

Android based Air Pollution Monitoring System

Pawan Shinde¹, Sonal Jagtap^{2*}, Shubham Sharma³, Suyog Sakpal⁴, Pallavi Sawant⁵

^{1,2,3,4,5} Dept. of E & TC Engg., Smt. Kashibai Navale College of Engineering, Savitribai Phule
Pune University, Pune

¹ amshindepawan@gmail.com

^{2*} skjagtap.skncoe@sinhgad.edu

³ sharmashubham6300@gmail.com

⁴ suyogsakpal786@gmail.com

⁵ pallavi.msawant_skncoe@sinhgad.edu

Abstract

Air pollution is one of the most important concern of the world. Air pollution may emerge from anthropogenic or natural sources. Air pollutants like CO, CO₂, SO₂, NO₂, and O₃ suspended particulate matter (SPM) in atmosphere, repairable suspended particulate matter (RSPM), and volatile organic compounds (VOCs) have considerable effect on the people health. Most of the cities in developing countries and most cities of the developed countries are suffering from this problem. Therefore a need to develop a real time air quality and pollution monitoring system is critical. An arduino based air pollution detector is developed which combined a small-sized, minimum-cost sensor to an arduino microcontroller unit along with this we are using machine learning. The advantage being that all the data of a location will be available to us on an android application platform. Features being that it is user-friendly, affordable, low-cost and minimum-power requirement hardware which is appropriate for mobile measurement, as well as comprehensible data collection. Software on android platform will be able to analyze, collected data with high precision. Simple instrument developed further can be commercially utilized.

Keywords—Air pollution; android; gas sensors; cloud; prediction.

I. INTRODUCTION

Air pollution means the presence of unwanted biological molecules, particles and other harmful components into the earth atmosphere. It constitutes a major cause of infections, allergies, and eventually results in death to some peoples.

It also harms other living inhabitants like that animals as well as irrigation crops, or the ecological or built environment. They are also responsible for various kinds respiratory diseases (like asthma), causes of different types of cancer in individuals, if they are unprotected to these toxic contents or chemicals for long period of time. For instance, carbon monoxide (CO) is extremely dangerous to people as it may cause serious asphyxiation, headaches because of the composition of carboxy-hemoglobin and thus a reason of death if unprotected for a long period of time.

The world health organization (WHO) in 2014 approximated those 7 million people deaths worldwide because of air pollution. The similar approximation roughly equalled by the International Energy Agency (IEA) also. These chemicals or pollutants are also responsible for various environmental calamities like acid rain and depletion of ozone layer. Because of a number of phylogenic actions, air pollution is on the steep growth and its controlling being humans is of utmost importance to alleviate actions to limit it.

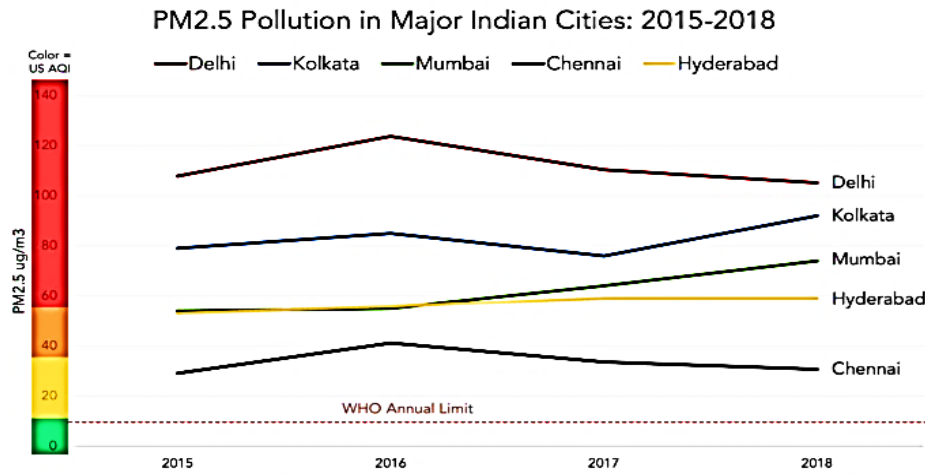


Fig.1 Pollution in Indian cities.

Fig.1 shows a report depicting the level of air pollution across various cities of a developing country like India. This data was recorded during the period of 2015-2018. Talking of past, the air quality measuring sensors were large, non-portable and costly. Presently, air pollution sensors are developed on five common air pollutants which are nitrous oxide, carbon monoxide, ozone, sulphur dioxide and particle matter. In today's world of technology, air pollution and quality monitoring are very important as it has a great effect on people health. It's known that eradicating it completely is hypothetical today but it can be appeased by being aware about it. The idea proposed is that to have pre knowledge of the level of pollution in any specific area so that a person opts for the safer route before going out. Say, for example if the journey is from A to B with multiple routes then with the knowledge of pollution the path with least pollution can be chosen.

II. RELATED WORK

The air pollution is one of the major problems the world is going through right now and it affects urban areas more because of density of automobiles and industries present there. The air is polluted to a level that it is not even safe to breathe in it. It is reported that the harmful gases have a fair share in causing casualties across the globe.

The air pollution monitoring system has flattered many mind to work on it and there is a considerable amount of projects done on this. Projects have been described in the literature that makes use of less expensive air pollution sensing devices that can be used by individuals or by the vehicles [1]. In a different study by De Nazelle et al. [2], the writers made use of environmental sensing approaches that reinvigorated the awareness and sympathy of individuals towards pollution. In two studies, [3] [4], the authors showed an environmental sensing method that recaptures attention and sympathy of citizens toward pollution. Exposure Sense [5] is a portable participatory sensing system that is made to screen one's daily movements.

In a study [6], the authors gave a cloud-based system that uses knowledge-based findings to find real-time pollution levels. The data is gathered by monitoring stations that are kept at different geo-locations. This has mobile devices for monitoring purposes. In [7] they presented an Android application which has users with information about air quality. By combining user area information and metropolitan air quality data gleaned by monitoring devices, this application presents a ubiquitous and unobtrusive monitoring [8] framework that is ready to advise users about their daily air pollution exposure.

Reshi et al. [9] designed a wireless network platform, called VehNode, that gave automobiles the ability to check the level of harmful matters in smoke of vehicle. Mujawar et. al. [10] figured an air pollution level measuring system with wireless sensor network for use in Solapur City. Micro-sensor nodes get and calculate the target gas by measuring the sensing layer's electrical conductivity. When the gases made contact with the surface of the sensor they are accumulated and the conductivity varies. Also, a semiconductor sensor is used at the emission port of the vehicle to sense the quantity of pollutants and send this level to the microcontroller [11]. The comparison table for the three algorithms discussed below for development of system.

TABLE I : Comparison of various algorithm

Method	Background Model	Advantages	Disadvantages
Multiple Regression	Uses multiple variables to build a model and hence the prediction is more reliable.	Better analysis of all the dependent features and more accurate.	Has to be too careful with data as small amount of false information leads to a whole wrong prediction
Logistic Regression	Logistic Regression is best used when the output needs to be binary in nature.	Works better with the categorical data which means the output can have only two categories	Can only predict a categorical outcome which is nominal in nature.
K Nearest Neighbour	Estimates the class of the data by considering its nearest neighbours. Euclidean distance is calculated.	Can work on Non linear classes easily.	Slow as it has to monitor all the data to find its nearest neighbours.

III. ARCHITECHURE AND METHODOLOGY

The air pollution is monitored with the help of embedded system. Arduino board is used which has the gas sensors embedded to it. The sensors like MQ2, MQ7, MQ135 are for sensing gases like Carbon Monoxide, Methane which are major pollutants of the air we breathe in. These sensors will be placed at various locations in the city and they will be constantly sending the data. Each arduino board will make use of wifi models to send the data and it will be first received at cloud storage. From the cloud storage then it will be send to the mobile application. The prediction part comes then where a machine learning model trained on the past records will be predicting pollution level or AQI for the upcoming days. Sensors are used to measure various gas levels. The sensors taken into use include MQ7 which has the property of sensing Carbon Monoxide while for sensing the methane content the MQ2 is used. MQ135 is highly sensitive to Ammonia NH_3 , Sulphide SO_2 . Gas sensor works on an operating voltage of 2.5V to 5.0V. Fig. 2 shows system architecture of air pollution monitoring.

The arduino board will read the data gathered by all these sensors and will combine their output to give the result in equivalent AQI. Arduino UNO is a microcontroller board based ATmega328P. It has 16 MHz quartz crystal, 14 digital input/output pins, 6 analog inputs, a USB connection, a power jack, an ICSP.

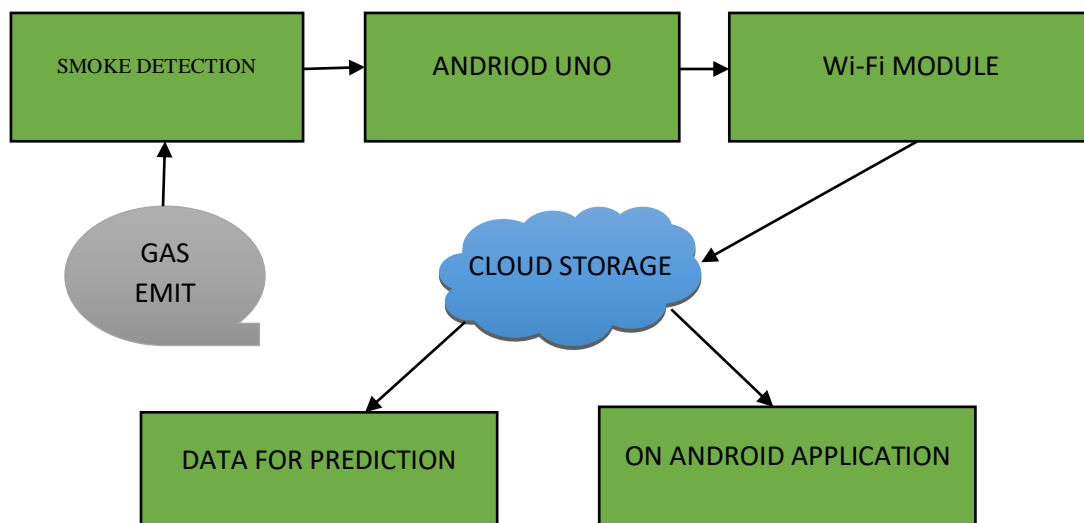


Fig.2 System architecture of air pollution monitoring

The Arduino platform is used for its coding and compilation the examining window. The data from the arduino is to be sent to the cloud storage and for that the wifi module is connected with the kit. ESP8266 WiFi Module is the WiFi module used and it a self contained SOC integrated with TCP/IP protocol stack that can give any microcontroller access to WiFi network. ESP8266 is capable of hosting an application or offloading all WiFi networking functions from an application processor. A cloud service is provided to users on demand by virtue of the Internet from a cloud computing provider's server as opposed to being provided from a company's server. The cloud service stores the data real time which is useful to access from anywhere in world. This data available on cloud can work for building data set for machine learning. The model built will collect the data from the cloud and will display on the android application.

A. Android Application

Android application will be a user friendly application to display the pollution level in any route on the map. Suppose, a person wants to travel from point A to point B then the application will show how much the route is polluted with help of various colors. Each color will represent various levels of pollution. Application will also show the prediction for the future which will be based on the past records.

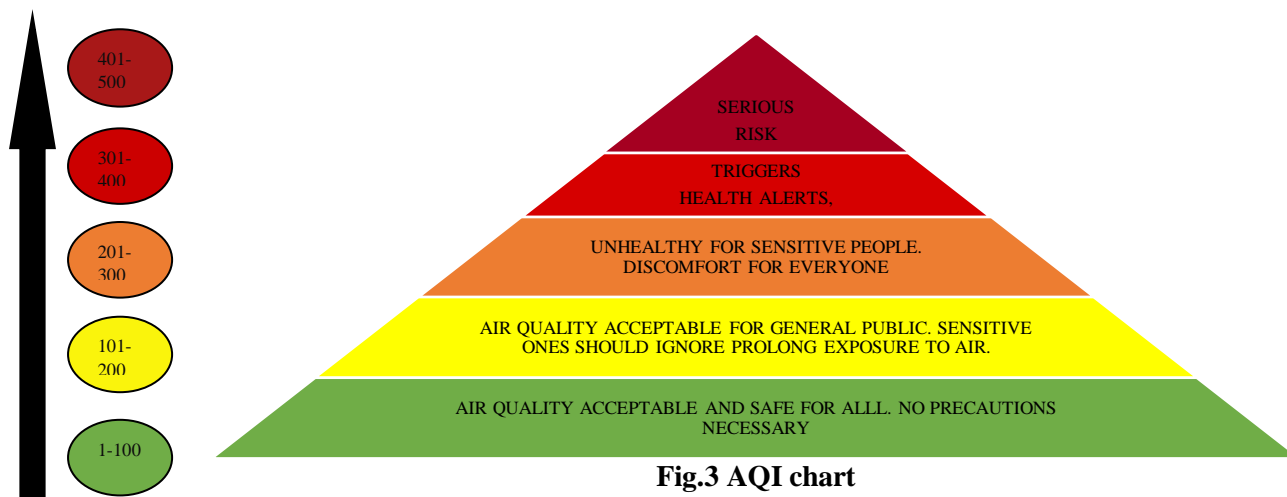


Fig.3 AQI chart

Air quality index (AQI) is a measuring unit for the pollution level in the air. Fig.3 represents AQI chart ranged from 0 to 500 and categorised in various levels. Each level says something about the air quality. It is depicted below with the help of the chart. Air quality index is calculated by a linear function by determining the concentration of the pollutant. The project predicts the quality of air based on its AQI range. The air quality index is composed of 8 pollutants (PM_{10} , $PM_{2.5}$, NO_2 , SO_2 , CO , O_3 , NH_3 , and Pb).

TABLE II: AQI category , pollutants and health breakpoints

AQI (Range)	$PM_{2.5}$	NO_2	SO_2	CO
Good(0-100)	0-100	0-80	0-80	0-2.0
Moderate(101-200)	101-250	101-168	81-380	2.1-10
Poor(201-300)	251-350	169-208	381-800	10-17
Very Poor(301-400)	351-430	209-748	801-1600	17-34
Risky(401-500)	430+	748+	1600+	34+

All these gases contribute greatly to the pollution of air however, in the project the main focus has been set on calculating the levels of SO_2 , NO_2 and CO .

B. Prediction

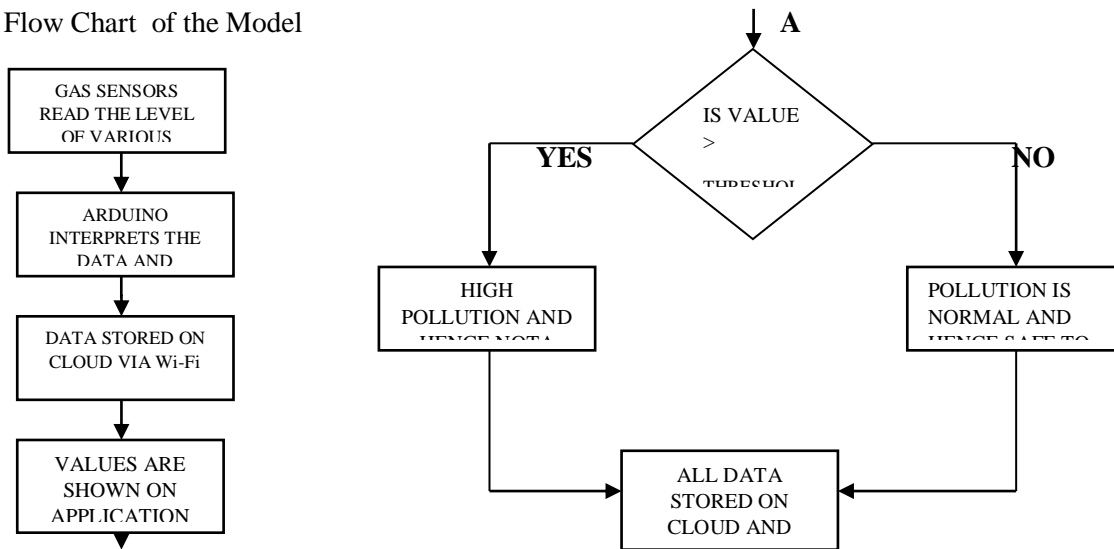
Machine learning is used to draw hidden insights from the dataset of the showing pollution in the city at various time and location. The old records are used to train the model and based on the algorithm used which is Multivariate Regression it will give the prediction. Since the dataset will not be single variant, the multiple variable regression is used as it will give the highest accuracy in this case.

The multiple regression equation explained above takes the following form:

$$Y = X_1b_1 + X_2b_2 + \dots + X_nb_n + C. \quad (1)$$

Here, b_i ($i=1,2,\dots,n$) are simply the regression coefficients, which represent the value at which the criterion variable changes when the predictor variable changes. $Y = Xb + C$ is the case for single variable dependency where it represents the linearity in the data. b denotes the slope of the equation while c is the y intercept which is the output when input is equal to zero. In this case the data will depend on multiple factors like the day, time and location and because of this multivariate regression is preferred.

C. Flow Chart of the Model



A

Fig. 4 Flowchart of the model

The flowchart above depicts the sequence of processes taking place to finally give the output on the mobile application. The data gathered from the sensors via the Arduino board will be sent on the cloud storage and from there it will be sent on the android application. The cloud data will be used to train the machine learning model and make predictions.

IV. RESULTS AND DISCUSSION

When the model was run gave the following results. The sensors sensed the gas fed to it. The results shows the output of the sensors readable on the LCD via the arduino board which is discussed. Fig. 5(a) shows the screenshot of UI of the app with the source (green) and destination (red) for indication. The pollution will be monitored on the route between these two locations. This app is analogous to the Google Maps with the main difference lying in its additional function of showing pollution levels too on the route. Fig. 5(b) shows the output of the sensors on LCD read by the arduino board. The gas sensors is embedded to the arduino board with LCD on the other side of the board. The data from the sensor is monitored by the arduino board and then displayed on the LCD continuously.

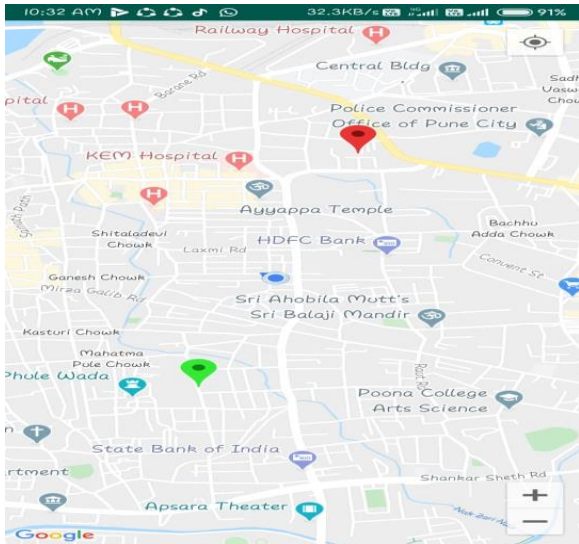


Fig.5a)UI of the application

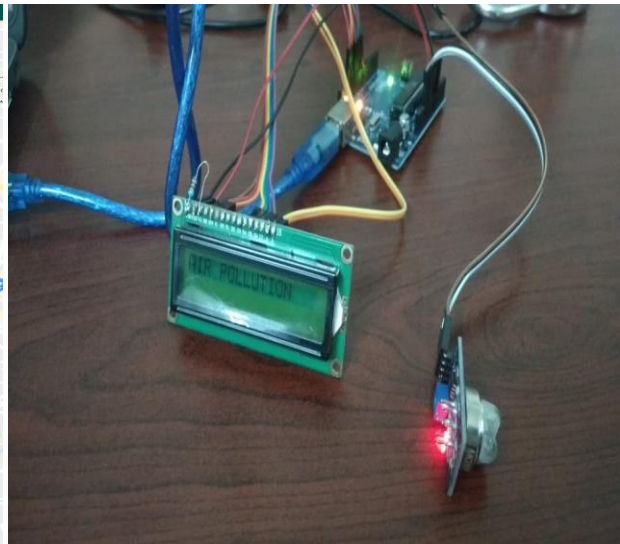


Fig.5(b) Data read via arduino on LCD

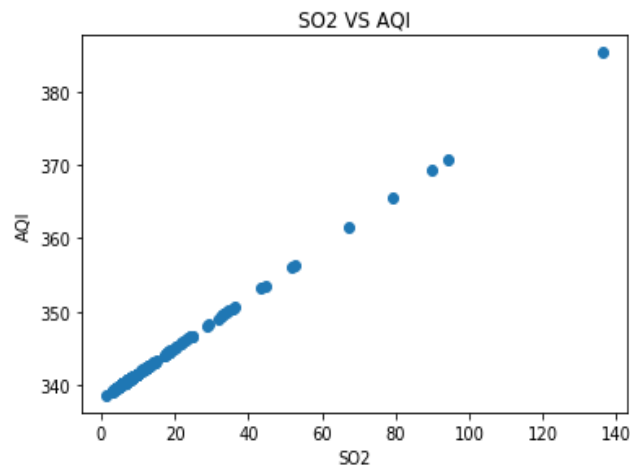


Fig.5 c) Variation of SO₂ with AQI

Fig. 5(d) shows a screenshot from the jupyternotebook where the data set has been trained. The multivariate machine learning algorithm was used to train on features like level of SO₂, NO₂ and Station code(STN_COD). The final output varies on all the features but only the relation between the SO₂ levels and the AQI is shown in fig 5 (c).

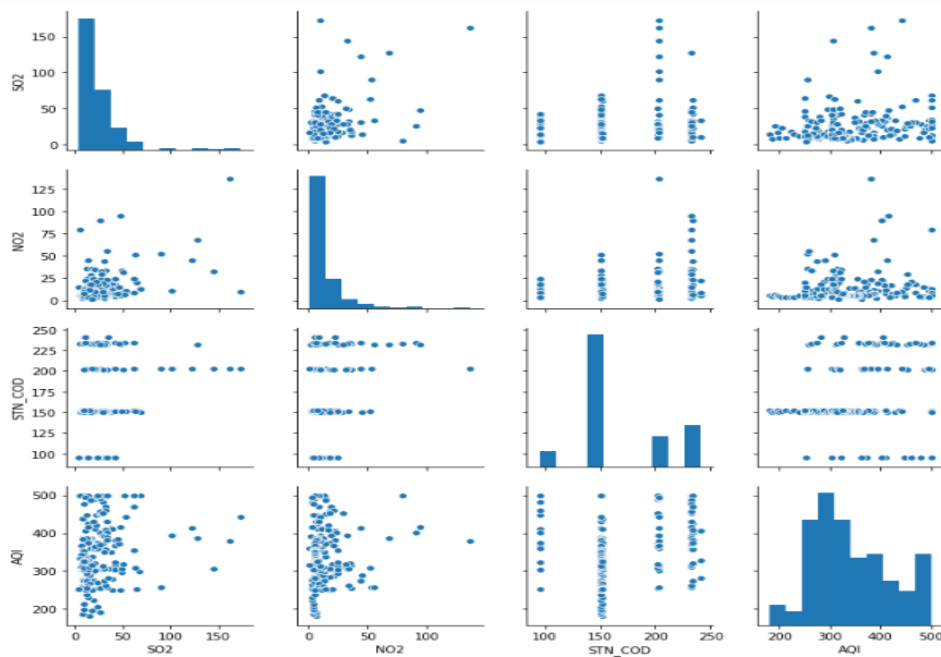


Fig.5 d) Inter relation of all features with each other.

This screenshot shows the inter relation between all the features of the dataset. For example, how SO_2 varies with different STN_COD, different levels of NO_2 and the AQI levels. As can be seen that has different levels of concentration at different station codes with maximum being at the STN_COD=250. SO_2 and NO_2 are almost linear to each other.

V. CONCLUSION

This air pollution monitoring system along with the integrated mobile application can be helpful to humanity by giving prior information about the air quality. The data gathered via the gas sensors will be interpreted by the arduino board and then the equivalent AQI will be calculated. This AQI will be stored in the cloud which will be further sent on the android application. On the application the current AQI and also the predicted AQI for future will be shown. The mobile application will be useful and handy in nature. The real time monitoring will help people to take the least polluted route.. It is a user friendly interface and on the basis of AQI(Air Quality Index) based on the concentration of SO_2 , NO_2 and CO and $\text{PM}_{2.5}$. Though there are many other gases responsible for the pollution but this project only considers SO_2 , NO_2 and CO and $\text{PM}_{2.5}$ as they have the most impact on pollution.

REFERENCES

- [1] G. Lo Re, D. Peri, and S. D. Vassallo, "Urban air quality monitoring using vehicular sensor networks," *Advances onto the Internet of Things*, Springer, pp. 311–323, 2014.
- [2] A. De Nazelle, E. Seto, D. Donaire-Gonzalez, M. Mendez, J. Matamala, M. J. Nieuwenhuijsen, and M. Jerrett, "Improving estimates of air pollution exposure through ubiquitous sensing technologies," *Environ. Pollut.*, vol. 176, pp. 92–99, 2013.
- [3] R. Peterová and J. Hybler, "Do-it-yourself environmental sensing," *ProcediaComput. Sci.*, vol. 7, pp. 303–304, 2011.
- [4] M. R. B. and A. J. B. Alok N. Bhatt, "Automation Testing Software that Aid in Efficiency Increase of Regression Process," *Recent Patents Comput. Sci.*, vol. 6, no. 2, pp. 107–114, 2013.
- [5] B. Predić, Z. Yan, J. Eberle, D. Stojanovic, and K. Aberer, "Exposuresense: Integrating daily activities with air quality using mobile participatory sensing," in *Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2013 IEEE International Conference on, pp. 303–305, 2013.

- [6] Y. Zheng, X. Chen, Q. Jin, Y. Chen, X. Qu, X. Liu, E. Chang, W.-Y. Ma, Y. Rui, and W. Sun, “A cloud-based knowledge discovery system for monitoring fine-grained air quality,” Prep. Microsoft Tech Report, <http://research.microsoft.com/apps/pubs/default.aspx>, 2014.
- [7] G. Lo Re, D. Peri, and S. D. Vassallo, “A mobile application for assessment of air pollution exposure,” in Proceedings of the 1st Conference on Mobile and Information Technologies in Medicine (MobileMed 2013), 2013.
- [8] A. Tamayo, C. Granell, and J. Huerta, “Using SWE standards for ubiquitous environmental sensing: a performance analysis,” *Sensors*, vol. 12, no. 9, pp. 12026–12051, 2012.
- [9] A. A. Reshi, S. Shafi, and A. Kumaravel, “VehNode: Wireless Sensor Network platform for automobile pollution control,” in Information & Communication Technologies (ICT), 2013 IEEE Conference on, pp. 963–966, 2013.
- [10] T. H. Mujawar, V. D. Bachuwar, and S. S. Suryavanshi, “Air Pollution Monitoring System in Solapur City using Wireless Sensor Network,” *Proc. Publ. by Int. J. Comput. Appl. CCSN-2013*, pp. 11–15, 2013.
- [11] S. Gaglio, G. Lo Re, G. Martorella, D. Peri, and S. D. Vassallo, “Development of an IoT environmental monitoring application with a novel middleware for resource constrained devices,” in Proceedings of the 2nd Conference on Mobile and Information Technologies in Medicine (MobileMed 2014), 2014.