



## EXPERIMENTAL STUDY ON APPLICATION OF SUPPORT VECTOR MODEL THROUGH PYTHON PROGRAMMING TO PREDICT THE AIR QUALITY PARAMETERS”

K. Teja<sup>1</sup>, M.Hari Krishna<sup>2</sup>, M.Siva Someswara Rao<sup>3</sup>, M.Raghu Aravind<sup>4</sup>, P.Sai Ram<sup>5</sup>, M.Sumanth Rayudu<sup>6</sup>, Y.Swetha<sup>7</sup>  
 Assistant Professor <sup>1</sup>,UG Student <sup>2</sup>, UG Student <sup>3</sup>, UG Student<sup>4</sup>, UG Student<sup>5</sup>, UG Student<sup>6</sup> UG Student<sup>7</sup>

Civil Engineering ,  
 NRI Institute Of Technology , Vijayawada ,India

**Abstract:** This research work presents a method of daily air pollution modeling by using support vector machine (SVM) technique in urban area at local scale. Hazardous air pollutants or toxic air contaminants refer to any substances that may cause or contribute to an increase in mortality or in serious illness, or that may pose a present or potential hazard to human health. In this work, based on the observed data of Sulphur Dioxide, Nitrogen Dioxide, Respiratory Suspended Particulate (RSPM) PM-10, Particulate Matters- PM 2.5 for the years 2017 to 2021. The support vector regression (SVR) technique is used to build the nonlinear dynamic model of the air quality in the urban area of the city of one main aim of this model was to make an initial preliminary estimate of the dependence between primary and secondary pollutants in the city. A second main aim was to determine the factors with the greatest bearing on air quality with a view to proposing health and lifestyle improvements. It is well-known that the United States National Ambient Air Quality Standards (NAAQS) establishes the limit values of the main pollutants in the atmosphere in order to ensure the health of healthy people. They are known as the criteria pollutants. This SVR fit captures the prime idea of statistical learning theory in order to obtain a good forecasting of the dependence among the main pollutants in the city's of Andhra Pradesh. Finally, on the basis of these numerical calculations using SVR technique, from the experimental data, conclusions of this study are exposed. As air pollution is a complex mixture of toxic components with considerable impact on humans, forecasting air pollution concentration emerges as a priority for improving life quality. So with the help of Python tools and some Machine Learning algorithms, we try to predict the air quality. In the 21<sup>st</sup> century, world is increasing its level of technology & engineering which have given potential rise to the spread of pollution in the environment. Air pollution is important among them as human couldn't survive without O<sub>2</sub>, such O<sub>2</sub> levels are getting deteriorated by deforestation, burning of fossil fuels etc. As per the studies made in the year 2019, air pollution is responsible for 5 million deaths every year which will be 9% of deaths globally varies from 2 to 15% by country. In the recent events of COVID-19 pandemic, NASA has observed air quality index globally using satellite measurements which drops 30 % of pollutants worldwide. As every country is busy in vaccination process in order to protect the human race from this pandemic, but before this situation all living have suffered from the air pollutants one way or the other. This scenario has been observed in our India also as major metropolitan cities such as Delhi, Mumbai, and Chennai etc major air pollutants of PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, and NO<sub>2</sub> etc have significant dip in their levels and there is need to propose strong policies that replicates protection of air quality after the COVID-19. Our current study is made on the Vijayawada city in order to assess the amount of air pollutants that made an impact.

**Index Terms** –RSPM, NAAQS, SVR, NASA .

### I. INTRODUCTION

Air pollution is one the most important concept when we discuss about the protection of the environment. Earth's atmosphere is composed of about 78 percent nitrogen, 21 percent oxygen, 0.9 percent argon, and 0.1 percent other gases (*National Geographic Society*). Trace amounts of carbon dioxide, methane, water vapor, and neon are some of the other gases that make up the remaining 0.1 percent. As, there are many ways to define the air pollution, increase in the amount of pollutants in the air with create the air pollution of that particular location. Due to the growing needs of humans which gradual increase in the development in all the fields such as engineering, technology, medicine etc. This is one the reason we experience air pollution, global warming across world because the use of chemicals, toxins in the manufacturing in industries for faster production and the use of fossil fuels which is a non-renewable energy to run the automobiles and burning of the waste products such as chemical waste, hazardous waste, E-waste which releases the toxins into the air and burning of crackers during the diwali festival where the crackers contains hazardous content that could harm the environment. The air pollution effect the health of all the living beings as the animal species will be vanished the growing air pollutants doesn't have any end to it. Plants are the most important living things that produce food, oxygen, shelter, products for industries, from past few decades these are getting effected by the addition

of pollutants directly, acid rain the most cause for this effect as Acid rain results when sulfur dioxide (SO<sub>2</sub>) and nitrogen oxides (NO<sub>x</sub>) are emitted into the atmosphere and transported by wind and air currents. The SO<sub>2</sub> and NO<sub>x</sub> react with water, oxygen and other chemicals to form sulfuric and nitric acids. These then mix with water and other materials before falling to the ground and into its soil porous where its roots and cell tissues on leaves of the plants and trees that will have an adverse effect the vegetation and crop production. The air pollutants such as sulphur dioxide and sulphates, nitrogen oxides and nitrates, chlorides, carbon dioxide and ozone. The materials most sensitive to pollutants are calcareous building stones and ferrous metals. Manifestations of damage include losses of mass, changes in porosity, discoloration and embrittlement (*Cambridge University Press, 2011*). Air pollution is a significant risk factor for a number of pollution-related diseases, including respiratory infections, heart disease, COPD, stroke and cancer. The human health effects of poor air quality are far reaching, but principally affect the body's respiratory system and the cardiovascular system. Individual reactions to air pollutants depend on the type of pollutant a person is exposed to, the degree of exposure, and the individual's health status and genetics (*Adel Ghorani-Azam, Bamdad Riahi-Zanjani, and Mahdi Balali-Mood 2016*). As per the studies made in the year 2019, air pollution is responsible for 5 million deaths every year which will be 9% of deaths globally varies from 2 to 15% by country (*Hannah Ritchie, Max Roser 2019*). In the recent events of COVID-19 pandemic, NASA has observed air quality index globally using satellite measurements which drops 30 % of pollutants worldwide. Air pollutants had effected the major metropolitan cities such as Delhi, Mumbai, and Chennai etc in India, major air pollutants of PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, and NO<sub>x</sub> etc had a worst scenario after China, USA and Russia. The study area, Vijayawada tops state in India (*Tejaswini Marthi, 2019*). Now there is a significant dip in their levels and there is need continue and improve the air quality index after the COVID-19.

#### Major primary pollutants produced by human activity include:

1. Sulphur oxides (SO<sub>x</sub>)
- 2 Nitrogen oxides (NO<sub>x</sub>)
- 3 Particulate matter (PM<sub>2.5</sub>, PM<sub>10</sub>)

#### SECONDARY POLLUTANTS:

- I. Particulate matter formed from gaseous primary pollutants and compounds in photochemical smog. Smog is a kind of air pollution; the word "smog" is a portmanteau of smoke and fog. Classic smog results from large amounts of coal burning in an area caused by a mixture of smoke and sulphur dioxide. Modern smog does not usually come from coal but from vehicular and industrial emissions that are acted on in the atmosphere by sunlight to form secondary pollutants that also combine with the primary emissions to form photochemical smog.
- II. Ground level ozone (O<sub>3</sub>) formed from NO<sub>x</sub> and VOCs. Ozone (O<sub>3</sub>) is a key constituent of the troposphere (it is also an important constituent of certain regions of the stratosphere commonly known as the Ozone layer). Photochemical and chemical reactions involving it drive many of the chemical processes that occur in the atmosphere by day and by night. At abnormally high concentrations brought about by human activities (largely the combustion of fossil fuel), it is a pollutant, and a constituent of smog.  
Peroxyacetyl nitrate (PAN) – similarly formed from NO<sub>x</sub> and VOCs

Table-1: Indian National Air Quality Standards (units: µg/m<sup>3</sup> unless mentioned otherwise)

Pollutant	SO <sub>2</sub>	NO <sub>2</sub>	PM <sub>2.5</sub>	PM <sub>10</sub>	O <sub>3</sub>		CO (mg/m <sup>3</sup> )		Pb	NH <sub>3</sub>
Averaging time (hr)	24	24	24	24	1	8	1	8	24	24
Standard	80	80	60	100	100	100	4	2	1	400

Table-2: INDIAN-AQI Category and Range

AQI Category	AQI Range
Good	0 – 50
Satisfactory	51 – 100
Moderate	101 – 200
Poor	201 – 300
Very Poor	301 – 400
Severe	401 - 500

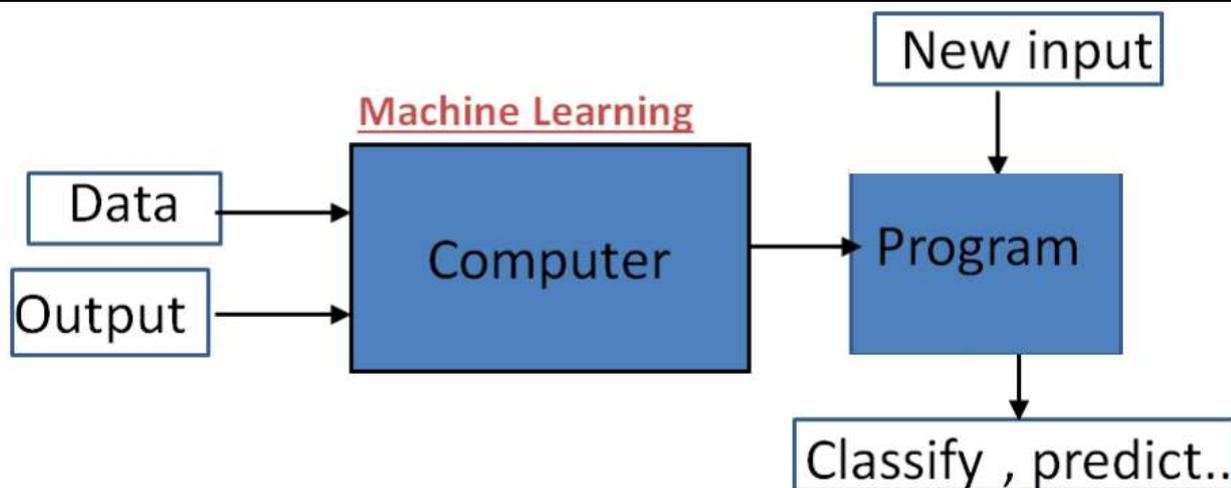


Fig-1: Machine Learning And Deep Learning

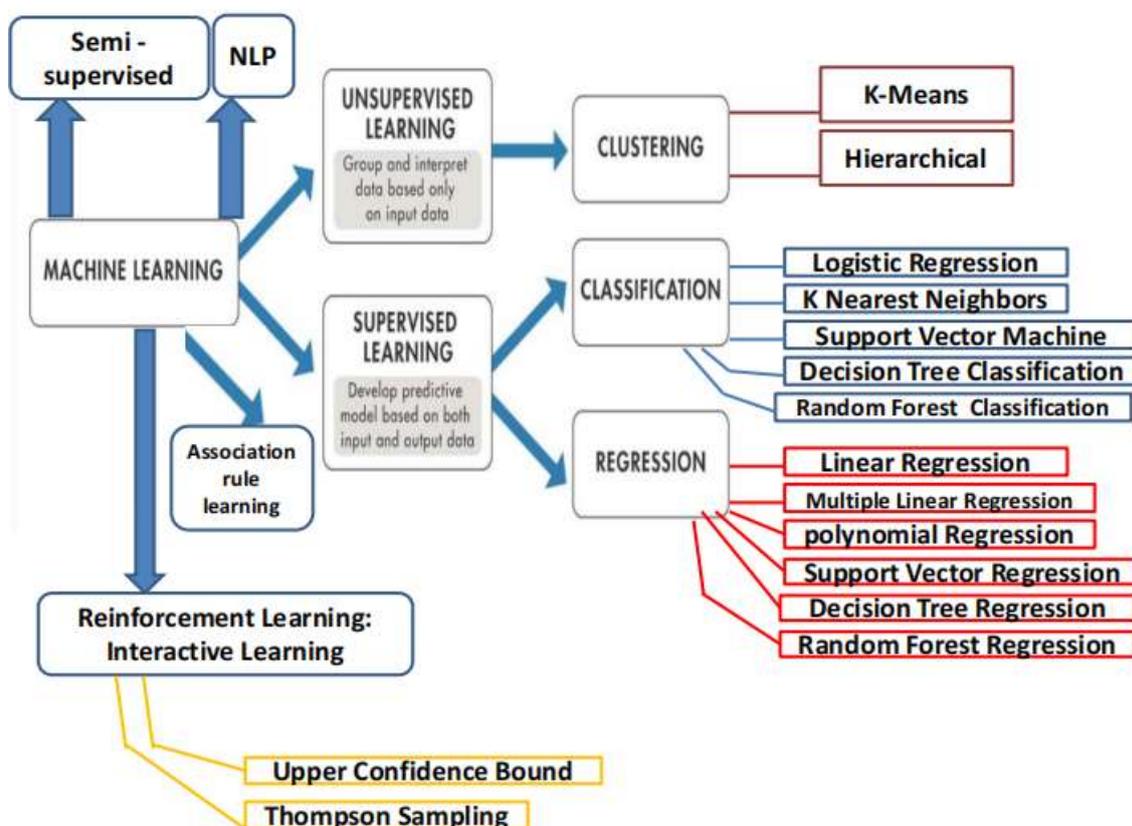


Fig-2: Different Types of Machine learning

**OBJECTIVES OF THE STUDY:**

- To analyse the air pollutants levels in the study area.
- To study on the pollutants during the COVID-19 pandemic.
- To propose to some methods based on the level of pollutants based literature study and article citation.

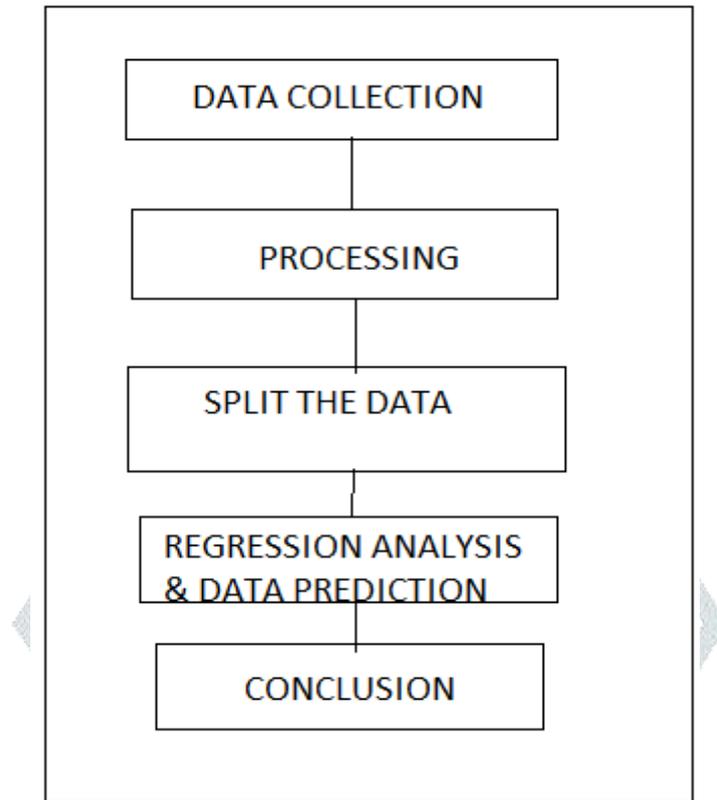
**MEDHODOLOGY:**

The primary data of air quality index (AQI) data from air pollution monitoring station located at 9 different locations across Vijayawada city from the Andhra Pradesh State Pollution Control Board(APSPCB). All the locations are categorized based on the locality such as residential, commercial, industrial and other area. The secondary data of National ambient air quality standards are obtained from the Central Pollution Control Board (CPCB). The map of Vijayawada has been generated from SDSM Software (Gregory Cletus Cigi and Sunny Agarwal ,2020).

The data is been analyzed in the tabular form and represented graphical manner. The current year 2021, air quality index data of the proposed study area is calculated according to the day-average of every pollutant on hourly basis and consolidated in monthly-average. The average annual all air pollutants for the year 2017 at station 7, 8, and 9 couldn't be analyzed as those were established in the year 2018.

**FLOW DIAGRAM:**

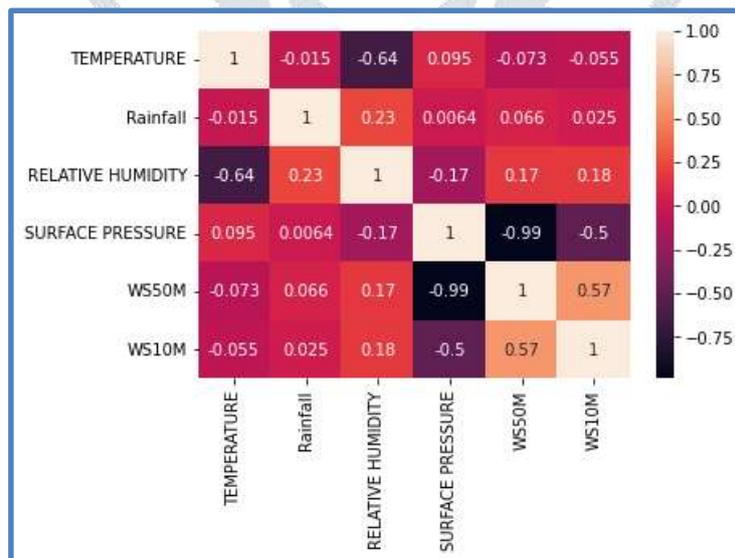
The Flow diagram represents the step by step process, from data preprocessing to air quality prediction.



**Fig-3.0 : Flow diagram**

**RESULTS**

In statistics, correlation or dependence is any statistical relationship, whether causal or not, between two random variables or bivariate data. In the broadest sense correlation is any statistical association, though it actually refers to the degree to which a pair of variables are linearly related. Familiar examples of dependent phenomena include the correlation between the height of parents and their offspring, and the correlation between the price of a good and the quantity the consumers are willing to purchase, as it is depicted in the so-called demand curve.



**Fig-4: Anathapur results**

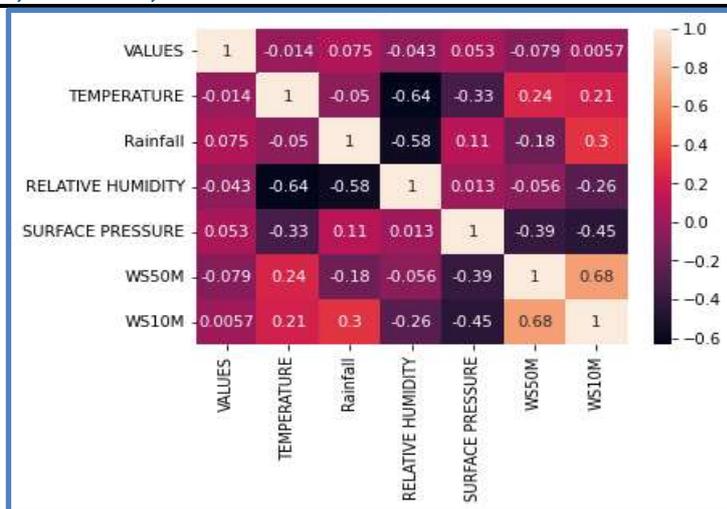


Fig-5: KADAPA RESULTS

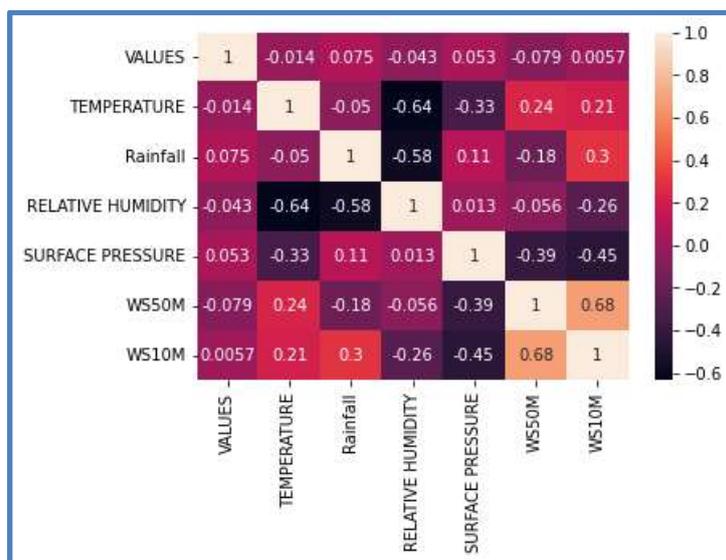


Fig-6: RAJAMUNDRY RESULTS

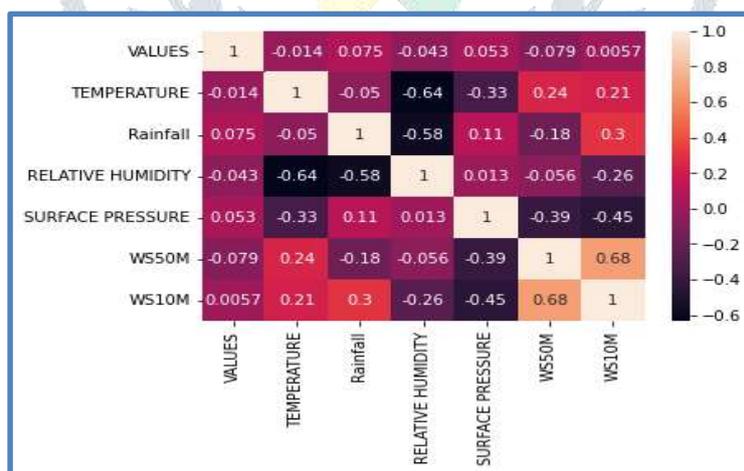


Fig-7.0 VIJAYWADA RESULTS

**CONCLUSIONS:**

Air pollution as one of the most serious forms of environmental pollutions poses huge threat to human life. Air pollution leads to environmental instability, and has harmful and undesirable effects on the environment. Air quality information such as the concentration of PM2.5, PM10 is of great significance for human health and city management. It affects the way of traveling, urban planning, government policies and so on. However, in major cities there is typically only a limited number of air quality monitoring stations. In the meantime, air quality varies in the urban areas and there can be large differences, even between closely neighboring regions. This study examines the performance of the feature selection in combination with multiple-linear regression (Step method, Backward&Forward methods), Correlation, Support Vector Method (Time sequence)'. The functional characteristics of the SVM are also investigated.

- Thus, our results suggest to proceed with a selection of chemical sensors based on the best ratio prediction/cost. For example, if only one trace gas sensor is affordable, the best performance can be reached with CO or NO<sub>2</sub> concentrations, while the use of two trace gases (SO<sub>2</sub> and NO<sub>2</sub>) are sufficient to get very close to the best possible accuracy. Urban air pollution is a major problem in both developed and developing countries, as atmospheric pollutants have a great effect on human health.
- The first finding shows that it is possible to get a reasonable prediction of pollutants concentrations only using the previous months data. There is slight decrease in the air pollutants in the 2020-2021. Thus, using an affordable meteorological station data further improves the prediction accuracy. In this case, a regression model tree gives a better prediction than a linear regression model.
- Air pollution is the “world’s largest environmental health threat”, causing 7 million deaths worldwide every year. Its major constituents are PM<sub>2.5</sub>, PM<sub>10</sub> and the harmful green house gases SO<sub>2</sub>, NO<sub>2</sub>, CO and other effluents from vehicles and factories affecting not only humans but also other living organisms both on land and sea. The only effective solution to this global issue is to implement machine learning algorithms to predict the AQI (Air Quality Index) that can make the people aware of the condition of the air of a certain region such that certain actions could be issued by the government for the improvement of the air quality in the future.
- The prime objective behind this project is to predict the AQI based on the concentration of PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, NO<sub>2</sub>, and CO as well as weather conditions like temperature, surface pressure, rainfall and wind pressure. Hence the data set is combined from various web sources like repository in order to bring accuracy in the prediction and to justify whether the Quality of air is suitable or not. This prediction will be brought about with the help of some supervised machine learning algorithms and the observation and the result will state which algorithm is giving better accuracy in prediction of AQI and which one is giving less error.

#### REFERENCES:

1. A.Aarathi, P.Gayathri, N.R.Gomathi, S.Kalaiselvi, Dr.V.Gomathi<sup>5</sup>, Air Quality Prediction Through Regression Model. INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 03, ISSN 2277-8616, MARCH 2020.
2. Adel Ghorani-Azam, Bamdad Riahi-Zanjani, and Mahdi Balali-Mood, Effects of air pollution on human health and practical measures for prevention in Iran. Journal of research in medical sciences, PMID: PMC5122104, 2016.
3. Air quality index data from Central control board.
4. Air quality index using satellite measurement, NASA DATA (2020).
5. Air Quality Management Process Article, Environmental Protection Agency.
6. Babar ZB, Shareefdeen Z (2014) Management and control of air emissions from electronic industries. Clean Techn Environ Policy 16(1):69–77.
7. Bergh J<sup>1</sup>, Van Den CJM<sup>2</sup>, Allen W<sup>3</sup> (2011) Environment versus growth — a criticism of ‘degrowth’ and a Plea for ‘a-growth’. Ecol Econ 70(5):881–890.
8. Bilal Asif<sup>1</sup>, Muhammad Zeeshan<sup>2</sup>, Sidra Iftekhar<sup>3</sup>, Mika Sillanpää<sup>4</sup>, Promoted three-way catalytic activity of the Co<sub>3</sub>O<sub>4</sub>/TiO<sub>2</sub> catalyst by doping of CeO<sub>2</sub> under real engine operating conditions. Atmospheric Pollution Research, Volume 12, Issue 7, July 2021, 101088.
9. Biographical data on Vijayawada City from Vijayawada Municipal Corporation (VMC).
10. Cariolet JM<sup>1</sup>, Colombert M<sup>2</sup>, Vuillet M<sup>3</sup>, Diab Y<sup>4</sup> (2018) Assessing the resilience of urban areas to traffic-related air pollution: application in greater Paris. Sci Total Environ 615(615):588–596.
11. Chae Y<sup>1</sup>, Park J<sup>2</sup> (2011) Quantifying costs and benefits of integrated environmental strategies of air quality management and greenhouse gas reduction in the Seoul metropolitan area. Energy Policy 39(9):5296–5308.
12. Conti ME<sup>1</sup>, Ciasullo R<sup>2</sup>, Tudino MB<sup>3</sup>, Matta EJ<sup>4</sup> (2015) “The industrial emissions trend and the problem of the implementation of the industrial emissions directive (IED)” *Air Quality*. Atmosph Health 8(2):151–161.
13. Contreras-Zarazúa G<sup>1</sup>, Villicaña-García E<sup>2</sup>, Cansino-Loeza B<sup>3</sup>, Vázquez-Castillo JA<sup>4</sup>, Ponce-Ortega JM<sup>5</sup>, Segovia-Hernández JG<sup>6</sup> (2018) Environmental impact and occupational Hazard evaluation on intensified processes to produce diphenyl carbonate. Comput Chem Eng: 1–12.