M.Fares Kanjo

INTELIGENT SYSTEM FOR AIR POLLUTION PREDICTION

NEU
2019

INTELIGENT SYSTEM FOR AIR POLLUTION PREDICTION

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES
OF
NEAR EAST UNIVERSITY

By
M. FARES KANJO

In Partial Fulfillment of the Requirements for the Degree of Master of Science
in
COMPUTER ENGINEERING

NICOSIA, 2019

INTELIGENT SYSTEM FOR AIR POLLUTION PREDICTION

A THESIS SUBMITTED TO THE GRADUATE

SCHOOL OF APPLIED SCIENCES
OF
NEAR EAST UNIVERSITY

By
M. FARES KANJO

In Partial Fulfillment of the Requirements for

the Degree of Master of Science
in
COMPUTER ENGINEERING

NICOSIA, 2019

M. Fares KANJO: INTELLIGENT SYSTEM FOR AIR POLLUTION PREDICTION

Approval of Director of Graduate School of

Applied Sciences

Prof. Dr. Nadire Cavus

We certify this thesis is satisfactory for the award of the degree of Master of Science in Computer Engineering

Examining Committee in Charge:

Assist. Prof. Dr. Besime ERİN	Department of Computer Engineering, NEU
Assist.Prof.Dr. Boran ŞEKEROĞLU	Department of Information Systems Engineering, NEU
Prof.Dr. Rahib H. Abiyev	Department of Computer Engineering, NEU

In this study all information in this document has been obtained and presented in conformity with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: M. Fares Kanjo

Signature:

Date:

To my parents...

ACKNOWLEDGEMENTS

First, I am grateful for the good health and wellbeing that were necessary to graduate and complete this project.

I wish to express my sincere thanks to all staff of the faculty of Computer engineering in Near East University, especially for who lead me and helped me that long way since I first started, Prof. Dr. Rahib Abiyev, for providing me with all the necessary information for completing the study and for his great supervision.

And to Mr.Murat Arslan, I’m grateful for him who helped me to develop and write this thesis and leading me to the right way.

I take this opportunity to express gratitude to my father Mr. Adel and my mother Jinan, they were always there cheering me up and stood by me through the good times and bad, I would like to thank my brothers Khaled, and Ahmad; and I appreciate their efforts with supporting me every time I need them.

I also place on record, my sense of gratitude to all my friends and colleagues especially.

Finally, I would like to thank one and all, who directly or indirectly, have lent their hand in this venture.

I dedicate this project to my country Syria and to all people that I mentioned before.

ABSTRACT

This thesis presents the design of an intelligent machine learning system for prediction of the air pollution. Different machine learning models have been studied, analysed and the neuro-fuzzy is proposed for the design of air pollution prediction. The neuro-fuzzy structure based on ANFIS structure and its learning algorithm have been described. The proposed neuro-fuzzy model has been tested with different parameters of the pollution from Istanbul and Bursa regions in order to estimate the performances and reliabilities of the models. The learning data has been achieved using artificial neural networks (ANN), NARX and ANFIS algorithms. The hourly pollution predication gained for Bursa and Istanbul area is used for training and testing of the models. The performances of the neural network and neuro-fuzzy models are tested using these data. The simulation results show that the neuro-fuzzy model predicted output match with the actual data in good accuracy. As a result of simulations of the ANFIS model it was found that the RMSE for training data was 0.0022, for testing data 0.0038. The results show that the ANFIS model is most fitted and suitable and acceptable for prediction of the air pollution. Comparisons of the results of different models show that the neuro-fuzzy model has the best performance in prediction of the hourly pollution data with the specific parameters than other considered models.

Keywords: machine learning; artificial neural networks; parameters; air pollution prediction

ÖZET

Bu araştırma, hava kirliliğini tahmin etmek için özel ve etkili bir makine öğrenme sistemini incelemektedir. Çalışma, İstanbul ve Bursa'daki kirliliğin farklı parametrelerine göre test edilmiş, aslında modelin performans ve güvenilirliğini tahmin etmek için nöral ağ ve nöro bulanık modeller bu tez üzerine odaklanmıştır. Öğrenme verileri, geri yayılım, NARX ve ANFIS algoritmaları kullanılarak gerçekleştirilmiştir. En iyi sonuca ulaşmak için, kirlilik verileri ANFIS ve ANN modelleri kullanılarak test edilmiş ve eğitilmiştir. Eğitim ve testten sonra kazanılan saatlik kirlilik önceliği, o modelin doğruluğunu sağlamak için tahmin edilen verileri gerçek verilerle eşleştiriyor. Bu çalışmanın etkisi ve iyi sonuç, modelin etkili olduğunu ve bu tahminle bu özel parametrelerle saatlik kirlilik verilerinin başarılı olduğunu göstermektedir. ANFIS modeli ile verileri eğitmek ve test etmek sadece RMSE'nin 0.0022, 0.0038 değerine kadar olan hataları azaltmakla kalmayıp, aynı zamanda bu tahmin verilerinin performansını ve güvenilirliğini de arttırmaktadır. Bu çalışmanın sonucu, ANFIS modelinin, kirlilik tahmini için yapay akıllı sistem ile en uyumlu ve uygun ve kabul edilebilir olduğunu ve çıkış verilerinin orijinal verilerle çıktı verilerinin karşılaştırılması yoluyla minimum hata ile en iyi doğruluğu sağladığını göstermektedir. ağ tarafından tahmin edilmiştir.

Anahtar kelimeler: makine öğrenimi; nöro-bulanık; nöral ağlar; parametreler; hava kirliliği tahmini.

TABLE OF CONTENTS

CHAPTER 6: DISCUSSION AND RESULT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial intelligence		
ANN	Artificial Neural Network		
NNBP	Neural Networks backpropagation		
ANFIS	Adaptive neuro-fuzzy inference system		
NARX	Nonlinear autoregressive exogenous		
FS	Feature selection		
GA	Genetic algorithm		
RMSE	Root Mean Squared Error		
MSE	Mean Squared Error		
AI	Artificial Intelligence		
API	AQI Value of The Next Day		
AQI	Air Quality Index		
CBPN	Cascade-forward back propagation neural network		
	PM2.5	µg/m3	Daily averaged concentration of PM2.5
POLLUTIANTS	PM10	µg/m3	Daily averaged concentration of PM10
	O3	µg/m3	Daily averaged concentration of O3
	NO2	µg/m3	Daily averaged concentration of NO2

CHAPTER 1
INTRODUCTION

The Ministry of Environment and Urban Planning show results with details of data in Turkey with transparency on their website. The pollution is playing important role in our world as it affects human beings, animals, our planet and all living things. Pollution can lead to unstable climate change which can disrupt the ecosystem. The average daily concentration value of the pollutant particulate matter 10 that allows to exceed the limit value only 35 times a year. According to the EU and the World Health Organization, this means urgent measures should be taken if the limit value (50 µg/m3) is exceeded more than 35 days a year in Turkey. This limit value is exceeded very often and no precautions are usually taken. In Istanbul, the air pollution has reached its highest level in recent years. Particularly in the districts of Yenibosna, Kadıköy and Esenyurt. Air pollution rate has increased by the effects of the using coal in urban transformation and transportations. The high dust amount, which is the source of the pollutant particulate matter 10 (PM10) is increasing and the environmental consequences of the urban transformation processes are not considered and being ignored (Güler, Ü, & Can, Kimyasal).

Air pollution is one of the world's biggest killers. Pollutions was directly involved in 6.4 million deaths in 2015. Furthermore, pollution was the cause of 19% of all cardiovascular deaths worldwide, 21% of stroke deaths and 23% for lung cancer deaths. Because of this, the authoritative predicting technique is needed to lead us to an important role in the danger of crisis response and emergency plans (Wang, H et al, 2016).

The first application for air pollutants concentrations predicting and modelling was conducted in 1993 by Boznar et al, which the ANN has been recognized as the most effective method in this field (Challoner, 2015). Most of the input parameters of established ANN network model were based on experience from the existing scientific literature and subjective inference or common sense (Fu et al., 2015; Perez & Salini, 2008). Some researchers who have tried to process feature selection before the training of ANN (Grivas et al. 2006) applied the genetic algorithm optimization operation for the selection of the input variables and compared the forecasted results with various linear regression models.

This study looks for predicting air pollution using a neuro-fuzzy and artificial neural network. Air pollution prediction will support analysing the changing patterns of types of pollutants. It will also help in arranging the protective measures in cases of pollution problems or disasters and managing them.

This study leads to regularization-based feature selection for the filtering model inputs to clear excrescent information, mitigate multicollinearity and enhance generalization ability. Two ordinary NN, models optimized by NARX and a BP model with feature selection are established as contrast, respectively. Statistical guide like mean absolute error Pearson correlation coefficient (R) and root mean square error (RMSE) are tested at the same time.

1.1 Aim of the Study

This study aims to predict air pollution using historical hourly data of Istanbul and Bursa using artificial intelligence methods such as artificial neural network and neuro-fuzzy system. The essence execution of algorithms will create an output data by classification of hourly data of Istanbul and Bursa according to the denomination using a neuro-fuzzy system. Comparable data will be classified for the specific and accurate information that has more precise findings. The suitable and exact predicted data will help in improving suitable strategies for the environment and Global Warming, I will also go over air quality level for humans and plant health to implement precautionary measures.

This study is based on historical hourly data from 1st of October 2017 to 1st of October 2018. descriptive statistic's for AQI during the observed period. The air quality ground measured hourly data pollute concentration data including PM 2.5, PM10, O3, NO2 and AQI are collected from the information bank of Turkey. This study is effective in predicting air pollution with great accuracy and thoroughness.

1.2 Significance of Study

Air pollution prediction is beneficial on macro level. This study is created value as it contributes to the field and takes into consideration the amount of pollution, agriculture, diseases, and keeps people updated with the level of air pollution in their cities so they will be aware and careful. Prediction of air pollution makes us understand how pollution affects human health. Reducing the quality of air leads to respiratory problems such as lung cancer and asthma. We can also predict how pollution can affect environmental degradation and detect how much the ozone layer is minimizing. The ozone layer is shield high up in the earths sky that can stop ultraviolet rays from crossing. This shield is located high in the sky as the result of human activities, chemicals and fluorocarbons. Also, we can predict how pollution affects infertility in lands. Soil becomes infertile because a lot of pollutants like ozone, carbon, and particulate matter all of these can affect the earth.

1.3 Limitations of the Study

i. Air pollution data of Bursa and Istanbul are hourly statistics. Therefore, the system will have the hourly output.

ii. The climate changes, pollution changes, amount of ozone and other elements affect the impact accuracy of the expected output.

iii. The system, in this study, will operate specifically with Matlab programme in windows 10 (R2018).

1.4 Problem Statement

To predict air pollution levels or conditions, it is crucial to handle and use historical data of parameters measured by considering the availability of large amounts of data to differentiate the type and extent of relationship for suitable and efficacious extraction of information. A lot of spatial data gained are dispersed in nature, like air quality, has different levels of pollution at different locations. The process for acquiring a persistent data collection from a sparse data repository for all intents and purposes are valuable. This study discusses the air pollution data in different locations in Istanbul and Bursa to predict and insert different data for different locations. In fact, and more accurate. This research uses techniques of artificial intelligence systems such as ANFIS, NARX, NNBP and data mining models for predicting air pollution.

The exact air pollution forecasting is somehow lacking which may help in various fields like ozone layer depletion and global warming forecasting. Formulating and creating calculation of air pollution prediction that would be instituted on similarities. It will give output predictions that are effective and reliable. The inexact prediction is loss of resources and wasting time and it can lead to ineffectual control crisis like poor air quality that can harm people and bring bad management of pollution. The need to create a good air pollution predicting system and more importantly creating a system that can be more accurate and have good accuracy as compared to the present air pollution predictor models is necessary.

1.5 Methodology

Creating the V0 with testing it to check out information from the test sample by Spiral model of programming, will get back the possible alterations. It This will be tested using the neural network and neuro-fuzzy system to get the accurate results.

1.6 The Study region and data

Metrological data included particulate matter PM10, matter PM2.5, Nitrogen Dioxide NO2, and, ozone O3. this will be analysed for 2 cities in Turkey; Istanbul and Bursa. These regions have experienced pollution with large population census that can bring cars and manufactures. The data was collected for the cities of Istanbul and Bursa. Data collected between 2017-2018 on an hourly basis to predict the air pollution.

1.7 Overview of the Study

This study has been intended as follows:

Chapter 1 Includes a preface on air pollution, an overview of the research, discussion the aim of this research.

Chapter 2 Reviews previous research in the suggested thesis as existing literature.

Chapter 3 Presents the air pollution prediction in general view and explains air pollution prediction in diverse fields; ozone layer depletion, human health, and global warming.

Chapter 4 Presents the descriptions of the artificial intelligence elements such as neural network and adaptive neuro-fuzzy inference system and different modelling techniques like NARX algorithms used for air pollution prediction. The models are explained with their specifications.

Chapter 5 Highlights the energizing of data and explain the pre-correlation and processing through the inputs and outputs with algorithms of the artificial neural network and adaptive neuro-fuzzy inference system.

Chapter 6 Highlights the confirmation of results and discussions of the air pollution prediction results. It discusses the root mean square error RMSE to assess the result that we analysed for accurate air pollution forecasting.

CHAPTER 2
LITERATURE REVIEW

2.1 Air pollution review

We live in an industrial world, where human activities and climate change affect the environment and air atmosphere. Since a big revolution in factories and vast human activities generate a massive amount of gases, particulates, and molecules to the air. Therefore, these types of elements generate air pollution which causes a different kind of disease, allergies and even it causes the death of million people around the world. According to the 2014 world health organization report, just in 2012, it caused the deaths of 7 million people around the (world World Health Organization,2003).

Based on the importance of the topic, academicians, researchers, and government officials in different countries started worrying about this challenge. As you may hear this as news headlines almost every week. This ends to a lot of trying to predict the air pollution or each substance separately since air pollution has a number of substances (Elbir et al., 2000; Tayanç, 2000) which emitted into the atmosphere like CO2, SOx, NOx, CO, VOC, particulates etc.

After the digital revolution, the rise of strong technologies and superpower computers with improvements in algorithms brought this opportunity to utilize these technologies for the prediction of air pollution in the future in a specific geographic area using relative specific geography air pollution historical data.

One of these technologies which are rapidly growing is machine learning or from a general perspective artificial intelligence. Therefore, researches conducted on optimizing machine learning algorithms for air pollution forecasting. What it does is to use from the previously stored data of a region, it tries to understand the signals that what factors affect air pollution and whether it’s possible to rise or goes down. These algorithms are very powerful which understands the sign of the change occurs in this regard.

Since one of the vital problems in cities especially metropolitan cities is air pollution, (Akkoyunlu and Ertürk, 2003; Karaca et al., 2004, 2005b) researchers tried to focus on each of them separately because each city pollution is different to another. The number of substances and their density is various. Therefore, the focus of the review will be city or country based to make it more accurate and purposeful. On the other hand, the focus for the technical part would be neural network since the algorithms used for the study is neural networks.

A study conducted by (Kurt et al., 2008), demonstrated on an online air pollution forecasting system. The target area is greater Istanbul. The prediction result is publishing in AirPolTool, it is a website (airpol.fatih.edu) which publishes the air pollutant of Istanbul for the next 3 days, and the data is updated twice a day for more accurate results.

The study used a neural network to predict three air pollutant indicators namely SO2, PM10 and CO levels for three coming days. The study claims that a simple neural network can predict air pollutant indicators level accurately. At the same time, the research presents some optimization techniques like different input parameters to enhance accuracy.

Their training method is quite simple, it uses the previous day's data to predict the next day and the merged result for the other days in the same sequence based. The best range of historical data is from 3-15 days which is achieved by multiple testing. Finally, it tested the effect of the day of the week as an input parameter to check whether it has any impact on accuracy or not. The result shows that it helps the algorithm to predict with higher accuracy. Which this property is also suggested for prediction.

Another study by (Wan & Lei, 2008) shows air pollution in Macau, a city in China is rising relative to the economic developments. Therefore, monitoring and predicting the Air Pollution Index (API) becomes increasingly important for the people of the city cause of harms and health effects. The study proposes an adaptive neuro-fuzzy based approach for the prediction of API on that city.

The model uses Sulphur dioxide and total suspended particular matter from the past and historical records with some other factors and concentration for the input of the model and prediction task. The study used backpropagation with the least square algorithm as a learning method. They used 10 years of historical data of the Macau city for training and performance improvement. The study claims that the experiments show a satisfactory performance in prediction of the air pollution in Macau.

A research by (Cai & Xie, 2009) presents an artificial neural network approach to predict air pollutant concentrations in an hourly based in Guangzhou, China. There are many factors which are affecting air pollutants concentrations near to urban areas, basically, they have a very strong and complex relationship. The factors are categorized into four parts traffic, background concentrations, meteorological and geographical.

All these factors which are categorized used as inputs for the ANN suggested approach. The prediction target was four pollutant concentrations CO, NO2, PM10, and O3. These pollutants are the most important elements of polluting the air which has been used for the testing of the model.

The data collected from two sites near the arterial via the equipments of vehicles which are functioning automatically. The purpose of the thesis is to predict the average hourly concentration in the mentioned cities. The range of prediction is until 10 hours maximum. The result shows that the back-propagation neural network can accurately predict the four mentioned pollutant elements.

Additionally, the model is compared with the other model's multiple linear regression models and the California line source model, the comparison conducted using the performance evaluation measurement methods namely MRE, MAE, RMSE, and Correlation coefficient. The experiments show that the presented approach outperform those models and predict concentrations more accurately.

(Pérez & Reyes, 2000) conducted an experiment to predict average concentrations of PM2.5 in hourly based for the Santiago, Chile. They focus on the mounts with high higher PM2.5 values. As the concentration is going to a high level from May to September. The data collected from this month for the years 1994 and 1995. The study claims that by fitting a function to measure the 24 hours of the previous day and considering the changes that are going to happen we can predict the concentrations of the day.

The data chosen to be used for training and testing is from 05/01 to 09/30 for both years. For each year 24 matrices were built which all of them consist of 25 columns. To make sure whether the prediction task is successful or not, three models being tested multilayer neural networks, linear regression, and persistence. The overall result shows that neural network outperformed other models. Actually, the accuracy level is different at different times of the day.

Prediction errors were 30% in the early hours of the day but it increases to 60% for late hours. The study discussed the reasons and causes of the low prediction accuracy. For instance, noise, not arranged data etc.

Another research has done by (Kolehmainen & Ruuskanen, 2001) to predict the hourly average of NO2 and basic meteorological variables. The data collected from Stockholm from 1994 to 1998. The study examined two fundamental and different models of neural network to assess their ability and possibility of this prediction.

Self-Organizing Maps (SOM) and Multi-Layer Perceptron (MLP) have been tested in different ways, using the periodic components, neural network methods to the residual values without periodic components, and applying only ANN. The overall result shows that MLP network with original data without processing achieves the best result.

CHAPTER 3
AIR POLLUTION

3.1 Introduction to Air pollution

The development and improvements of the world in each sector create new challenges to the humans of 21th century. Multiple revolutions have changed the world and serious problems that people never thought of appeared. The industrial revolution, traffic, population increase, gradual climate change, and many more issues cause the many danger to the environment and atmosphere. Air pollution is one of these challenges which has got the attention of countries, united nation, and although researchers around the world.

Air pollution is the result of the dangerous and massive amount of different substances which includes gases, particulates, and biological molecules (Seaton et al, 1995). These harmful substances are produced, mixed and introduced to the Earth’s atmosphere. Like it is mentioned in Chapter 2, it causes diseases, allergies and the deaths of millions in the world. Which the statistic of 2012 shows the number of death due to air pollution were around 7 million (Reed, 2016). This number can compete with almost any other problem like hanger, cancer, terrorism etc. in the world.

Due to the fact of more death around the world from air pollution, the responsibility of fighting these phenomena goes to all governments, universities, individuals and whoever can take part to decrease the density and prevent from rapidly growing pollutant substances. The concept of globalization brought new sights to the world. Basically, the world is now a mutual home for humans which both bad and good actions simultaneously affects everybody in this era. Therefore, mutual interest, benefit, and loss require unity and share strategies to combat these universal challenges.

Figure 3. 1: Air pollution production from factories (thehill.com, 2018)

.

Due to the importance of the topic, combating this challenge needs clarifications and research. Many chemical elements contributed in air pollution as the source of air pollution is various from region to the region or from a country to another country. These elements are necessary to be addressed if we are going to fight with air pollution, these elements should be addressed separately for their property and specifications. The following is a brief description of these elements.

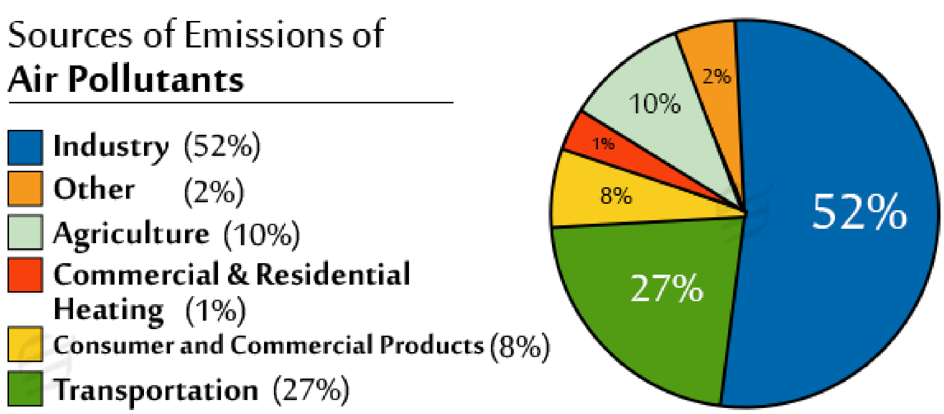

Figure 3. 2: Air pollution contribution percentages (essaycorp.com, 2018)

3.2 Air pollutants

An air pollutant is a material in air or atmosphere which has harmful effects on humans. It affects negatively to the ecosystem. We have a different form of pollutants like solid particles, gases, or liquid droplets. Their origins are divided into two parts, man-made and natural. The pollutant can be categorized as primary and secondary pollutants.

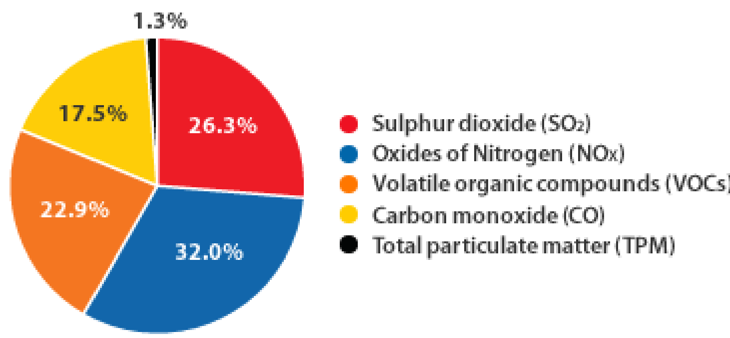

Figure 3. 3: Most common pollutants (askiitians.com, 2018)

3.2.1 Primary pollutants

Primary pollutants which are dominated pollutants are produced and spread by a volcanic eruption. Another one is carbon monoxide gas that is mostly generated from exhausts of a variety of vehicles. Although the factories processes which produce a massive amount of sulfur dioxide. Primary pollutants are as follows.

3.2.1.1 Carbon dioxide (CO_2)

This pollutant is one the most significant pollutants among others which is rapidly produced by many factors (Eldering et al, 2017) and it is harmful than others. Usually, whenever there is a discussion about CO2, it has been described as "the leading pollutant" (Seaton et al, 1995) and "the worst climate pollution" (Vaidyanathan, ClimateWire, Gayathri).

Carbon dioxide naturally exists excessively in the atmosphere and essential for plants. Nowadays, CO2 contributes about 410 parts per million (ppm) of the earth's atmosphere while it was 280 ppm in pre-industrial times (Sundquist & Keeling, 2009). Possible to claim billions of metric tons of CO2 are emitted annually by the burning of fossil fuels.

3.2.1.2 Sulphur Oxides (SO_x)

This pollutant and the most popular form of it which is SO2 is likely produced in many industrial processes by volcanoes. It is found in coal and petroleum. Therefore, using these fuels especially for the power system is much of environmental concern.

3.2.1.3 Nitrogen Oxides (NO_2)

Like before the popular one is Nitrogen dioxide, which is a chemical toxic gas which is the result of high temperature and other issues. It is a prominent pollutant in air pollution formation. This gas has sharp and biting odour characteristics.

3.2.1.4 Carbon monoxide (CO)

This pollutant is a colourless, odourless toxic gas. The production and spreading of carbon monoxide are from the combustion of natural gas, coal, or wood. Although, the exhaust of vehicles produces a lot of carbon monoxide into the atmosphere. Surveys in 2013 showed that vehicle traffic produces more than half of the carbon monoxide into the atmosphere. One gallon of gas produces more than 20 pounds of carbon monoxide into the earth (Hansen et al, 2013).

3.2.1.5 Volatile Organic Compounds (VOC)

VOCs are one of the prominent air pollutants. These pollutants are categorized into methane and non-methane parts. Researchers claim that methane (CH4) is very much of efficient gas for global warming enhancement. The fragrant non-methane (NMVOCs) benzene, toluene, and xylene are suspected cancer-causing agents and may prompt leukaemia with a delayed introduction.

3.2.1.6 Particulates Matter (PM)

Tiny particles of solid or liquid suspended in a gas are usually referred to Particulate Matters. Meanwhile, particles and gas together called aerosol. Some of the particles are originating from nature. For instance, volcanoes, dust storms, forest fires, sea spray etc. Furthermore, some of the human activities also generate a lot of aerosols like burning fossil fuels, power plants, and industrial processes etc. these particles are increased rapidly and caused health hazards like heart disease, lung function, and cancer and although asthma (Balmes & Sheppard, 1987).

There are other pollutants that have an effect on air which takes part to decrease the air quality. These pollutants are persistent free radicals, toxic metals, chlorofluorocarbons, Ammonia, Odours, and Radioactive pollutants.

3.2.2 Secondary pollutants

These pollutants are formed in the air as the consequence of primary pollutants react or we can say interact. These pollutants are not released directly. For saying, ground-level ozone is one of these pollutants. Not to be forgotten some of the pollutants can fit in both primary and secondary category. Secondary pollutants are divided into three categories.

- Particulate matters made from vaporous essential pollutants and mixes in photochemical smog. Smog is a sort of air contamination. Exemplary exhaust cloud results from a lot of coal consuming in a territory caused by a blend of smoke and sulphur dioxide.
- Ground level ozone (O3) shaped from NOx and VOCs. Ozone (O3) is a key constituent of the troposphere. It is additionally an essential constituent of specific areas of the stratosphere generally known as the Ozone layer.
- Peroxyacetyl nitrate (C2H3NO5) – similarly formed from NOx and VOCs.

3.3 Air pollution Sources

According to what has mentioned, there are multiple elements which are counted as responsible factors for released pollutants into the air. These factors or sources are categorized into two parts anthropogenic (man-made) and Natural sources.

Man-made sources:

- Stationary sources such as fossil fuel power, factories, waste incinerators, wood, crop waste and dung.
- Mobile sources like vehicles, marines and aircraft.
- Controlled Burn in agriculture
- Fumes took from varnish, aerosol sprays, paint, hairspray etc.
- Waste deposition in landfills, consequences to methane.
- germ warfare, rocketry, Nuclear weapons and toxic gases which are used in Military.
- Fertilized farmland which produces NO_x

Natural Sources:

- Dust from natural earth sources
- Methane, which is from animals' food
- Radon gas, coming from radioactive decay
- CO and Smoke from wildfires
- Vegetation
- Volcanic activity

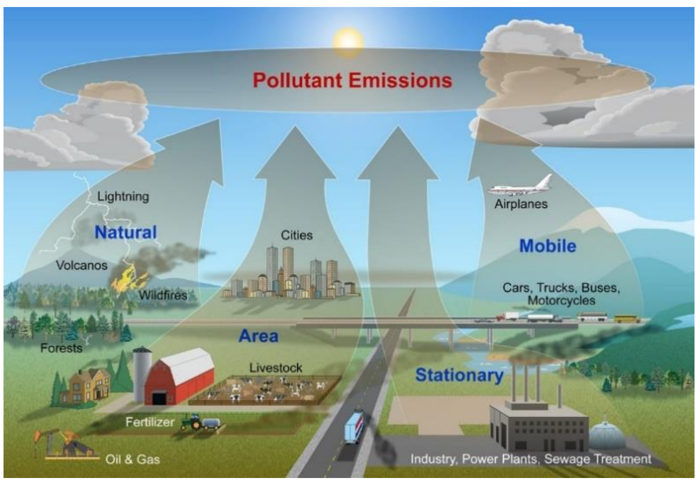

Figure 3. 4: Air pollution sources (nps.gov, 2018)

3.4 Air pollution effects

The air pollution affects the earth in various ways as we mentioned before. From health to agriculture and economics. Each of them should be addressed separately.

3.4.1 Health effects

The contribution of the above pollutants into our atmosphere creates harmful risks. There are many pollution-related diseases and even to the death of the humans. The air pollution causes many health problems, breathing hard, wheezing, coughing, asthma and cardiac problems. These things affect the human body and generally the body ecosystem (Boubel et al, 2013). To make it more precise the following is a list of the health effects of air pollution.

- Mortality
- Cardiovascular disease
- Lung disease
- Lung cancer
- Infants
- Central nervous system

3.4.2 Agriculture & Economic effects

This global challenge has some impacts on agriculture and other economic factors. The experiment shows in India that crop yields are reduced by half in most polluted areas. Meanwhile, according to a study by World Bank and the Institute for Health Metrics and Evaluation (IHME) at the University of Washington, air pollution costs the world economy 5$ trillion dollar each year due to the losses in productivity and quality of life (Bank, 2016). This was a brief insight into what it has brought to the world in the 21th century, the researcher believes that it has greater effects and harms in lower layers which needs to be discussed.

3.5 Reduction Strategies

The developments and revolutions in technology create opportunities to fight against air pollution. There are many strategies available to combat this problem (Fensterstock et al, 1971). The cause is those pollutants, which the task of a strategy is to decrease or eliminate the pollutants whether by replacing the functionality by something else or decreasing in the usage of whatever generate those pollutants (Hagevik, 1972). Some strategies briefly mentioned below.

- Reduction in using fossil fuels by replacing with other technologies
- Spreading Titanium dioxide which is able to reduce air pollution
- Transition to renewable energy
- Using different control devices

CHAPTER 4
MACHINE LEARNING TECHNOLOGIES

4.1 Introduction to machine learning

Machine learning is a first-class ticket to the most thrilling career in data predicting and analysing. It is an idea to learn from examples and specifications or experience data without being frankly programmed, without writing any code. We build a logic depends on the data given and we feed it in that genic algorithm. Machine learning also can be referred to the alteration in network systems that implement tasks related and linked with artificial intelligence systems. These tasks include recognition, diagnosis, prediction planning, and robot control system. We can say that the machine learning is training the computer for sure with different algorithms to test the machine in automatic intelligent data processing.

For example, in one kind classification algorithm we can put data in different groups. it can detect handwriting of alphabet, or identify faces in the image for example. machine learning is a field which raised from artificial intelligence (AI). It seeks to build better smart and intelligent machines. And the only way to achieve this task is to let the machine learn from itself and this sound similar to a little child learning from himself in human childhoods. Machine learning developed as a new ability. Now, machine learning exists in many divisions of technology that we didn't even realize it while we are using it. Machine learning is correlating with the study of the algorithms to increase the effectiveness of machine spontaneous through testing and training of that machine using the algorithm with different data.

ML improve and evolve rules that help the machine learning to process the similar conditions every relation efficaciously in the hybrid model. Understanding input variables, how its moving into vectors is such important thing. ML has minimized the manual job offers for people which may have a size for other jobs (Smola & Vishwanathan, 2008).

4.1.1 Artificial neural networks

ANN has shown that its very powerful pattern with strong classification and recognition capabilities by following the logic systems particularly the human brain. Artificial neural networks able to learn from experience presently. So, its the computer network systems that can process very large and intelligent tasks. ANN is a parallel system which fulfils the most complicated operations or tasks of realization in different fields of business industry and science. It predicts and detects without increasing the complexity of the problem. ANN has hidden layers in the middle of one input and one output that will process the information data to the next layer and each layer to the next layer by forwarding the result until it arrives at the final layer which is output layer. ANN is the most used algorithm nowadays, it’s the most popular machine learning algorithm in artificial intelligence. It uses particularly for different processing as FBP feed forward back propagation, NARX Nonlinear autoregressive exogenous model, each model with a different function from others in ANN. These effective machines solve complex problems every second. Therefore, claims that it made the people’s life much easier. (Yegnanarayana, 2009).

4.1.1.1 Neurons

Our brain includes set of biological neurons connected as network structures. We have interconnected set of neural networks realizing our thinking, reading, breathing, and motion. Some of these neural structures were at birth and other parts have been learned by experience. Artificial neural network proceeds these huge and complex tasks that fed into the layers of the neural network and process the data as an artificial neural network just like the neurons that operate in human brains for processing tasks. In ANN neurons, we should train them with old data to obtain the future forecasting data.

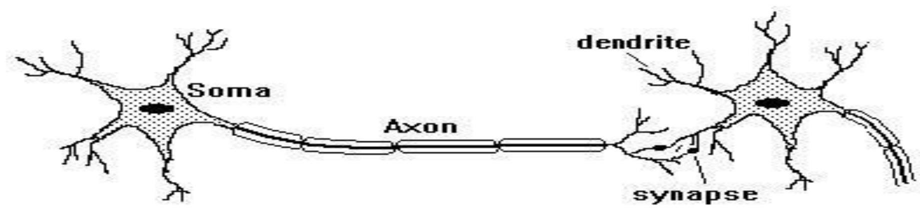

Figure 4. 1: Neuron scheme (Skorpil & Stastny, 2006)

Next, the testing and training data are holding out to test the result with other data to gain the difference by feeding the network system with numbers of neural network neurons, which we can say the Number of neurons is changeable and depend on processing complexity and to the data that we are going to feed it in the network system. Subsequently, depending on the output and input complication and the layers on the network system. Thus, the architecture may vary from one to another (Demuth, Beale, Jess, & Hagan, 2014).

4.1.1.2 Structure of Artificial neural network

Neural network includes a large number of units arranged in a series layer which are the artificial intelligence neurons.

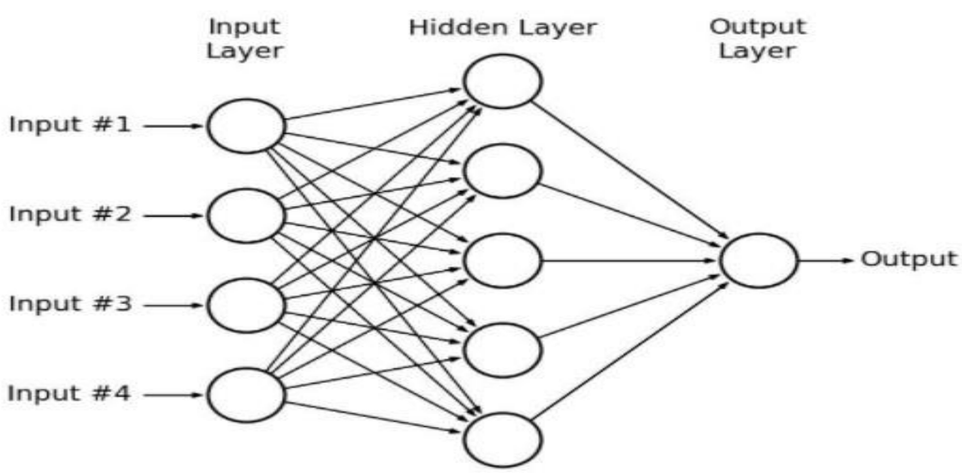

Figure 4. 2: Structure of ANN (Ahn, 2017)

INPUT LAYER: it is the first layer that contains artificial neurons which receive the input data from outside in order to learn to recognize or processes.

HIDDEN LAYER: the middle layers between the input layer and the output layer. The job of these layers is to process data and transform the input data through the network neurons to the output. For fineness and validity, the weights are continuously updated to the output of the hidden layer.

OUTPUT LAYER: the final layer in the structure of ANN contains units that respond to the data through learning to obtain the final result.

Most neural networks are fully linked and connected that means the hidden layer fully linked between each neuron in the next output layer and to the previous layer or input layer at first.

4.1.1.3 Weights

When neural network takes the large dataset, split data into tiny fragments, then transmit these fragments through all neurons. The neurons take the data, process them using the stored weight, then send the results to the output. So, in ANN architecture the information and data are stored in memory storages. The weight also modified at each step during the training, testing and validation. so, the output accuracy is carried out and the data is saved for any feature operation.

4.1.1.4 Feedforward neural network

It is also often called feedforward neural network or multilayer perceptron, it’s called feedforward because the data flow through the function that evaluated from the input. Feedforward neural network is a biologically impacted classification algorithm. It depends on a large number of normal neurons like in a process the units and orderly in a layer with all previous layers. The input layer process the data that receive and send the obtained result to the next layer. Each linked layer may have a different weight’s or strength. The result can gain through the processing of each layer. lastly, it can be gain from the output layer. any layer that is not an output layer or input layer is a hidden layer. The artificial neurons work like a human brain that process the input data work in the neurons. Neurons in the layers send the data or information through a channel called connection and the layer only connect to the previous layers.

4.1.1.5 Backpropagation algorithm

Backpropagation is an algorithm using gradient descent of artificial neural network. The method calculates with consideration of weight and the gradient of the error function. So, it's used to detect the errors in order to highlight the performance of the network using inputs. The accuracy obtained from the output and the number of neurons for validity checking.

Backpropagation is easy to understand and simple yet productive algorithm. With the calculation function, it consists of elements of n processes (Y.H.Zweiri, J.F.Whidborne, & L.D.Seneviratne, 2002).

$$Y = G(X, W) \tag{4.1}$$

The equation above describes that W is the error weight propagation matrix, X as an input vector, Y as an output vector.

By the equation 4.2 the later matrix is shown

$$\mathbf{W} = (w_1^{\mathrm{T}}, w_2^{\mathrm{T}}, \ldots, w_N^{T})^{\mathrm{T}} \tag{4.2}$$

The equation above w1, w2, w3…by equation below the individual vector are given

$$w_i = \begin{bmatrix} w_{i1} \\ w_{i2} \\ \vdots \\ w_{in} \end{bmatrix}, \qquad i = 1, 2, .., N. \tag{4.3}$$

4.1.1.6 Nonlinear autoregressive exogenous model (NARX)

NARX is a popular network model that identify and recognize the tasks. Forecasting and prediction can be done by NARX model. This model sends information to various layers of the network. NARX is feedback NN which is effective in predicting the accurate outputs result (Khamis, Nabilah, & Abdullah, 2014).

$$y(t+1) = f\,[y(t),\ldots,x(t\text{-}d_y+1);\ u(t\text{-}k),\ u(t\text{-}k\text{-}1),\ldots,\ u(t\text{-}k\text{-}d_u+1)] \qquad (4.4)$$

The above equation explains the algebraic expression of NARX.

$$y(t+1) = f\,[y(t);\ u(t)] \qquad (4.5)$$

The Nonlinear autoregressive exogenous model is used generally for the recognition tasks and for identification. The prediction may also be made effectively by using the Nonlinear autoregressive exogenous models.

NARX uses feedback connections (sending information from neurons to other neurons) in the various layers of the network to enhance the accuracy. It is based on the ARX model. The nonlinearity estimator comprises both nonlinear and linear functions that work on the model regressors to give the model output. The linear function is used to forecast the time series usually (Khamis, Nabilah, & Abdullah, 2014).

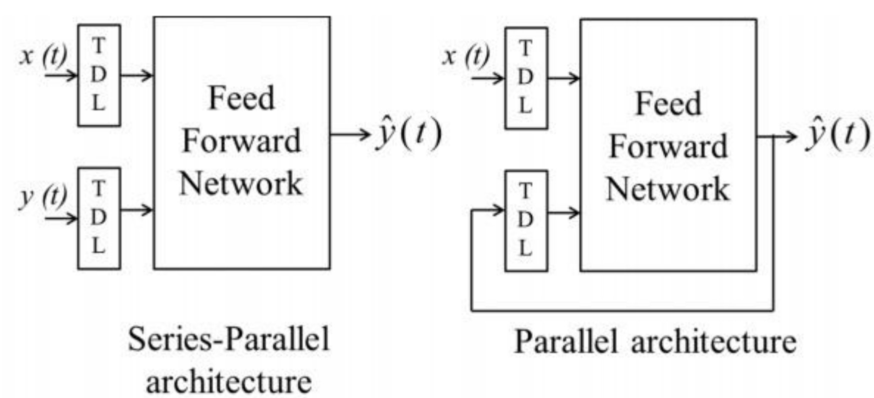

Figure 4. 3*:* Architecture of NARX neural network (Khamis, Nabilah, & Abdullah, 2014)

4.1.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is adaptive neuro-fuzzy inference system and an efficient ML algorithm. ANFIS is a hybrid system that integrates neural networks and fuzzy-logic. The hybrid algorithm aims to simplify computing of output desired result. The network also aims to reduce the complexity of the operation. The neurons in the hybrid algorithm work as nodes. It uses neurons for processing data. The adaptive network concept is using particular techniques to operate the desired outputs.

The result depends on updating inputs and their parameters. The node is a processing unit of the neuro-fuzzy. The ANFIS design the rules using different optimization techniques. For each operation NF system gives set of rules (the given rules depend on the input and output), the neuro-fuzzy system stores the information and data for feature processes (Wahyuni, Mahmudy, & Iriany, 2017).

4.1.2.1 Adaptive Neuro-Fuzzy Inference System architecture

We consider simply the fuzzy interference system, it basically has two inputs and one output and the rules that generated through ANFIS contains If and Then type Takagi and Sugeno rules as follow:

If x is A and y is B then z is f(x,y)

Where A and B the values of input variables. For Takagi and Sugeno type rules F(x,y) is a deterministic function. The output in this system has a linear collection of input variables created by constant term. The weighted average for each rule is the last output. The hybrid system obtains the rules while operating the data for precise and accurate results. The rules help to process future information and data for good efficiency. The ANFIS structure is as follows.

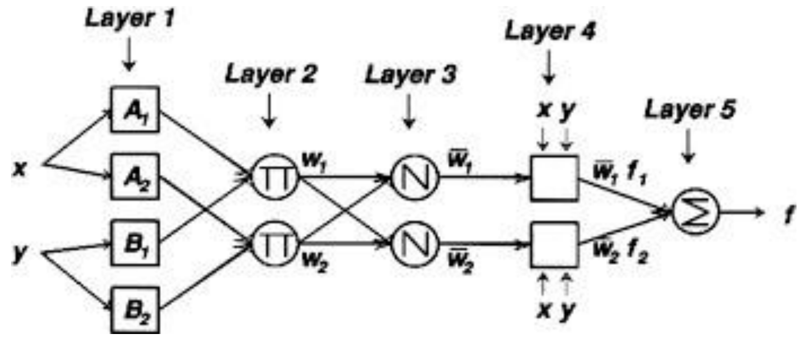

Figure 4. 4: ANFIS architecture (MRINAL BURAGOHAIN)

The descriptions of the layers are given below:

Layer 1: each i–th node calculates the output of first layer as follows

$$O_{1,i} = \mu A_i(x), \quad for\ i = 1,2\ and \qquad (4.6)$$
$$O_{1,i} = \mu B_{i-2}(y), \ for\ i = 3,4$$

X is the input of i-th node, A_i is the linguistic variable. $\mu_{Ai}(x)$ is the membership function of x. Usually μ_{Ai} (x) is chosen as

$$\mu Ai\ (x) = \begin{cases} 0; & x \le ai\ or\ x \ge c \\ \frac{x-ai}{bi-ai}; & ai \le x \le bi \\ \frac{bi-x}{ci-bi}; & bi \le x \le ci \end{cases} \qquad (4.7)$$

X is the input, ai, bi, ci the premise parameter set.

Layer 2: Each node in this layer is a fixed node. The result obtained by this layer processes of the output of previous layer. The output is calculated as:

$$\mu_{A_i}(x) = exp\ \{-(\frac{x - c_i}{a_i})^2\} \qquad (4.8)$$

Layer 3: Each node in this layer is a fixed node. The output from ith node is the normalized firing strength. Each ith node calculates the ratio of ith rule's firing strength to the sum of firing strengths of all rules. given by:

$$O_i^3 = \overline{w_i} = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \tag{4.9}$$

Layer 4: each layer has nodes and those nodes are adaptive with the function bellow:

$$O_i^4 = \overline{w_i} f_i = \overline{w_i}\,(p_i x + q_i y + r_i), \quad i = 1, 2 \tag{4.10}$$

Where the output is wi and {pi, qi, ri } is the consequent parameter.

Layer 5: calculating the overall output from one fixed node that in this layer alone as the last of all incoming signals.

$$O_i^5 = overall\ output = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{4.11}$$

(MRINAL BURAGOHAIN).

4.1.2.2 ANFIS learning algorithm

ANFIS architecture has five layers. The first layer is nonlinear while the fourth layer is linear. The fourth and the first layers include parameters that can be updated continuously time to time. Thus, the first and the fourth layer need to be updated through learning algorithm. ANFIS system is one that can train two layers at the same time (Faulina & Suhartono, 2013). To train the layer 1 and 4 together ANFIS uses descent gradient through propagation the errors backwards. Through the hybrid system the ANFIS network is trained. The error can be measured by the equation below:

$$E_p = \sum_{m=1}^{\#(L)} (T_{m,p} - O^L_{m,p})^2 \tag{4.12}$$

$T_{m,p}$ is the *m*th element of the *p*th target

$O^L_{m,p}$ is the *m*th element of our output vector

The overall error is as follows,

$$E = \sum_{p=1}^{P} E_p \tag{4.13}$$

CHAPTER 5
SIMULATION

5.1 Data processing

The purpose of this research is the prediction of air pollution using elements of AI. The hourly data Bursa and Istanbul cities are taken from the metrological department of Turkey. The Bursa and Istanbul data includes considerable attributes of air pollution like ozone, particulate matter 10, particulate matter2.5 and nitrogen dioxide. The data studied and tested in this research will be the input for predicting air pollution. The prediction will be performed with machine learning algorithms such as neural network, Adaptive Neuro-Fuzzy Inference System, and nonlinear autoregressive exogenous models.

The air pollution index is suggesting as a new regulation, which is a function of various sub-indicators like S.1 and S.2. The indicators include pollutants like PM10, PM2.5, NO2 and OZONE(O3) that are categorized by the mass concentration respectively. It is important to list the pollutants for network inputs.

$$AQI = max\{IAQI_1, IAQI_2, IAQI_3, \ldots IAQI_n\} \tag{S.1}$$

$$IAQI_n = \frac{IAQI_{Hf} - IAQI_{Lo}}{BP_{Hf} - BP_{Lo}} (C_n - BP_{Lo}) + IAQI_{Lo} \tag{S.2}$$

In equation 5.1 IAQI1, IAQI2, IAQI3, IAQI4, and IAQI5 are the values for each pollution. The pollution raw that we measured are converted into separate AQI value for each pollutant (particulate matter PM10, particulate matter 2.5, ozone O3, nitrogen dioxide NO2) using the standard equation 1 by EPA. The highest value of these pollutants AQI is reported as the AQI for that day. For big cities, states and local agencies are required to report the AQI level on that city for Public health and awareness (Index, A. Q. (2009). USA: EPA).

In equation 2 the air quality index is a piecewise linear function of the pollution concentration (Elshout, Bartelds, Heich, , & Léger, 2012) [32] (CAQI–2012).

AQI = the Air Quality index,

C_n= pollutant concentration,

BP_{low}=concentration breakpoint that is ≤ Cn ,

BP_{Hf}= concentration breakpoint that is ≥ Cn ,

I_{low}= index breakpoint corresponding to BP_{low},

I_{high}= index breakpoint corresponding to BP_{high} .

Qualitative name	Index or sub-index	Pollutant (hourly) density in μg/m^3			
		NO_2	PM_{10}	O_3	$PM_{2.5}$(optional)
Very low	0–25	0–50	0–25	0–60	0–15
Low	25–50	50–100	25–50	60–120	15–30
Medium	50–75	100–200	50–90	120–180	30–55
High	75–100	200–400	90–180	180–240	55–110
Very high	>100	>400	>180	>240	>110

Figure 5. 1: Five CAQI ranges and AQI "CAQI Air quality index — (Comparing Urban Air Quality across Borders)

Istanbul city:

Istanbul is a big city in Turkey that straddles Europe and Asia across the Bosporus strait. The city is located in the northwest of the country, with 1539 km2 and 15 million population. Air pollution has reached highest level in recent years. Specialy in the district of Yenibosna, Kadıköy, and Esenyurt. The pollution has been increasing in these region continously and the amount of pollution and source of PM10 is increasing becousep the environmental consequences of urban transportation process are being ignored. It is visible that Istanbul has an air pollution problem which becomes chronic. PM 10 and PM 2.5 are the most important pollutants in Turkey and exceed the limit values. The local authorities need to take measurements by evaluating this risk map and informing the public (Akkoyunlu, & Erturk, 2002).

	Limit Values (24-hour average)		Permitted excesses each year (total number of days)	
	EU	Turkey	EU	Turkey
PM10 (Particulate Matter)	50 µg/m³	90 µg/m³	35 times/year	----

Figure 5. 2: Comparison of the limit values between Europe and Istanbul determined by Turkey and Europe

Bursa city:

Bursa in Marmara Region, located in north-western Anatolia. It is the fourth most populous city in Turkey and one of the most industrialized metropolitan centres in the country with 1036 km2 and 1.8 million population. One of the biggest problems in Bursa is the air pollution due to dense industry and high population which is over 1.2 million and more things like home heating industry and transportation. Those are crucial elements that affect the air of the city. There are three industrial districts in Bursa including car transportation, leather energy, chemistry metal and also the residence in this district use firewood and coal for heating in the winter season. Air pollution is one of the most important environmental problems, located in the western part of Turkey, during the winter periods, PM 10 and PM 2.5 are the most important pollutants in Turkey and exceed the limit values (Tasdemir , Cindoruk, & Esen. 2005) [36].

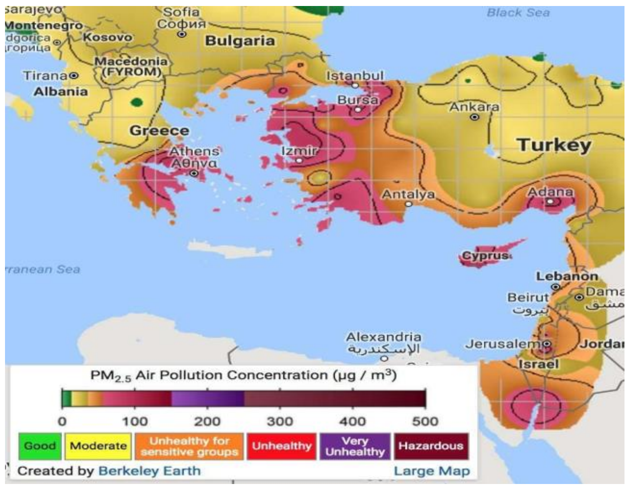

Figure 5. 3: Air quality in the Istanbul and Bursa regions (berkeleyearth)

The air quality system in Bursa needs to focus on informing people about the air quality level in their city. To make people life carefully, the government should be more serious in avoiding pollution and decrease the number of dense industries inside the city. Appropriate steps must be taken to minimize the air pollution level. So, analysing the air pollution prediction for this region become too important.

5.1.1 Data Pre-Processing for Istanbul region

The hourly air pollution dataset collected from 2017.8.1-2018.8.1 for Istanbul. The attributes of air pollution that are to estimate in this study have been observed with their varying trends. The data was taken from the metrological department of Turkey describes air pollution cycle for each month in hourly bases.

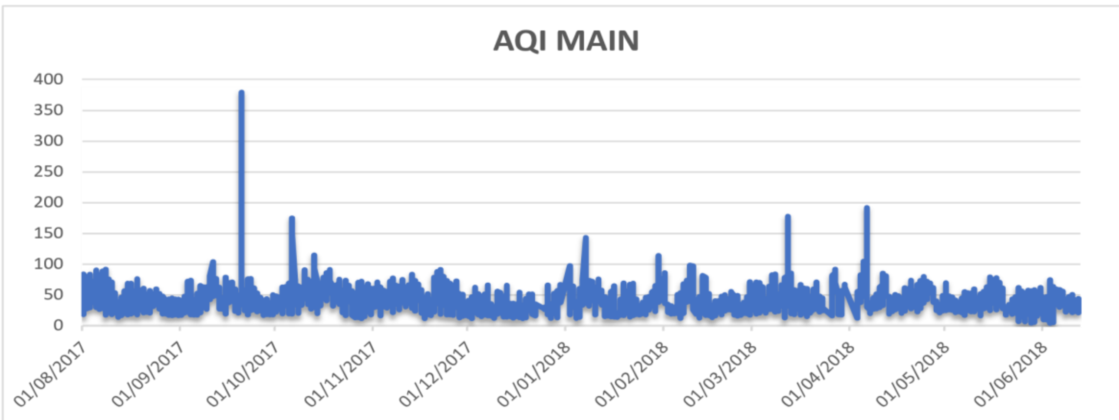

Figure 5. 4: Distribution of air quality in Istanbul

Figure 5.4 above is showing air pollution cycle for each month hourly, while climate change is the global process. It has a deep impact that can affect societies temperature. The pollution increases directly connected with poor quality, which can affect our health, hearts and exacerbate cardiovascular disease. Average AQI in Istanbul is 40.16, which is between 25 and 50 that is just low level depending on air quality index level in Europe. And the peak value is 378.9 which is very high pollution level.

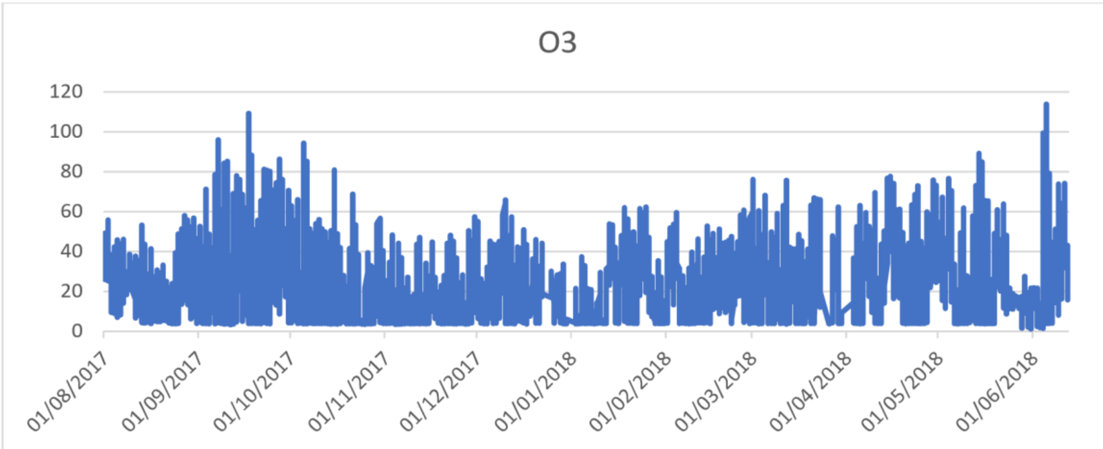

Figure 5. 5: Distribution of ozone for the Istanbul region

Figure 5.5 show the ozone hole is a severe depletion of Earth's protective ozone layer. It can be caused by chemical reactions that take place firstly on the surface of polar stratospheric clouds, ice particles or liquid droplets which form at high altitudes in extreme cold. Above graph explaining air pollution and how ozone dealing with a time period for each month for Istanbul pre-data processing. As we see here each big ozone cycle start from September and end in December in the winter season. So the air pollution affect to the temperature variations in the upper atmosphere. In colder years, more ice particles will freeze, allowing more chemical destruction of the ozone layer. To quantify the impact of changing emission of ozone-depleting chemicals we should understand the variability of the ozone hole that leads us to predict feature of the ozone hole. Average O3 in Istanbul is 23 which is between 0 and 60 that is very low level depending on the air quality index level in Europe. And the peak value is 113.9 between 120 and 180 which is a medium level of air quality.

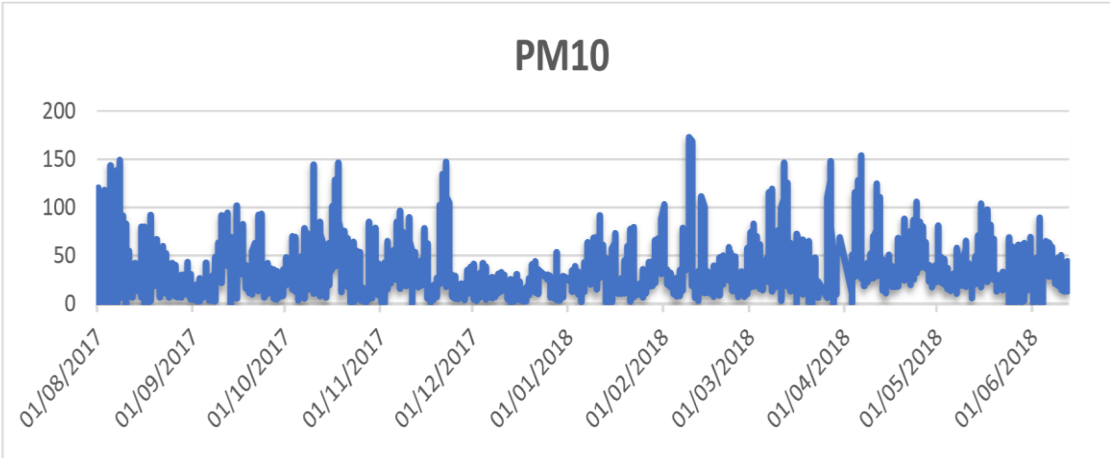

Figure 5. 6: Distribution of particulate matter 10 in the Istanbul region

Figure 5.6 shows the distribution of particulate matter, PM10 in the diameter of 2.5 to 10 micrometres. It is caused by crushing or grinding operations and dust stirred up by transportation vehicles on roads. Particulate matter its tiny particles which are about 30 times smaller than hair width and small enough to get inheld past our defensive nose hair and into our lungs. The graph shows us PM10 cycle in Istanbul. In the summer when tourism start, the PM10 pollution is starting. A lot of transportation and vehicles in the streets increases the pollution in Istanbul. Average PM10 in Istanbul is 33.3 which is between 25 and 50 in low level depending on the air quality index level in Europe. And the peak value is 171.96 between 90 and 180 which is on high pollution level of air quality.

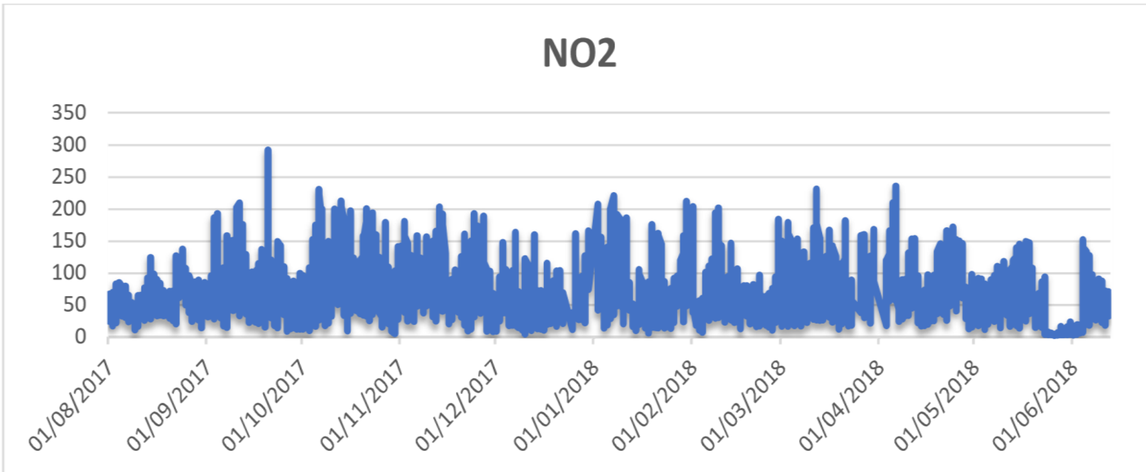

Figure 5. 7: Distribution of nitrogen dioxide in Istanbul region

Figure 5.7 shows nitrogen dioxide distribution, NO2 is burning of fossil fuels from vehicles and mostly impacts the health of the people. They are generated from diesel or gasoline trucks loaders mobile cranes. NO2 effects to the healthy people with adverse respiratory effects including airway inflammation and increases respiratory symptoms in people with asthma. Graph showing how NO2 is varying between 2017-2018. Average NO2 in Istanbul is 65.8 which is between 50 and 100 on low level depending on the air quality index level in Europe. And the peak value is 291.18 between 200 and 400 which is high pollution level of air quality.

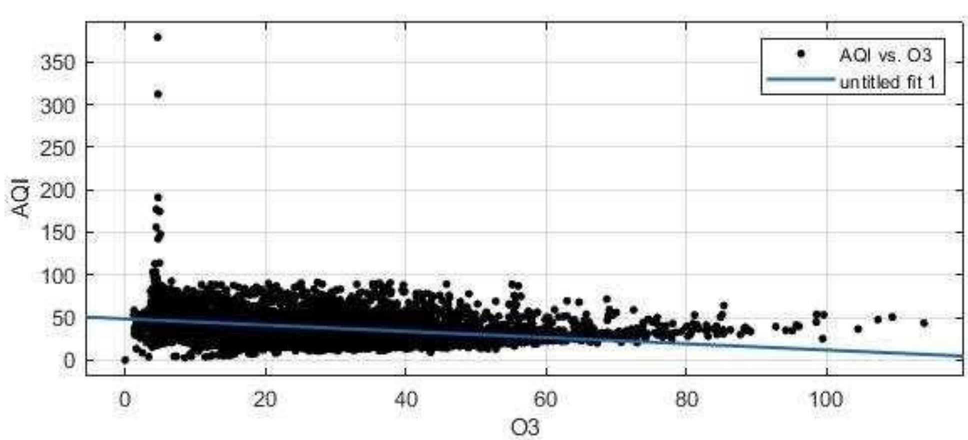

Figure 5. 8: Correlation between air quality and ozone

Figure 5.8 above explains the correlation between AQI and O3. Correlation is 0.35 which is not higher between O3 and AQI. This shows that O3 affecting air pollution quality but not so much as other pollutants.

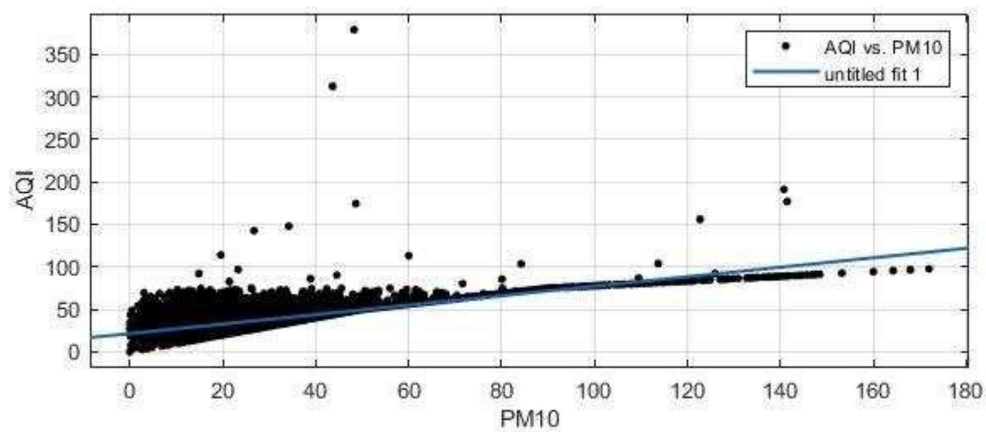

Figure 5. 9: Correlation between air quality and particulate matter 10

Figure 5.9 shows a correlation between AQI and PM10 which is equal 0.73. This value is close to 1, which demonstrate strong relationship between AQI and PM10. This means that the air pollution is strongly affecting with a particular matter.

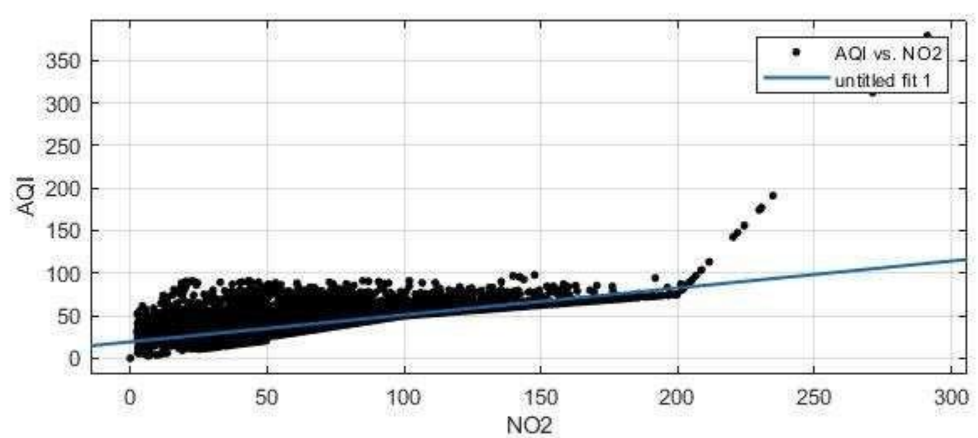

Figure 5. 10: Correlation between air quality and nitrogen dioxide

Figure 5.10 highlights the correlation between air pollution index and nitrogen dioxide. The correlation is 0.66 which is close to 1. So, the graph shows that NO2 is affecting the air pollution in Istanbul.

5.1.2 Data Pre-Processing for Bursa region

The hourly air pollution dataset collected from 2017.8.1-2018.8.1 for Bursa. The attributes of air pollution that are to estimate in this study have been observed with their varying trends. The data was taken from the metrological department of Turkey and it's showing the air pollution cycle for each month hourly.

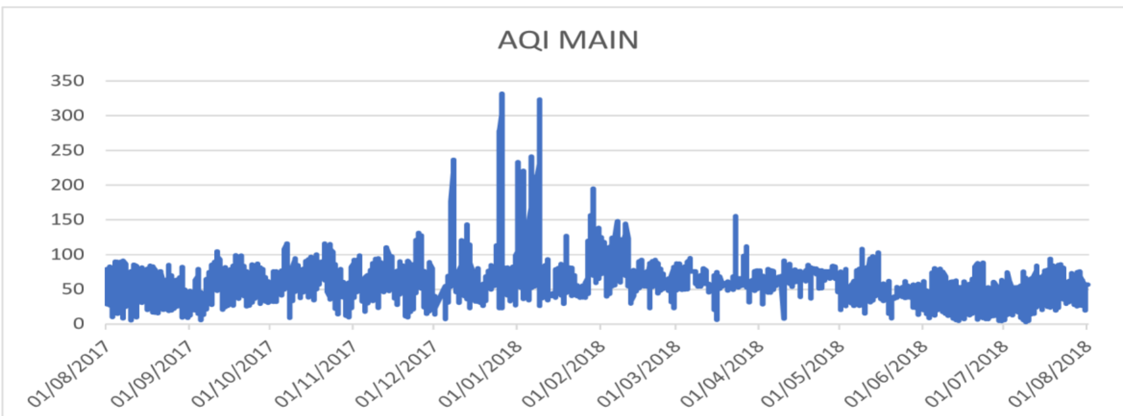

Figure 5. 11: Distribution of air quality in Bursa region

Figure 5.11 shows the distribution of air pollution in last 2 years for Bursa. The air pollution index data explain the changing pollution cycle. Bursa is crowded city and it’s always going bigger that makes it worse. While climate change is a global process. It has a deep impact that can affect societies temperature increasing are directly connected with poor quality. Which can affect our health and hearts and exacerbate cardiovascular disease. It shows that Bursa has a high value of air pollution but highest pollution values in winter and it almost very high value of pollution that arrive to very high pollution level. The time that people use unhealthy heaters and many other pollutants things. Average AQI in Bursa is 57.32 which is between 50 and 75 on medium level depending on the air quality index level in Europe. And the peak value is 331.75 which is very high pollution level.

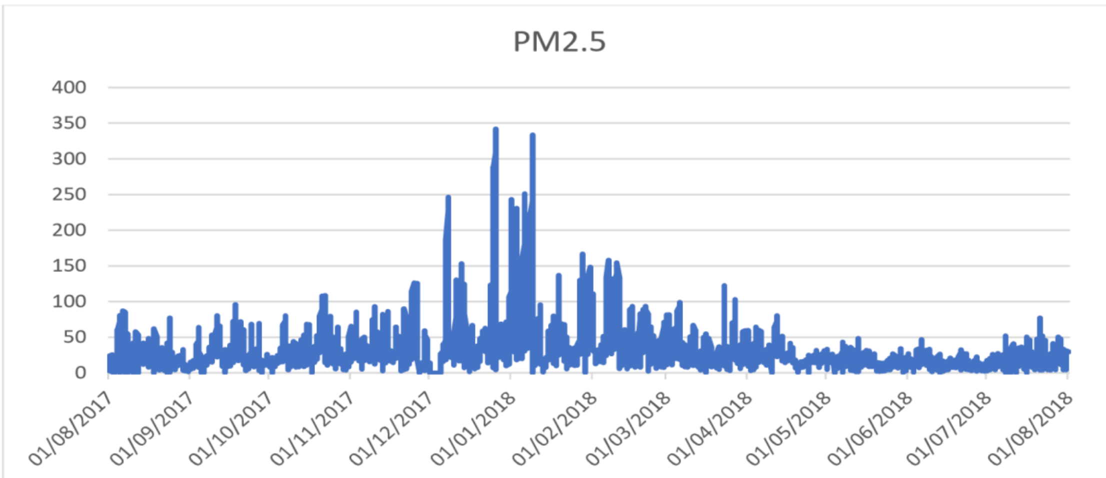

Figure 5. 12: Distribution of particulate matter in Bursa region

Figure 5.12 shows particulate matter distribution, it shows the particulate matter that has diameter of fewer than 2.5 microammeters, which is about 3% the diameter of a human hair. They include motor vehicles, residential wood burning, airplanes, forest fires, and volcanic eruptions. PM2.5 pollution is mostly happening in winter when people use unhealthy heaters by burning plastic and another pollutant emission. Average PM2.5 in Bursa is 28.4 which is between 25 and 50 that is just in low level depending on the air quality index level in Europe. And the peak value is 341.75 more than 180 which is very high pollution level of air quality.

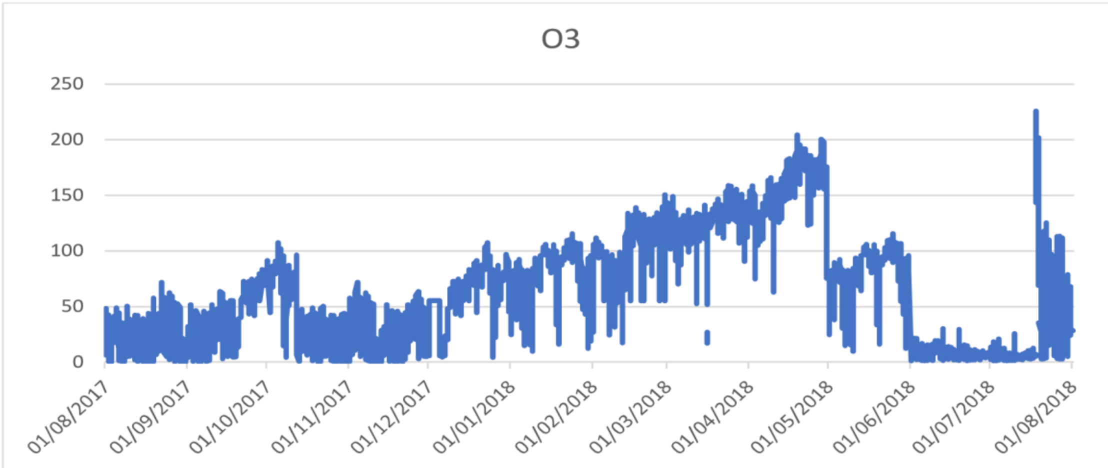

Figure 5. 13: Distribution of ozone in Bursa region

Figure 5.13 shows the ozone distribution; the ozone hole is a severe depletion of Earth's protective ozone layer. it can be caused by chemical reactions that take place firstly on the surface of polar stratospheric clouds, ice particles or liquid droplets which form at high altitudes in extreme cold. Above graph explaining air pollution and how ozone dealing with a time period for each month for Bursa pre-data processing. To quantify impact of changing emission of ozone-depleting chemicals we should understand variability of the ozone hole that leads us to predict feature of ozone hole. Average O3 in Bursa is 64.9 which is between 60 and 120 that is very low level depending on the air quality index level in Europe. And the peak value is 255.7 between 180 and 240 which is medium level of air quality.

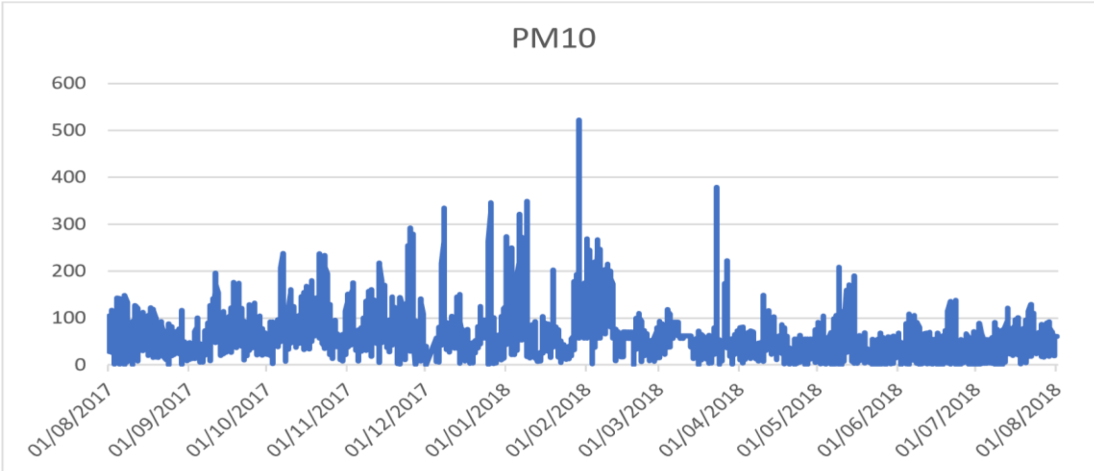

Figure 5. 14: Distribution of particulate matter 10 in Bursa region

Figure 5.14 shows particulate matter distribution, PM10 that is in diameter of 2.5 to 10 microammeters. It causes of grinding or crushing operations and dust pollutant stirred up by vehicles transportation on roads. Its small enough to get inheld past our defensive nose hair and into our body and lungs. It shows us PM10 cycle in Bursa. Average PM10 in Bursa is 57.12 which is between 50 and 90 that is medium level depending on the air quality index level in Europe. And the peak value is 522.34 which is very high pollution level of air quality.

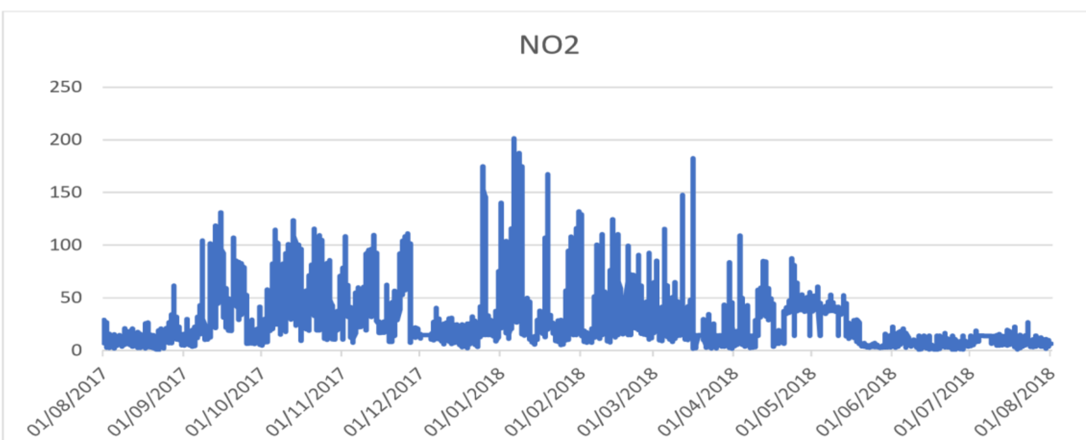

Figure 5. 15: Distribution of nitrogen dioxide in Bursa region

Figure 5.15 shows nitrogen dioxide distribution NO2 is burning of fossil fuels from vehicles. It’s most concern because of the impact on health. Being in a place can NO2 effect you with adverse respiratory effects including airway inflammation in healthy people and increased respiratory symptoms in people with asthma. Its showing how NO2 is varying in 2017-2018. Also. Average NO2 in Bursa is 26.13 which is between 0 and 50 that is very low level depending on the air quality index level in Europe. And the peak value is 201.4 between 200 and 400 which is high pollution level of air quality.

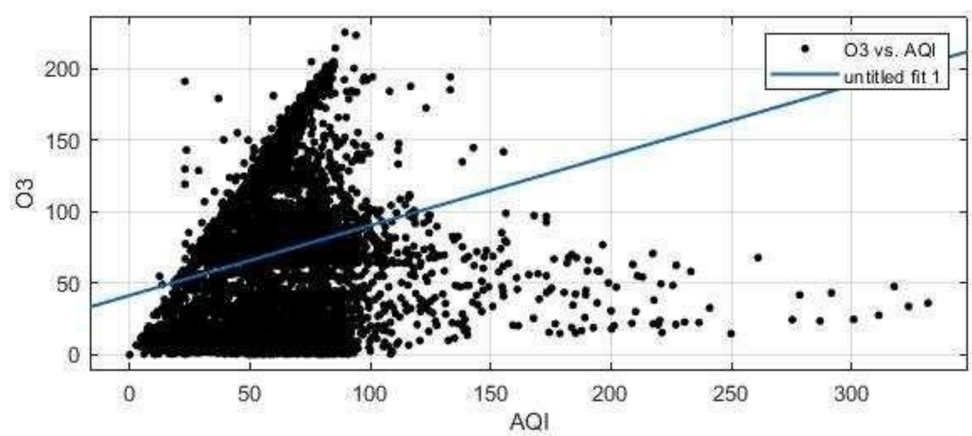

Figure 5. 16: Correlation between air quality and ozone in Bursa

Figure 5.16 highlight and explain correlation between AQI and ozone O3. Between O3 and AQI correlation 0.18 and its far from 1. Which it has a low relationship. It’s not that much that we can say that the air quality index related to ozone value in Bursa.

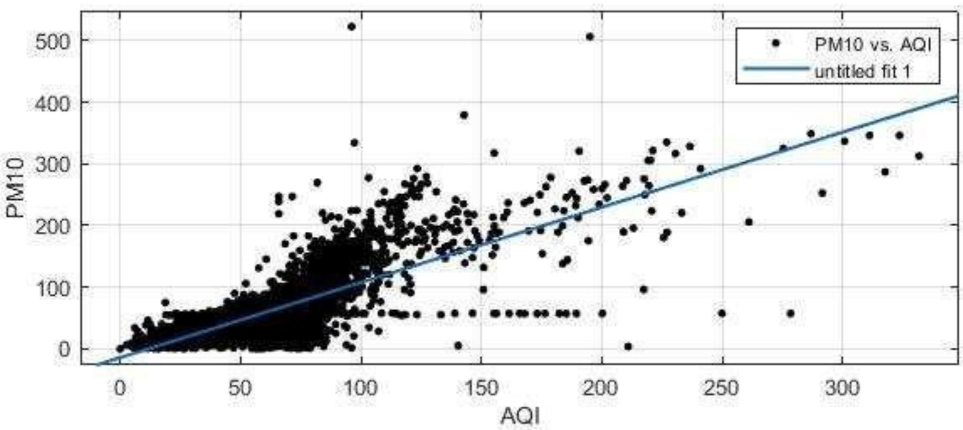

Figure 5. 17: Correlation between air quality and particulate matter 10 in Bursa

Figure 5.17 above shows correlation between air quality index and particulate matter. Correlation between AQI and PM10 is 0.835 which is near to 1. which means air pollution in Bursa is very related to PM10. Which is relationship with AQI and it's affecting the air pollution and correlation shows air pollution is increasing with increasing the particular matter strongly.

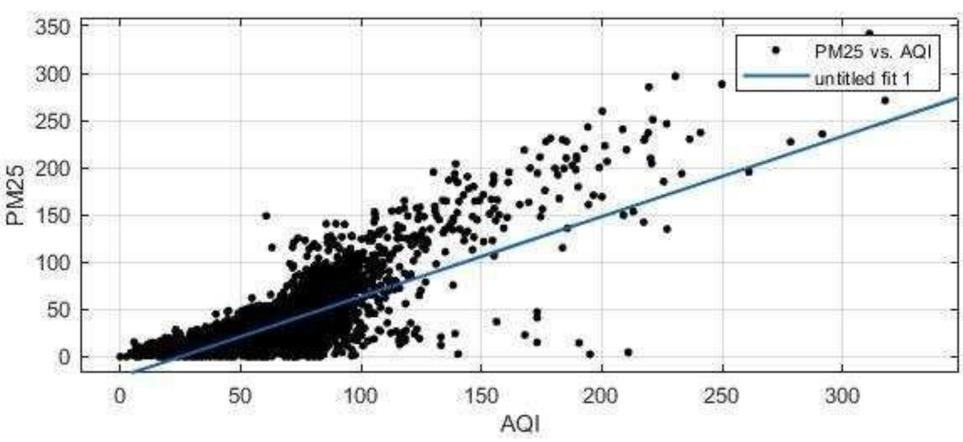

Figure 5. 18: Correlation between air quality and particulate matter 2.5 in Bursa

Figure 5.18 above shows that correlation between the air quality index and particulate matter. Correlation between AQI and PM2.5 is 0.832 near to 1. which is very relationship with AQI and its one of the most affecting AQI. The correlation shows that air pollution is increasing with increasing particular matter.

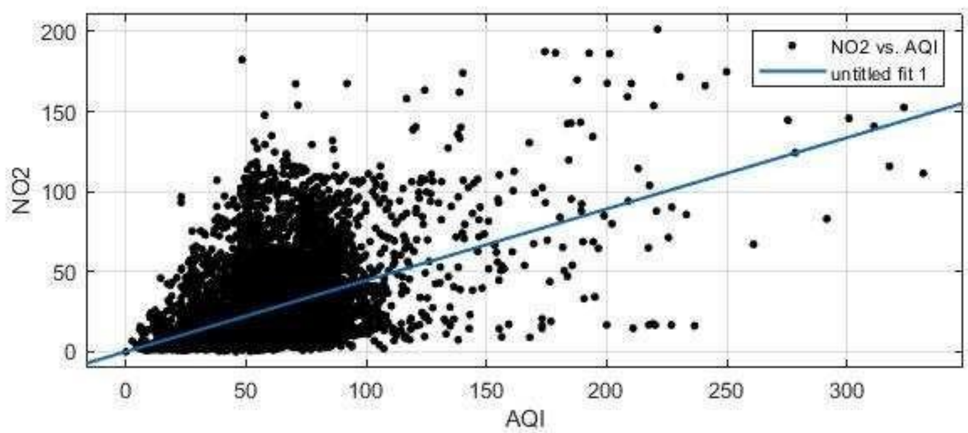

Figure 5. 19: Correlation between air quality and nitrogen dioxide in Bursa

Figure 5.19 above shows that correlation between air quality index and nitrogen dioxide. Correlation between AQI and NO2 is 0.49. Correlation between AQI and NO2 is not affecting that much as PM 10 and PM2.5 on air pollution in Bursa. But air pollution also relation to nitrogen dioxide that affecting AQI.

5.2 Flowchart for air pollution prediction

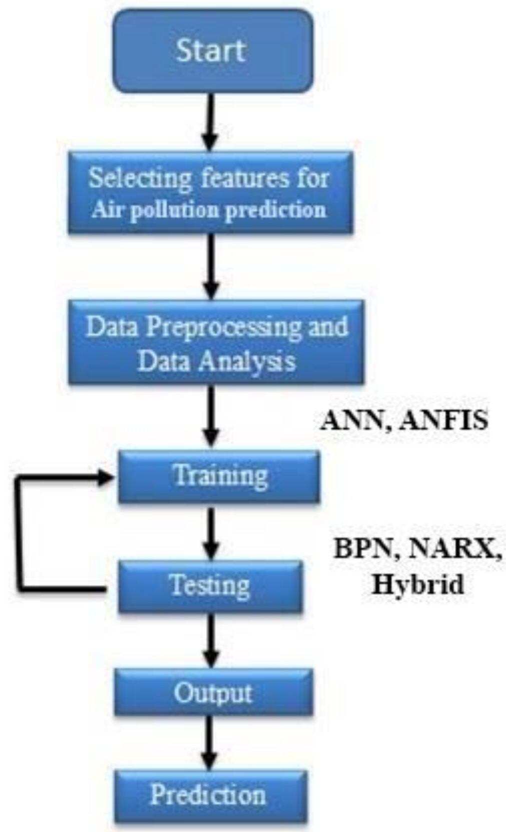

Figure 5. 20: Flowchart of predicting using ANFIS, NARX and BP

Fig 5.20 depicts the flow chart of the program used for prediction of air pollution. The second flow select the parameters that air pollution will be predicted. Second block implements testing pre-processing, sealing operation, the training and testing implement the design of prediction system. In this chart, the algorithm is implemented as an optimization for weight and sloping of ANFIS to avert the local minimum and improve the forecasting accuracy. The root-mean-square- deviation. To look for the better fitness function the AQI RMSE of the predicted daily mean is taken. the schematic representation of the hybrid algorithm is above.

5.3 Selection of the inputs and output data

In this research, the dataset is collected from the metrological department to test input and the output. The input data for Istanbul include the exact values that were measured hourly by metrological stations in Turkey, the attributes that we have in this data are particulate matter 2.5 (PM2.5), particulate matter 10 (PM10), ozone (O3) and nitrogen dioxide (NO2). The output data of the air pollution index for 2017 and 2018, the attributes PM2.5, PM10, O3, and NO2 are selected for 2017 and 2018 for Bursa by the metrological department. There is no air pollution season for each year but most of the pollution happening in Turkey in the summer season. Thus, to forecast the air pollution for these two cities and as we mentioned the parameters will be reflected as considerable inputs to predict the pollution quality.

5.4 Feature Extraction

This research based on the hourly values of the air quality index parameters particulate matter 10, particulate matter 2.5, ozone, nitrogen dioxide. The parameters are explained as:

5.4.1 PARTICULATE MATTER 2.5

The PM that has a diameter of smaller than 2.5 microammeters, which is about 3% the diameter of a human hair. The pollution can be caused by the power plants, motor vehicles, forest fires, residential wood burning, agricultural burning. As much it’s small that can increase chances of human and animals inhaling them into their bodies, they are able to bypass the nose to go deep in lungs and somehow it can penetrate the circulatory system. It’s a total of all liquid and solid particles pendent in the air, which many of them are dangerous and hazardous. The two main components of particulate matter in Europe are sulphate and organic matter (World Health Organization 2003).

5.4.2 PARTICULATE MATTER 10

As we said the PM2.5 have a diameter less than 2.5 but PM10 have a diameter between 2.5 and 10. And it caused by dust stirred up by transportations and vehicles on roads. The most majority of this pollution was attributable to fugitive dust from roads. PM10 is smaller than the hair with about 10 times it can also inheld our nose and enter in lungs. We can say that particulate matter 10 is hazardous (World Health Organization. 2003).

5.4.3 OZONE O3

O3 is founded in air that we are breathing, ozone may be bad or good depending on where it's happening. In our research, we discuss about the bad ozone. Its released by cars powers plants, industrial boilers, refineries, and chemical plants. Ozone pollution is more happening summer or in warmer months. It can cause people lungs disease such as asthma and emphysema. By increasing the ozone pollution, the air quality index becomes very hazardous (Shindell, Wong, & Rind. 1997).

5.4.4 NITROGEN DIOXIDE NO2

NO2 is created as a result of road traffic and other fossil fuel burning processes. The pollution caused by motor vehicles and, in some places, by energy generators. It's a part of a group of gaseous air pollutants. Study on people populations explains that long-term exposure to NO2 levels actually observed in Europe, and this may reduce lung function and rise level of the risk of respiratory symptoms such as sharp bronchitis and cough and phlegm, especially in children. (World Health Organization. (2003).

5.5 Training, Testing and Validation

We obtained air pollution data set from the metrological department of Turkey. The next step is to use database for training NNBP back-propagation neural network algorithm, NARX nonlinear autoregressive artificial neural network, and ANFIS. The inputs data has been trained, tested, and validated with MATLAB 2018 software programme by using neural network tools and hybrid system tools. Predicting start from training then testing of the inputs with agreeable error is carried out. The neural network tool standard holds seventy percent for training the data

30% for testing. But in ANFIS tool request to manual attaching of the inputs data for train, test and validation. ANFIS hold 70% for training and 30% for testing and checking.

5.6 Artificial Neural Network

There are different features of NN but the most known is that NN can learn from monitoring data sets. The artificial neural network is used as a random function parataxis tool. That can evaluate the cost function and obtain the solution. ANN takes data samples, not the entire data to have the best result. NN it can save our time and money for running and operate the big and heavy data. Artificial neural network includes three layers they are linked together. The first layer contains inputs neurons and these neurons after processing sending the data to the next hidden layer than to the last layer which is the output layer. Training of the artificial neural network includes the air pollution of network parameters.

5.6.2 Applying back-propagation neural network model for Istanbul

BPNN is used for prediction of pollution Istanbul region to take out the best accuracy for training, validation, and testing from inputs data and outputs data.

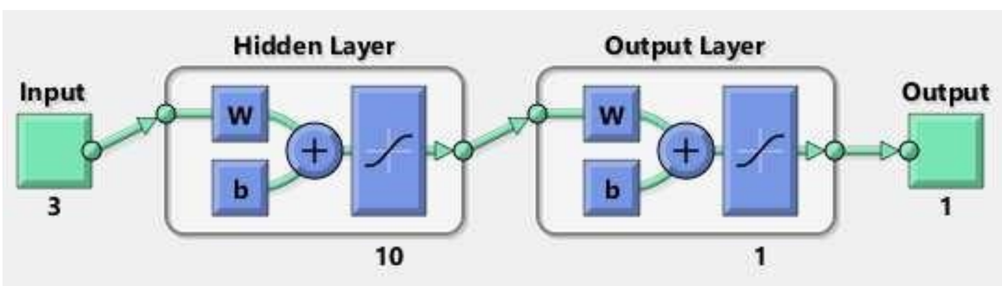

Figure 5. 21: Backpropagation network architecture for Istanbul

Figure 5.21 highlights the BPNN architecture sued for Istanbul region. In BPNN architecture have 3 inputs two hidden layers and one output layer. The prediction was done using particulate matter 10, ozone O3 and nitrogen dioxide NO2, particulate matter 2.5 for predicting the air quality index is not available in the metrological department of Turkey. The BPNN designed using 10 neurons in 2 hidden layers. The output is used for prediction air quality index.

Levenberg Marquardt (trainlm) is used for training. The algorithm can update the biases and the weights of NN. As gives good accuracy and high effectiveness.

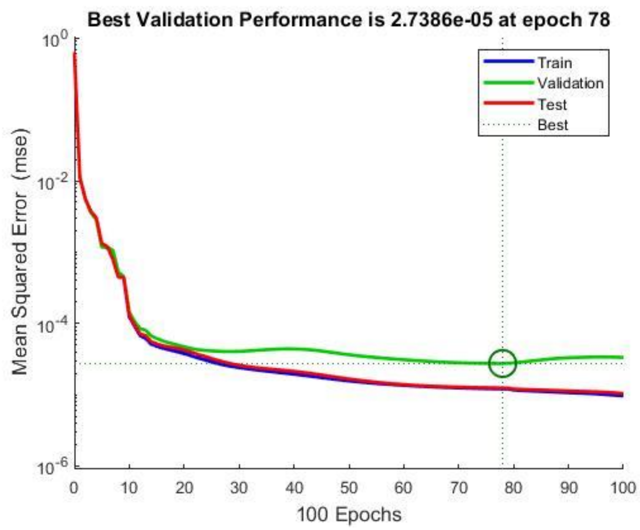

Figure 5. 22: Performance of using NNTOOL

For training the data we used 100 epochs. The back-propagation result for training testing and validation is listed in the figure.

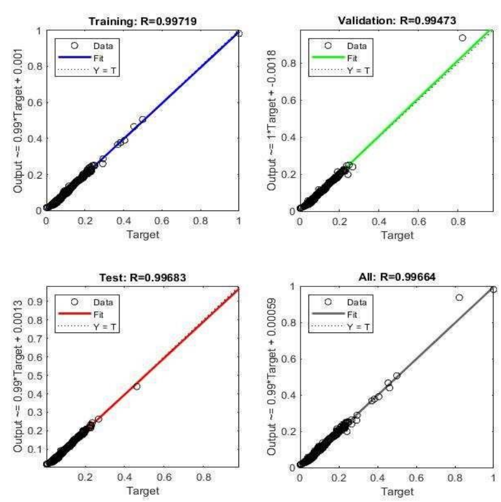

Figure 5. 23: Back propagation regression using NNTOOL

Figure 5.23 highlights the training, testing and validation accuracy; the accuracy of training is 0.99 and it's so close to 1 which is very accurate and exact result. The testing accuracy was 0.99. The validation accuracy is 0.99.

5.6.3 Applying NARX neural network model for Istanbul

The nonlinear autoregressive exogenous (NARX) network architecture contain an input layer, two hidden layers and an output layer. The NARX network has a feedback connection. In this network, there are 3 inputs particulate matter 10, ozone o3 and nitrogen dioxide, particulate matter 2.5 data set is not available in the metrological department of Turkey. The best-fitted number of neurons was 10 neurons for NARX. The obtained result of output is a prediction of air quality index of Istanbul region. Tapped delay lines as we see below TDL that are; 0:1 and 1:2 at inputs side. Concurrently reducing the complexity of the network by sorting the predicted values during the training phase.

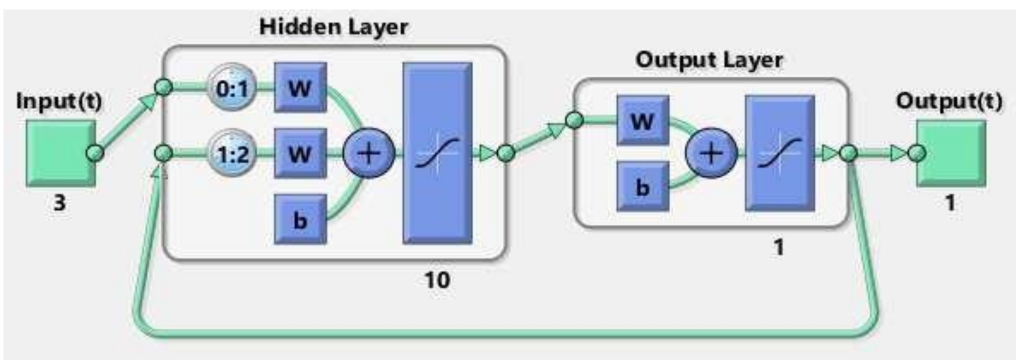

Figure 5. 24: Shows nonlinear autoregressive exogenous neural network architecture for Istanbul

Figure 5.26 highlight, using the Levenberg Marquardt (trainlm) function we trained the data using a nonlinear autoregressive exogenous algorithm. The gradient that we achieved is 4.26 with 100 epochs. The Mu value is 1e-08. we have the validation checks is equal to 33 out 100 epochs.

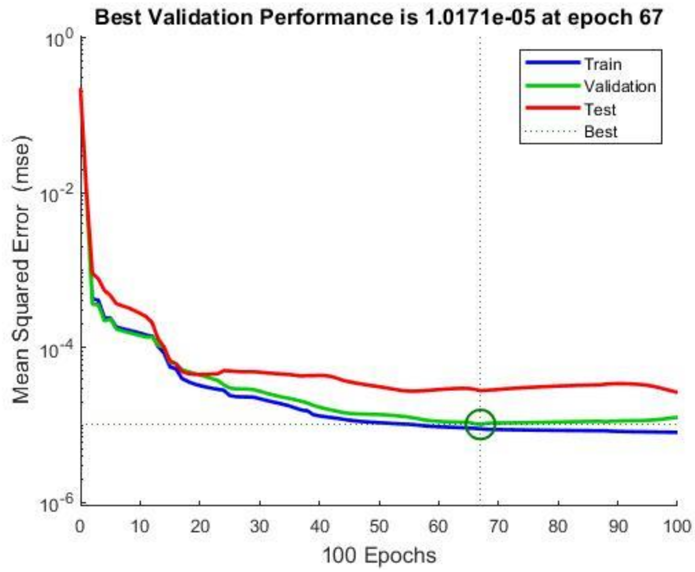

Figure 5. 25: Performance using NARX

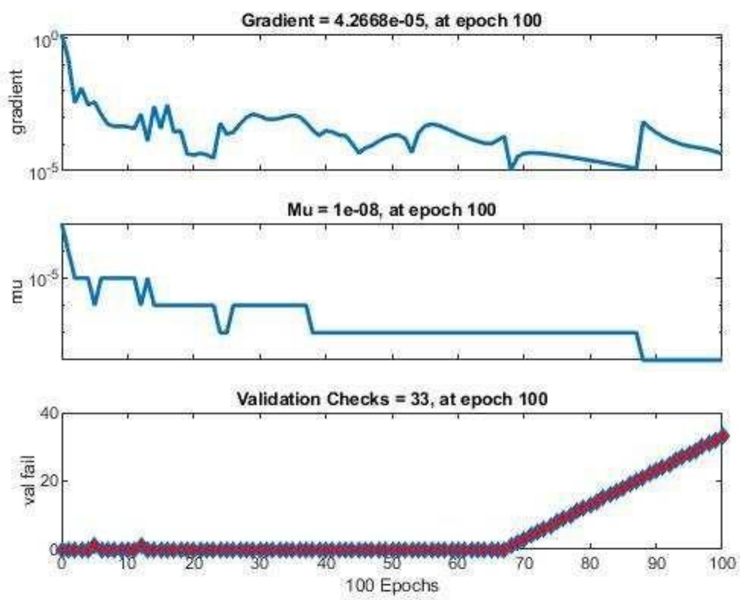

Figure 5. 26: Training state using nonlinear autoregressive exogenous for Istanbul

5.6.4 Applying back-propagation neural network model for Bursa

BPNN is used for prediction of pollution Bursa region to take out the best accuracy for training, validation, and testing from inputs data and outputs data.

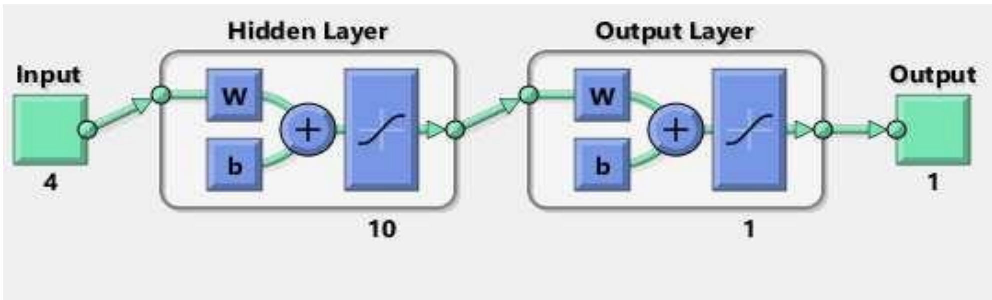

Figure 5. 27: Shows back propagation network architecture for Bursa

Figure 5.27 highlights the BPNN architecture used for Istanbul region. In BPNN architecture have 3 inputs two hidden layers and one output layer. The prediction was done using particulate matter 10, ozone o3 and nitrogen dioxide, particulate matter 2.5 for predicting the air quality index is not available in the metrological department of Turkey. It’s also an optional parameter for predicting the air quality index. The BPNN designed using 10 neurons in 2 hidden layers. The output is used for prediction air quality index. Levenberg Marquardt (trainlm) is used for training. The algorithm can update the biases and the weights of NN. As gives good accuracy and high effectiveness.

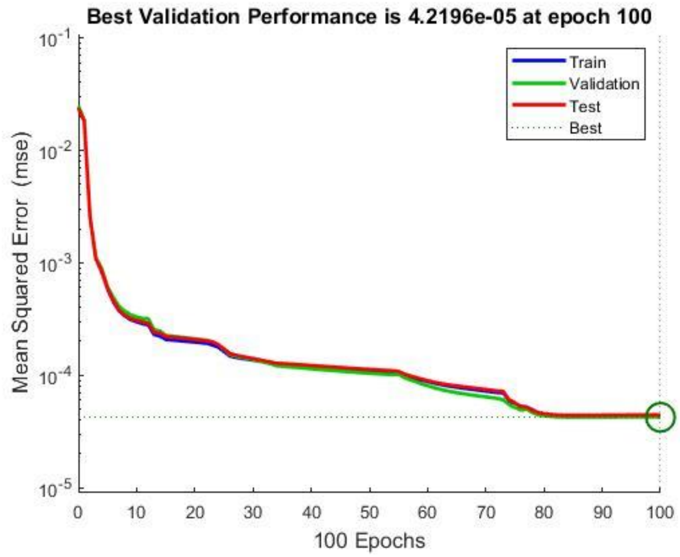

Figure 5. 28: Performance using NNTOOL

For training the data we used 100 epochs. The back-propagation result for training testing and validation is listed in the figure.

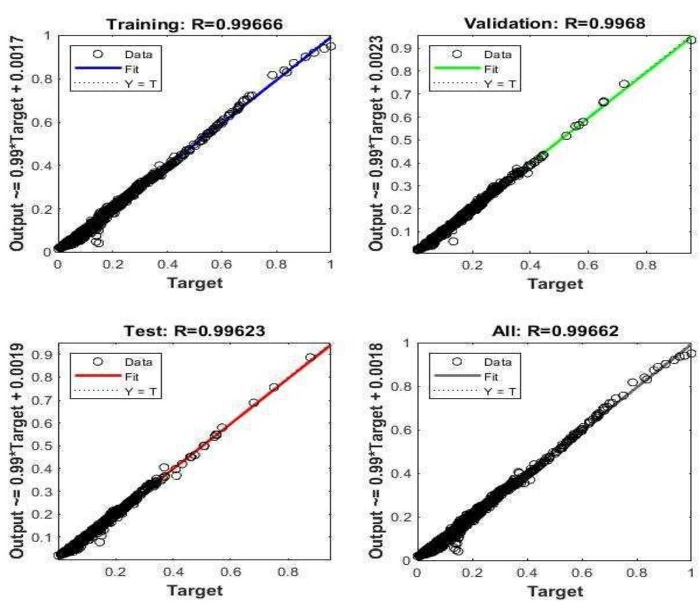

Figure 5. 29: Back propagation regression using NNTOOL

Figure 5.29 highlights the training, testing and validation accuracy; the accuracy of training is 0.99 and it's so close to 1 which is very accurate and exact result. The testing accuracy was 0.99. The validation accuracy is 0.99.

5.6.5 Applying NARX neural network model for Bursa

The nonlinear autoregressive exogenous (NARX) network architecture contain input layer, two hidden layers and an output layer. The NARX network has a feedback connection. In this network, there are 4 inputs particulate matter 10, ozone o3 and nitrogen dioxide, particulate matter 2.5 data set is not available in metrological department of Turkey. The best-fitted number of neurons was 10 neurons for NARX. The obtained result of output is prediction of air quality index of Istanbul region. Tapped delay lines as we see below TDL that are; 0:1 and 1:2 at inputs side. Concurrently reducing the complexity of the network by sorting the predicted values during the training phase.

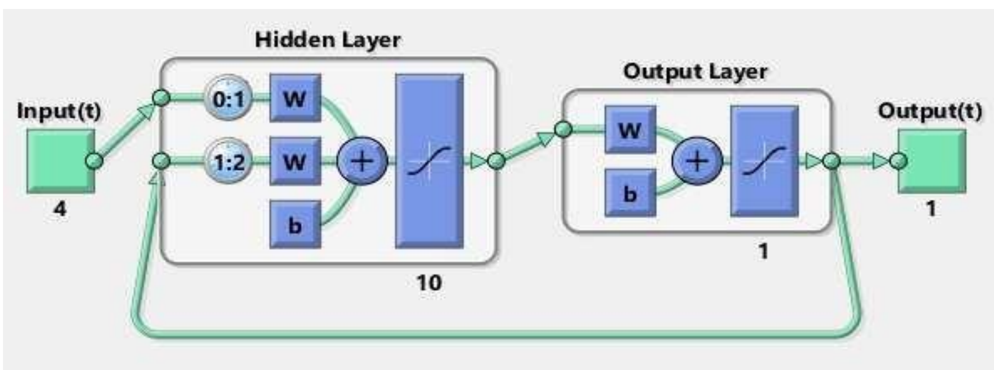

Figure 5. 30: Shows nonlinear autoregressive exogenous neural network architecture for Bursa

Figure 5.32 shows that using the Levenberg Marquardt (trainlm) function we trained the data using a nonlinear autoregressive exogenous algorithm. The gradient that we achieved is 0.00022 with 100 epochs. The Mu value is 1e-06. we have the validation checks that is equal to 0 out 100 epochs.

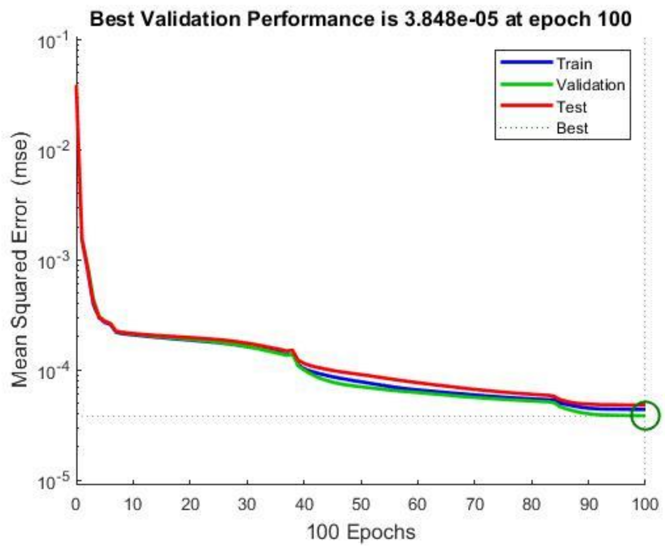

Figure 5. 31: Performance using NARX

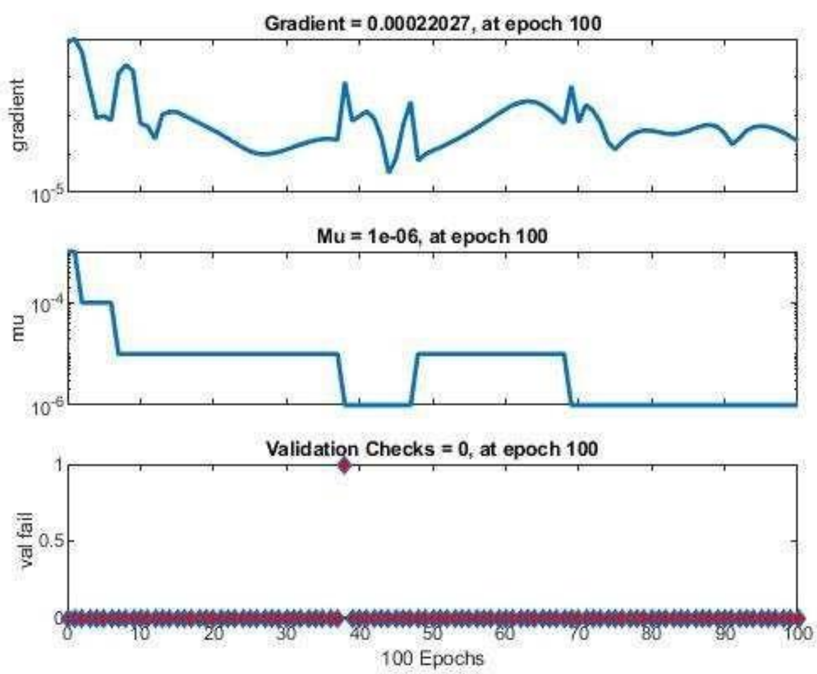

Figure 5. 32: Training state using nonlinear autoregressive exogenous for Bursa

5.7 ANFIS

The adaptive neuro-fuzzy inference system is focusing on predicting air quality index using data from the metrological department of Turkey from the past years 2017 - 2018. Forecasting the air quality index is nonlinear system. ANFIS is suitable through training this algorithm. Therefore, the network generated and tested with different FIS algorithm and with confirmed error tolerance and epochs. The parameters are analysed to have better RMSE.

5.7.1 Applying ANFIS for prediction air pollution of Istanbul

The ANFIS architecture used for Istanbul region. The data trained using ANFIS model for each city separately. The data parameters through ANFIS algorithm is manually fed, and its specified 70% for training and 30% for testing. The prediction was done using particulate matter 10, ozone o3 and nitrogen dioxide NO2.

ANFIS algorithm trained using 10 epochs and 0 error tolerance. The generated FIS is gauss2mf type and FIS linear type. According to this parameters, inputs data and output data, trained tested and checked. The error obtained using ANFIS model is 0.0022 which it definitely accurate for making the air quality forecasting.

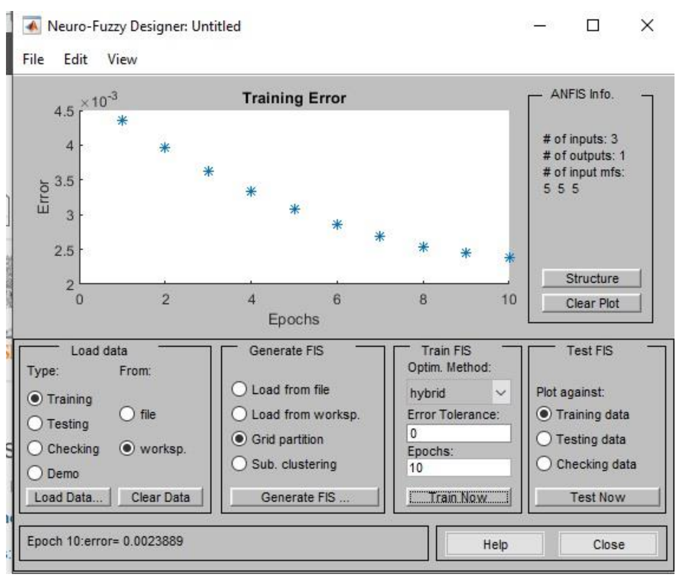

Figure 5. 33: Performance using ANFIS

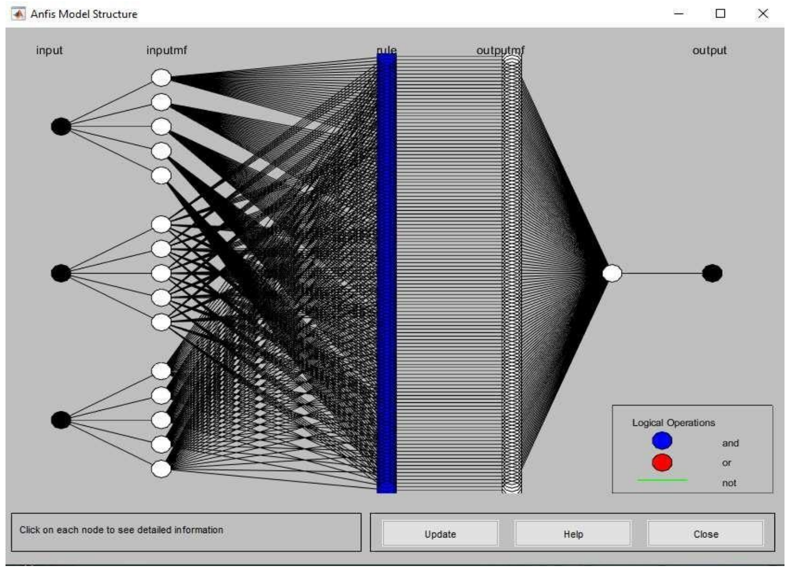

Figure 5. 34: This figure explains ANFIS architecture for Istanbul

Figure 5. 35: This figure explains ANFIS rules viewer for Istanbul

After feeding ANFIS model with training data, testing data. The ANFIS automatically produce the rules as we see in the figure 5.35. Improving the rules by the input data. It obtains the output data as air quality index prediction. Figure 5.35 shows that each input, it has a various set of rules based on it.

5.7.2 Applying ANFIS for prediction air pollution of Bursa

The ANFIS architecture used for Bursa region. The data trained using ANFIS model for each city separately. The data parameters through ANFIS algorithm is manually fed, and its specified 70% for training and 30% for testing. The prediction was done using particulate matter 10, ozone o3, nitrogen dioxide NO2, and particulate matter 2.5.

ANFIS algorithm trained using 10 epochs and 0 error tolerance. The generated FIS is gauss2mf type and FIS linear type. According to this parameters, inputs data and output data, trained tested and checked. The error obtained using ANFIS model is 0.003847 which it definitely accurate for making the air quality forecasting.

Figure 5. 36: Performance using ANFIS

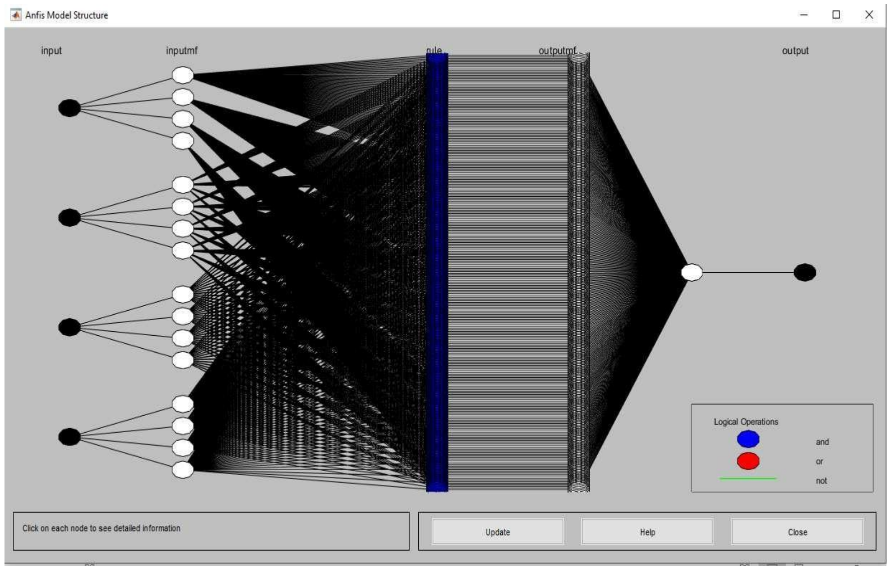

Figure 5. 37: This figure explains ANFIS architecture for Bursa

After training and testing data, ANFIS automatically produce the rules. It obtains the outputs data as air quality index prediction.

Figure 5. 38: This figure explains ANFIS rules viewer for Bursa

The Table 5.1 below explain the training, testing and validation results using backpropagation NNBP for Istanbul and Bursa.

Table 5. 1: Training, testing and validation results using BP model for Istanbul and Bursa

Location	Neurons	Epoch	Inputs	Output	Function	Training	Testing	Validation
Istanbul	10	100	O3, PM10 and NO2.	AQI	Trainlm	0.9971	0.9968	0.9947
Bursa	10	100	PM2.5, O3, PM10 and NO2.	AQI	Trainlm	0.9966	0.9962	0.9968

Table 5.2 below explain the training, testing and validation results using none linear autoregressive exogenous NARX for Istanbul and Bursa.

Table 5. 2: Training, testing and validation results using NARX model for Istanbul and Bursa

Location	Neurons	Epoch	Inputs	Output	Function	Gradient	Mu	Validation checks	Accuracy
Istanbul	10	100	O3, PM10 and NO2.	AQI	Trainlm	4.62	1e-08	33	0.9971
Bursa	10	100	PM2.5, O3, PM10 and NO2.	AQI	Trainlm	0.0002	1e-06	0	0.9959

The Table 5.3 below explain the training, testing and validation results using ANFIS model for Istanbul and Bursa data.

Table 5. 3: Training, testing and validation results using ANFIS model for Istanbul and Bursa

Location	**MF type**	**Epoch**	**Inputs**	**Output**	**Function**	**Error tolerance**	**Obtained error**	**Accuracy**
Istanbul	Linear MF	10	O3, PM10 and NO2.	AQI	gauss2mf	0.00	0.0022	0.9986
Bursa	Linear MF	10	PM2.5, O3, PM10 and NO2.	AQI	gauss2mf	0.00	0.0038	0.9975

CHAPTER 6
DISCUSSION AND RESULT

The thesis has devoted the use of artificial intelligence elements for forecasting the air quality index. The hourly data were taken for 2017-2018 for two different locations- Istanbul and Bursa cities from the Metrological department of Turkey. ANFIS, backpropagation neural network (NNBP) and nonlinear autoregressive exogenous (NARX) models are used to design air pollution prediction problem. After training, validation and testing process are performed in order to make comparative analysis and to select best prediction model.

Table 6. 1: R squared

City	**Number of neurons**	**R^2**		**R^2 HYBRID**
		BPNN	**NARX**	
Istanbul	10	0.9933	0.9945	0.9973
Bursa	10	0.9933	0.9931	0.9951

Table 6.1 shows R squared (R^2) obtained for Istanbul and Bursa. The most efficient model is with the R squared result equal to 0.9973 for Istanbul and 0.9951 for bursa, which obtained with the ANFIS model.

Table 6. 2: Root-mean-square-error (RMSE) after training and testing

City	Number of neurons	RMSE		RMSE HYBRID
		BPNN	NARX	
Istanbul	10	0.0038	0.0035	0.0022
Bursa	10	0.0065	0.0066	0.0038

Table 6.2 shows root mean square error value (RMSE) obtained for Istanbul and Bursa. The most efficient model is with the RMSE result equal to 0.0022 for Istanbul and 0.0038 for Bursa, which obtained with the ANFIS model.

6.1 Predicted and Actual data for Istanbul

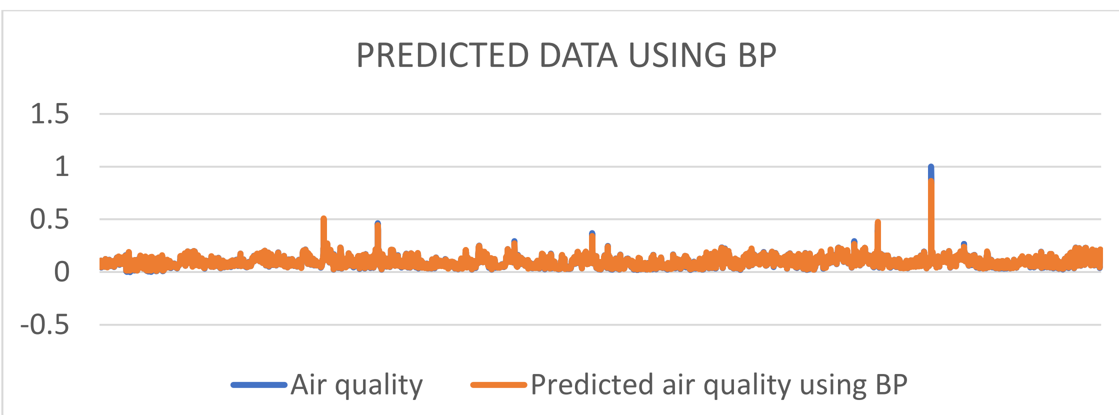

Figure 6. 1: Comparison of actual and predicted data for Istanbul using Backpropagation Neural Network

Figure 6.1 above shows the comparison of actual historical air quality index and predicted air quality index using Backpropagation Neural Network for Istanbul. There is some similarity and accuracy between the actual and predicted data.

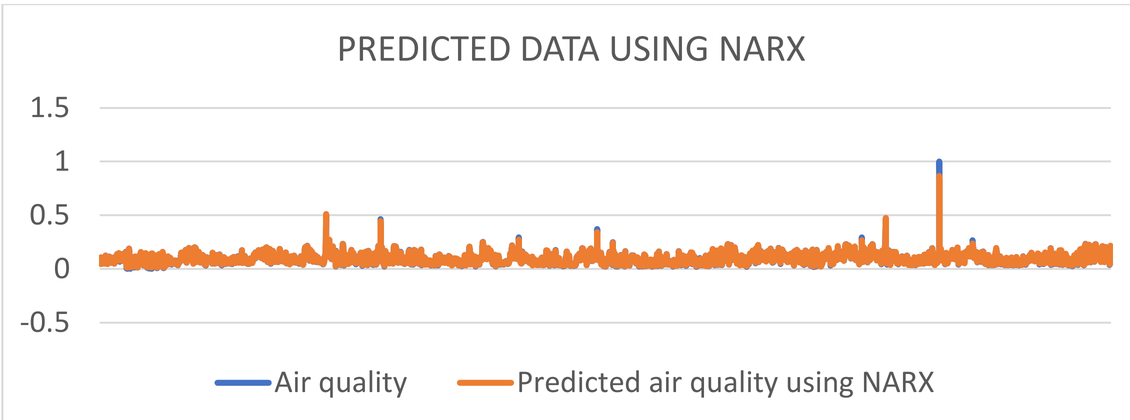

Figure 6. 2: Comparison of actual and predicted data for Istanbul using nonlinear autoregressive exogenous

Figure 6.2 shows the comparison of the actual air quality index and predicted air quality index using NARX Neural Network for Istanbul. There is some similarity and accuracy between the actual and predicted data and it's close to NNBP result.

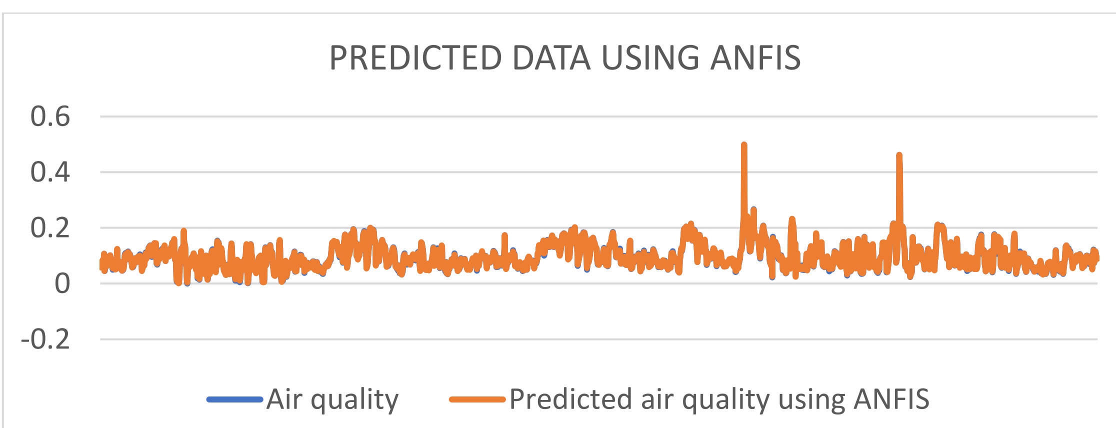

Figure 6. 3: Comparison actual and predicted data for Bursa using ANFIS

Figure 6.3 shows the comparison of actual air quality index and predicted air quality index using ANFIS. There is close relation between the actual and predicted data with maximum accuracy.

6.2 Predicted and Actual data for Bursa

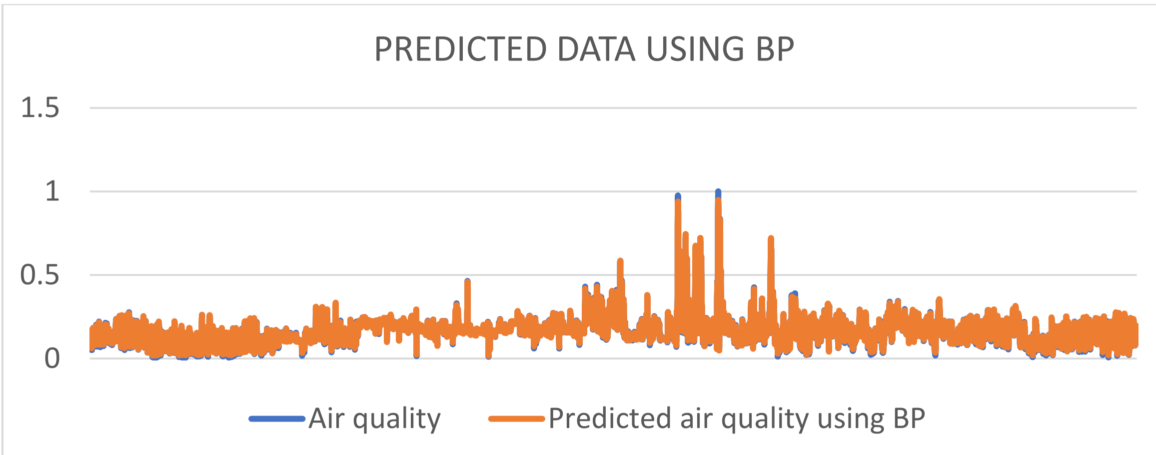

Figure 6. 4: Comparison actual and predicted data for Bursa using Backpropagation Neural Network

Figure 6.4 shows the comparison of actual and predicted air quality index using Backpropagation Neural Network for Bursa. There is some similarity and accuracy between the actual and predicted data.

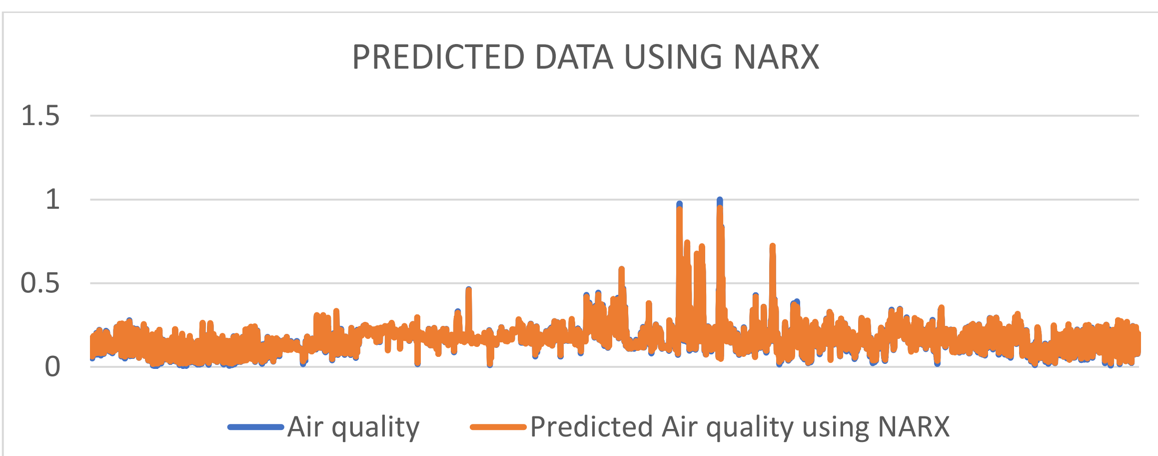

Figure 6. 5: Comparison actual and predicted data for Bursa using nonlinear autoregressive exogenous

Figure 6.5 shows the comparison of actual and predicted air quality index using NARX Neural Network for Bursa. There is some similarity and accuracy between the actual and predicted data.

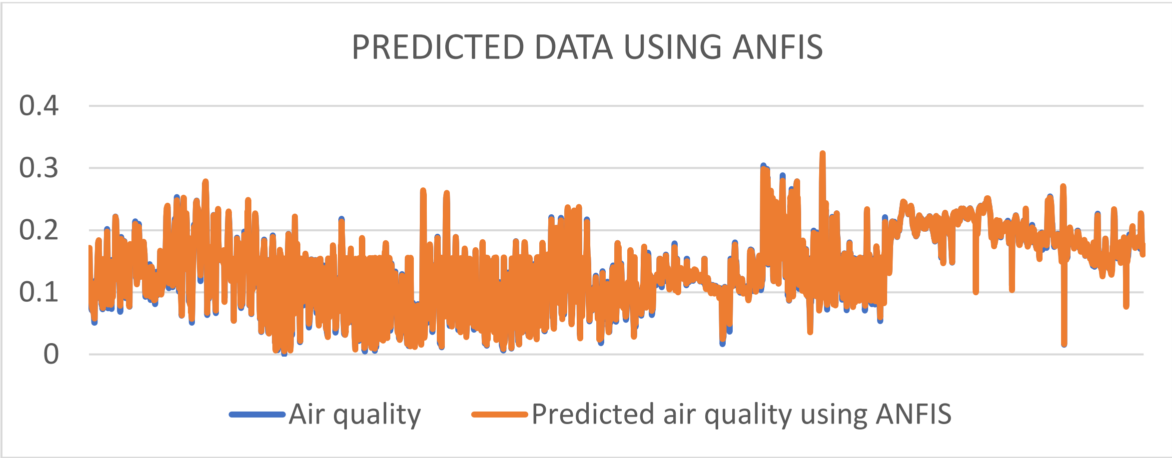

Figure 6. 6: Comparison actual and predicted data for Bursa using ANFIS

Figure 6.6 shows the comparison of actual and predicted air quality index using ANFIS hybrid network for Bursa. There is close relation between the actual and predicted data with maximum accuracy.

CONCLUSION

Air pollution is worthy phenomenon that is so important for human living and existence. Because of the changing climatic conditions, air quality cycle is changing and pollution of the ground is changing from low levels of pollution to a high level of pollutions. Precise air quality predicting can be helpful for a lot of issues and for the most important early air quality alert. It gives the people to be more careful and take precautionary measures like avoid direct exposures polluted air. Burning of Fossil Fuels Industry, transportation, exhaust from factories. Without a good air quality, the human life is so hard which is full of diseases. To control this issue, ANFIS model is designed to predict air quality.

In this study, an BPNN, NARX and ANFIS models are designed for predicting the daily Air Quality Index (AQI). The goal is to design an intelligence system for air pollution prediction that could predict hourly air quality effectively and accurately with a minimal error of prediction. The research includes the prediction of air quality index of two cities using a different intelligent algorithm such as BPNN, NARX and ANFIS. The data is applied to train, validate and test the models. The simulation results of presented models after train and test stages are compared in order to find more accurate ones. The comparisons were done using the root-mean-square error. The results have shown ANFIS model has more accurate result than other models. The effectiveness of ANFIS model has been proved by the simulation results obtained for air pollution prediction of Istanbul and Bursa cities.

REFERENCES

Helwan, A., Ozsahin, D. U., Abiyev, R., Bush, J. (2017). One-Year Survival Prediction of Myocardial Infarction. International Journal of Advanced Computer Science and Applications, Vol. 8, No. 6, 173-178.

Abiyev, R. H. (2006, September). Time series prediction using fuzzy wavelet neural network model. In *International Conference on Artificial Neural Networks* (pp. 191-200). Springer, Berlin, Heidelberg.

Abiyev, R. H. (2010, June). A type-2 fuzzy wavelet neural network for time series prediction. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems* (pp. 518-527). Springer, Berlin, Heidelberg.

Abiyev, R. H., & Abiyev, V. H. (2012). Differential evaluation learning of fuzzy wavelet neural networks for stock price prediction. *Journal of Information and Computing Science*, *7*(2), 121-130.

Abiyev, R., Abiyev, V. H., & Ardil, C. (2005). Electricity consumption prediction model using neuro-fuzzy system. *Development*, *149*, 1653.

Abraham, A. (2005). Artificial Neural Networks. John Wiley and Sons, Ltd.

Ahn, C. H., Hong, V. A., Park, W., Yang, Y., Chen, Y., Ng, E. J., ... & Kenny, T. W. (2015, June). On-chip ovenization of encapsulated disk resonator gyroscope (DRG). In *Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS), 2015 Transducers-2015 18th International Conference on* (pp. 39-42).

Akkoyunlu, A., & Erturk, F. (2002). Evaluation of air pollution trends in Istanbul. *International journal of environment and pollution*, *18*(4), 388-398.

Balmes, J. R., Fine, J. M., & Sheppard, D. (1987). Symptomatic Bronchoconstriction after Short-Term Inhalation of Sulfur Dioxide1, 2. *Am Rev Respir Dis*, *136*, 1117-1121.

Bank, W. (2016). The cost of air pollution: strengthening the economic case for action. *Washington: World Bank Group*.

Bergkamp, L., & Hanekamp, J. C. (2015). Climate change litigation against states: the perils of court-made climate policies. *European Energy and Environmental Law Review*, *24*(5), 102-114.

Boubel, R. W., Vallero, D., Fox, D. L., Turner, B., & Stern, A. C. (2013). *Fundamentals of air pollution*. Elsevier.

Buragohain, M., & Mahanta, C. (2008). A novel approach for ANFIS modelling based on full factorial design. *Applied soft computing*, *8*(1), 609-625.

Cai, M., Yin, Y., & Xie, M. (2009). Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. *Transportation Research Part D: Transport and Environment*, *14*(1), 32-41.

Challoner, A., Pilla, F., & Gill, L. (2015). Prediction of indoor air exposure from outdoor air quality using an artificial neural network model for inner city commercial buildings. *International journal of environmental research and public health*, *12*(12), 15233-15253.

Demuth, H. B., Beale, M. H., De Jess, O., & Hagan, M. T. (2014). *Neural network design* Martin Hagan.

Dye, T. S. (2003). Guidelines for developing an air quality (Ozone and PM2. 5) forecasting program.

Elbir T, Muezzinoglu A, Bayram A. (2000) Evaluation of some air pollution indicators in Turkey. Environment 2000;26(1–2):5–10.

Eldering, A., Wennberg, P. O., Viatte, C., Frankenberg, C., Roehl, C. M., & Wunch, D. (2017). The Orbiting Carbon Observatory-2: First 18 months of science data products. *Atmospheric Measurement Techniques*, *10*(2), 549-563.

Elshout, S., Bartelds, H., Heich, H., & Léger, K. (2012). CAQI Air quality index–Comparing Urban Air Quality across Borders–2012.

Feigin, V. L., Abajobir, A. A., Abate, K. H., Abd-Allah, F., Abdulle, A. M., Abera, S. F., ... & Aichour, M. T. E. (2017). Global, regional, and national burden of neurological disorders during 1990-2015. *Lancet Neurology*.

Fensterstock, J. C., Kurtzweg, J. A., & Ozolins, G. (1971). Reduction of air pollution potential through environmental planning. *Journal of the Air Pollution Control Association*, *21*(7), 395-399.

Fu, M., Wang, W., Le, Z., & Khorram, M. S. (2015). Prediction of particular matter concentrations by developed feed-forward neural network with rolling mechanism and gray model. *Neural Computing and Applications*, *26*(8), 1789-1797.

Grivas, G., & Chaloulakou, A. (2006). Artificial neural network models for prediction of PM10 hourly concentrations, in the Greater Area of Athens, Greece. *Atmospheric environment*, *40*(7), 1216-1229.

Güler, Ü. A., & Can, Ö. P. (2017). Kimyasal Kontaminantların Çevre Sağlığı ve Gıdalar Üzerine Etkileri. *Sinop Üniversitesi Fen Bilimleri Dergisi*, *2*(1), 170-195.

Hagevik, G. (1972, May). The Relationship of Land Use and Transportation Planning to Air Quality Management. Rutgers University, Center for Urban Policy Research and Conferences Department, University Extension Division, New Brunswick, New Jersey.

Hansen, J., Kharecha, P., Sato, M., Masson-Delmotte, V., Ackerman, F., Beerling, D. J., ... & Rockstrom, J. (2013). Assessing "dangerous climate change": required reduction of carbon emissions to protect young people, future generations and nature. *PloS one*, *8*(12), e81648.

Index, A. Q. (2009). A guide to air quality and your health. *USA: EPA.*

Khamis, A., & Abdullah, S. N. S. B. (2014). Forecasting wheat price using backpropagation and NARX neural network. *The International Journal of Engineering and Science*, *3*(11), 19-26.

Kim, J., & Kasabov, N. (1999). HyFIS: adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems. *Neural Networks*, *12*(9), 1301-1319.

Kolehmainen, M., Martikainen, H., & Ruuskanen, J. (2001). Neural networks and periodic components used in air quality forecasting. *Atmospheric Environment*, *35*(5), 815-825.

Kurt, A., Gulbagci, B., Karaca, F., & Alagha, O. (2008). An online air pollution forecasting system using neural networks. *Environment International*, *34*(5), 592-598.

Metrological department of Turkey. (2018). The air pollution data ,2018 from http://www.havaizleme.gov.tr/STN/STN_Report/StationDataDownload

Pérez, P., Trier, A., & Reyes, J. (2000). Prediction of PM2. 5 concentrations several hours in advance using neural networks in Santiago, Chile. *Atmospheric Environment*, *34*(8), 1189-1196.

Qin, S., Liu, F., Wang, J., & Sun, B. (2014). Analysis and forecasting of the particulate matter (PM) concentration levels over four major cities of China using hybrid models. *Atmospheric environment*, *98*, 665-675.

Reed, S. T. A. N. L. E. Y. (2016). Study Links 6.5 Million Deaths Each Year to Air Pollution. *New York Times. http://www. nytimes. com/2016/06/27/business/energy-environment/study-links-6-5-million-deaths-each-year-to-air-pollution. html*.

Seaton, A., Godden, D., MacNee, W., & Donaldson, K. (1995). Particulate air pollution and acute health effects. *The lancet*, *345*(8943), 176-178.

Shindell, D. T., Wong, S., & Rind, D. (1997). Interannual variability of the Antarctic ozone hole in a GCM. Part I: The influence of tropospheric wave variability. *Journal of the atmospheric sciences*, *54*(18), 2308-2319.

Song, Y., Qin, S., Qu, J., & Liu, F. (2015). The forecasting research of early warning systems for atmospheric pollutants: A case in Yangtze River Delta region. *Atmospheric Environment*, *118*, 58-69.

Sundquist, E. T., & Keeling, R. F. (2009). The Mauna Loa carbon dioxide record: Lessons for long-term earth observations. *Geophysical Monograph Series*, *183*, 27-35.

Tasdemir, Y., Cindoruk, S. S., & Esen, F. (2005). Monitoring of criteria air pollutants in Bursa, Turkey. *Environmental Monitoring and Assessment*, *110*(1-3), 227-241.

Wahyuni, I., Mahmudy, W. F., & Iriany, A. (2017). Rainfall Prediction using Hybrid Adaptive Neuro Fuzzy Inference System (ANFIS) and Genetic Algorithm. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(2-8), 51-56.

Wan, F., & Lei, K. S. (2008). Adaptive Neuro-Fuzzy Approach to Air Pollution Prediction in Macau.

Wang, H., Naghavi, M., Allen, C., Barber, R. M., Bhutta, Z. A., Carter, A., ... & Coggeshall, M. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The lancet*, *388*(10053), 1459-1544.

World Health Organization. (2003). Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide: report on a WHO working group, Bonn, Germany 13-15 January 2003.

World Health Organization. (2014). 7 million premature deaths annually linked to air pollution. *World Health Organization, Geneva, Switzerland*.

Yegnanarayana, B., & Murty, K. S. R. (2009). Event-based instantaneous fundamental frequency estimation from speech signals. *IEEE Transactions on Audio, Speech, and Language Processing*, *17*(4), 614-624.

Zhai, B., & Chen, J. (2014 Dec 26). Research on the forecasting of Air Quality Index (AQI) based on FS-GA-BPNN: A case study of Beijing, China.

Zweiri, Y. H., Whidborne, J. F., & Seneviratne, L. D. (2003). A three-term backpropagation algorithm. *Neurocomputing*, *50*, 305-318.