

Article

Classifying Sources Influencing Indoor Air Quality (IAQ) Using Artificial Neural Network (ANN)

Shaharil Mad Saad ^{1,*}, Allan Melvin Andrew ¹, Ali Yeon Md Shakaff ¹,
Abdul Rahman Mohd Saad ², Azman Muhamad Yusof @ Kamarudin ² and Ammar Zakaria ¹

¹ Center of Excellence for Advanced Sensor Technology (CEASTech),
Universiti Malaysia Perlis (UniMAP), Taman Muhibbah, Jejawi, 02600 Arau, Perlis, Malaysia;
E-Mails: allanmelvin.andrew@gmail.com (A.M.A.); aliyeon@unimap.edu.my (A.Y.M.S.);
ammazakaria@unimap.edu.my (A.Z.)

² Faculty of Engineering Technology, Universiti Malaysia Perlis (UniMAP), Kampus UniCITI Alam,
02100 Sungai Chuchuh, Padang Besar, Perlis, Malaysia;
E-Mails: abd.rahman@unimap.edu.my (A.R.M.S.); azman@unimap.edu.my (A.M.Y.K.)

* Author to whom correspondence should be addressed; E-Mail: shaharil85@gmail.com;
Tel.: +60-4-979-8897.

Academic Editor: Leonhard M. Reindl

Received: 19 January 2015 / Accepted: 22 April 2015 / Published: 20 May 2015

Abstract: Monitoring indoor air quality (IAQ) is deemed important nowadays. A sophisticated IAQ monitoring system which could classify the source influencing the IAQ is definitely going to be very helpful to the users. Therefore, in this paper, an IAQ monitoring system has been proposed with a newly added feature which enables the system to identify the sources influencing the level of IAQ. In order to achieve this, the data collected has been trained with artificial neural network or ANN—a proven method for pattern recognition. Basically, the proposed system consists of sensor module cloud (SMC), base station and service-oriented client. The SMC contain collections of sensor modules that measure the air quality data and transmit the captured data to base station through wireless network. The IAQ monitoring system is also equipped with IAQ Index and thermal comfort index which could tell the users about the room's conditions. The results showed that the system is able to measure the level of air quality and successfully classify the sources influencing IAQ in various environments like ambient air, chemical presence, fragrance presence, foods and beverages and human activity.

Keywords: indoor air quality; artificial neural network (ANN); pattern recognition

1. Introduction

People normally spend most of their time in indoor environments. Therefore, their health depends heavily on the indoor environment in which they live. Hence, meticulous attention should be given to make sure the indoor environment is safe and comfortable. As a major part of the indoor environment, attention should also be given to indoor air quality (IAQ). Continuous monitoring of IAQ is important to make sure people breathe in a healthy and safe air. Real-time IAQ monitoring keeps people alert to any pollution that might be present in an indoor environment right as it happens. A good IAQ monitoring system should also be able to tell the users about the source of pollutants (for example: volatile organic compounds (VOCs) emitting from a chemical product). A better IAQ monitoring system with enhanced features is proposed in this paper—a smart IAQ monitoring system. This smart IAQ monitoring system could identify and inform the users about the source influencing the IAQ level. For example, when there is smoke in a room, this system could identify the instance and inform the users. The information is sent through wireless network to a database where the users can access it from anywhere.

In order to make sure that the level of IAQ is within the acceptable level, many parties especially building administrators have made a considerable effort. Some researchers found that low level of IAQ could affect the quality of life of the occupants and may result in low productivity [1]. Modern buildings, especially, need more attention on IAQ because these buildings have been built with the concern of energy conservation [2]. Following the oil crisis in the late 70s, the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) had come out with certain guidelines for new buildings designed to be energy efficient [3]. This change in the practice of modern buildings however has increased the possibility for building occupants to develop Sick Building Syndrome (SBS) and some other Building Related Illness (BRI).

The standards for IAQ have been issued by various parties such as ASHRAE, WHO, US EPA and by individual country like Malaysia, Singapore and Hong Kong. For the purpose of this study, the IAQ standards issued by US EPA and ASHRAE have been followed closely. IAQ, as defined by the ASHRAE in Standard 62-2001: Ventilation for Acceptable Indoor Air Quality, is air in which there are no known contaminants, as harmful concentrations are determined by cognizant authorities and with which a substantial majority (80% or more) of the people exposed to it do not express dissatisfaction. Even though IAQ problems happen occasionally, it could be severe when they do happen. The studies conducted by the United States Environment Protection Agency (US EPA) showed that pollutants present in indoor environments sometimes can be more dangerous than pollutants in outdoor environments. In most cases, the concentration of pollution was usually about five times higher than the concentration level found outside. However, in severe cases, the concentration of pollution could be 100 times higher [1]. Among the most common indoor air pollutants include gases, chemicals and living organisms like mold and pest. These pollutants could impair the health of the building occupants. Sore eyes, headaches and fatigue are among the common complaints associated with low

IAQ level. In serious cases, occupants always complain about respiratory related illness, heart disease, and some may even suffer from cancer and other serious health problems [4,5]. At a very high concentration, pollutants like carbon monoxide can lead to death [6].

This study aims to propose a smart IAQ monitoring system where the sources influencing IAQ can be identified. This study uses ANN to train the system to recognize the sources influencing the IAQ level in five conditions: ambient air, chemical presence, fragrance presence, foods and beverages presence and human activity. In artificial intelligence (AI) field, ANN is one of the major components which use a modeling technique inspired by the human brain (memory-less processing elements known as neurons or nodes which are nonlinearly interconnected) and which can be trained from examples. When ANN is properly trained, it can cleverly classify a data to a population. This attribute of ANN is so special that it may give better results as compared to other techniques in the same situation. Therefore, many researchers have opted for the use of ANN including in environmental sciences—usually to predict atmospheric concentration of NO₂, Ozone, benzene and PM₁₀ [7–9].

This IAQ monitoring system consists of three parts: sensor module cloud, base station and service-oriented client. The sensor module cloud (SMC) contains collections of sensor modules that measure the air quality data and transmit the captured data to base station through wireless network. Each sensor modules includes an integrated sensor array that can measure indoor air parameters like Nitrogen Dioxide (NO₂), Carbon Dioxide (CO₂), Ozone (O₃), Carbon Monoxide (CO), Oxygen (O₂), Volatile Organic Compounds (VOCs) and Particulate Matter (PM₁₀) along with temperature and humidity.

2. Related Works

Some researchers have proposed a framework or a system to monitor IAQ using wireless sensor networks such as [10–12], while others have incorporated index calculation to determine the IAQ level such as [13–15].

Hui Xie *et al.* [7] exploited the ANN in order to predict the SBS. The results were compared to the other linear regression models. It was concluded that ANN gave better predictions and could be useful in other areas as well. Ghazali *et al.* [16] also developed ANN air quality prediction model using feed-forward network. The result is similar to Hui Xie *et al.* [7] where the model with ANN structure produced the best prediction. Abd Rahman *et al.* [17] proposed a forecasting of air pollution index. The forecasting was based on 10-year monthly data of Air Pollution Index (API) located in industrial and residential monitoring stations area in Malaysia. The autoregressive integrated moving average (ARIMA), fuzzy time-series (FTS) and ANN were used to forecast the API values. The results showed that ANNs gave the smallest forecasting error as compared to the other two methods. Clearly, ANN is useful in decision making processes for air quality control and management.

Previous researches on IAQ monitoring system usually used the ANN to predict or forecast either the value of air quality or the value of air pollution. This project is different from previous researches because this project aimed at training the ANN to classify and identify the sources that influence the indoor environment. The functionality of this system and the methodology of this project are explained in the next section.

At present, odour has also been included as part of sources influencing IAQ [18]. Therefore, many studies had been carried out to use the IAQ monitoring system for odour detection in real time. In fact,

the IAQ monitoring system is the only system available to get continuous, quick and reliable information about the presence of odours in ambient air [18]. However, the odour recognition principle was mainly used to investigate the odour in outdoor environments [19–22]. In indoor environments, odour recognition was usually used to detect odour from one single category. For example, the use of odour recognition to recognize types of mushrooms, oil flowers and pure chemicals only [8,15,23]. Nonetheless, odour recognition from sources of indoor air pollutants has not been carried out.

3. Selection of IAQ Parameters

There are various parameters involved in measuring IAQ. These parameters are divided into four categories: physical condition, chemical contaminants, biological contaminants, and other common contaminants. However, for the purpose of this project, only nine parameters have been chosen. These parameters have been identified as the parameters that are most used in measuring the IAQ [24–27]. The parameters that are used in this project include parameters of common indoor air contaminants (CO_2 , CO , O_3 , NO_2 , VOCs, O_2 and PM_{10}) and thermal comfort parameters (temperature and humidity).

4. System Architecture

The proposed system has been designed in such a way that it is able to monitor air quality in both indoor and outdoor environments. Nine sensors have been used to capture data for the nine parameters required in this project. Figure 1 shows our proposed system architecture for real time IAQ monitoring. The proposed system consists of sensor module cloud (SMC), base station and service-oriented client. The SMC contains an array of sensor modules that captures the air quality data. The data are then transmitted to the base station through wireless connection where the data are stored in a server. The server functions as data logger to keep track of data received in base station. It stores the data in database, processes the data, performs analysis and provides information about IAQ information through web service. The web service enables the clients or users to be informed about the IAQ level in real-time.

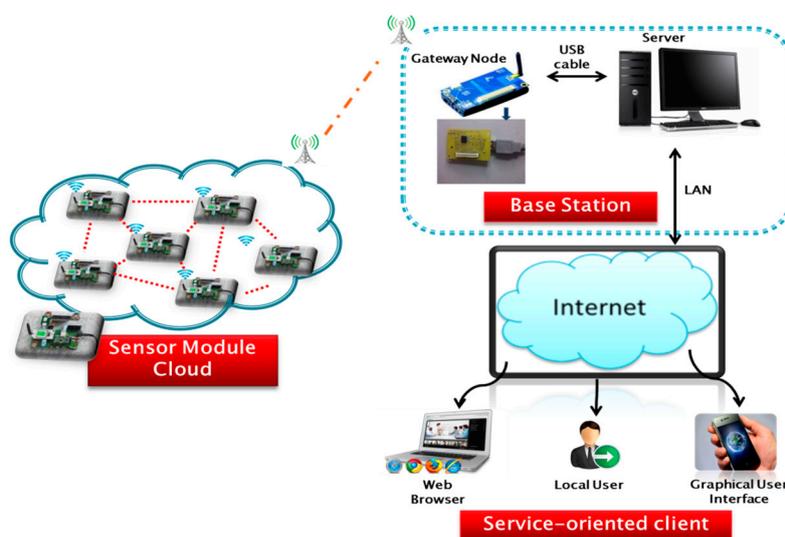


Figure 1. System architecture for real-time IAQ monitoring.

4.1. Sensor Module

The sensor module is used to sense and measure air quality in indoor environment. Figure 2 depicts a schematic diagram of the major components of the sensing module. Basically, there are three major components: sensor components, microcontroller and wireless transceiver. This entire component is powered-up by using wall adaptor for continuous power supply.

In this study, a microcontroller unit (MCU) named STC12C5A32S2 with built in 10 Bits Analog-To-Digital Converter (ADC) converts the analogue reading for each sensor response to digital value. STC12C5A32S2 is chosen because this 8-bit microcontroller processes 12 times faster than general 80C51 (it can go up to 35 MHz operation frequency) [28]. After the conversion of the sampling voltage by ADC, the microcontroller then is responsible for encapsulating the converted data into packets and passing the packets via serial port to the wireless transceiver.

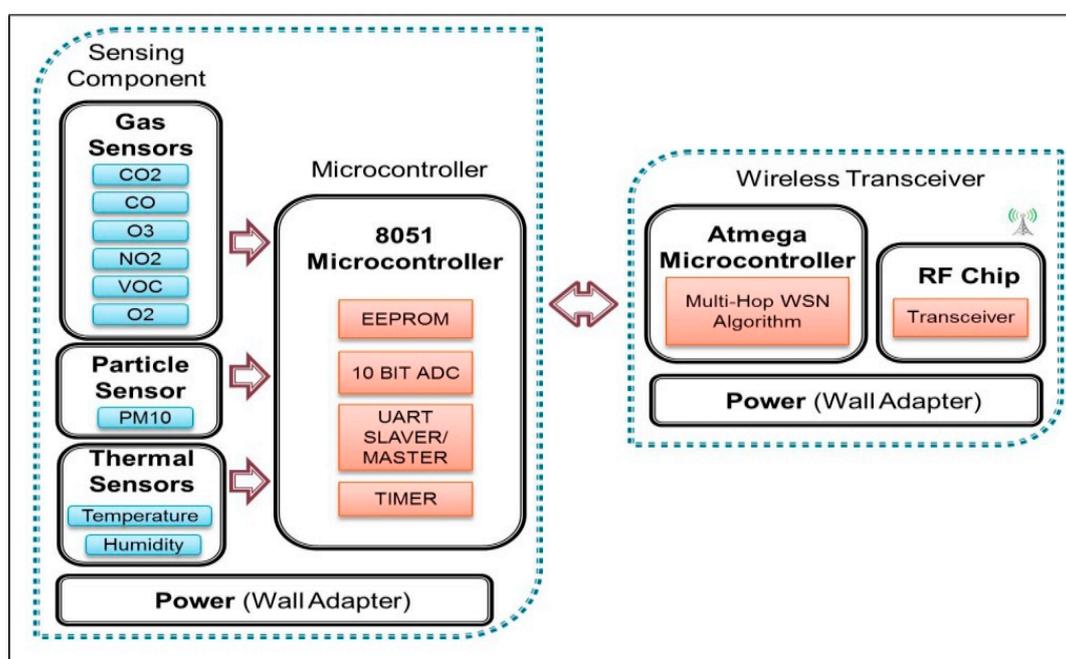


Figure 2. Block diagram of the sensing node.

The wireless transceiver unit transmits the data packets by using Multi-Hop Wireless Sensor Network (WSN) algorithm through the Radio Frequency (RF) chip. Implementation of WSNs in IAQ monitoring has reduced the installation costs due to wire depositions. The IRIS mote from Memsic Company has been selected as the wireless transceiver unit for its low power consumption, low price and complies to the IEEE 802.15.4 wireless protocol [29]. The operating system used in this mote is based on TinyOS which allows the user to quickly implement the communication network [30]. In this system, each IRIS mote is responsible to receive data from sensor components through microcontroller. The data from IRIS is sent to the base station. However, if there is another IRIS mote nearer to the base station, the data is sent to its neighbour (the other IRIS mote) before sending the data to the base station. This data transmission process is known as multi-hopping process. Figure 3 shows the prototype sensor module, consisting of nine gas sensors, and wireless transceiver with microcontroller attached to the wireless transceiver.

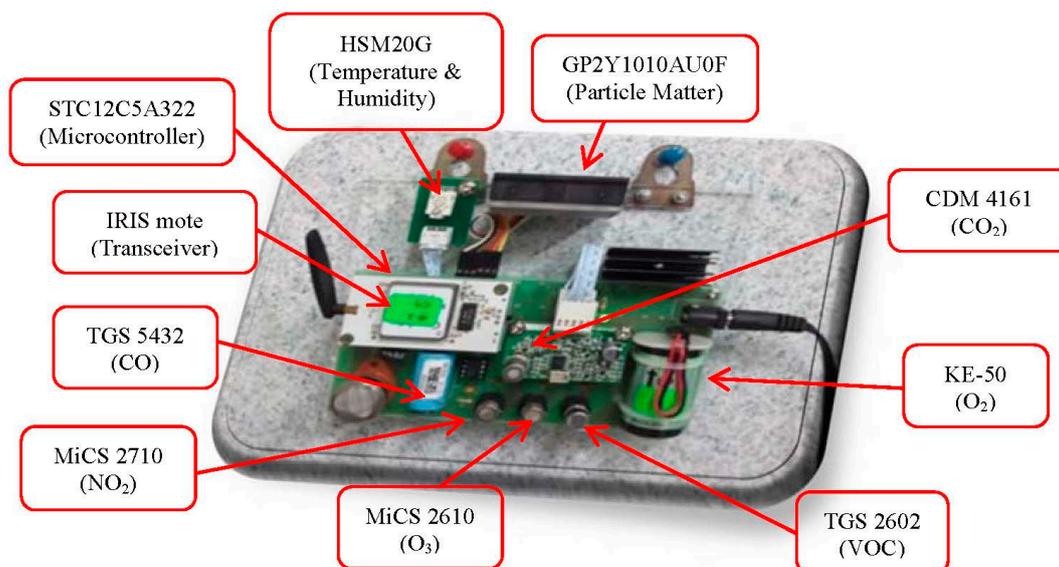


Figure 3. Prototype of sensor module.

The sensor components consist of three types of sensors: gas sensor, particle sensor and thermal sensor. Lists of sensors used in the proposed system along with their operational range are presented in Table 1 below. Each sensor generates voltage signal based on current environment. These sampling voltage levels are read by microcontroller periodically.

Table 1. List of sensors used in the system.

No	Sensor Name	Sensor Type	Manufacturer	Target Gas	Typical Detection Range
1	CDM 4161	MOS	Figaro	CO ₂	400–2000 ppm
2	TGS 5342	Electrochemical	Figaro	CO	0–100 ppm
3	TGS 2602	MOS	Figaro	VOCs	0–30 ppm
4	MiCS-2610	MOS	SGX Sensortech Limited	O ₃	0.01–1 ppm
5	MiCS-2710	MOS	SGX Sensortech Limited	NO ₂	0.01–5 ppm
6	KE-25	Electrochemical	Figaro	O ₂	0%–100%
7	HSM20G	Thermal	GeeTech	Humidity	20%–95% RH
		Thermal		Temperature	0–50 °C
8	GP2Y1010AU0F	Optical	SHARP	PM ₁₀	0–0.5 mg/m ³

Selecting a proper gas sensor is a relatively complicated issue as many factors need to be taken into consideration. For this study, most of the gas sensors are metal oxide based, while the rest are electrochemical based. These Metal Oxide Semiconductor (MOS) sensors contain tin dioxide (SnO₂) sensing element that responds to the gas molecules, which are typically volatile compounds [31]. It consists of two major parts, namely the heater and sensor substrate. The substrate has two terminals and its resistance is measured as a representation of the amount of gas concentration in the environment while the heater provides the stabilized temperature needed for the measurement [32]. Due to its long lifetime, high sensitivity response and low cost, this type of sensor is commonly used in many indoor applications such as homes, offices and factories appliances. The second type of gas sensor used in this study is electrochemical based sensor. This type of sensor has high sensitivity to environmental change and it does not need power to operate.

However, this type of sensor has its limitations. These low cost sensors cannot provide accurate readings of the gas presence in the air since they are strongly affected by the temperature, humidity and the presence of other gases. To compensate this limitation, the patterns are observed instead of the actual data measured. The sensor resistance and the gas concentration in the environment interact in the following expression.

$$R_s = A(C) - \alpha \quad (1)$$

where R_s is the sensor resistance of the sensor, A is constant, C is the gas concentration and α is the slope of R_s curve [33].

To detect the physical contaminant like dust or particulate matter, Sharp GP2Y1010 optical dust sensor from Sharp has been chosen. This sensor contains an infrared emitting diode and a phototransistor which are arranged diagonally. Any dust that gets into the sensor makes the infrared light reflect which can be detected by phototransistor. The advantage of this dust sensor is it can detect small particles like cigarette smoke. It is commonly used in an air purifier system because it is small, cheap, robust and uses low current consumption [34]. For thermal sensor, HSM-20G sensor is used to detect temperature and humidity. HSM-20G is an analog sensor that converts the ambient temperature and relative humidity into standard output voltages (1–3.3 V). It can measure the temperature between the ranges of 0–50 °C. It is also able to detect the relative humidity (RH) from 20% to 95% with accuracy $\pm 0.5\%$ at 25 °C [35]. With the calibration curve provided in the datasheet, these analog voltages can be converted to the unit of temperature (°C) and relative humidity (%).

4.2. Base Station

The base station in IAQ monitoring system contains two components which are wireless transceiver and a server that act as a data logger. The base station is responsible for managing, collecting and recording the data before displaying it on a computer screen and on the web service. The wireless transceiver unit is similar to the sensor module, which contains IRIS mote—ATmega1281 low power microcontroller and AT86RF230 radio frequency (RF) [29]. It also contains Future Technology Devices International (FTDI) device which emulates RS 232 transmission protocol and communicates with Data Processing Module (DPM). DPM is responsible for processing and writing the air quality data into the database. At the same time, DPM sends the data to the Web Service which allows users to access the information in real-time. A simple database is based on a SQLite format which is used to log the data for further processing (if necessary) on the server system. SQLite has been selected instead of alternatives like MySQL simply due to the fact that it is easier to setup and it uses single file storage. However, due to its limitation of storage, the DPM has been programmed to create one SQLite file which contains one week of data, and this process is repeated for the following weeks. Figure 4 shows the block diagram of the base station.

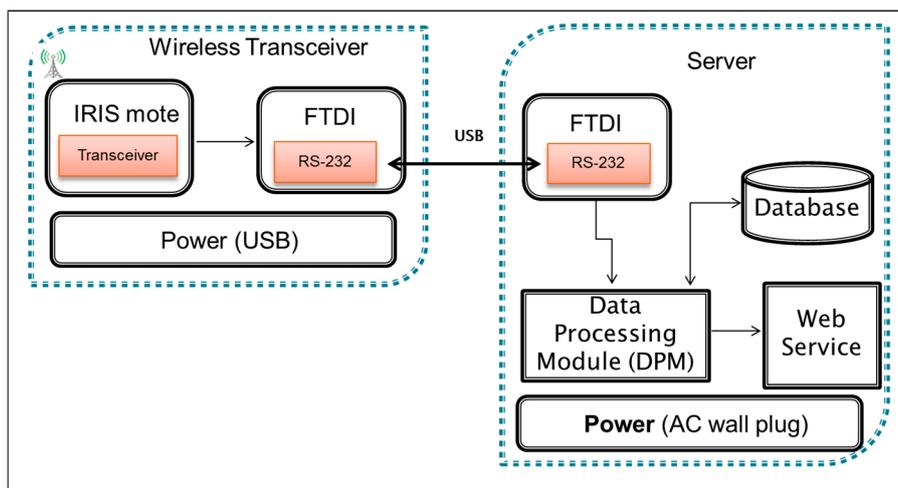


Figure 4. Block diagram of base station.

4.3. Service-Oriented Client

Service-oriented client provides all information on IAQ accessible in real-time. The data is shared through self-developed Graphical User Interface (GUI). The GUI is the interface that facilitates the users to interact with the programs. In this research, the GUI was developed by using LABVIEW software program. In order to stream sensor data for web service that was located in the server, the GUI adopted websocket technology. Figure 5 shows the visual of GUI that has been developed.

This GUI provides a map location of the sensor modules which are placed at different locations such as meeting room, lecture room and postgraduate room (this study uses CEASTech Institute in UNIMAP, Perlis as its base location). It also provides the current value IAQ parameters with color-coded graph based on IAQ index level. The system receives the data in the forms of voltages. These voltages are compared to the specified ANN output model that has been trained and then the data are classified according to the training set (any one of the five categories trained). The GUI then displays the information of the source of activity influencing IAQ level.

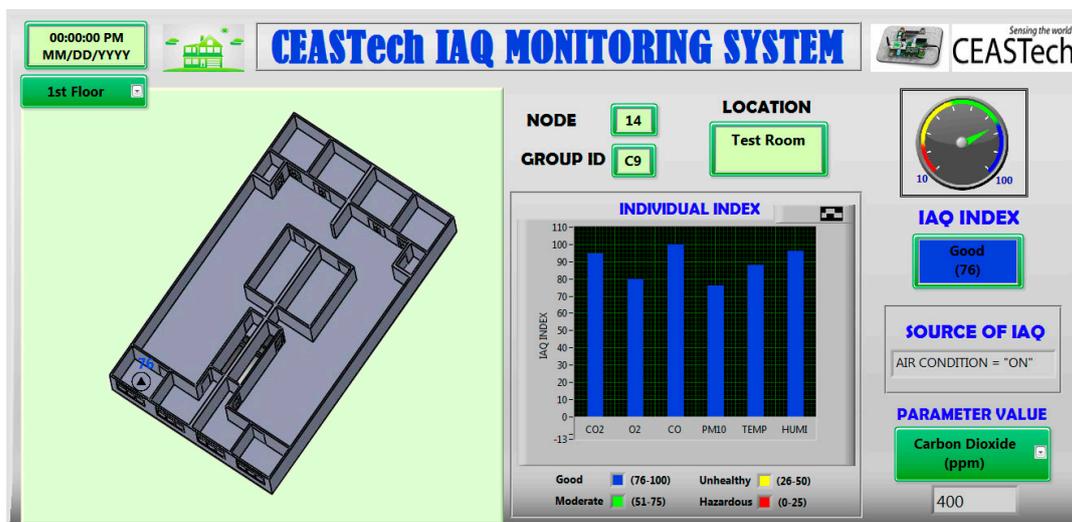


Figure 5. GUI visualization for IAQ monitoring system.

5. Development of IAQ Index (IAQI)

Indoor environmental index has been developed to identify the quality of indoor air as well as the comfort level of the occupants inside a building. To achieve this, two separate indices have been developed: indoor air quality index (IAQI) and thermal comfort index (TCI). Both indices are divided into four status categories as shown in Table 2 below. The indoor environmental index was actually modified from the U.S. EPA Air Quality Index (AQI) which was used for outdoor air so that it fits with the indoor air parameters.

Table 2. Index values and status categories.

Index Value	Status	
	IAQI	TCI
76–100	GOOD	MOST COMFORT
51–75	MODERATE	COMFORT
26–50	UNHEALTHY	NOT COMFORT
0–25	HAZARDOUS	LEAST COMFORT

6. Methodology

6.1. Calibration and Validation

Calibration of gas sensors is one of the main challenges during the development phase. Although calibration was carried out for all sensors, the discussion was limited only to the calibration of CO₂. Calibration of the CO₂ sensor was performed in a laboratory environment with each sensor mounting in a completely sealed gas test box SR3 (manufactured by Figaro) specially designed for the testing of gas sensors to avoid any other gases affecting the experiment. The size of the box is 235 mm × 180 mm × 210 mm. Figure 6 shows calibration setup for CO₂ gas sensor which was set in the test box.

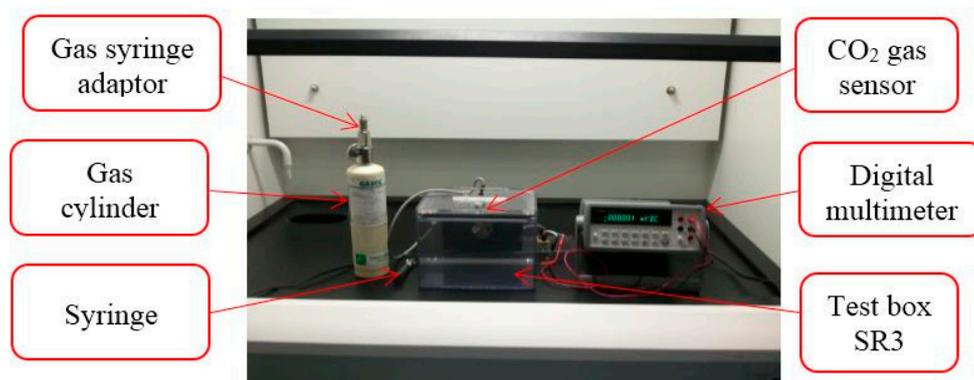


Figure 6. Calibration setup.

For the use of the SR3 box, the method specified by the manufacturer was used. Initially, the test box was left open in a clean environment and mixing fan was turned on for 3 min to ensure that all contaminants had been removed. After that, a lid was put on the box. Subsequently, the syringe was filled with a volume of CO₂ extracted from a gas syringe adaptor of the gas cylinder. High purity gas

($\geq 99.995\%$) was used for the calibration. CO₂ gas was injected in the box through a silicon septum and the mixing fan was turned on for 30 s. A time lapse of 30 s was allowed before reading the sensor output. The lid of the box was removed so that it could return to the 3 min cleaning cycle. During the experiment, room temperature of the calibration environment was maintained at 25 °C. Figure 7 shows the scatterplot of CO₂ sensor calibration result at different CO₂ gas concentrations.

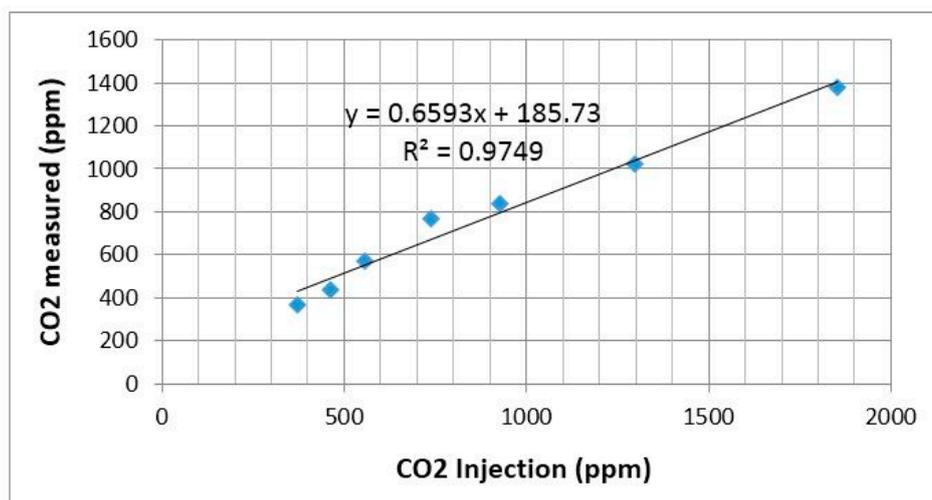


Figure 7. Scatterplot of CO₂ sensor calibration.

It was observed that the output of the gas sensor was linear to the gas sample concentration. Simple linear regression (LR) method was used as the calibration method for the gas sensors. The coefficient of determination (R^2) was 0.97 indicating that the value was a good correlation between the measured data to the gas injected. This data was used to recalibrate the equations provided by the manufacturer in the datasheet.

Validation procedure had also been carried out to make sure that the data collected by the self-developed sensor node was similar to the data collected by a commercialized device. The discussion of this procedure was limited only to the NO₂, temperature and humidity sensors. The sensor validation was carried out with Aeroqual portable indoor monitor device (a professional grade air quality measurement system) which had been pre-calibrated [36]. Three sensor nodes and the Aeroqual device were placed in a clear sealed glass container of 100 × 40 × 30 cm which was completely sealed. Then, the gas concentration inside the sealed container was made varied by injecting the particular gas of interest. The outputs of the sensors were recorded continuously for 1 h and plotted (Figure 8). Figure 8a shows that the result of NO₂ sensor when 35 ppb of NO₂ gas concentration was injected to the sealed container. It can be observed that the value for all sensor nodes including Aeroqual device gave relatively similar readings for 1 h period. During the experiment, the same room temperature setting at 25 °C was applied as shown in Figure 8b while Figure 8c shows the readings for humidity.

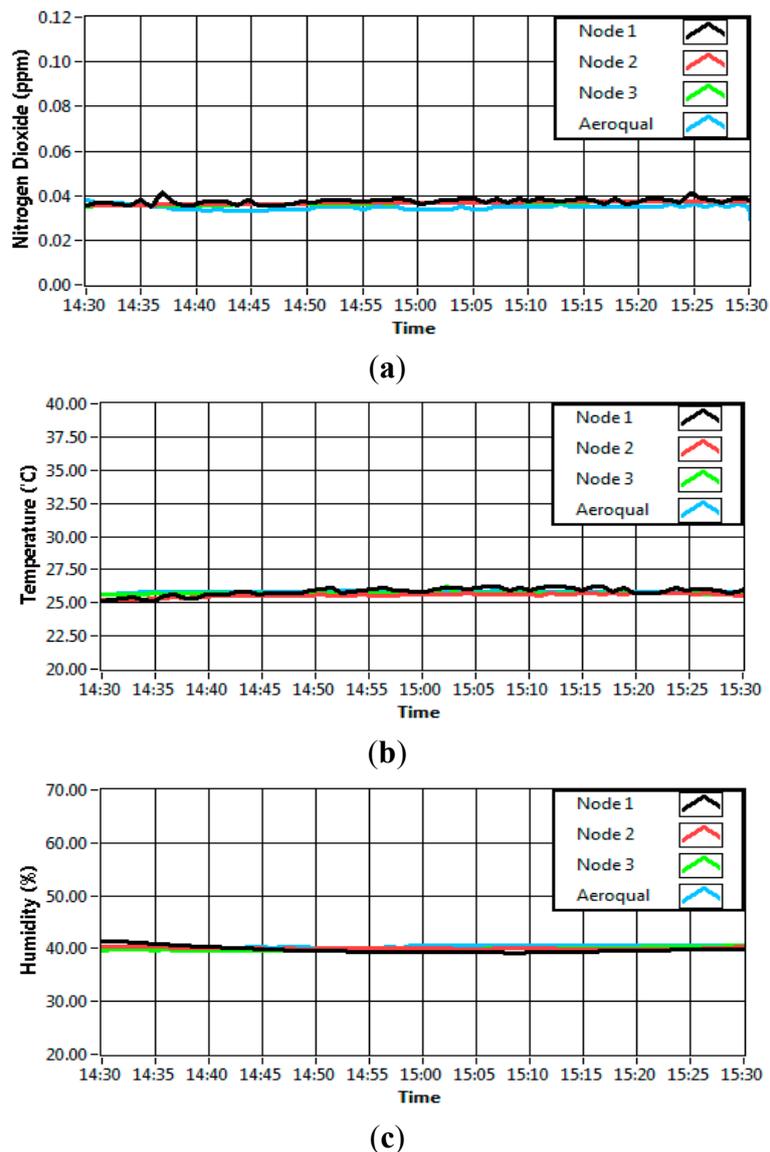


Figure 8. (a) NO₂ data; (b) Temperature data; (c) Humidity data.

For validation purposes, means and standard deviations for all three parameters were calculated as shown in Table 3 below. Also shown, the mean value for NO₂, temperature and humidity of three nodes (Node 1, Node 2 and Node 3) did not differ substantially between Aeroqual (a pre-calibrated device). This shows that the data measured for each developed sensor modules provide a similar response with the calibrated device. The standard deviation (Sd) from the table shows how the data differed from the mean value for each node. Overall, it shows that the developed system provides reliable data.

Table 3. Means and standard deviations for gas and the sensors' calibration.

Parameter	Node 1		Node 2		Node 3		Aeroqual	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
NO ₂ (ppb)	35.9	2.0	34.9	1.7	34.6	1.7	35.5	2.2
Temperature (°C)	25.6	0.3	26.5	0.1	25.7	0.1	25.7	0.1
Humidity (%)	39.7	0.8	39.6	0.6	39.3	0.6	40.2	0.1

6.2. Experimental Setup and Data Collection

The experiment was conducted in a medium-sized room of 4.5 m × 2.4 m × 2.6 m. There was an air-conditioner located at the center of the room at a height of 2.2 m from the ground. The sensor module was installed hanging up to the wall of the room at a height of 1.1 m from the ground. According to the Malaysian Standard on IAQ, the monitoring device or instrument should be positioned at a height between 75 cm and 120 cm, preferably 110 cm from the floor [37]. This position is considered as a breathing zone for the occupants. The node was powered up by using adaptor 7.5 V and was programmed to send the data to the base station every 1 minute. The data collection was conducted in 22 days between 9.00 a.m. and 5.00 p.m. with the room temperature set at 22 °C. Every day, after each experiment, the air in the room was purged out by opening windows to clean the air. Figure 9 shows the process of data collection for all five conditions from day 1–22.

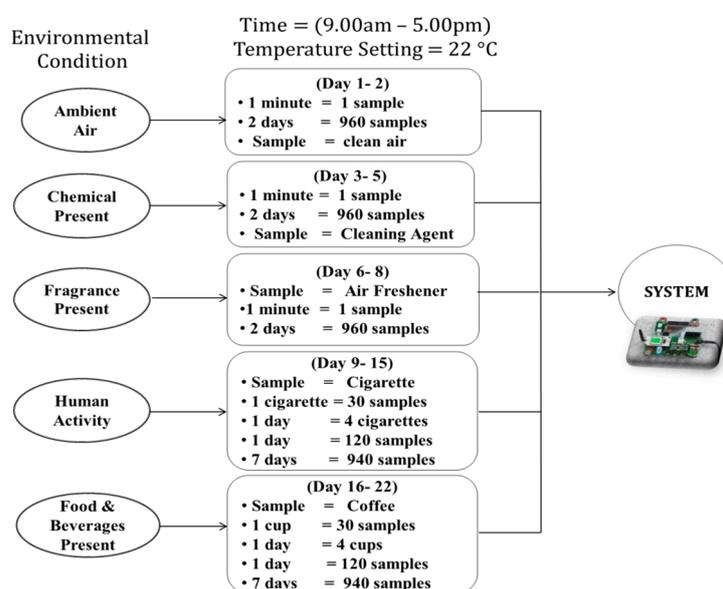


Figure 9. Data collection process.

The first condition was the ambient air environment. The purpose of this experiment was to collect the data of clean air for the room with the assumption that the ambient air was not contaminated. For the first environment, the data collection process took about two days. Thus, at the end of day 2, there were 960 samples collected for ambient air. The second environment was the environment with the presence of chemical substance. In this experiment, a cleaning agent was used as a proxy of the chemical substance. About 100 mL of chemical was put in a beaker and placed inside the room—at the centre of the room. The experiment was repeated for two days and 960 samples were collected during that period. For the third environment, an air freshener was used as a proxy to the fragrance. An automatic air freshener which released fragrance every 15 min was placed inside the room. It was hung up on the wall at a height of 2 m from the floor and about 2 m from the sensing node. The data was collected for two days with 960 samples. For the fourth condition of room environment with human activity, a person smoking a cigarette was chosen as a proxy. A person was asked to smoke in the room so that the real data of a person smoking cigarette in a room could be collected. That person smoked one cigarette at the centre of the room. Every cigarette produced data for approximately 30 min. The experiment was

repeated four times a day for seven days. The amount of data collected for environment with human activity was 940 samples. Lastly, for the room environment with the presence of food and beverages, coffee had been selected to represent this category. A cup of coffee was placed in the middle of the room. Each cup of coffee manufactured 30 min of aroma. The experiment was repeated four times a day for seven days with 940 samples collected.

7. Result and Discussion

7.1. Sensor Response

Figure 10 below shows the sensors response toward five different environments which are ambient air, human activity, chemical presence, fragrance product (air freshener) and foods and beverages (coffee). The data was taken from one of samples of the experiments. They showed that the values of the sensors of the module give different responses to the environment. The response of the sensors indicated that the sensing module of the system functions accordingly to the varying sample concentration of different environment conditions. From the graphs, it is clear that the sensor node gave different reactions towards different environments. In Figure 10a, the sensors gave a relatively steady reading throughout the time. The sensors' response was as expected since there was no substance which could interrupt the ambient air concentration. On the other hand, in Figure 10b, with the presence of chemical substance, it can be observed that certain gas sensors such as VOC, NO₂ and O₃ reacted differently as compared to ambient environment. The reading of VOC gas sensor particularly, rose sharply when the chemical was present in the room. Figure 10c represents the response of the sensors when the automatic air freshener released fragrance into the room every 15 min. The fragrance of the air freshener, however, vaporized quickly into the air after it was released. Thus, these changes of high and low concentration of fragrance in the air could be observed from the disturbed graph. Meanwhile, for the last two environments, 30 min of data were recorded instead of 8 h because these two environments had impact on the sensors for a short amount of time. Figure 10d illustrates the effect of the cigarette smoking activity on the sensor while Figure 10e denotes the presence of food and beverages (coffee) in the room. Notably, in all graphs, different sets of gas sensors reacted differently to different environments.

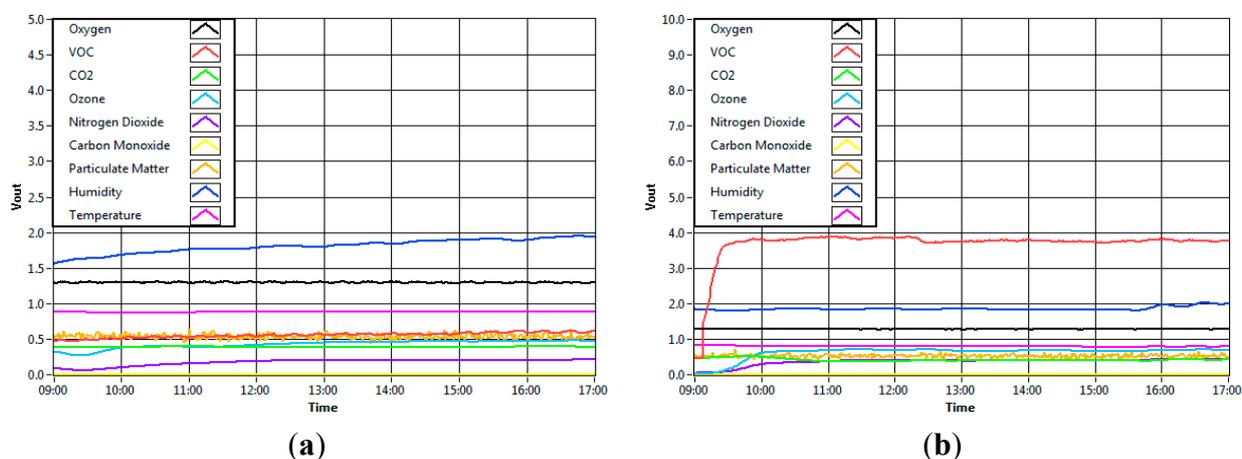


Figure 10. Cont.

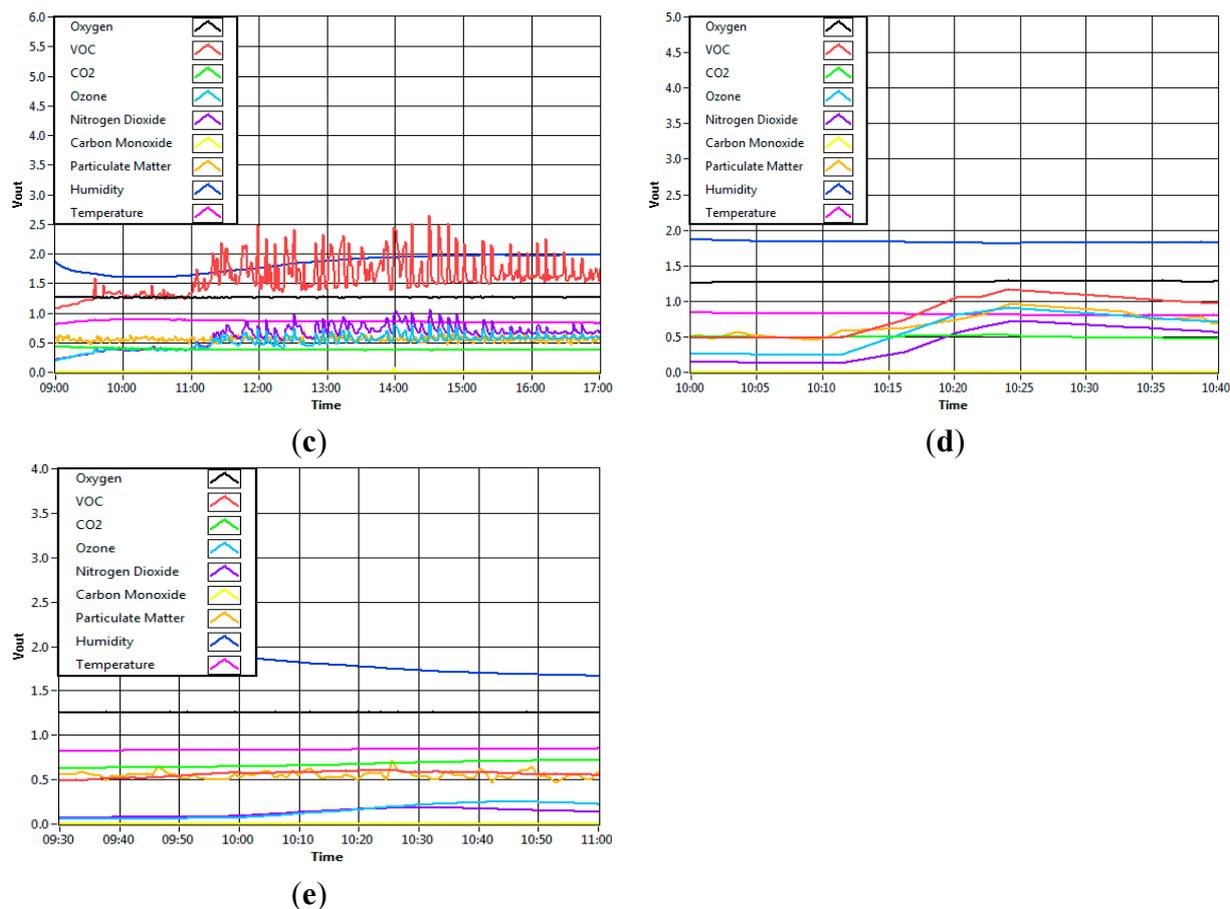


Figure 10. (a) Ambient environment; (b) Chemical presence; (c) Fragrance presence; (d) Human activity; (e) Food and beverages

7.2. Principal Component Analysis (PCA)

The PCA was implemented to distinguish different varieties of the samples used in all environments. It is an unsupervised pattern recognition technique used to cluster the data according to groups. The techniques will reduce the size of data variable without losing the information [38]. Each principal component is a linear combination of the original variables as defined by the equation given below:

$$PC_p = w_{1p}x_1 + w_{2p}x_2 + \dots + w_{np}x_n \quad (2)$$

where PC_p is the notation for p -th order principal component for the overall n number of data, w_{np} is regression coefficient (or weight) determined by PCA while X_n is adjusted matrix. The result of the PCA analysis is shown in Figure 11.

The scores of the five groups of sources of pollutants are plotted for principal component 2 (PC2) versus principal component 1 (PC1). Both components have two great discrimination of 85.330% and 7.717% (or total cumulative variance of 93.047%). From the PCA plot, the samples can be qualitatively clustered into five different groups based on different sources of pollutant.

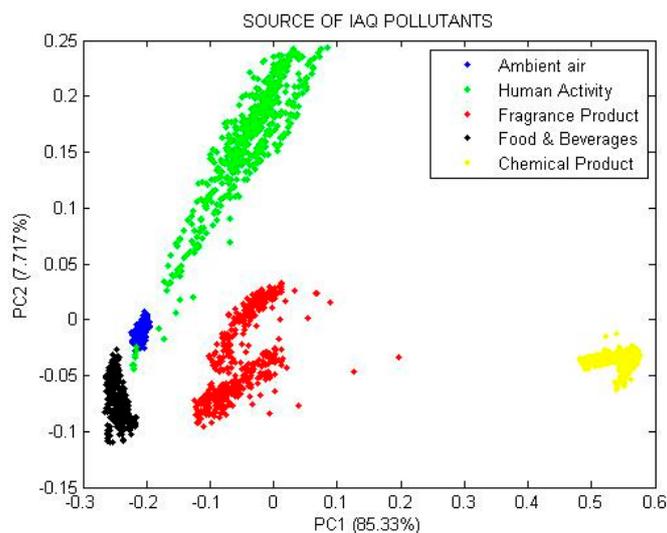


Figure 11. PCA plot of five different sources of IAQ pollutant.

7.3. ANN Analysis

In order to identify the source affecting the IAQ, this study used the multilayer feed-forward neural network which consists of three layers: input layer, hidden layer and output layer. As network architecture, a three-layer perceptron model as shown in Figure 12 was used.

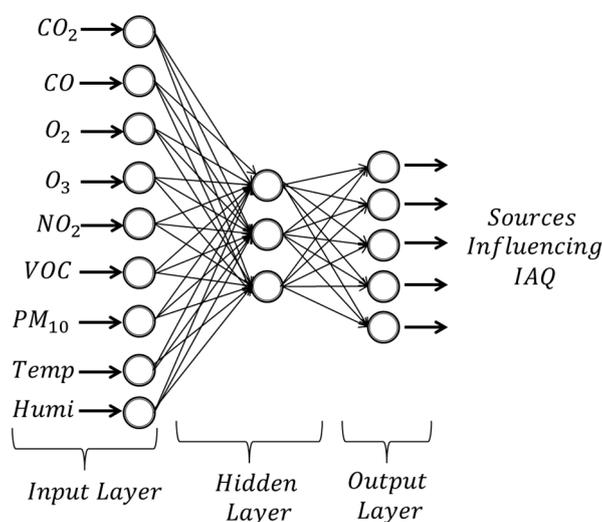


Figure 12. ANN model for source influencing IAQ.

The first input layer contains the input variables for the network. Figure 12 illustrates that the input layer contains nine neurons of IAQ parameters which are CO₂, CO, O₃, NO₂, O₂, VOC, PM₁₀, temperature and humidity. There is one hidden layer used and the number of hidden neurons is to be chosen in the model. The last layer of the model is the output layer which consists of five target outputs. After the system has been trained, it is expected that the system is able to identify the source influencing IAQ based on the five conditions. The ANN analysis was done using LabVIEW Machine Learning Toolkit (MLT) from National Instruments. The toolkit provides various machine learning algorithms in LabVIEW usually classified as supervised or unsupervised method. It is a powerful tool

for pattern recognition, cluster identification and visualization of high-dimensional data. Some of classifiers function based on different ANN-algorithms such as multi-layer perceptron (MLP), self-organizing map (SOM), radial basis function (RB) and support vector machine (SVM) are available in this toolkit. For this research, the ANN-algorithm based on MLP with Back Propagation algorithm has been used to evaluate the system performance in the classification source influencing IAQ. The detailed parameter for ANN training is given in Table 4.

There are 4760 data collected and these data are gathered in a database as the training set. The database contains two data sets: 80% of data were used in training the network and 20% were designated as a testing set. The optimum structure of ANN model is determined by a trial and error method. The number of input neurons is 9, which corresponds to the number of sensors and the number of outputs is five neurons, which corresponds to the five environmental conditions. The number of hidden neurons was adjusted until the desired performance was achieved. In order to ensure convergence of the model, the input data was normalized within the range of [0, 1] based on Equation (3).

$$X_s = \frac{X - X_{min}}{(X_{max} - X_{min})} \quad (3)$$

where X_s is the normalized value, X_{min} and X_{max} are the minimum and maximum value of input, respectively. The normalized data then was randomized and organized in as matrix input for training. An activation function was used to calculate the output response of a neuron. The sum of the input signal was applied with an activation to obtain the response. There are a number of common activation functions in use with neural networks. For this model, the hyperbolic tangent function was employed as the activation function. For each model, the network was trained 10 times. The mean, minimum and maximum classification rates are observed and recorded.

Table 4. Parameters for ANN training.

Training Parameter	Value
Sample	
• Number of samples used for training: 3808	4760
• Number of samples used for testing: 952	
Input	9
Hidden neurons	Flexible
Output neurons	5
Performance	MSE
Goal	0.0001
Learning rate	0.01
Momentum constant	0.5

The results for the network model are tabulated and shown in Table 5. From Table 5, it can be observed that the model with network structure 9-15-5 gives the best accuracy with minimum and maximum classification accuracies of 98.8% and 100%, respectively. The result showed that the test classifications highly correlate with the sample data. The ANN model was able to classify the source that influenced IAQ and the classification rate was 99.1%.

Table 5. Results for network model.

Model Number	Model Structure	Mean Classification Accuracy		
		Minimum Classification (%)	Maximum Classification (%)	Mean Classification (%)
1	9-3-5	29.4	55.0	45.0
2	9-6-5	52.6	65.6	57.7
3	9-9-5	70.0	81.0	75.3
4	9-12-5	76.0	97.0	89.5
5	9-15-5	98.8	100.0	99.1

The confusion matrix for the model is shown in Table 6. Rows and columns represent actual and predicted values, respectively. While observing the confusion level of five different source of IAQ pollutant given in Table 6, it is observed that the chemical, fragrance, food and beverages presence has no confusion level. The human activity environment which is the presence of cigarette smoking has the highest confusion level compared to the other sources.

Table 6. Confusion matrix for five different sources of IAQ pollutant.

Sources of IAQ Pollutant		Predicted					Accuracy (%)
		Ambient	Human Activity	Chemical	Fragrance	Food & Beverage	
Actual	Ambient	189	1	0	0	0	99.47
	Human Activity	6	184	0	0	0	96.84
	Chemical	0	0	191	0	0	100.00
	Fragrance	0	0	0	191	0	100.00
	Food & Beverages	0	0	0	0	190	100.00

8. Conclusions

In this paper, an indoor air quality monitoring system has been developed with the additional function of classifying sources influencing the IAQ based on five different environments such as ambient air, chemical presence, fragrance presence, foods and beverages, and human activity. The data collection has been completed in 26 days in different environment simulations to obtain the desired effects of the five environments. The data for the ANN classification training was collected using a self-developed IAQ monitoring system. Each sensor module in the system contains nine sensors of gases and other parameters that are usually used in IAQ measurement like Carbon Dioxide, Carbon Monoxide, Ozone, Nitrogen Dioxide, Oxygen, Volatile Organic Compound and Particulate Matter, Temperature and Humidity. Based on the results for the network models, the ANN model modelled with network structure 9-15-5 was able to classify the sources influencing the IAQ with a minimum and maximum of 98.8% and 100%, respectively. On average, the system was about 99.1% correct. Overall, it can be concluded that the system delivered a high classification rate based on ANN.

Acknowledgments

The equipment used in this project was provided by Universiti Malaysia Perlis (UniMAP). Special thanks to the members of the Centre of Excellence for Advanced Sensor Technology (CEATech), UniMAP, Malaysia. Shaharil acknowledges the financial sponsorship provided by Malaysian

Technical University Network (MTUN) grant and the Malaysian Ministry of Higher Education (MOHE) under myBrain15 scheme.

Author Contributions

The presented work is the product of cooperative effort by the whole group members of CEASTech, UniMAP. In particular, Shaharil Mad Saad has performed the experiments, analyzed the data and drafted the manuscript. Allan Melvin Andrew, Ali Yeon Md Shakaff, Abdul Rahman Mohd Saad, Azman Muhamad Yusof @ Kamarudin and Ammar Zakaria contributed to the ideas, developments of certain parts of the project and made critical revisions to the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. EPA Buildings and Their Impact on the Environment: A Statistical Summary. Available online: <http://www.epa.gov/greenbuilding/pubs/gbstatpdf> (accessed on 26 November 2014).
2. Law, A.K.Y. *Development of the Indoor Air Quality Index for Commercial Buildings in Hong Kong*; The Hong Kong Polytechnic University: Hong Kong, China, 2002.
3. ASHRAE. *Standard 62-1989, Ventilation for Acceptable Indoor Air Quality*; American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.: Atlanta, GA, USA, 1999.
4. Mohave, L. Indoor air quality and health. In Proceedings of the Conference Healthy Buildings 2000, Espoo, Finland, 6–10 August 2000.
5. Maroni, M.; Seifert, B.; Lindwall, T. *Indoor Air Quality: A Comprehensive Reference Book*; Elsevier: New York, NY, USA, 1995.
6. Raub, J.A.; Mathieu-Nolf, M.; Hampson, N.B.; Thom, S.R. Carbon monoxide poisoning—A public health perspective. *Toxicology* **2000**, *145*, 1–14.
7. Xie, H.X.H.; Ma, F.M.F.; Bai, Q.B.Q. Prediction of Indoor Air Quality Using Artificial Neural Networks. *2009 Fifth Int. Conf. Nat. Comput.* **2009**, *2*, 414–418.
8. Archie, L.W. *Prediction of Odor Pleasantness Using Electronic Nose Technology and Artificial Neural Networks*; The Pennsylvania State University: University Park, PA, USA, 2006.
9. Mumyalmaz, B.; Karabacak, K. An E-Nose-based indoor air quality monitoring system: Prediction of combustibles. *Turkish J. Electr. Eng. Comput. Sci.* **2015**, doi:10.3906/elk-1304-210.
10. Choi, G.H.; Choi, G.S.; Jang, J.H. A framework for wireless sensor network in web-based monitoring and control of indoor air quality (IAQ) in subway stations. In Proceedings of the 2009 2nd IEEE International Conference on Computer Science and Information Technology, Beijing, China, 8–11 August 2009; pp. 378–382.
11. Bhattacharya, S.; Sridevi, S.; Pitchiah, R. Indoor air quality monitoring using wireless sensor network. In Proceedings of the Sixth International Conference on Sensing Technology, Rome, Italy, 19–24 August 2012; pp. 422–427.

12. Preethichandra, D.M.G. Design of a smart indoor air quality monitoring wireless sensor network for assisted living. In Proceedings of the IEEE International Instrumentation and Measurement Technology Conference, Minneapolis, MN, USA, 6–9 May 2013; pp. 1306–1310.
13. Postolache, O.A.; Pereira, J.M.D.; Girao, P.M.B.S. Smart sensors network for air quality monitoring applications. *IEEE Trans. Instrum. Meas.* **2009**, *58*, 3253–3262.
14. Al-Ali, A.R.; Zualkernan, I.; Aloul, F. A mobile GPRS-sensors array for air pollution monitoring. *IEEE Sens. J.* **2010**, *10*, 1666–1671.
15. Kim, J.; Chu, C.; Shin, S. ISSAQ: An integrated sensing systems for real-time indoor air quality monitoring. *IEEE Sens. J.* **2014**, *14*, 4230–4244.
16. Ghazali, S.; Ismail, L.H. Air quality prediction using artificial neural network. In Proceedings of the International Conference on Civil Environmental Engineering Sustainability, Johor Bahru, Malaysia, 3–5 April 2012; pp. 1–5.
17. Haizum, N.; Rahman, A.; Hisyam, M.; Talib, M. Forecasting of air pollution index with artificial neural network. *J. Teknol.* **2013**, *63*, 59–64.
18. Capelli, L.; Sironi, S.; Rosso, R. Del Electronic noses for environmental monitoring applications. *Sensors* **2014**, *14*, 19979–20007.
19. Nicolas, J.; Romain, A.C.; Wiertz, V.; Maternova, J.; André, P. Using the classification model of an electronic nose to assign unknown malodours to environmental sources and to monitor them continuously. *Sens. Actuators B Chem.* **2000**, *69*, 366–371.
20. Sironi, S.; Capelli, L.; Céntola, P.; Del Rosso, R.; Il Grande, M. Continuous monitoring of odours from a composting plant using electronic noses. *Waste Manag.* **2007**, *27*, 389–397.
21. Sohn, J.H.; Pioggia, G.; Craig, I.P.; Stuetz, R.M.; Atzeni, M.G. Identifying major contributing sources to odour annoyance using a non-specific gas sensor array. *Biosyst. Eng.* **2009**, *102*, 305–312.
22. Dentoni, L.; Capelli, L.; Sironi, S.; del Rosso, R.; Zanetti, S.; Torre, M. Della Development of an electronic nose for environmental odour monitoring. *Sensors* **2012**, *12*, 14363–14381.
23. Loutfi, A.; Coradeschi, S. Odor recognition for intelligent systems. *IEEE Intell. Syst.* **2008**, *23*, 41–48.
24. Gang, C.; Ma, W.J. Monitoring and analyse of indoor air quality at several important public spaces in HuaiNan. In Proceedings of the 5th International Conference on Bioinformatics and Biomedical Engineering, Wuhan, China, 10–12 May 2011; pp. 1–4.
25. Borkar, C. Development of Wireless Sensor Network System for Indoor Air Quality Monitoring. Master's Thesis, University of North Texas, Denton, TX, USA, 2012.
26. Guo, H.; Lee, S.C.; Chan, L.Y. Indoor air quality in ice skating rinks in Hong Kong. *Environ. Res.* **2004**, *94*, 327–335.
27. Sekhar, S.C.; Ching, C.S. Indoor air quality and thermal comfort studies of an under-floor air-conditioning system in the tropics. *Energy Build.* **2002**, *34*, 431–444.
28. STC12C5A60S2 series MCU. Available online: <http://www.stmcu.com/datasheet/stc/stc-ad-pdf/stc12c5a60s2-english.pdf> (accessed on 24 April 2015).
29. MEMSIC Cooperation. *IRIS Sensor Node Module*. Available online: http://www.memsic.com/userfiles/files/datasheets/wsn/iris_datasheet.pdf (accessed on 24 April 2015).
30. De Vito, S.; Fattoruso, G.; Liguoro, R.; Oliviero, A.; Massera, E.; Sansone, C.; Casola, V.; di Francia, G. Cooperative 3D air quality assessment with wireless chemical sensing networks. *Procedia Eng.* **2011**, *25*, 84–87.

31. Martin, Y.G.; Concepci, M.; Oliveros, C.; Luis, P.; Cordero, B.M. Electronic nose based on metal oxide semiconductor sensors and pattern recognition techniques: Characterisation of vegetable oils. *Anal. Chim. Acta* **2001**, *449*, 69–80.
32. TGS 2602. Available online: <http://biakom.com/pdf/TGS2602.pdf> (accessed on 24 April 2015).
33. Figaro. Technical Information for Air Quality Control Sensors. Available online: <http://www.figarosensor.com/products/2602Dtl.pdf> (accessed on 24 April 2015).
34. Budde, M.; Busse, M.; Beigl, M. Investigating the use of commodity dust sensors for the embedded measurement of particulate matter. In Proceedings of the 2012 Ninth International Conference on Networked Sensing Systems (INSS), Antwerp, Belgium, 11–14 June 2012; pp. 1–4.
35. SN-HMD Humidity Sensor User's Manual. <https://www.yumpu.com/en/document/view/24407013/sn-hmd-humidity-sensor-users-manual-electro-tech-online> (accessed on 24 April 2015).
36. Aeroqual. Series-200–300–500-Portable-Monitor-User-Guide-11–14. Available online: <http://www.aeroqual.com/wp-content/uploads/Series-200-300-500-Portable-Monitor-User-Guide-11-14.pdf> (accessed on 24 April 2015).
37. DOSH. *Industry Code of Practice on Indoor Air Quality*; Department of Occupational Safety and Health (DOSH): Kuala Lumpur, Malaysia, 2010.
38. Jolliffe, I.T. *Principal Component Analysis*, 2nd ed.; Springer-Verlag: New York, NY, USA, 2006.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).