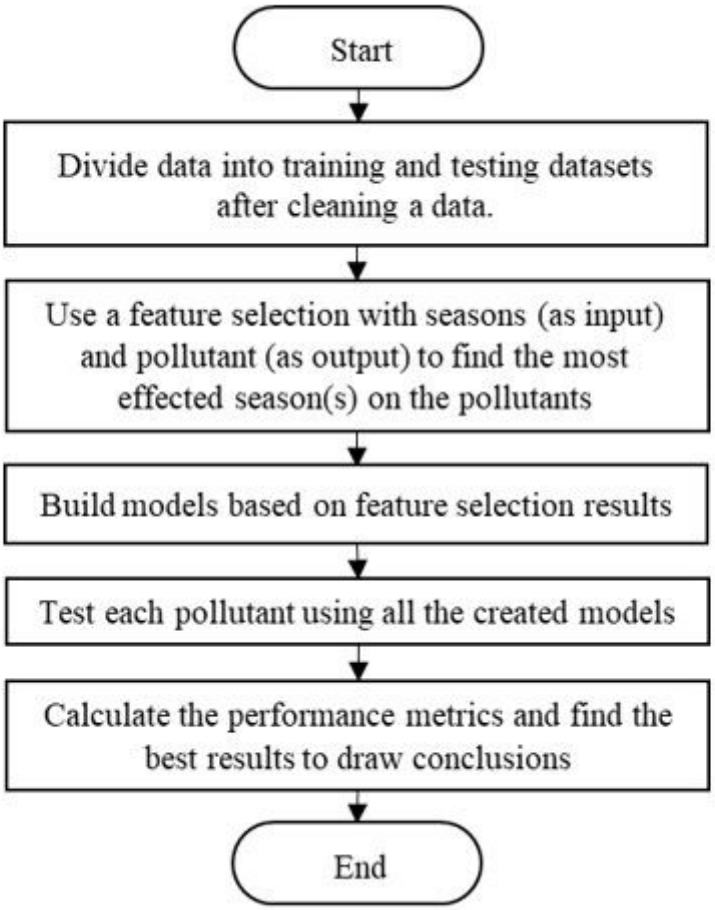
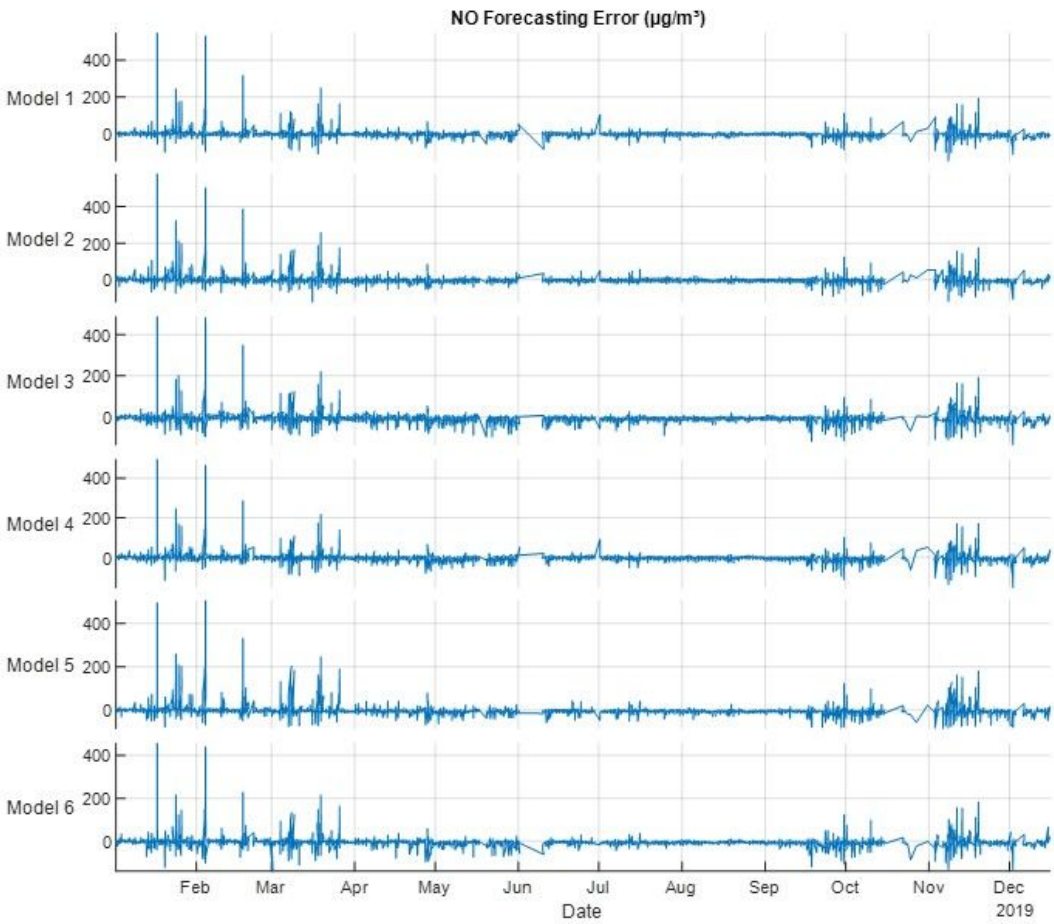


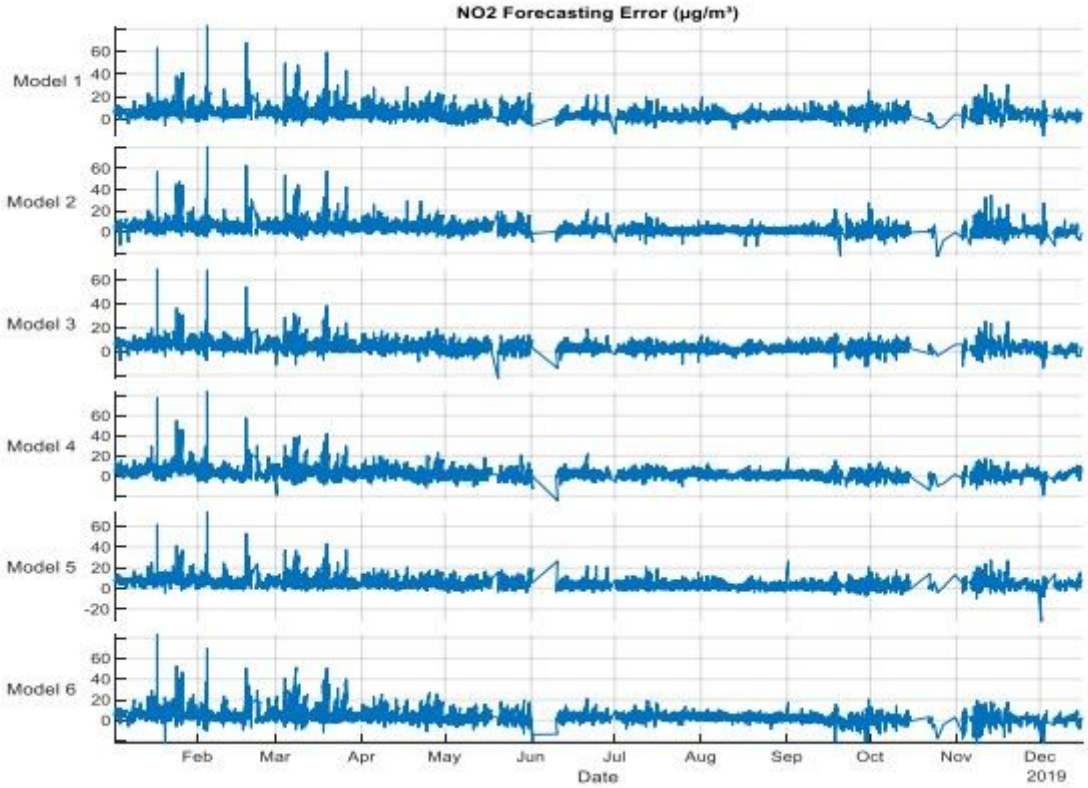

**Figure 3**

The used methodology to build models.



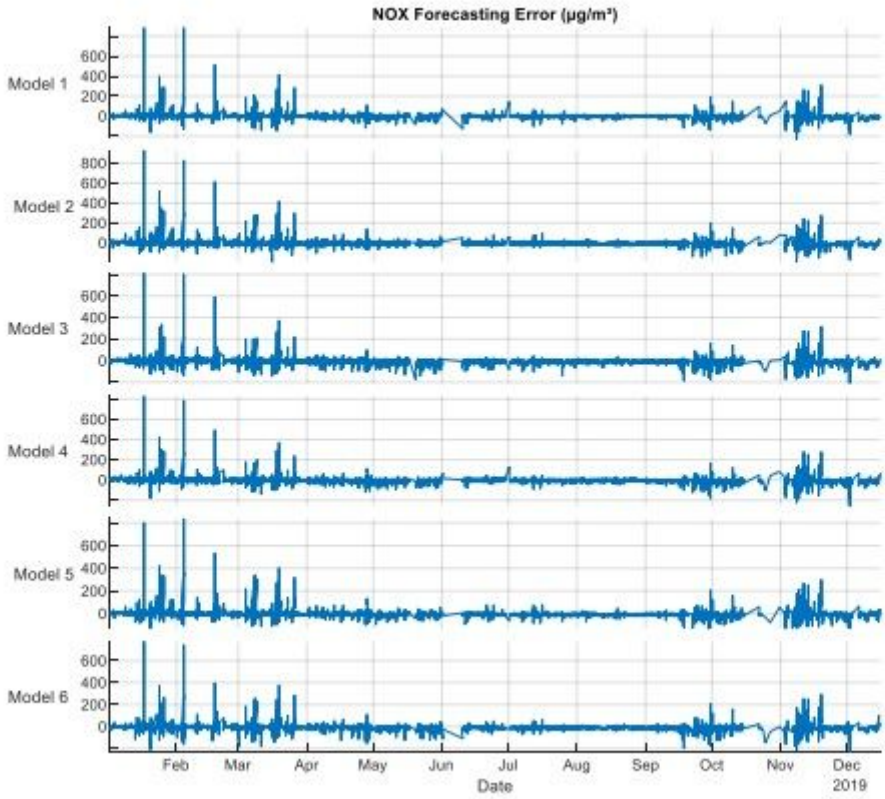

## Figure 4

NO error (μg/m³) forecasting for 2019 using NARX.



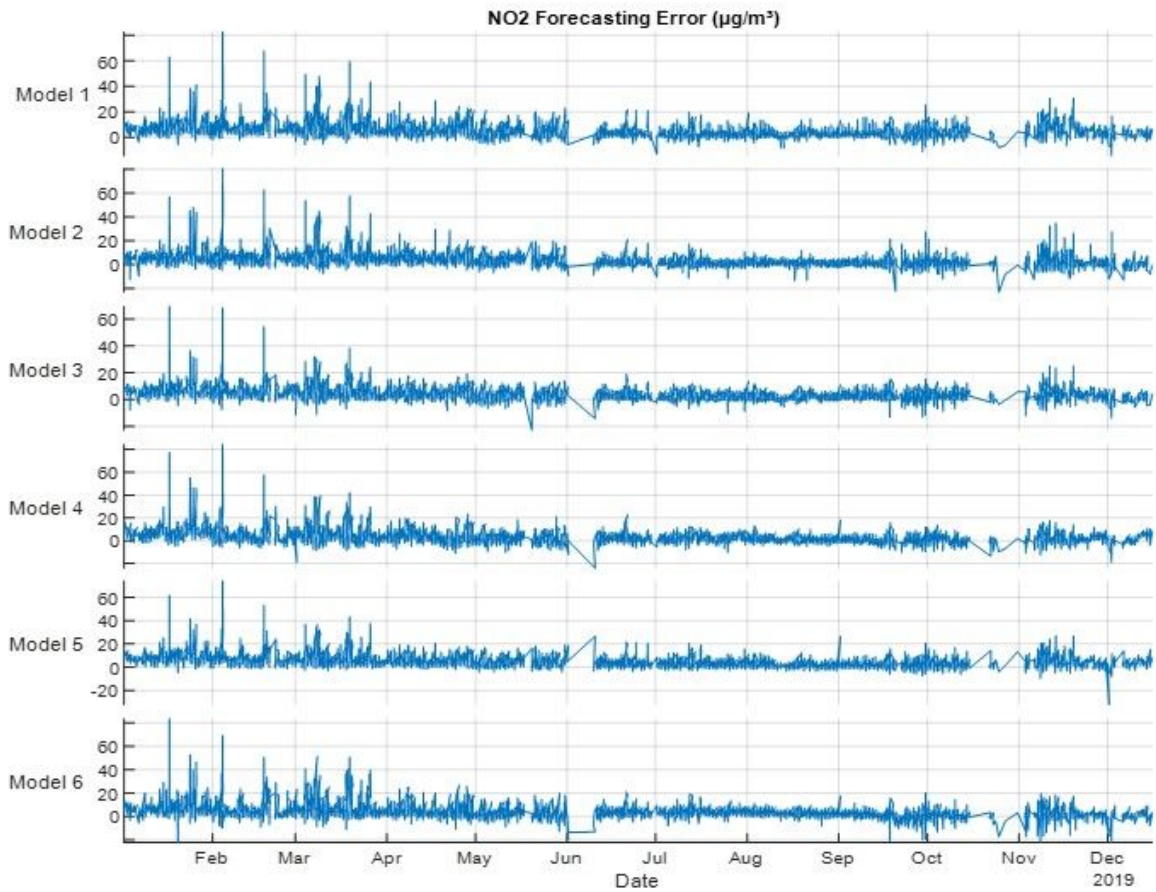

## Figure 5

NO2 error (μg/m³) forecasting for 2019 using NARX.



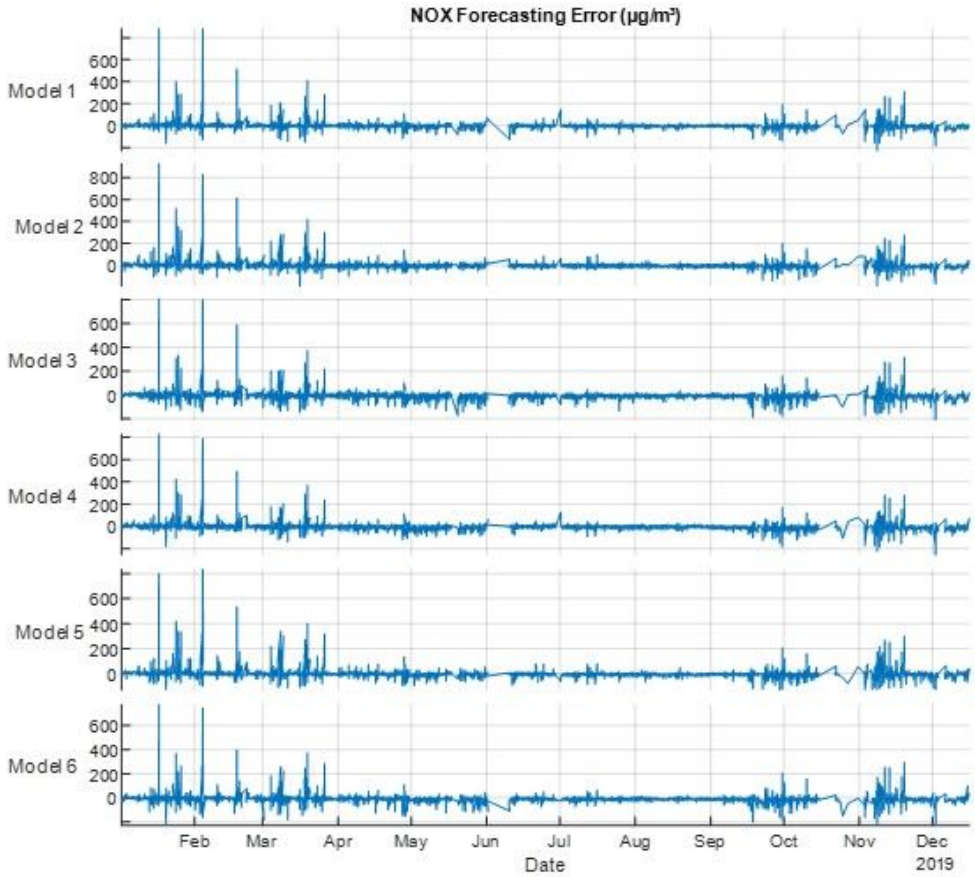

## Figure 6

NOX error (µg/m³) forecasting errors for 2019 using NARX.



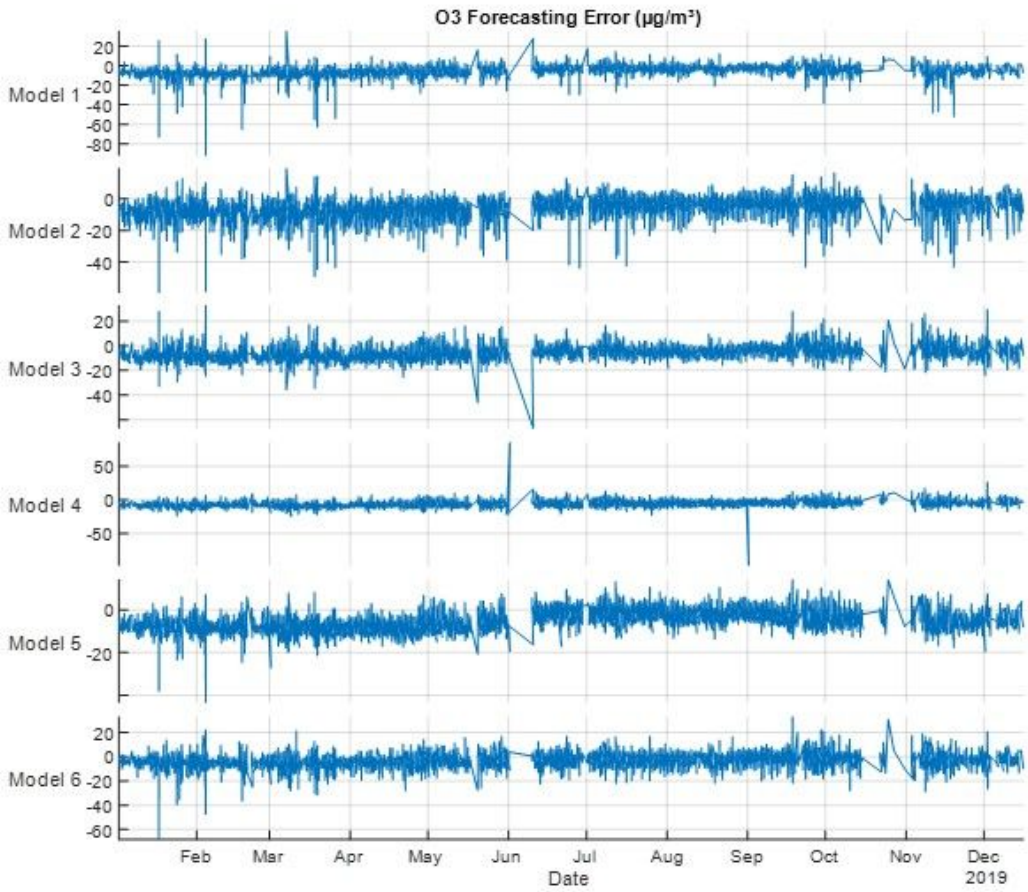

## Figure 7

NO2 error (µg/m³) forecasting for 2019 using NARX.




## Figure 8

NOX error (µg/m³) forecasting errors for 2019 using NARX.




## Figure 9

O3 error (μg/m³) forecasting errors for 2019 using NARX.

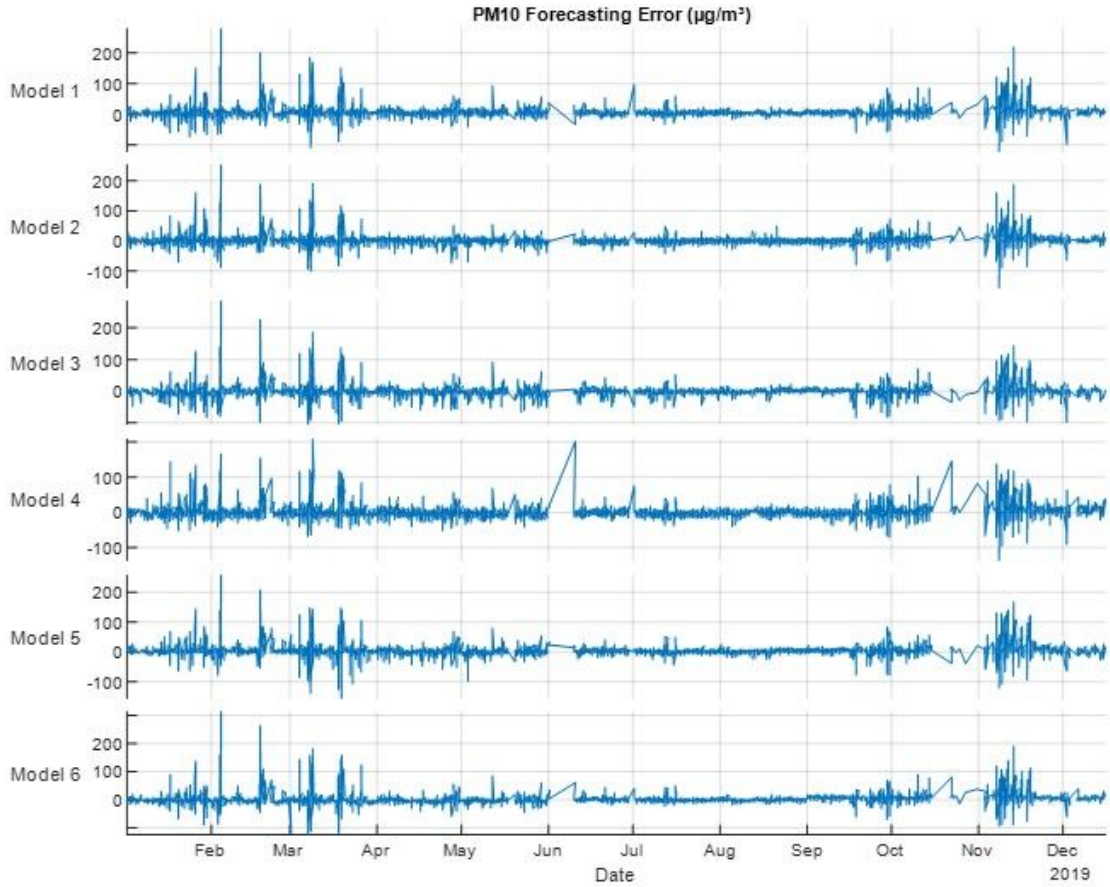



**Figure 10**

PM10 error (µg/m³) forecasting errors for 2019 using NARX.

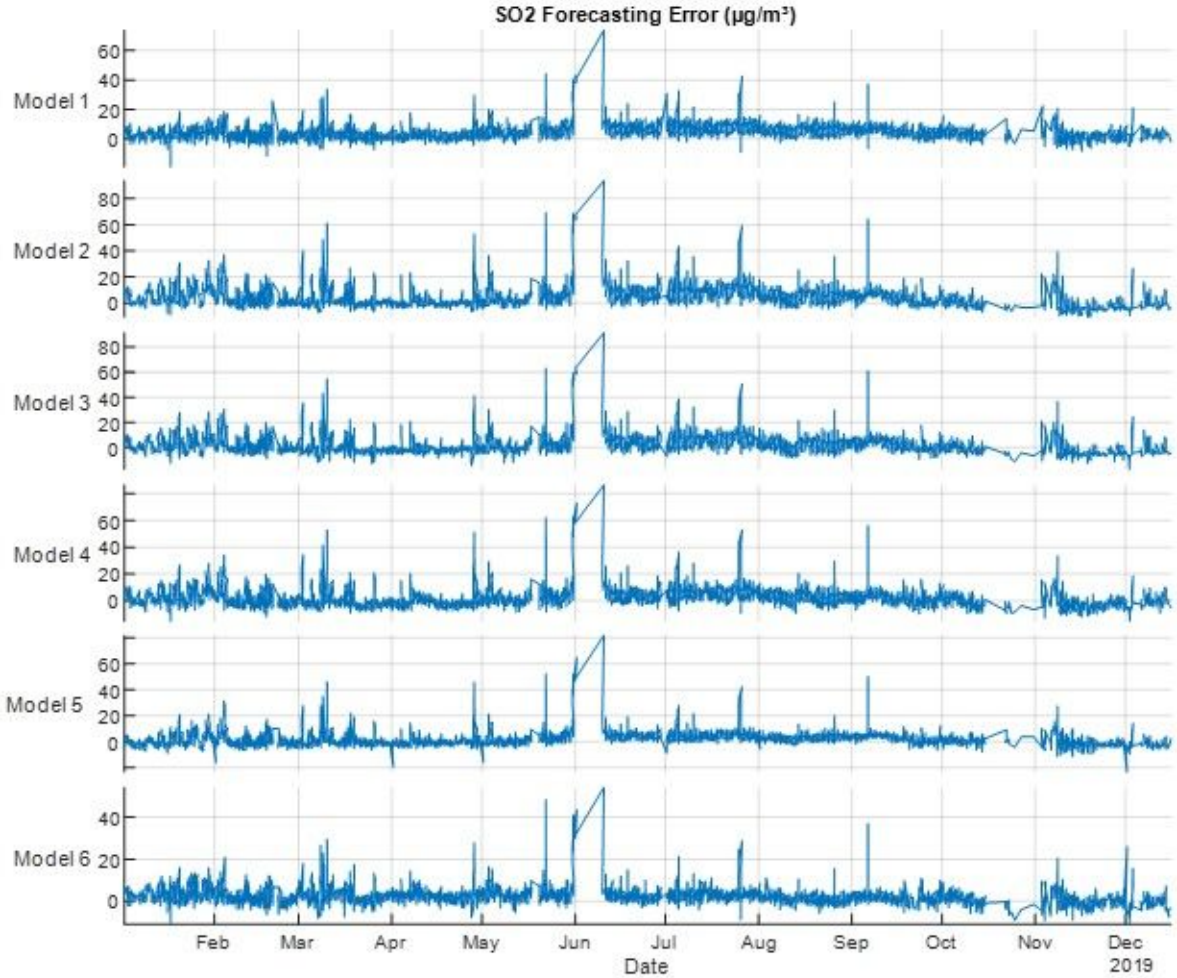



**Figure 11**

SO2 error (µg/m³) forecasting errors for 2019 using NARX.