

Research Article

Multilevel RNN-Based PM10 Air Quality Prediction for Industrial Internet of Things Applications in Cleanroom Environment

Himawan Nurcahyanto , Aji Teguh Prihatno , Md Morshed Alam ,
Md Habibur Rahman , Israt Jahan , Md. Shahjalal , and Yeong Min Jang 

Department of Electronics Engineering, Kookmin University, Seoul, Republic of Korea

Correspondence should be addressed to Yeong Min Jang; yjang@kookmin.ac.kr

Received 9 July 2021; Revised 14 September 2021; Accepted 1 February 2022; Published 22 March 2022

Academic Editor: Dimitrios Katsaros

Copyright © 2022 Himawan Nurcahyanto et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Adequate air ventilation systems and maintaining environmental air quality are essential things that need to be considered to create an excellent industrial environment in a company. Operating or maintaining ecological air quality is a challenge for companies to optimize the existing work environment. Especially at the economic and business level in the company are facing the main problems. In this case, monitoring and predicting future air quality are needed to maintain air quality conditions. PM10 concentration, humidity, and temperature were used to predict annual and seasonal indoor air quality using a recurrent neural network (RNN) and long short-term memory (LSTM). In this paper, we propose the IAQP method for air quality management systems that combines indoor air quality forecasting based on real-time data. To measure indoor air quality, we predict the outcome from the IIoT sensor and LoRa sensor. The result of prediction is that the multilevel RNN model outperformed the LSTM, as the model demonstrated excellent results and feasibility.

1. Introduction

Atmospheric dynamic and meteorological conditions play an influential role in determining air quality conditions from one season to another. The World Health Organization explained that 30% of newly constructed and industrial companies cause serious indoor air quality (IAQ) problems [1]. Production losses and medical costs should be considered, as they are among the highest expenses that can degrade a company's performance [2]. In industrial companies, pollution is a high-impact problem that can create unhealthy working environments. A company's IAQ should be maintained efficiently to provide employees with healthy and comfortable working environments [3]. In the industrial sector, IAQ is generally assessed by measuring particulate matter 10 (PM10) particles, temperatures, and humidity [4].

In the cleanroom of a semiconductor factory, air quality is among the primary concerns. Entry and exit into the cleanroom can change particle amounts, humidity, and temperature. In industrialized countries such as South

Korea, 80%-90% of the population spends their time inside buildings and companies [5]. The consequences of persistent air pollution, such as nose, eye, and throat irritation, which manifests shortly after exposure to pollutants, are harmful to human health. Therefore, such environments can generate harmful indoor pollutants. The primary pollutants inside a cleanroom are PM10, which should be considered to maintain the environment quality. Other pollutants that should be considered include carbon monoxide (CO), sulfur dioxide (SO₂), carbon dioxide (CO₂), formaldehyde (HCHO), nitrogen dioxide (NO₂), ozone (O₃), total suspended particles, and total volatile organic compounds (TVOC) [6]. Several studies have established the importance of good air quality and strong correlations between exposures to PM [7, 8]. Hung et al. proposed a monitoring system with low-cost materials called uSense and showed real-time concentrations of pollutants around a city [9]. Specifically, uSense can distribute a considerable amount of data and is profitable. Moreover, the system was implemented through an onsite experiment [10]. The measurements of PM2.5 exert

significant effects on investigation results. Typical PM_{2.5} sources are industrial exhaust, coal-burning exhaust, and machine exhaust released outdoors. Thus, several researchers focused on the design architecture of sensors for measuring PM_{2.5} concentrations [11]. Elbayoumi et al. proposed an IAQ system for detecting sources that can influence air quality conditions [12]. In [13, 14], the IAQ monitoring models were introduced and the results showed the conditions of indoor and outdoor air quality containing PM₁₀ and PM_{2.5}. Several optional tests for testing a cleanroom are available, such as the airborne particle count for ultrafine or microparticles; airflow test; air pressure difference tests; installed filter system leakage test; airflow direction tests and visualization; temperature, humidity, and electrostatics tests; particle deposition tests; recovery tests; and containment leak tests [15, 16]. However, most of the manuscripts only focused on monitoring IAQ systems without considering the depth and breadth applications for predicting environmental conditions. Therefore, the demand for IAQ prediction (IAQP) increased, thereby leading to the prediction of air pollutants in living and working spaces. The primary purpose is to predict a situation before it worsens and causes problems.

In this study, the proposed framework elucidated the real-time monitoring of multiple sensors and communication modes and their ability to meet various IAQ prediction requirements in the IoT environment. Using multiple sensors for monitoring was more effective than using single or few sensors. In addition, different communication interfaces were illustrated for cleanroom and changing room-monitoring scenarios. However, researchers developed PM_{1.0}, PM_{2.5}, and PM₁₀ modelling to predict IAQ. Based on difficulties, the prediction model with acceptable accuracy was expanded [14, 17]. Certain parameters, such as air exchange rates, outdoor air pollutants, and meteorological parameters, are considered to realize IAQP [18]. This manuscript is aimed at providing prediction results using deep learning methods, such as a multilevel recurrent neural network (multilevel RNN) and long short-term memory (LSTM). The time series model uses the historical data of the corresponding parameters relating the seasonal and environment aspect.

IAQ is the air quality within and around buildings and structures. The main factors that are responsible for IAQ are particles, microbes, pets and pests, humidity, ventilation, and temperature. Among those factors, we have discussed monitoring of temperature, humidity, and particles which are considered the strong factors of IAQ. In [19], the authors observed a positive correlation between the daily variations in relative humidity (RH) and PM₁₀ concentrations. A higher value was observed during the night time and early morning, with lower values occurring during the afternoons. This rapid drop in PM₁₀ tends to coincide with when RH reaches a value of around 75%. Similarly, in the case of temperature, there is a negative correlation with the PM₁₀. When the temperature reaches to maximum during the daytime, the PM₁₀ concentrations tend to be at their lowest. Conversely, while temperature reaches to minimum overnight, the PM₁₀ levels are at their highest [19].

Since the air quality is affected by those factors, it is very important to monitor and predict them for ensuring the better quality of air in the room. This model uses continuous measurements repeatedly until the data are sufficient for prediction. In terms of predicting seasonal situations, first, this study focuses on predicting the PM₁₀ status inside the cleanroom of a semiconductor factory. Second, the results are compared with the observed PM₁₀, humidity, and temperature data. This work has been designed and developed integrated sensors and communication modes for measuring the indoor and outdoor air quality. Finally, the study conclusions are presented. This paper is arranged as follows: the methodology is described in Section 2, and the experimental setup for indoor and outdoor conditions is comprehensively addressed in Section 3. The paper result and discussion for the performance evaluation of the suggested technique is demonstrated in Section 4. Finally, this paper is concluded in Section 5.

2. Methodology

To provide people with comfortable indoor environments and improve IAQ, a new IAQP scheme is developed for air quality management. Several communication techniques were used in the proposed IAQDs because of different locations of the cleanroom from the gateway. A time series prediction model is used to provide comprehensive results in terms of daily working time, thereby managing the IAQ inside a semiconductor cleanroom. This study conducts three sets of experiments, specifically, two classes of indoor and one class of outdoor experiments. Industrial Internet of things (IIoT) devices are used to monitor the air quality inside the cleanroom, and a LoRa IoT device is used to monitor the outdoor air quality.

2.1. Communication Interface. In this section, different communication techniques are discussed. The advantages of RS485 are low hardware cost, shortest time delay, and low packet drop rate. But, for the long-distance communication, the overall performance is relatively bad due to high-cost wired connection. However, LoRa is a physical-layer system, which is a long-distance wireless transmission [20] system focusing on the distribution of the spectrum connection. In addition, it has advantages in terms of low energy consumption; LoRa is application-based technology and does not need a direct physically wired connection. Table 1 shows the differences among various communication interfaces based on their performance parameters.

In the field of integrated wireless communications, LoRa's spread spectrum architecture can shift the balance between the transmission power consumption and transmission distance and completely change conditions [21]. Table 2 shows the overall and provisional analyses of the proposed IAQDs with various communication interfaces and systems in the recent literature.

2.2. Cleanroom Environment Classification. Cleanrooms are closed spaces or areas built specifically to control and collect a limited amount of airborne particles (air contaminants)

TABLE 1: Experimental scenario situation.

| Interface | Frequency band | Rate | Range | Topology |
|------------|-------------------------------|---------------------------------|--------------------|----------------|
| Modbus RTU | 900 MHz 928 MHz | 19200 kbps | 1 km | Serial bus |
| LoRa | 433 MHz 868 MHz 915 MHz | 0.3-50 kbps | 15-20 km | Star |
| Wi-Fi | 2.4 GHz 5 GHz | 450-600 Mbps Up to 1300 Mbps | 0.46 km 0.24 km | P2P connection |

T = temperature; H = humidity; VOC = volatile organic compounds; PM = particulate matter.

TABLE 2: Comparative interface implementation.

| Ref. | Year | Application | Interface | Sensor |
|------------|------|---------------------|--------------------|-----------------------------------|
| [22] | 2013 | Standalone | Zigbee | CO ₂ |
| [23] | 2014 | LAN, WAN | Zigbee, Wi-Fi | CO, CO ₂ , VOC |
| [24] | 2017 | LAN, WAN | Wi-Fi, Bluetooth | O ₂ |
| [25] | 2018 | Standalone | Zigbee | CO, CO ₂ , PM2.5, T, H |
| This paper | 2021 | Standalone LAN, WAN | RS485, LoRa, Wi-Fi | T, H, PM2.5, PM10 |

and regulate temperature, humidity, air pressure, airflow patterns and air pressure, air movement, vibrations, noise environment, and form and number of microorganisms. Airborne particle contamination is conducted, regulated, and used inside a cleanroom to reduce the introduction, generation, and retention of particles in the room. Generally, the cleanroom technology is divided into three parts, namely, design, testing, and operation [26]. The semiconductor manufacturing industry, which manufactures processors for computers, automobiles, and other machines, is a huge consumer of cleanrooms. Semiconductors are manufactured in cleanrooms with extremely high cleanliness levels to mitigate contamination issues. Contaminants from molecules can be caused by outgassing, oil vapor, alcohols, paints, glues, epoxies, aromatics, and so on [22].

However, there are two types of cleanrooms: turbulent-ventilated cleanrooms and unidirectional-ventilated cleanrooms. Turbulent- and unidirectional-ventilated cleanrooms are also known as “nonunidirectional flow” and “laminar flow” cleanrooms, respectively. In the unidirectional-ventilated model, air sweeps through the room and exits through the floor, thereby cleaning the room of air pollution. In the turbulent-ventilated model, significant considerations include the number and location of air supply diffusers [22], which is shown in Figure 1. Generally, in an air-conditioned room, it is located where the air supply enters the room as an air-dispersing system to ensure satisfactory air mixing in the room and reduce the draft caused by high air velocity.

As shown in Figure 2, in some conventional ventilated cleanrooms, diffusers are not employed owing to the discharging of supplied air directly into the cleanroom. To obtain unidirectional flow and satisfactory conditions for controlling air contamination, this approach was chosen under a filter [22]. As illustrated in Figure 1, improved con-

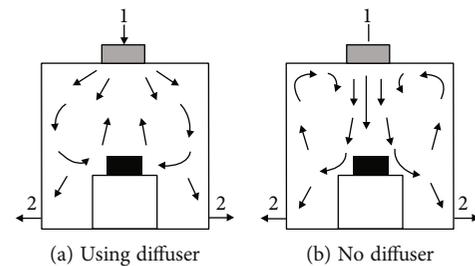


FIGURE 1: Cleanroom pressurization and air movement control.

ditions can be obtained using the “dump” technique below the air supply area. However, the adequate number and size of the diffusers should be determined to promote improved air mixing. If conditions in critical areas must be improved, then ensuring satisfactory air mixing in the cleanroom through diffusers and using unidirectional cabinets or workstations in the critical areas would be safe. If the “dump” method is selected, then ideally, the filters should be uniformly distributed throughout the space. Moreover, the filters can be classified to maintain the environment to be kept clean. However, the dirtiest part determines the quality of the cleanroom, which can indicate a low PM classification [26].

2.3. Particle Classification. A widespread distribution of airborne PM can be observed in the atmosphere. The concentration, particle size, and chemical characteristics of PM can vary considerably over time and space [23]. Therefore, the origin of particles within a certain range should also be considered. Although produced nanoparticles are not associated with ambient PM, they should be included as inhalable particles [24]. As shown in Figure 3, the aerodynamic

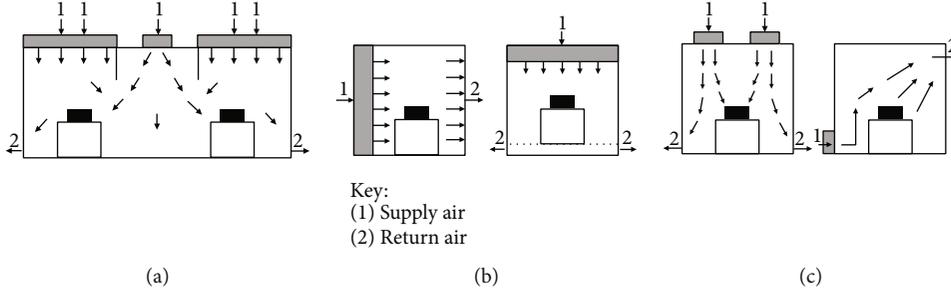


FIGURE 2: Types of cleanroom: (a) mixed airflow, (b) unidirectional airflow, and (c) nonunidirectional airflow.

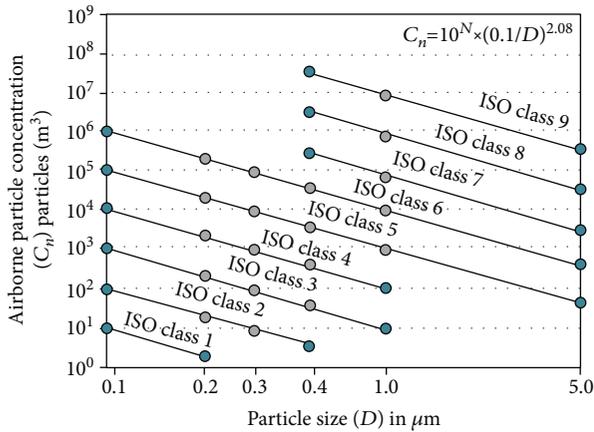


FIGURE 3: ISO classification.

diameter, which is defined as the diameter of a spherical particle with a density of 1 g/cm^3 and the same rate of settling as the particle to be characterized, is a feasible expression of particle size [25].

Airborne PM classifications are described under three room conditions, namely, as-built, at-rest, and operational conditions. As-built conditions imply a completed installation and the absence of manufacturing equipment, materials, and workers. At-rest conditions indicate the absence of staff, and the operational status means that a factory is operating under the prescribed manner and specified work, with numerous staff present. For every particle size D considered, the maximum permissible particle concentration is C_n [22].

$$C_n = 10^N \times \left(\frac{0.1}{D}\right)^{2.08}. \quad (1)$$

Moreover, N is the ISO classification number, which should not exceed the value of 9, and is the maximum allowable concentration of airborne particles equal to or greater than the particle size considered which is described in Table 3. In addition, it is rounded to the nearest integer. Intermediate ISO classification numbers can be defined using 0.1, which is the smallest allowed increment of N , and D in μm is the particle size considered. A constant with a dimension of μm is 0.1 [22]. Cleanroom classification and

preservation must be established to ensure that dirty air from unclean adjacent areas does not enter the cleanroom. Air should flow from the cleanroom to adjacent less clean areas, which means that the manufacturing area should have a higher pressure. If the cleanroom has a higher pressure than the adjacent areas, air can flow to the adjacent areas from the cleanroom.

2.4. Sensor Selection. A general term used to describe consolidated solid particles and liquid droplets in the air is PM, which consists of atmospheric chemical reactions (e.g., from motor vehicles, power plants, and industrial installations burning fuel). Aerosols, smoke, ashes, and pollen are also typically included in PM. As shown in Figure 4, environmental PM can be graded based on an aerodynamic diameter, such as $10 \mu\text{m}$ or less (PM10), $2.5 \mu\text{m}$ or less (PM2.5), or $0.1 \mu\text{m}$ or less (PM0.1). Airborne PM contamination can have serious effects on human health. Exposure to small PM is linked to hospital visits and serious health problems, including premature death [23].

2.5. IAQP Based on LSTM. LSTM is an RNN architecture, in which memory controllers are introduced to determine when to remember, forget, and output [27]. This enables the extension of the training method to learn long-term dependencies. LSTM units are construction units for RNN layers and designed to solve the disappearing gradient problem that occurs when training a traditional RNN. A standard LSTM unit comprises the input gate, forget gate, output gate, and cell unit. The vanilla LSTM unit shown in Figure 5(a) is a cell that acts as a memory and remembers values and includes an input gate, forget gate, and output gate. The gates regulate the flow of input and output information from the cell. LSTM operations can be explained as follows. (1) After the activation of the input gate, new input information accumulates in the memory cell. The first step is to determine the data from the cell state to be forgotten, which is called the “forget gate.” (2) When the forget gate is triggered, the previous status is forgotten by the cell. The second step is to determine the new data to be stored in the cell, which is achieved by the “input gate,” which specifies the values to be changed and the tanh that produces new values for the candidate. (3) The latest cell output propagates the ultimate state when the output gate is triggered. The most essential component is the cell unit C_t , which has a

TABLE 3: Class differentiation of ISO.

| ISO classification number (N) | $\geq 0.1 \mu\text{m}$ | $\geq 0.2 \mu\text{m}$ | $\geq 0.3 \mu\text{m}$ | $\geq 0.5 \mu\text{m}$ | $\geq 1 \mu\text{m}$ | $\geq 5 \mu\text{m}$ |
|-----------------------------------|------------------------|------------------------|------------------------|------------------------|----------------------|----------------------|
| ISO class 1 | 10 | 2 | | | | |
| ISO class 2 | 100 | 24 | 10 | 4 | | |
| ISO class 3 | 1000 | 237 | 102 | 35 | 8 | |
| ISO class 4 | 10000 | 2370 | 1020 | 352 | 83 | |
| ISO class 5 | 100000 | 23700 | 10200 | 3520 | 832 | 29 |
| ISO class 6 | | 237000 | 102000 | 35200 | 8320 | 293 |
| ISO class 7 | | | | 352000 | 83200 | 2930 |
| ISO class 8 | | | | 3520000 | 832000 | 29300 |
| ISO class 9 | | | | 35200000 | 8320000 | 293000 |

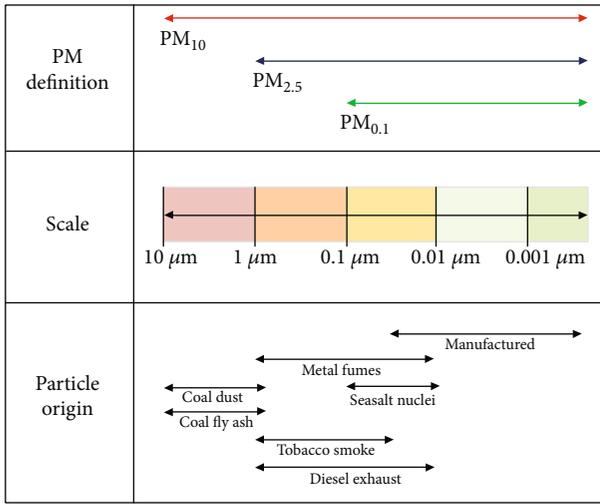


FIGURE 4: Classification of particulate matter (PM) and sources of particles.

linear self-loop regulated by the forget gate unit f_t , which sets the forward contribution of C_t to a value between 0 and 1 [28].

LSTM can solve the gradient loss problem caused by the gradual decrease in the gradient's backward propagation process [28]. Each small unit has three sigmoid functions, namely, the forget gate, input gate, and output gate, given by the LSTM. The three gates perform gating functions to control the entry and exit of data. To achieve various functions and solve various problems, each unit adds a different nonlinear function to different positions in a variant RNN, such as an LSTM. In a conventional RNN, the chain structure unit contains only one nonlinear function for nonlinear data conversion, such as the tanh function [29]. The cells will judge the data in the LSTM. The data that satisfy the rules will be maintained, whereas the information that does not meet the requirements will be forgotten. This model can be used to address the issue of long sequence dependence in neural networks. The forget gate is the first gate in the LSTM architecture. This gate determines the data in the memory cell to be discarded and how much data

from the previous cell will enter and continue to the C_t memory cell.

$$\begin{aligned}
 f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f), \\
 i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i), \\
 o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o), \\
 \tilde{c}_t &= \tanh(W_c[h_{t-1}, x_t] + b_c), \\
 c_t &= f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t, \\
 h_t &= o_t \otimes \tanh(c_t).
 \end{aligned} \tag{2}$$

LSTM is suitable for managing problems closely related to a time series, as it can train long-term dependency knowledge and solve the problem of gradient disappearance caused by the gradual reduction of the backpropagation process. LSTM is an updated version of the RNN and inherits features from most of the RNN versions [29]. Therefore, throughout this framework, LSTM neural networks are used in the predictive data to solve the first layer of the air pollution concentration problem.

2.6. IAQP Based on Multilevel RNN. A neural network with a feedback system whose output is connected not only to the input and weight of the current network but also to the input of the former network is called an RNN. RNN research was influential in recent years, and an RNN was proposed as an effective technique for implementing nonlinear adaptive filtering owing to its promising ability to model nonlinear dynamic systems [30]. The RNN architecture is shown in Figure 5(b), in which each node in the network consists of input from the previous node, followed by feedback. Each node generates the current hidden state and output using the current input as well as the previous hidden state. An RNN is generally used for sequential data types:

$$\begin{aligned}
 h_t &= f(W_h h_{t-1} + V_h x_t + b_h), \\
 o_t &= f(W_o h_t + b_o),
 \end{aligned} \tag{3}$$

where h_t represents the hidden block at each time (t), the weights of the hidden layers in the recurrent connection

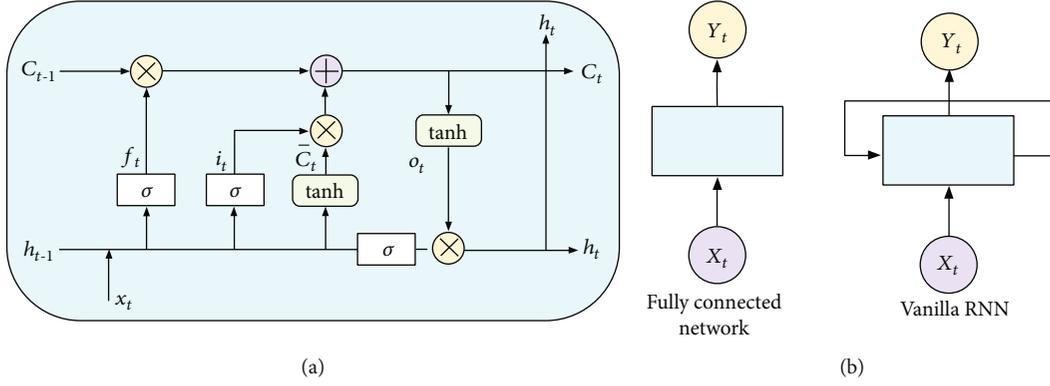


FIGURE 5: Graphical representation of (a) vanilla LSTM architecture and (b) vanilla RNN architecture.

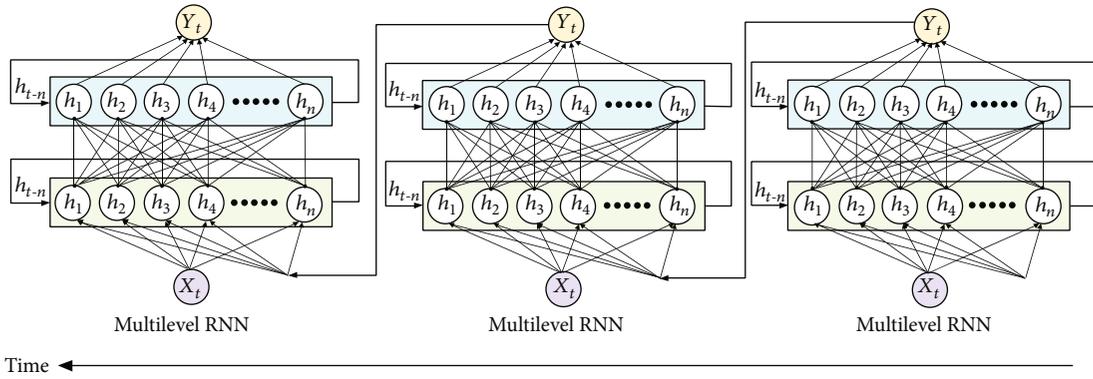


FIGURE 6: Multilevel recurrent neural network based on time series modelling.

are represented by W and V , b represents the bias condition in the hidden state, and F represents the activation function used at each node in the entire network [31]. Figure 6 describes a multilevel RNN with multihidden layers; this scheme could be stacked on top of each other. Due to the integration of numerous simple layers, this results in a flexible mechanism. Data may be important at multiple levels of the stack, in particular.

In the multilevel RNN case, this study uses several layers of RNN to preprocess and predict the amount of data. The multilevel RNN is more complex and consumes more time compared with the traditional RNN. This study tried to enhance the performance of its algorithm by the arrangement of the algorithm. For integrating the features from such blocks with the input to succeeding layers, we evaluate three fusion options: forward, sum, and concatenation. Traditional feedforward topologies are referred to as forward:

$$X^{L+1} = F(X^L, W), \quad (4)$$

i.e., the block simply becomes a new layer; sum denotes the original residual network approach.

$$X^{L+1} = X^L + F(X^L, W), \quad (5)$$

so that the multilevel RNN module operates as a residual block, while, in concatenation, features from multiple layers (the same spatial sizes) are concatenated:

$$X^{L+1} = [X^L; F(X^L, W)] (;) \text{ refers to concatenation.} \quad (6)$$

The number of layers is denoted by the letter L . As a result, the resulting feature map's channels will equal the sum of the two concatenated layers' channels (the number of parameters will be increased for the next layers) [32].

2.7. Workflow of the Proposed System. The workflow of this study consists of several steps. As shown in Figure 7, first, the sensor was installed for sensing the IAQ and OAQ data. Then, the amount of data was stored in the IoT platform. The IoT platform used is the OneM2M standard. After collecting the data into the IoT Platform, then the data was prepared for AI programming using some data preprocessing techniques.

We try to resample the data and also scale the data to get the best prediction value. Then, data was split into a training set and testing set. We are trying to use several algorithms such as RNN, multilevel RNN, and LSTM. The training process was held for several times, around 1 hour until 5 hour in one case. Optimization and error calculation were implemented to measure the performance of the AI algorithm. If

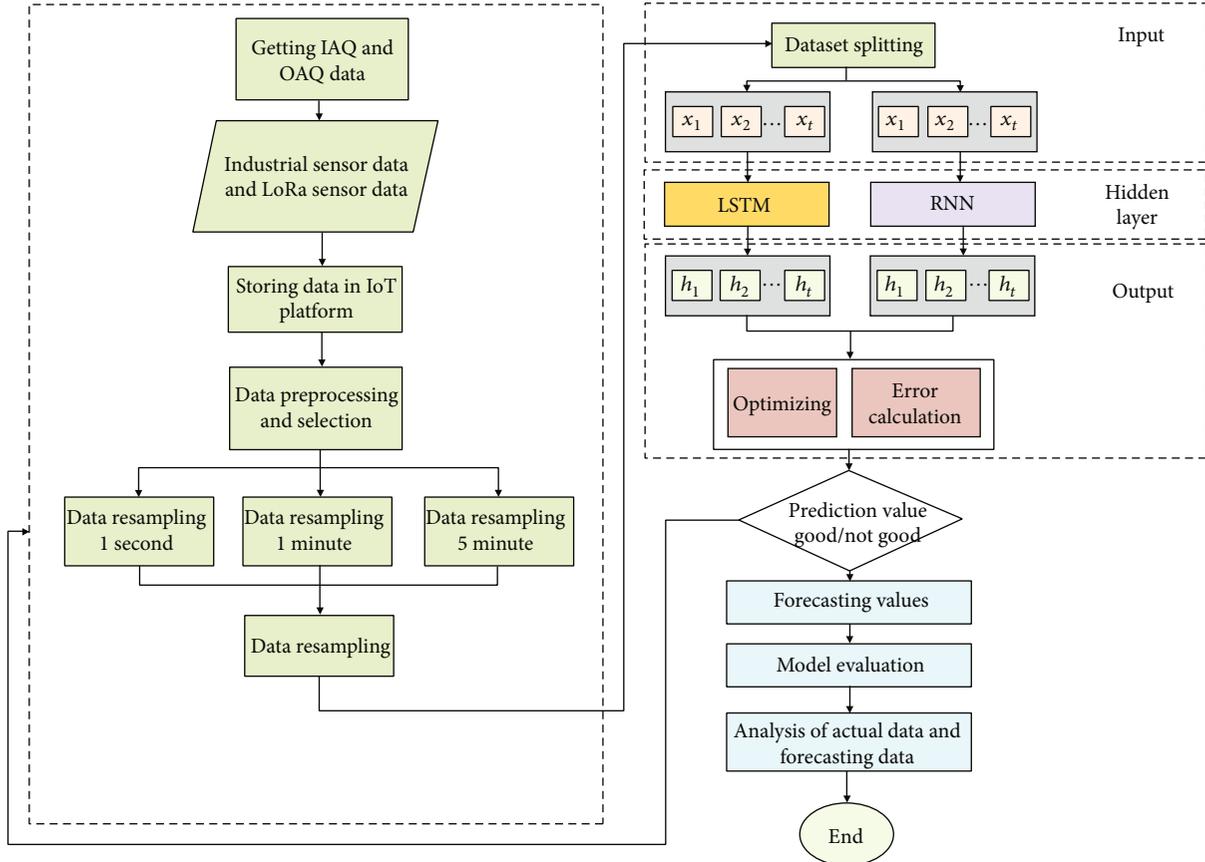


FIGURE 7: Indoor air quality prediction workflow.

the prediction is good, then we can continue to fix the forecasting value and then we can analyze the result of the forecasting. However, if the forecasting result is not good enough, we need to tune the parameter whether it is in the data preprocessing part or setting the parameter of the algorithm part.

3. Experimental Setup

3.1. Data Collection. This study uses two rooms to simulate the conditions such as the cleanroom and changing room, as shown in Figure 8. A space with dimensions $11.6\text{ m} \times 8\text{ m} \times 2.64\text{ m}$ is used for the cleanroom, and a space with dimensions $8.8\text{ m} \times 3.2\text{ m} \times 2.64\text{ m}$ is used for the changing room to predict the air quality in the cleanroom. Each room has an IIoT sensor that will sense the room environment. To obtain precise and reliable data, a PSU650 industrial sensor containing three sensors, that is, a temperature sensor, humidity sensor, and two-particle sensor (i.e., PM10 and PM2.5), is used in this study. For the outdoor environment measurements, a LoRa IoT sensor is used to compare the indoor and outdoor air quality. Moreover, for the communication gateway, RS485 serial communication is used to obtain reliable and accurate data. A cable connection has a satisfactory data transfer rate. The AX72432 is used to convert the RS485 device connection to a USB. To avoid missing links when data are obtained, the data are sent to the local

server, which uses an edge computing system. Edge computing can enable isolated virtualized computing based on a distributed paradigm, local storage, and communication resources by end users. Contrarily, centralized computing systems far from end users are involved in cloud computing. Thus, cloud computing is inefficient in processing intensive amounts of data [18].

An industrial-grade sensor, namely, the PSU650, which uses the Modbus protocol, is used in this study. In addition, three sensors are placed in three locations. The LoRa sensor was used to measure the outdoor air quality that measured dust particle concentration using the PM10 sensor. Besides, to compare between the indoor and outdoor air quality, we used the same LoRa sensor and recorded the indoor particle concentration data. Industrial-grade sensors are also used in the changing room to determine the IAQ conditions in the room. A temperature test is performed 1 h after an air conditioner is turned on. In addition, the number of measurement locations should be at least two to compare the air quality. The humidity level test, which is typically represented by relative humidity, assesses the ability to maintain air humidity levels, in which humidity conditions are closely related to contaminated air particle conditions. Furthermore, the number of particles was tested to measure the number of particles deposited on the surface or in the cleanroom area. In this work, we collected five air quality variables, namely, PM10, PM2.5, temperature, humidity,

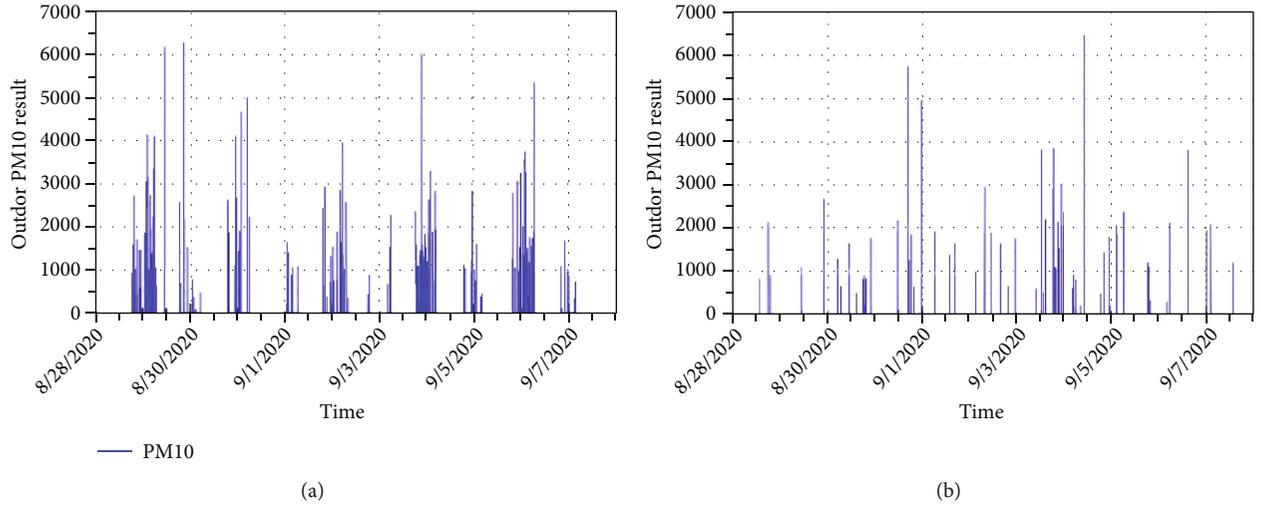


FIGURE 9: LoRa data acquisition: (a) LoRa outdoor and (b) LoRa indoor.

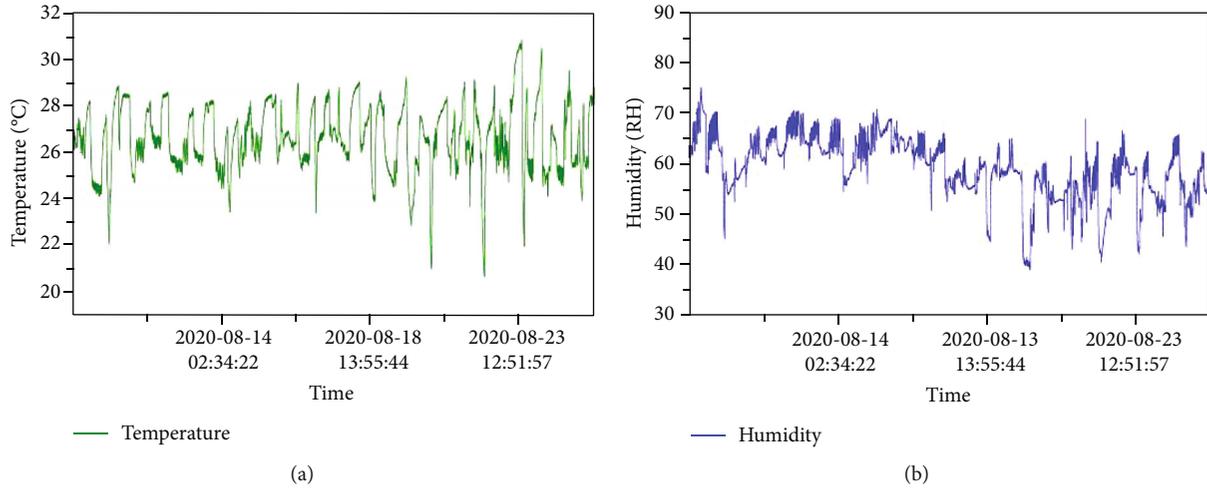


FIGURE 10: Indoor air quality monitoring: (a) temperature and (b) humidity.

where f denotes the predictive multilevel RNN model and n denotes the time step back up to which x_{t+1} has a good correlation.

4. Result and Discussion

The data from the sensor are more accurate due to higher sensitivity of the PSU650. For this reason, the proposed IAQDs will satisfy the accuracy and reliability requirements. An insignificant deviation appeared during the measurement of PM10 values at certain intervals because of changing the negligible amount of moving PM10. Therefore, another two sensors which are used in the LoRa gateway are deployed in the changing room and outside the room.

For comparing the changing room’s air quality and the outdoor air quality, both of the sensors are deployed in the particular position. In Figures 9(a) and 9(b), the changes of PM10 in indoor and outdoor are presented for sensing particle concentration. The LoRa indoor and outdoor consist of

TABLE 6: Comparison of the configuration model.

| Hyperparameter | Multilevel RNN | LSTM |
|---------------------------|-----------------------|-----------------------|
| Recurrent layer | 2 RNN layers | 2 LSTM layers |
| Hidden units | 256 of each layer | 256 of each layer |
| Gate activation | tanh | tanh |
| Dropout layer | 0.1 | 0.1 |
| Time distribution process | 1 time distributed | 1 time distributed |
| Processing | 22 of step prediction | 28 of step prediction |
| Loss function | MSE | |
| Optimizer | Adaptive momentum | |

a PPD42 sensor that could sense PM10 particle sizes. This study also measured the outside particle, as the LoRa sensor Node 2 was installed outside the building. The changes of

TABLE 7: RNN-LSTM performance comparison.

| Performance | RNN train | LSTM train | RNN test | LSTM test |
|-------------|-----------|------------|----------|-----------|
| MSE | 0.4400 | 0.6711 | 0.1296 | 0.2592 |
| R2 | 0.7328 | 0.5429 | 0.6093 | 0.4230 |
| RMSE | 0.6633 | 0.8192 | 0.3600 | 0.5092 |
| MAE | 0.3735 | 0.5291 | 0.2634 | 0.326 |

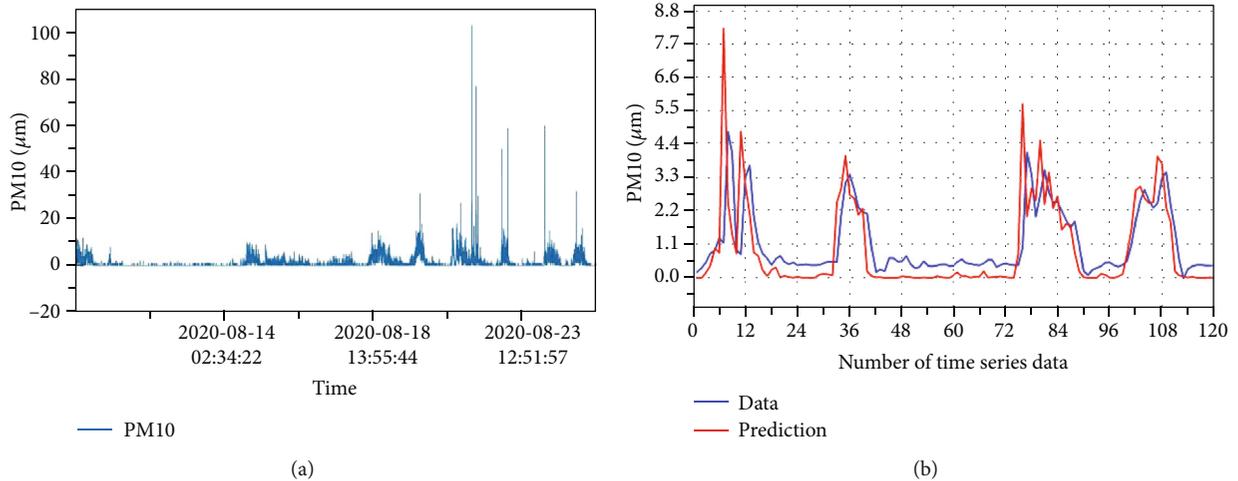


FIGURE 11: Indoor air quality data: (a) PM10 monitoring and (b) IoT PM10 prediction by using multilevel RNN.

the outdoor PM10 have been shown in Figure 9(a). Ventilation preservation was performed every afternoon to maintain the IAQ inside the room. The graph in Figure 9(b) shows that depending on the acceleration of the airflow, there exist fluctuations of different values of PM and the measured sensor values were quite small compared with those of the outdoor condition. We can see the difference condition between Figures 9(a) and 9(b). In Figure 9(a), the concentration of PM10 is high, whereas the condition of the LoRa sensor in the indoor room recorded has less PM10 as shown in Figure 9(b). The LoRa sensor has monitored that the air quality at the indoor room is better than the outdoor air quality condition.

Therefore, the humidity and temperature are related for changing the concentration of particle in both cases. There is a clear effect of relative humidity on the concentrations of particulate matter (PM10), where there is a correlation; when the humidity of the room is high, then the less PM10 contamination was detected in the indoor room. For this purpose, the proposed system also focused on monitoring the variation of temperature and humidity. In Figures 10(a) and 10(b), the changes of humidity and temperature of the indoor room have been presented. The experiment was conducted in the summer time in South Korea, so the condition of the RH is quite high as we can see in Figure 10(b); the highest humidity value was 74%. It was conducted over the course of two weeks, from August 11, 2020 (00:00), to August 26, 2020 (23:58). The IAQDs sent data to the cloud server every second. First, with machine learning and deep learning algorithms, we implemented model (multilevel RNN) prediction (in terms of enhancing output accuracy). For the out-

come analysis, we described the visualization of the real and predicted scenarios in the different figures. Finally, we depicted our experimental results to evaluate and compare the optimal time-step size of the multilevel RNN with that of other approaches.

We considered the most precise multilevel RNN and LSTM in our experiment to verify the results. We applied both algorithms to a model with two hidden layers with a total of 256 hidden nodes, as defined in Table 6. The time-step dimensions varied from 256 to 1, and the time-step sizes were 22 epochs (the best cost learning performance was 1.4) for the highest learning performance and 3.79 epochs for the lowest training performance.

Table 7 presents the results under different parameter settings of the multilevel RNN and LSTM models. If the difference between the real and expected values is close to zero, then the prediction is correct. From the results, it can be concluded that the multilevel RNN model outperformed the LSTM model. Specifically, the best performance of the multilevel RNN model (0.12) was substantially better than that of the LSTM model (0.2). Moreover, the multilevel RNN worked well with large and shallow networks in terms of network architecture, without a deep architecture overhead. In terms of learning time and efficiency, our approach was substantially better than the LSTM method. In other words, the best time-step size was automatically and efficiently found, thereby resulting in improved output compared with that of other processes.

This paper focuses on the PM10 prediction. As shown in Figure 11(a), the PM10 concentration was monitored. Then,

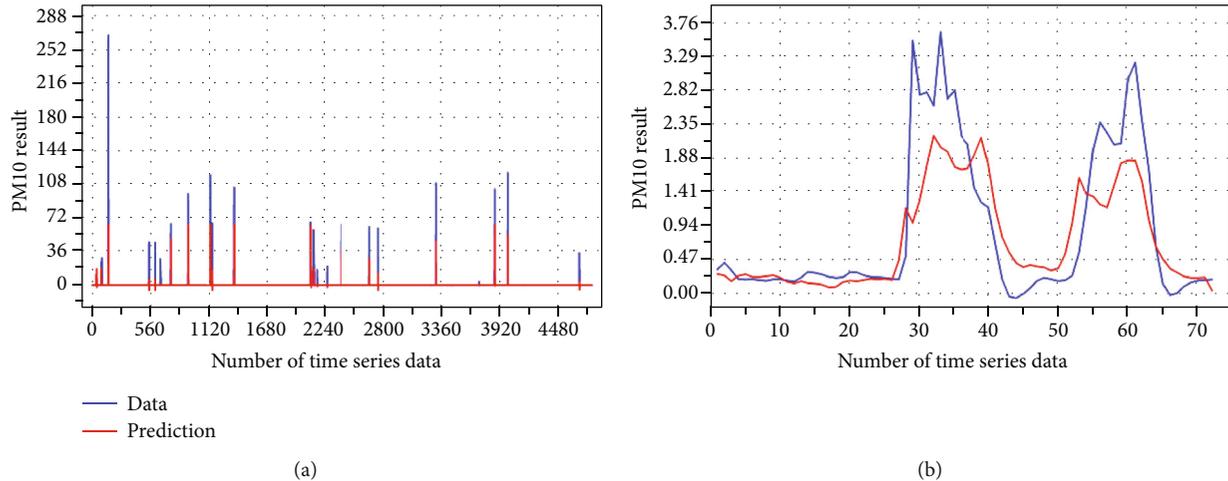


FIGURE 12: Prediction result: (a) LoRa PM10 prediction by using LSTM and (b) IIoT PM10 prediction by using LSTM.

the amount of data is used for the air quality prediction in order to maintain the indoor air quality condition.

Prediction of PM10 using multilevel RNN is represented in Figure 11(b). The two graphs demonstrate similar shape and patterns, as the model exhibited high precision. Figure 12(a) shows the outdoor air quality prediction by using the LoRa sensor outside the building. It is representing the prediction condition as the example of several-day data from the LoRa and can be the algorithm consideration for predicting outdoor air quality. Then, Figure 12(b) shows the prediction of PM10 by using LSTM in the indoor environment.

5. Conclusion

IIoT sensors were developed to predict IAQ by enhancing the accuracy of the deep learning method. A combination model was presented in this study to obtain the most suitable results. Moreover, an air quality prediction method using IIoT sensor data with deep learning was proposed. The data of different sensors were predicted using both models, namely, the multilevel RNN and vanilla LSTM. Under the assumption that the model implemented in this study was more efficient than the vanilla LSTM system in terms of its prediction capacity, its efficiency was verified. In several experiments under various parameter settings, our model demonstrated excellent results and feasibility. Thus, in future research, we plan to combine additional sensors and implement other algorithms for machine learning and deep learning techniques.

Data Availability

The data are available on request by contacting Himawan Nurcahyanto (himawanurcahyanto@gmail.com).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was financially supported by the Ministry of Trade, Industry and Energy (MOTIE) and Korea Institute for Advancement of Technology (KIAT) through the International Cooperative R&D program (Project ID P0011880). And this work was also supported by the Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (No. 2017-0-00824, Development of Intelligent and Hybrid OCC-LiFi Systems for Next Generation Optical Wireless Communications).

References

- [1] K. Zheng, S. Zhao, Z. Yang, X. Xiong, and W. Xiang, "Design and implementation of LPWA-based air quality monitoring system," *IEEE Access*, vol. 4, pp. 3238–3245, 2016.
- [2] S. Tong, T. Wang, Y. Li, and B. Chen, "A combined backstepping and stochastic small-gain approach to robust adaptive fuzzy output feedback control," *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 2, pp. 314–327, 2013.
- [3] J. M. Daisey, W. J. Angell, and M. G. Apte, "Indoor air quality, ventilation and health symptoms in schools: an analysis of existing information," *Indoor Air*, vol. 13, no. 1, pp. 53–64, 2003.
- [4] O. A. Postolache, J. M. D. Pereira, and P. M. B. S. Girao, "Smart sensors network for air quality monitoring applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 9, pp. 3253–3262, 2009.
- [5] T. V. Tran, N. T. Dang, and W. Y. Chung, "Battery-free smart-sensor system for real-time indoor air quality monitoring," *Sensors and Actuators B: Chemical*, vol. 248, pp. 930–939, 2017.
- [6] Q. P. Ha, S. Metia, and M. D. Phung, "Sensing data fusion for enhanced indoor air quality monitoring," *IEEE Sensors Journal*, vol. 20, no. 8, pp. 4430–4441, 2020.
- [7] D. D. Pope CA 3rd and M. Ezzati, "Fine-particulate air pollution and life expectancy in the United States," *New England Journal of Medicine*, vol. 360, no. 4, pp. 376–386, 2009.

- [8] S. Brienza, A. Galli, G. Anastasi, and P. Bruschi, "A low-cost sensing system for cooperative air quality monitoring in urban areas," *Sensors*, vol. 15, no. 6, pp. 12242–12259, 2015.
- [9] F. H. Hung, K. F. Tsang, C. K. Wu et al., "An adaptive indoor air quality control scheme for minimizing volatile organic compounds density," *IEEE Access*, vol. 8, pp. 22357–22365, 2020.
- [10] H. Zheng, K. Xiong, P. Fan, and Z. Zhong, "Data analysis on outdoor–indoor air quality variation: buildings' producing dynamic filter effects," *IEEE Systems Journal*, vol. 13, no. 4, pp. 4386–4397, 2019.
- [11] S. Mad Saad, A. M. Andrew, A. Y. M. Shakaff, A. R. Mohd Saad, A. M. Y. Kamarudin, and A. Zakaria, "Classifying sources influencing indoor air quality (IAQ) using artificial neural network (ANN)," *Sensors*, vol. 15, no. 5, pp. 11665–11684, 2015.
- [12] M. Elbayoumi, N. A. Ramli, N. F. F. Md Yusof, and W. Al Madhoun, "Spatial and seasonal variation of particulate matter (PM₁₀ and PM_{2.5}) in Middle Eastern classrooms," *Atmospheric Environment*, vol. 80, pp. 389–397, 2013.
- [13] W. Wei, O. Ramalho, L. Malingre, S. Sivanantham, J. C. Little, and C. Mandin, "Machine learning and statistical models for predicting indoor air quality," *Indoor Air*, vol. 29, no. 5, pp. 704–726, 2019.
- [14] L. Zhao, W. Wu, and S. Li, "Design and implementation of an IoT-based indoor air quality detector with multiple communication interfaces," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9621–9632, 2019.
- [15] J. Bhatt and H. Verma, "Design and development of wired building automation systems," *Energy and Buildings*, vol. 103, pp. 396–413, 2015.
- [16] M. Elbayoumi, N. A. Ramli, N. F. F. Md Yusof, A. S. Bin Yahaya, W. Al Madhoun, and A. Z. Ul-Saufie, "Multivariate methods for indoor PM₁₀ and PM_{2.5} modelling in naturally ventilated schools buildings," *Atmospheric Environment*, vol. 94, pp. 11–21, 2014.
- [17] M. Elbayoumi, N. A. Ramli, and N. F. F. M. Yusof, "Development and comparison of regression models and feedforward backpropagation neural network models to predict seasonal indoor PM_{2.5-10} and PM_{2.5} concentrations in naturally ventilated schools," *Atmospheric Pollution Research*, vol. 6, no. 6, pp. 1013–1023, 2015.
- [18] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: a survey," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 450–465, 2018.
- [19] R. Cristóvão, C. Botelho, R. Martins, and R. Boaventura, "Pollution prevention and wastewater treatment in fish canning industries of Northern Portugal," in *International Conference on Environment Science and Engineering IPCBEE*, vol. 32, pp. 12–16, Bangkok, Thailand, 2012.
- [20] R. Liang, L. Zhao, and P. Wang, "Performance evaluations of LoRa wireless communication in building environments," *Sensors*, vol. 20, no. 14, pp. 1–19, 2020.
- [21] X. Zhang, M. Zhang, F. Meng, Y. Qiao, S. Xu, and S. Hour, "A low-power wide-area network information monitoring system by combining NB-IoT and LoRa," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 590–598, 2019.
- [22] W. Whyte, *Cleanroom Technology: Fundamentals of Design, Testing and Operation*, John Wiley & Sons, 2001.
- [23] K. Sulovcova, R. Nosek, J. Jnadacka, and M. Holubcik, "Geometrical optimization of the flue gas path with regard to the reduction of particulate matter," *Emission Control Science and Technology*, vol. 4, no. 1, pp. 40–44, 2018.
- [24] C. Mühlfeld, B. Rothen-Rutishauser, F. Blank, D. Vanhecke, M. Ochs, and P. Gehr, "Interactions of nanoparticles with pulmonary structures and cellular responses," *American Journal of Physiology—Lung Cellular and Molecular Physiology*, vol. 294, no. 5, pp. L817–L829, 2008.
- [25] T. de Kok, H. Driec, J. Hogervorst, and J. Briedé, "Toxicological assessment of ambient and traffic-related particulate matter: a review of recent studies," *Mutation Research*, vol. 613, no. 2–3, pp. 103–122, 2006.
- [26] S. H. Huang, H. Y. Shih, S. N. Li, S. C. Chen, and C. J. Tsai, "Spatial and temporal distributions of a gaseous pollutant during simulated preventive maintenance and pipe leaking events in a working cleanroom," *IEEE Transactions on Semiconductor Manufacturing*, vol. 22, no. 3, pp. 391–398, 2009.
- [27] R. A. Rajagukguk, R. A. A. Ramadhan, and H.-J. Lee, "A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power," *Energies*, vol. 13, no. 24, p. 6623, 2020.
- [28] K. Khalil, O. Eldash, A. Kumar, and M. Bayoumi, "Economic LSTM approach for recurrent neural networks," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 11, pp. 1885–1889, 2019.
- [29] B. Wang, W. Kong, H. Guan, and N. N. Xiong, "Air quality forecasting based on gated recurrent long short term memory model in Internet of things," *IEEE Access*, vol. 7, pp. 69524–69534, 2019.
- [30] Q. Song, Y. Wu, and Y. C. Soh, "Robust adaptive gradient-descent training algorithm for recurrent neural networks in discrete time domain," *IEEE Transactions on Neural Networks*, vol. 19, no. 11, pp. 1841–1853, 2008.
- [31] H. Lynn, S. Pan, and P. Kim, "A deep bidirectional GRU network model for biometric electrocardiogram classification based on recurrent neural networks," *IEEE Access*, vol. 7, pp. 145395–145405, 2019.
- [32] H. K. Kwan, "Multilayer recurrent neural networks," *Midwest Symposium on Circuits and Systems*, vol. 3, pp. 1–17, 2002.
- [33] B. P. De Moraes, *Conversational AI: Automated Visualization of Complex Analytic Answers from Bots*, p. 37, Computer Science, 2018.
- [34] V.-S. Nguyen, N. Im, and S. Lee, "The interpolation method for the missing AIS data of ship," *Journal of Navigation and Port Research*, vol. 39, no. 5, pp. 377–384, 2015.
- [35] H. Zhen, D. Niu, K. Wang, Y. Shi, Z. Ji, and X. Xu, "Photovoltaic power forecasting based on GA improved Bi-LSTM in microgrid without meteorological information," *Energy*, vol. 231, p. 120908, 2021.