

# A Study on the Development of a Light Scattering Particulate Matter Sensor and Monitoring System\*

Jun-hee Choi<sup>1</sup> and Hyun-Sug Cho<sup>2</sup>

<sup>1</sup> Department of Disaster Prevention, Graduate School, Daejeon University  
62, Daehak-ro, Dong-gu, Daejeon, Korea  
junhee3020@hanmail.net

<sup>2</sup> Department of Fire & Disaster Prevention, Daejeon University  
62, Daehak-ro, Dong-gu, Daejeon, Korea  
chojo@dju.kr

**Abstract.** The gravimetric method, which is mainly used among particulate matter (PM) measurement methods, includes the disadvantages that it cannot measure PM in real time and it requires expensive equipment. To overcome these disadvantages, we have developed a light scattering type PM sensor that can be manufactured at low cost and can measure PM in real time. We have built a big data system that can systematically store and analyze the data collected through the developed sensor, as well as an environment where PM states can be monitored mobile in real time using such data. In addition, additional studies were conducted to analyze and correct the collected big data to overcome the problem of low accuracy, which is a disadvantage of the light scattering type PM sensor. We used a linear correction method and proceeded to adopt the most suitable value based on error and accuracy.

**Keywords:** Sensor Development, Fine Particles, Mobile Monitoring, Big Data, Light Scattering.

## 1. Introduction

As studies on the harmfulness of PM to the human body progress, PM has attracted great public interest, and as the times during which PM concentrations are high have increased, PM has become a serious social problem. In general, PM refers to particulate matter with particle sizes not larger than  $10\ \mu\text{m}$ , and occurs due to natural factors and artificial factors. In modern times, the ratio of PM generated due to artificial factors is high, and mainly those substances that have not been completely burned during the process of combustion of fossil fuels pollute the atmosphere. Fine particulate matter (FPM) is dust consisting of particles smaller than those of PM, and means fine particles not larger than  $PM\ 2.5\ \mu\text{m}$ . As with PM, FPM is also produced by artificial factors, and because the particle sizes are smaller than

---

\* Jun-hee Choi, Hyun-Sug Cho, "A Study on the Improvement of the Accuracy of Low-Cost Light Scattering Type Particulate Matter Sensors," Asia-pacific Society of Convergent Research Interchange, 2019.

those of PM, it may not be filtered out in the respiratory organs and can have fatal effects on the human body[1-2].

In order to solve the problems of PM, various efforts are being made both domestically and abroad [3-5], and studies regarding the measurement of PM and FPM are being conducted. PM measurement methods are divided into concentration measurement and size classification. In South Korea, PM is measured with concentration measurement methods. The concentration measurement methods comprise the gravimetric method, which collects PM for a certain time and measures the weight, and a method where  $\beta$ -rays are emitted to PM in the atmosphere and the  $\beta$ -rays absorbed by the PM are measured to determine the concentration of PM. These two methods are currently used interchangeably in South Korea, and show the highest accuracy among the measurement methods, but they require expensive equipment and provide data once an hour because they require preliminary work for measurement such as collecting and drying PM. To overcome these problems, light scattering type PM measurement methods have been studied[6]. The light scattering type methods use light for measurement and can measure the amount of PM in the atmosphere in real time. In addition, they have the advantage of enabling mass production at low prices since the light emitting part and the light receiving part play the largest roles among the components of the sensor. However, they have a vulnerable point in that the accuracy is relatively low because there is unit conversion two times with the measurement method. Due to this problem, PM data measured using light scattering type sensors have not been officially recognized until recently. However, the Ministry of Environment provided guidelines for light scattering type PM sensors in October 2018. The grades of data were divided according to the accuracy of measurements using PM sensors, and the data were divided with usable ranges, such as data for official measurement and those for research[7]. Undergoing correction work to meet such standards was recommended, and data with low rates of error based on the gravimetric method can be used as officially measured data. Correcting errors or increasing accuracy are applied in various ways in the big data field[8-17].

In this study, we developed a light scattering type PM sensor and built a system that stores data received from the sensor in a big data server. Based on these data, we established a mobile service that can monitor PM situations and conducted a study to correct the accuracy of the PM sensor[18]. we aimed to solve the problem of low accuracy in low-cost light scattering PM sensors.

## 2. Related Works

### 2.1 $\beta$ -ray absorption method

The  $\beta$ -ray absorption method measures the amount of  $\beta$ -rays absorbed in the filter paper (filter) used to collect PM and obtains the concentration of PM from the measured value using the property of  $\beta$ -rays that are absorbed more when they pass through substances with

larger masses[19]. The  $\beta$ -ray absorber removes particles not smaller than  $10 \mu m$  because a separator is installed in its air inhaling unit, and inhales air at a set flow rate to collect dust not larger than  $10 \mu m$  on the filter paper to measure the mass concentration of PM10. The PM10 separator is built in a certain structure that can separate particles by size using inertia and can be a cyclone type or an impactor type according to the shape of the particle size separation structure. If installed with an additional separator, the  $\beta$ -ray absorber can also measure the mass concentration of PM2.5 by collecting particles not larger than  $2.5 \mu m$ . In the case of PM2.5, particles are separately collected by size through a primary separator (removing particles not smaller than  $10 \mu m$ ) and a secondary separator (removing particles not smaller than  $2.5 \mu m$ ) using the impactor method. The minimum detection limits of the measured mass concentrations of PM2.5 and PM10 are known to be 5 and  $10 \mu g/m^3$  or less, respectively, and the ranges of measured concentrations are  $0\sim 1,000 \mu g/m^3$  in the case of PM2.5 and  $0\sim 1,000, 0\sim 5,000, 0\sim 10,000 \mu g/m^3$ , etc. in the case of PM10.

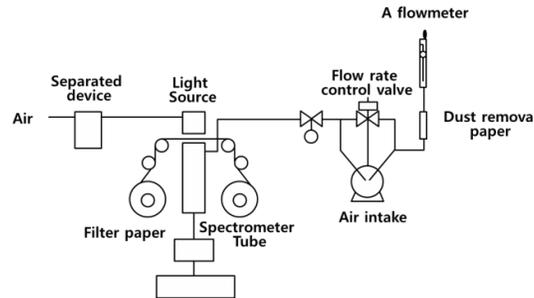


Fig. 1.  $\beta$ -ray measurement method operation principle

The  $\beta$ -ray absorption method measures PM concentrations by calculating the  $\beta$ -rays absorbed and dissipated when the  $\beta$ -rays irradiated from the light source that emits  $\beta$ -rays pass through the dust collected on the filter paper following Equations (1) and (2).

$$I = I_0 \times \exp(-\mu X) \quad (1)$$

where  $I$  refers to the intensity of the  $\beta$ -rays that penetrated through the dust collected on the filter paper and  $I_0$  refers to the intensity of the  $\beta$ -rays that penetrated through the blank filter paper. That is,  $I_0$  means the intensity of the  $\beta$ -rays that penetrated through the filter paper on which no dust has been collected.  $\mu$  is the  $\beta$ -rays mass absorption extinction coefficient ( $cm^3/\mu g$ ) by PM, and can be said to be constant when the component and particle size distribution are constant.  $X$  refers to the mass ( $\mu g/cm^3$ ) of the dust collected per unit area. The  $\beta$ -ray intensity calculated through Equation (1) can be applied to Equation (2) to calculate the final PM concentration.

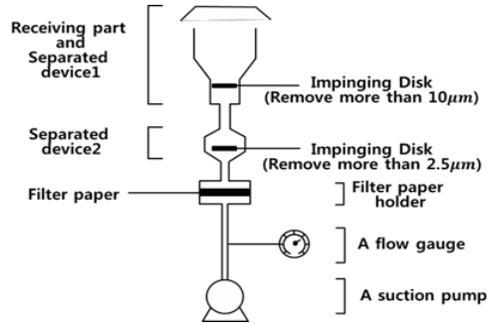
$$C(\mu g/m^3) = \frac{S}{\mu \cdot Q \cdot \Delta t} \ln \left( \frac{I}{I_0} \right) \quad (2)$$

where  $C$  refers to PM concentration,  $S$  refers to the area ( $m^2$ ) of the filter paper on which PM was collected,  $V$  refers to amount of air inhaled ( $m^3$ ), and  $\Delta t$  refers to the sampling time (min). The reason for using  $\beta$ -rays, which are a sort of radiation, as a light source is that they have a property where they further lose their intensity when they pass through substances with larger masses. In addition,  $\beta$ -rays are safe because they cannot pass through metal plates or plastics, and enable comparison using initial filter paper as a background concentration because they pass through the collecting filter paper and reach the  $\beta$ -ray sensing unit. On the other hand, particle separators use inertial force, and even though their initial error range is very low, their separation capacity cannot but vary according to the weather conditions when foreign matter in the atmosphere adheres to them or the PM is highly viscous. Therefore, the principal limitation of a separator using a cyclone or impactor cannot but exist.

## 2.2 Gravimetric method

The gravimetric method collects samples for 24 hours and measures the mass of PM smaller than  $2.5 \mu m$  among the substances collected on the filter paper. The light scattering method measures the amount of scattered light using the principle that when light is cast on an object, the light collides with the object and is scattered in various directions and the concentration of PM is obtained from the measured value.

The gravimetric method is one of the methods of measuring the mass concentration of PM10 or PM2.5 in the atmospheric environment. It collects PM samples in the atmosphere using a sampler, and calculates the weight difference of the filter before and after the collection of PM samples as the mass concentration of the PM. The collected samples have the advantage of enabling physical and chemical analysis later, but substances such as sulfur dioxide ( $SO_2$ ) or nitric acid that undergo chemical reactions to be oxidized into sulphate or nitrate have disadvantages of leading to the indication of excessive mass concentrations because solid salts are generated or resulting in errors in measurement results due to a reduction in the mass concentration as a result of the dissociation process. The limit of detection of the measured mass concentration should be within  $5 \mu g/m^3$  when the range of the measured mass concentrations is  $80 \mu g/m^3$  or lower and within 7% of the measured mass concentrations when the range exceeds  $80 \mu g/m^3$  on the basis of aerodynamic diameter in the case of PM10. The gravimetric method's limit of quantification of PM2.5 is  $3 \mu g/m^3$ . Figure 2 shows a block diagram of the sampler used in the gravimetric method for PM2.5. First filter the PM10 samples by inhaling air from the atmosphere, and then measure the PM2.5 samples.



**Fig. 2.** Operation principle of the gravimetric method

The collected samples have the advantage of enabling physical and chemical analysis later, but in the case of substances such as sulfur dioxide ( $\text{SO}_2$ ) or nitric acid that undergo chemical reactions to be oxidized into sulphate or nitrate, the PM concentration measured is indicated in units of  $\mu\text{g}/\text{m}^3$ , which represents the weight of PM in  $1 \text{ m}^3$  ( $\mu\text{g}$  means one millionth gram). The weight concentration is indicated in units of  $\mu\text{g}/\text{m}^3$  and is calculated by the following equation (3).

$$\text{PM} = \frac{W_f - W_i}{V_a} \quad (3)$$

Both PM10 and PM2.5 are calculated with the same method.  $W_f$  refers to the weight of filter paper after sampling ( $\mu\text{g}$ ),  $W_i$  refers to the weight of filter paper before sampling ( $\mu\text{g}$ ), and  $V_a$  refers to the total sampling volume ( $\text{m}^3$ ). Here, the flow rate is corrected using the standard temperature and air pressure ( $0^\circ\text{C}$ ,  $1 \text{ atm}$ ) and is calculated as shown in equation (4).

$$V_a = Q \times \frac{293}{273+T} \times \frac{P}{760\text{mmHg}} \times t \quad (4)$$

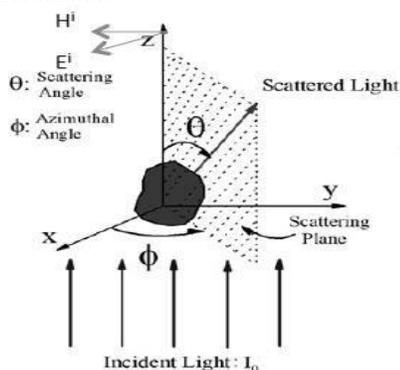
Recently, various PM measuring devices have been developed using the light scattering measurement method because this method has various advantages such as the downsizing of measuring devices, real-time measurement, and dramatic reductions in the cost of device manufacturing.  $T$  and  $P$  represent the temperature and pressure at the time of measurement, respectively, and  $t$  refers to the total time of sampling. Particles not smaller than  $10 \mu\text{m}$  in the air sucked into the device by the suction pump collide with the impactor at the inlet due to inertial force and are separated from the main air stream. Thereafter, the PM particles that passed through the impactor are collected on the filter paper, and the weight is measured later to indicate the weight concentration of the PM particles in comparison with the air volume. The filter paper used in this case is nitrocellulose membrane filter paper, quartz fiber, or a Teflon material with an initial collection rate of at least 99% for  $0.3 \mu\text{m}$  particles, and the effect of moisture in the air should be minimized. The gravimetric method is divided into high-volume air samplers and low volume air samplers according to the volume of air collected. In the case of high-volume samplers, the suction flow should be about  $2 \text{ m}^3/\text{min}$  at no load and the sampler should be capable of continuous measurement for 24 h. The low volume air sampler can measure in a range of  $10\text{--}30 \text{ L}/\text{min}$ .

### 2.3 Light scattering measurement method

The light scattering method obtains the quantity of particulate matter using the principle that when light is cast on particulate matter (PM) suspended in the atmosphere, the light is scattered by the particles, and that when light is cast on particulate matter with the same physical properties, the amounts of scattered light are proportional to the mass concentrations of the particulate matter[6].

In order to understand climate change caused by the scattering effects of PM in the air, experiments have been carried out worldwide using a device called a Nephelometer. The three-wavelength scatterometer Nephelometer, which is the one most commonly used in the field, yields total scattering coefficients and back scattering coefficients for three wavelengths ( $\lambda$ ) of 450, 550, and 700 nm. Photomultiplier tubes (PMT), which are the main element of the measuring unit, measure the quantities of scattered particles (number of photons) coming in through the scattering unit for three wavelengths of blue (450 nm), green (550 nm), and red (700 nm), respectively.

The spectrometer inhales samples in the air together with the air in the atmosphere at a constant flow rate and irradiates the laser to the particles passing through the measuring space inside the monitor. Then the laser beam is scattered by the particulate matter in the air flow. The light scattered this way is collected by the light concentrator and sent to the photodetector. The photodetector generates electrical signals in proportion to the amount of light concentrated, and the generated electrical signals are measured in pulses. The heights and numbers of the pulses can be converted into the diameters and numbers of particles for measurement and the outcome can be divided into 32 channels in a range from 0.23 to 32  $\mu$ m depending on the specification of the monitor. Let us take a closer look at the principle of the spectrometer. Both the hardware and software required to make a particle size measurement have gone through many revisions to improve accuracy, precision, reliability, and ease of use. At the very heart of the laser diffraction technique is the relationship between light and surfaces (which can be freely interchanged with "particle" for our purposes). When light strikes a surface, it is either diffracted, reflected, refracted, absorbed or reradiated.



$$I = \frac{I_0 F(\theta, \phi)}{k^2 r^2} \quad (5)$$

$I_0$ : Intensity of incident light

$I$ : Intensity of scattered light

$\theta$ : Scattering angle

$\phi$ : Azimuthal angle

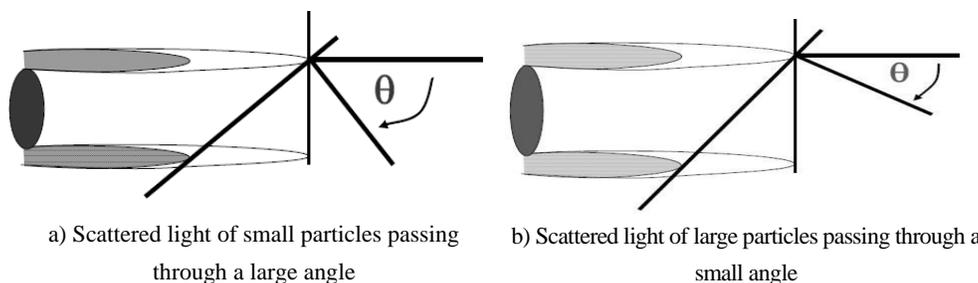
$k$ : Wave number

$r$ : Distance from scatter to detector

Fig. 3. Coordinate geometry for Rayleigh and Mie scattering[20]

For particles larger than a certain size, the vast majority of light is scattered by diffraction. The scattered light is at a relatively high intensity and low angle for these larger particles[21]. The "certain size" is determined as a multiple of the wavelength of light used for the measurement and is typically approximated at 20 microns. Particles larger than this size communicate useful size information through diffraction and not refraction. This means the measurement will not benefit from the use of a refractive index to accurately interpret refracted light.

For particles smaller than 20 microns, refracted light becomes increasingly important to calculate an accurate particle size. The scattered light is at a relatively low intensity and wide angle for these smaller particles.



**Fig. 4.** Relationship between distribution width and angular light intensity scattering patterns

The use of a refractive index and the Mie scattering theory directly affects accuracy in this size range. The basic workflow of a laser diffraction particle size analysis breaks down into two parts:

- Measure scattered light angle and intensity
- Transform that scattering data into a particle size distribution

Measurement quality refers to the analyzer itself: the quality of components, engineering refinement, and a fundamental design that reflects basic principles.

## 2.4 Decimation-in-Frequency (DIF) Radix-2 FFT Algorithm

The section considers radix-2 decimation in the frequency FFT algorithm and the Radix-4 Complex FFT algorithm that is applied to the developed measurement system[22].

### 2.4.1. Decimation-in-Frequency FFT Algorithm

In the context of fast Fourier transform algorithms, a butterfly is a portion of the computation that combines the results of smaller discrete Fourier transforms (DFTs) into a larger DFT, or vice versa (breaking a larger DFT up into sub-transforms). The name "butterfly" comes from the shape of the data-flow diagram in the radix-2 case. Most

commonly, the term "butterfly" appears in the context of the Cooley–Tukey FFT algorithm, which recursively breaks down a DFT of composite size  $n = rm$  into  $r$  smaller transforms of size  $m$  where  $r$  is the "radix" of the transform.

Let  $N$  samples of the input signal,  $S(n)$ ,  $n=0, \dots, N-1$  be, at the same time  $N$  is the whole degree of two  $N=2^L$ . In the decimation-in-frequency algorithm the input  $S(n)$ ,  $n=0, \dots, N-1$ , is halved, i.e.  $S_0(m)=S(m)$  and  $S_1(m)=S(m+N/2)$ ,  $m=0, 1, \dots, N/2-1$ . Spectral samples  $S(2 \cdot p)$ ,  $p=0, \dots, N/2-1$ , with even indexes are  $N/2$ -point DFT of the signal  $X_1(m)=S(m) + S(m+N/2)$ . Similarly consider the signal  $S(n)$  for spectral samples  $S(2 \cdot p+1)$ ,  $p=0, \dots, N/2-1$ , with odd indexes.

Thus, the divide procedure includes calculating signals of half duration  $X_0(m)$  and  $X_1(m)$ ,  $m=0, \dots, N/2-1$ . Then it is possible to replace  $N$ -points DFT with two  $N/2$ -points DFT of signals  $X_0(m)$  and  $X_1(m)$ . The procedure for calculating signals of half duration can be presented in the form of a butterfly diagram as in Figure 5.

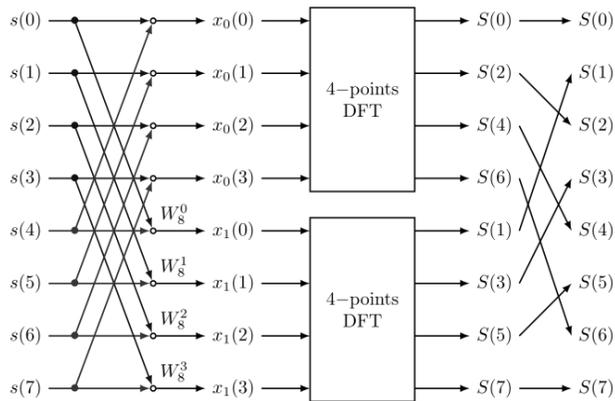


Fig. 5. Diagram of Decimation-in-Frequency FFT Algorithm for  $N=8$

A decimation-in-time radix-2 FFT breaks a length- $N$  DFT into two length- $N/2$  DFTs followed by a combining stage consisting of many butterfly operations.

### 2.4.2. Radix-4 Complex FFT Algorithm

This paper used Radix-4 Complex FFT Functions in the firmware algorithm for transferring from the electronic level of signal to frequency level of signal.

Complex fast Fourier transform (CFFT) and complex inverse fast Fourier transform (CIFFT) is an efficient algorithm to compute discrete Fourier transform (DFT) and inverse discrete Fourier transform (IDFT). The computational complexity of CFFT is reduced dramatically when compared to DFT. This set of functions implements CFFT/CIFFT for Q15, Q31, and floating-point data types. The functions operate on an in-place buffer that

uses the same buffer for input and output. Complex input is stored in the input buffer in an interleaved fashion. The functions operate on blocks of input and output data and each call to the function processes  $2 \cdot \text{fft Len}$  samples through the transform. pSrc points to In-place arrays containing  $2 \cdot \text{fft Len}$  values.

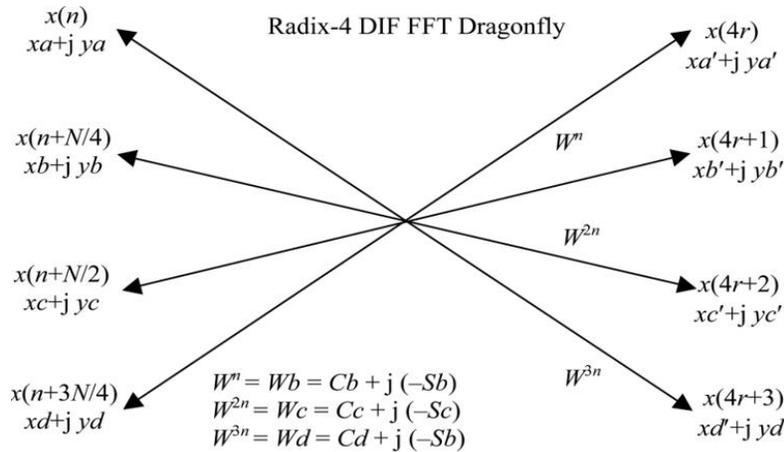


Fig. 6. Radix-4 Decimation-in Frequency Complex FFT

**Algorithm: Complex Fast Fourier Transform**

**Input real and imaginary data:**

$$\begin{aligned} x(n) &= xa + j * ya \\ x(n+N/4) &= xb + j * yb \\ x(n+N/2) &= xc + j * yc \\ x(n+3N/4) &= xd + j * yd \end{aligned}$$

**where N is the length of FFT Output real and imaginary data:**

$$\begin{aligned} x(4r) &= xa' + j * ya' \\ x(4r+1) &= xb' + j * yb' \\ x(4r+2) &= xc' + j * yc' \\ x(4r+3) &= xd' + j * yd' \end{aligned}$$

**Twiddle factors for radix-4 FFT:**

$$\begin{aligned} W^n &= Cb + j * (-Sb) \\ W^{2n} &= Cc + j * (-Sc) \\ W^{3n} &= Cd + j * (-Sb) \end{aligned}$$

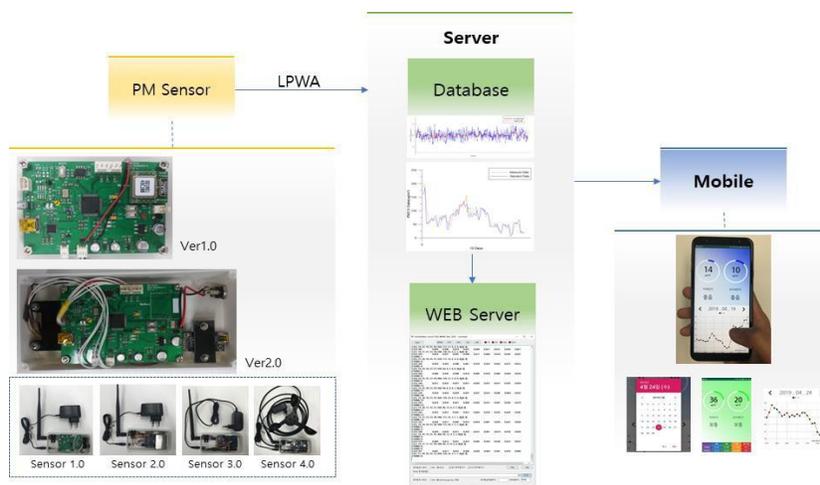
Different FFT algorithms provide different benefits, but there is always a trade-off between computation speed and FPGA area used. Reducing an FFT's computation time reduces hardware complexity. Different decomposition methods are available, such as Radix-2 (Radix-2 nodes are known as Butterfly nodes) and Radix-4 (Radix-4 nodes are known as Dragonfly nodes).

**DIF Radix-4 CFFT Process Code:**

```

for (i = 0; i < SAMPLES; i += 2) {
  /* Each 22us ~ 45kHz sample rate */
  HAL_ADC_Start_IT(&hadc1); Delay(21);
  /* Real part, must be between -1 and 1 */
  Input[(uint16_t)i] = (float32_t)(ADC_val -
    (float32_t)2048.0) / (float32_t)2048.0;
  /* Imaginary part */
  Input[(uint16_t)(i + 1)] = 0; }
/* Initialize the CFFT/CIFFT module, intFlag = 0, doBitReverse = 1 */
arm_cfft_radix4_init_f32(&S, FFT_SIZE, 0, 1);
/* Process the data through the CFFT/CIFFT module */
arm_cfft_radix4_f32(&S, Input);
/* Process the data through the Complex Magnitudes Module for calculating
the magnitude at each bin */
arm_cmplx_mag_f32(Input, Output, FFT_SIZE);
/* Calculates maxValue and returns corresponding value */
arm_max_f32(Output, FFT_SIZE, &maxValue, &maxIndex);

```

**3. PM sensor development**

**Fig. 7.** Conceptual diagrams of the PM sensor, big data, and mobile monitoring service

The developed sensor is a complex sensor that measures all the factors that can affect PM measurement, and measures temperatures, humidity, and wind speeds based on the PM

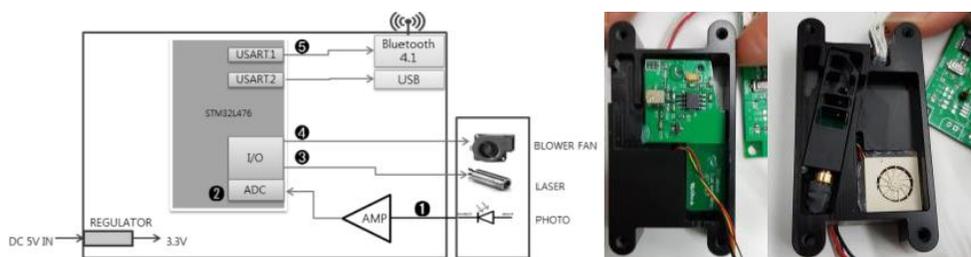
measuring function. The measured data are transmitted to the big data server through the long-range (LoRa) network of the low-power wide area (LPWA). The correction value of the PM sensor was calculated using the data stored in the server and was applied to the measurement sensor to improve accuracy, and finally, a service that can monitor the PM data on the desired date in the mobile terminal was built. An overall system block diagram is shown in Figure 7.

### 3.1 Sensor board development and modularization

The study developed a sensor board, which is the basis of the sensor. A block diagram is shown in Figure 8. The sensor board consists of a light emitting unit that shoots laser light at the sensing position, a light receiving unit that converts the received light into electrical signals and delivers the signals to the processor, and a heating unit that introduces air into the dust measurement sensor. Internally, a processing unit for converting electric signals into FFT to detect the PM that corresponds to the frequency band is included.

As for the process to detect PM, the light emitting unit shoots laser light at the air flowing into the sensor so that scattered light is generated by the PM included in the air. At this time, the fan introduces air into the sensor. The light receiving unit receives the light from the light emitting unit that is scattered by the PM, converts the light into electrical signals, amplifies the signals, and transmits the amplified signals to the processing unit. The processing unit identifies the frequency characteristics of the received electrical signals through FFT processing and measures those PM particles that correspond to individual frequency characteristics.

The LPWA module was connected to the rear of the sensor board to enable long-distance, low-power communication. Since the sensor board and LPWA module are included together in the case, they were connected by wire considering the issue of interference of wireless communication.



**Fig. 8.** The developed PM sensor board and sensor



**Fig. 9.** The sensor board with an LPWA communication module applied, and the inside of the PM sensor module

### 3.2 Sensor housing

The exterior of the PM sensor was fabricated with a 3D printer. The exterior is a plastic case to contain sensor elements, cables, and the control board. The case is made for fabrication of the prototype in a size useful to mount Arduino and Raspberry Pi, and the sensor board was built in the case later.

Inside the sensor case, a PM sensor, a PM sensor board, an LPWA communication module, various cables were built in, and the exterior was fabricated to enable the connection of the antenna and the power supply.

### 3.3 Low power mode implementation

The PM sensor must be usable with a battery because it should be able to measure PM anywhere in real time. To realize the foregoing, a processor low power mode was implemented by reducing the permanent power with a view to securing hours of use not shorter than one month when general batteries are used. At the beginning, the low power mode was implemented so that the PM sensor would measure data every second and transmit the data immediately after measurement. However, since the power consumed for wireless communicating to store the data in the server was very large, a method of reducing the number of transmission times was chosen. The low power mode was implemented so that the PM sensor would identically operate every second, but the data would be transmitted every 10 minutes or only when requested by the server, and as a result, the power consumption decreased to a level of 20%.



Fig. 10. PM sensor in which the low power mode was implemented.

### 3.4 Sensor advancement

In cases where PM is measured with light scattering type sensors, the measurement accuracy is affected by external environmental factors. To measure these variables and reflect them in the correction, the sensor was advanced by adding additional sensors that can measure various factors. Since additional power is necessary when additional sensors are attached, the advanced sensor is more suitable for installation types than for portable types. Additional sensor elements are included in Table 1. The composite sensors according to the combinations of sensor components are as follows.

- Sensor module A: Arduino UNO + additional dust sensor + temperature and humidity sensor + LPWA communication module, 9~12V power supply
- Sensor module B: Arduino Pro Mini + additional dust sensor + LPWA communication module, 5V power supply
- Sensor module C: Dust sensor self-developed in the first year + LPWA communication module, 5V power supply
- Sensor module D: Arduino UNO + wind speed sensor + atmospheric pressure/altitude sensor + communication module, 12V power supply

**Table 1.** Sensor components

Division	Shape	Specification	Remark
Anemometer		<p><b>Operating voltage:</b> 12V~ 24V Analog Output-Use after ADC Conversion</p> <p><b>Power Consumption:</b> Below 0.3W</p> <p><b>Resolution:</b> 0.1 m/s</p> <p><b>Effective measurement range:</b> 0 ~ 30m/s</p> <p><b>Operating Temperature:</b> -40 ~ 80 °C</p>	Mounted on wind speed / pressure sensor module
Barometer/ Altimeter	 Air pressure,height module	<p><b>Operating voltage:</b> 3.3V</p> <p>Barometer Altitude measurement</p> <p>TTL 232 Communications (fixed at 11520 bps) set to 0.1 Hz~10 Hz</p>	Mounted on wind speed / pressure sensor module
Dust sensor self-developed in the first year		<p><b>Operating voltage:</b> 5V</p> <p>Combustion floating matter reaction</p> <p>LPWA communication module combined light scattering type sensor applied Cortex-M4 core</p> <p>LED on / off, Fan on / off functions</p>	Specialized into combustibles sensor
Additional PM sensor		<p><b>Operating voltage:</b> 5V</p> <p>Particle Count Function Light scattering type</p> <p>TTL-232 communication</p> <p>Passive / Active mode switching</p>	Mounted on PM and temperature/ humidity sensor module
LPWA Communications module		<p><b>Operating voltage:</b> 3.3V</p> <p>Low power broadband communication device</p> <p>1:1 pairing mode</p> <p>1:N mode support</p> <p>Use 400 MHz band</p>	Mount on all sensor modules

## 4. Improvement of PM measurement accuracy of light scattering type sensors

To improve the accuracy of the sensor, a “data-based correction algorithm” was applied. PM standard data provided by Air Korea were used as the standard for accuracy and the data measured PM using the gravimetric method and  $\beta$ -rays. The PM concentration is indicated using units of  $\mu\text{g}/\text{m}^3$  and represents the weight of PM. In the experiment, data for 10 days in the Daejeon region in 2019 were used, and a data-based correction algorithm was applied using the standard data and measured data.

### 4.1 Dataset

Whereas the light scattering type PM sensor can measure PM in real time, the standard data are output every hour because the gravimetric method and  $\beta$ -rays are used. Therefore, the data collected using the light scattering type sensor were used in the correction work using the 1-hour average values.

**Table 2.** Data set used in the experiment

Division	Daejeon standard data	Data measured by the sensor
Period	10 days (240h)	10 days (240h)
Unit (Time)	1 Hour	1 Second $\rightarrow$ 1 Hour(average)
Unit (PM)	10 / 2.5 $\mu\text{g}/\text{m}^3$	10 / 2.5 / 1 $\mu\text{g}/\text{m}^3$
Method	Mass Concentration (Manual)	Sensor (Auto)

The dataset and test set were divided at a ratio of 8:2 to avoid overfitting in the data-based correction training results. The data for the first 8 days were used to carry out correction learning. Thereafter, the results were applied to the data for remaining 2 days of data and the results were checked.

### 4.2 Evaluation indicator

To quantify the difference between standard data and measured data, root mean square error (RMSE), which is a value that shows the difference between the two data, was used as an evaluation index. RMSE can be expressed as shown in Equation 6, where  $\hat{S}_i$  is the data measured using the sensor,  $d_i$  is the standard data, and  $n$  is the total number of data.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n |\hat{S}_i - d_i|^2}{n}} \quad (6)$$

Accuracy means the ratio to which the data are identical to the four levels of PM provided by the Ministry of Environment. A value of 1 is given when the levels of the two data are identical, and 0 is given otherwise. Therefore, the data will have a value of 1 when all the levels are identical. The PM indicator is as shown in Table 3.

**Table 3.** PM indicator( $\mu\text{g}/\text{m}^3$ )

	Good	Moderate	Bad	Very bad
<b>PM10</b>	~30	~80	~150	151~
<b>PM2.5</b>	~15	~35	~75	76~

### 4.3 Experiment

In order to correct the data collected with the sensor, the slope value of the data was changed or parallel translated and corrected using the combination with which the RMSE becomes the smallest. The slope was changed from -10 to 10 in units of 0.1 and the parallel translation was calculated from -50 to 120 in the changed units of 1.

**Table 4.** Performance evaluation results (Test data)

Data		RMSE	Accuracy
Measure Data	PM10	26.3764	0.1666
	PM2.5	12.8299	0.8333
Revise Data (Slope)	PM10	18.0323	0.3750
	PM2.5	9.1544	0.8958
Revise Data (Shift)	PM10	9.5686	0.8750
	PM2.5	8.7376	0.8333
Revise Data (Slope + Shift)	PM10	9.5686	0.8750
	PM2.5	10.0550	0.8541

The resulting values according to the performance evaluation indicator are shown in Table 4. It can be seen that the corrected data generally show higher performance values than the measured data. Although the particulate matter (PM10) showed high RMSE values and low accuracy when there was no correction, it could be identified that the performance was greatly improved by the correction work. The case of parallel translation showed the highest performance value, and given that there was no change even when the slope was applied, it could be predicted that the data were measured to be low in general. From the viewpoint of accuracy, the highest accuracy value was shown when the slope was applied. The FPM (PM2.5) also showed the highest RMSE performance when corrected by parallel translation, but from the viewpoint of accuracy, it showed the highest accuracy when corrected using slopes. In addition to slopes and parallel translation, various methods such as secondary correction equations and leveling-off were used, but it was identified that there was no big difference in terms of performance. With these results, the limits of mathematical correction were recognized, and it is expected that if the correlations of factors that affect PM are analyzed and the results are reflected, there will be additional performance improvement.

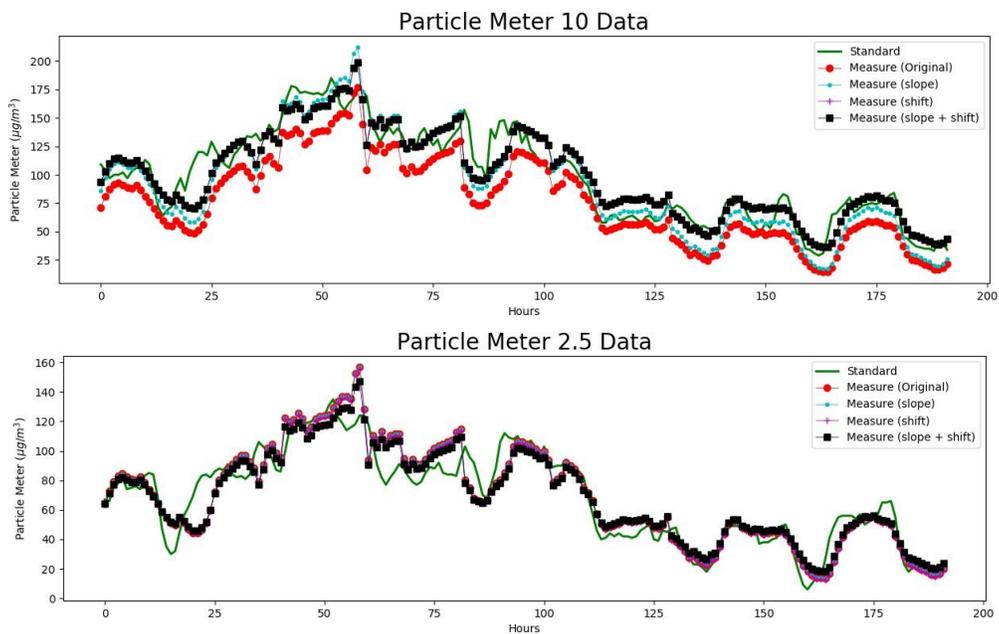


Fig. 11. PM(PM10) and FPM(PM2.5) data for 8 days

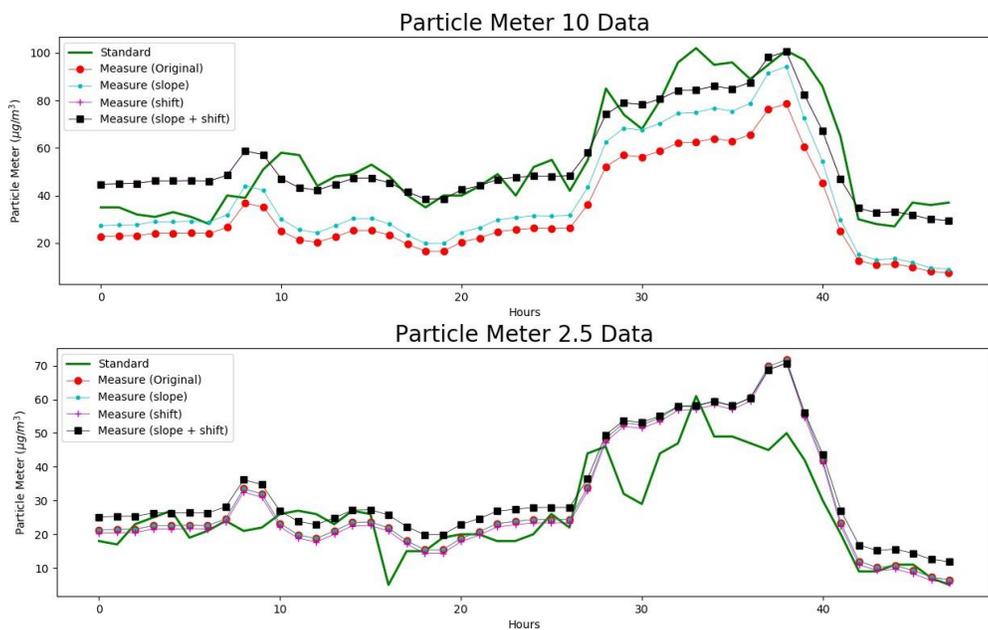


Fig. 12. PM(PM10) and FPM(PM2.5) data for 2 days

## 5. Mobile monitoring system

Since the data collected by the sensor must be transmitted to a server at a long distance, long-distance communication must be possible while the power is managed. Therefore, the LPWA communication-based LoRa method was applied to store the data in the server, which were then provided to the mobile application to build the PM monitoring system.

### 5.1 Low power technology (LoRa)

LoRa, which is one of the LPWA technologies, is a communication technology suitable for cases where low data transmission rates ( $<1\text{Mb/s}$ ) that require low power consumption are required. LoRa's physical layer uses spread spectrum modulation (SSM) and also uses a way to encode the fundamental signal into a higher frequency sequence. This enables the fundamental signals to be intentionally spread over a wider bandwidth, reducing power consumption and improving tolerance to electromagnetic interference.

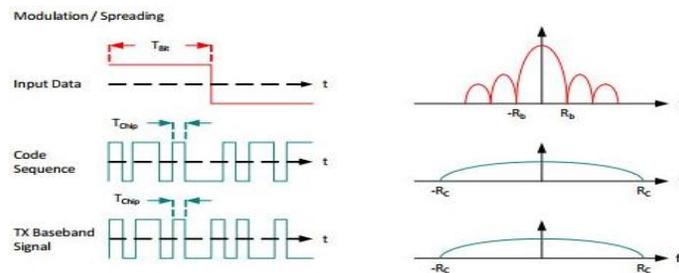


Fig. 13. Frequency characteristics according to signal lengths

In addition, SSM-based chirp spread spectrum (CCS) modulation is fundamentally sinusoidal signals that can be modulated into a wideband frequency that increases or decreases over time because it encodes data through chirp.

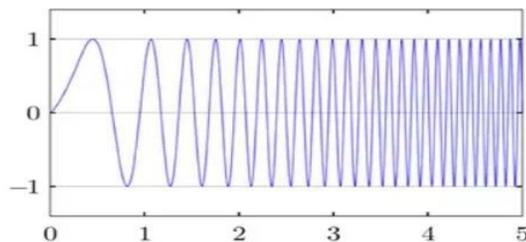


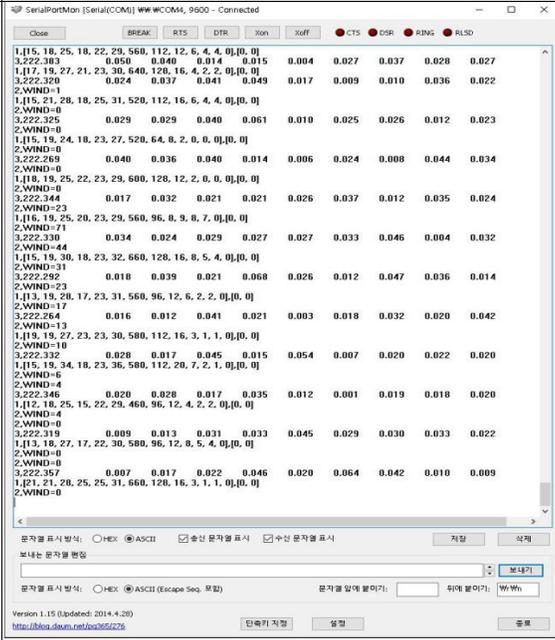
Fig. 14. Chirp spread spectrum (CSS) modulation

The network server is connected to multiple gateways through a secure TCP / IP connection and uses a way to eliminate duplicate messages and determine the gateway that should respond to end node messages. In addition, end-node data transmission rates are managed through the adaptive data transmission rate (ADR) method so that the network capacity can be maximized and the end-node battery life can be extended.

### 5.2 1: N star mode network of LPWA modules

Each PM sensor has an LPWA communication module that is connected to the big data server with 1: N star mode networks. By pairing, the band, channel and group ID of each transmission/receiving module are matched, address numbers are assigned sequentially, and the modules are implemented to carry out auto pairing when the power supply is reset later.

Table 5. Implementation of 1:N star mode networks

Carry out pairing	Form 1:3 LPWA networks
<pre> AT.START AT+ACODE=00000000 OK AT+PAIR=1  REG.START TGT:2,30 04 FA GID:60 05 1F CH:1,DR:2,M.ADD:0,D.ADD:2,BCST:12 REG.OK AT+RST=1 OK RDY SW:DJ12.0 BAND:3,CHN:1,DRATE:2,MODE:1,ADD:0,BCST:0  1,[16, 21, 28, 20, 25, 31, 520, 96, 12, 3, 1, 1, 0],[0, 0] 1,[18, 21, 28, 22, 25, 31, 540, 112, 12, 3, 1, 1, 0],[0, 0] 1,[18, 21, 30, 22, 25, 32, 540, 112, 16, 5, 1, 1, 0],[0, 0] 1,[18, 21, 28, 22, 25, 31, 500, 112, 16, 3, 1, 1, 0],[0, 0] 1,[16, 21, 28, 20, 25, 31, 480, 112, 12, 3, 1, 1, 0],[0, 0] 1,[16, 21, 28, 20, 25, 31, 480, 112, 12, 3, 1, 1, 0],[0, 0] 1,[15, 19, 25, 18, 23, 29, 500, 96, 12, 3, 1, 1, 0],[0, 0] 1,[13, 19, 25, 17, 23, 29, 460, 96, 12, 2, 0, 0, 0],[0, 0] 1,[15, 19, 25, 18, 23, 29, 480, 96, 12, 2, 0, 0, 0],[0, 0] 1,[13, 19, 22, 17, 23, 26, 440, 80, 12, 0, 0, 0, 0],[0, 0] 1,[13, 16, 19, 17, 20, 23, 460, 64, 8, 0, 0, 0, 0],[0, 0] 1,[15, 16, 19, 18, 20, 23, 500, 80, 8, 0, 0, 0, 0],[0, 0] 1,[15, 16, 19, 18, 20, 23, 520, 80, 8, 0, 0, 0, 0],[0, 0] 1,[13, 16, 19, 17, 20, 23, 500, 64, 8, 0, 0, 0, 0],[0, 0] 1,[15, 16, 19, 18, 20, 22, 480, 64, 8, 0, 0, 0, 0],[0, 0] 2,sglfjgkjdglsgjdfg 1,[13, 16, 18, 17, 20, 22, 460, 64, 4, 0, 0, 0, 0],[0, 0] 2,                     </pre>	

### 5.3 Big data server and monitoring software

The big data server stores PM /wind speed /atmospheric pressure /temperature and humidity data received through the LPWA module in the database and processes the statistical data and delivers the outcome to an Android app. Separate data visualization software was developed to analyze the data stored in the big data server, shown in Figure 15.

As for the process through which the PM data of the server are transmitted to the mobile app, when the PM sensor data monitoring request has been delivered to the HTTP protocol, the Apache web server accesses MySQL via JDBC, requests the query from the server, and transmits the collected PM data to the Android app as a web response.

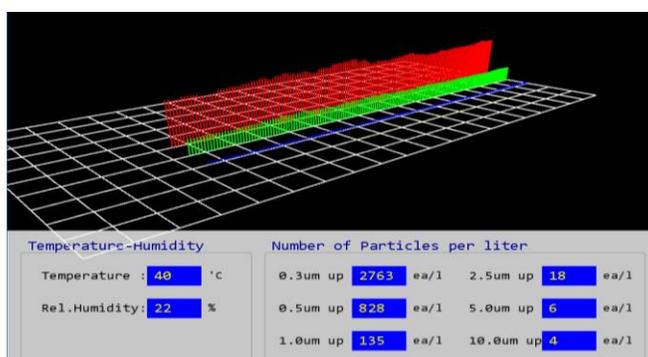


Fig. 15. Real time data visualization

### 5.4 Development of mobile application

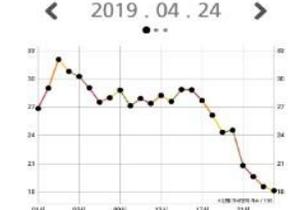
We developed an application that receives data by making a web connection in the data server, sending requests, and receiving web responses and outputting the data as a graph. The application was implemented to enable stable use by connecting and disconnecting the socket for each response when receiving data with a web connection and terminating the socket when the necessary data were transmitted.

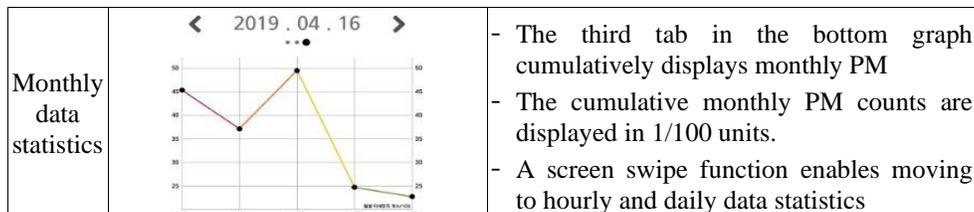
The application outputs real-time PM and FPM concentration values and enables selecting desired dates to identify the PM concentration in the time zone. The PM and FPM concentration values are output to meet the four level standards according to the PM / FPM concentrations provided by the Ministry of Environment, and monthly, daily and hourly statistical data can also be output together and checked. This application works with Android 9.0 and above.

To ensure the quality of the software, it was registered. The PM10 and PM2.5 data to be displayed were correctly output by date and time, and the response time was within 1 second. The program is based on the Android operating system, so it is compatible with various

mobile devices. However, in the process of importing the data, the loading time took several seconds for the massive amount of data. To prevent this, it was improved by efficiently loading the data. But access by a large number of users can lead to the same problems, and efficient allocation of resources is required to address these problems[23-24]. Research has also been conducted to select and analyze information that users need from big data[25-26].

**Table 6.** Description of the monitoring app user interface

Division	Screen	Remark
Real-time PM status		<ul style="list-style-type: none"> <li>- The current status can be checked with the air environment standard for PM (Framework Act on Environmental Policy)</li> <li>- PM concentration standard 0-30 (good), 31-80 (moderate), 81-150 (bad), 151- (very bad)</li> <li>- FPM concentration standard 0-15 (good), 16-50 (moderate), 51-100 (bad), 101- (very bad)</li> <li>- Screen color conversion (background / circular) Good (blue / blue), moderate (green / green), Bad (orange / red), Very bad (red / black)</li> </ul>
Change statistics check date		<ul style="list-style-type: none"> <li>- The calendar can be called up by tapping the date in the middle of the screen</li> <li>- The date can be changed by day by clicking the left and right arrows in the middle of the screen</li> <li>- The hourly, daily and monthly statistical data can be checked by selecting the desired date from the calendar.</li> </ul>
Hourly data statistics		<ul style="list-style-type: none"> <li>- Hourly PM is cumulatively displayed in the first tab on the bottom graph</li> <li>- The cumulative PM count per hour is displayed in 1/100 units</li> <li>- A screen swipe function enables moving to daily and monthly data statistics</li> </ul>
Daily data statistics		<ul style="list-style-type: none"> <li>- The second tab in the bottom graph cumulatively displays daily PM</li> <li>- The cumulative daily PM counts are displayed in 1/100 units</li> <li>- A screen swipe function enables moving to hourly and monthly data statistics</li> </ul>



## 6. Conclusion and future works

Among the methods of measuring PM, the  $\beta$ -rays method and the gravimetric method are used as official data because they have high accuracy. However, the problem is that the equipment is expensive and measurement in real time is impossible. In order to solve this problem, a light scattering method for measuring PM has been proposed, but it has the disadvantage of low accuracy. However, it could be used for official data if it satisfied high accuracy. In general, hardware performance was improved during the manufacturing stage of the sensor, but there were no studies that performed the process of correcting the sensor based on data.

The measured values were linearly corrected to satisfy high accuracy of the light scattering sensor, and the best performance was selected based on the RMSE and accuracy. The correction result is approximately five times more accurate than before correction, and the RMSE is much improved. However, since the collection method of the mathematical model is based on PM data of the gravimetric method, there is a limit to performance improvement. Since the light scattering type sensor is greatly affected by external factors when measuring PM, factors that affect PM measurement and their correlations will be analyzed for better correction in future studies.

We developed a light scattering type PM measurement sensor that can be supplied at a low price and enables measurements in real time. A low power mode is implemented in the sensor. It is powered by commercial batteries for use in a mobile environment, and enters a low-power mode during standby to allow long-term use. In addition to the PM sensor, wind speed, barometric pressure, and altitude sensors are attached. An LPWA communication module was attached to the PM sensor and LPWA LoRa was used as the communication method to transmit to the server. It connects to a 1: N star mode network and collects several PM sensor data in one server.

The mobile monitoring system was developed to check real-time PM data and recorded data through a mobile application. The data measured in real time are transmitted to the server at certain intervals and stored there. The data are analyzed and applied with a sensor correction value. It provides a PM monitoring service using the data stored in the big data server, and users can access all data such as monthly, daily and desired date in the server. Since the PM sensor was developed and a system to monitor it was built, it is expected that the PM information of the area where the PM sensors are installed can be easily accessed.

Although the accuracy was improved by an error correction algorithm, it is necessary to analyze the cause of the error. It will further analyze the causes that affect the concentration of PM by using data such as moisture or wind speed in the atmosphere among the data collected through sensors.

**Acknowledgements.** This research was supported by the Daejeon University Research Grants (2016)

## References

1. Loomis, D.; Huang, W.; Chen, G. The International Agency for Research on Cancer (IARC) evaluation of the carcinogenicity of outdoor air pollution: Focus on China. *Chin. J. Cancer* 2014, 33, 189–196.
2. Ki-Hyun Kim, Ehsanul Kabir, Shamin Kabir, "A review on the human health impact of airborne particulate matter," *Environment International*, Vol. 74, pp. 136-143, Jan., 2015.
3. Kwasny, F.; Madl, P.; Hofmann, W. Correlation of air quality data to ultrafine particles (UFP) concentration and size distribution in ambient air. *Atmosphere* 2010, 1, 3–14.
4. Costa, M.A.M.; Carvalho, J.A.; Neto, T.G.S.; Anselmo, E.; Lima, B.A.; Kura, L.T.U.; Santos, J.C. Real-time sampling of particulate matter smaller than 2.5  $\mu\text{m}$  from Amazon forest biomass combustion. *Atmos. Environ.* 2012, 54, 480–489.
5. França, D.D.A.; Longo, K.M.; Neto, T.G.S.; Santos, J.C.; Freitas, S.R.; Rudorff, B.F.T.; Cortez, E.V.; Anselmo, E.; Carvalho, J.A. Pre-Harvest Sugarcane Burning: Determination of Emission Factors through Laboratory Measurements. *Atmosphere* 2012, 3, 164–180.
6. PM2.5/PM10 Particle Sensor Analog Front-End for Air Quality Monitoring Design (2016)
7. Fine Particulate Matter (PM-2.5) Simple Meter GuideBook (2018)
8. Yu, Zhou; Yu, Jun; Xiang, Chenchao; Fan, Jianping; Tao, Dacheng.: Beyond Bilinear: Generalized Multimodal Factorized High-Order Pooling for Visual Question Answering. *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, 2018, 29(12): 5947-5959.
9. Yu, Jun; Guo, Yukun; Tao, Dapeng; Wan, Jian.: Human pose recovery by supervised spectral embedding. *NEUROCOMPUTING*, 2015, 166: 301-308.
10. Yuyu Yin, Wenting Xu, Yueshen Xu, He Li, Lifeng Yu: Collaborative QoS Prediction for Mobile Service with Data Filtering and SlopeOne Model. *Mobile Information Systems* 2017: 7356213:1-7356213:14 (2017)
11. Yuyu Yin, Yueshen Xu, Wenting Xu, Min Gao, Lifeng Yu, Yujie Pei: Collaborative Service Selection via Ensemble Learning in Mixed Mobile Network Environments. *Entropy* 19(7): 358 (2017)
13. Y. Yin, W. Zhang, Y. Xu, H. Zhang, Z. Mai and L. Yu, QoS Prediction for Mobile Edge Service Recommendation with Auto-encoder, *IEEE Access*. doi: 10.1109/ACCESS.2019.2914737
13. Debnath Bhattacharyya, "Garlic Expert Advisory System using Parallel ABC Algorithm", *International Journal of Disaster Recovery and Business Continuity*, NADIA, ISSN: 2005-4289 (Print); 2207-6425 (Online), vol.8, November (2018), (pp. 11-20), <http://dx.doi.org/10.14257/ijdrbc.2018.9.02>.
14. N. Thirupathi Rao, "A Review on Industrial Applications of Machine Learning", *International Journal of Disaster Recovery and Business Continuity*, NADIA, ISSN: 2005-4289 (Print); 2207-6425 (Online), vol.8, November (2018), (pp. 1-10), <http://dx.doi.org/10.14257/ijdrbc.2018.9.01>.

15. G. M. Faruk Ahmed, "Systematic Literature Review: Disaster Recovery and Business Continuity Plan for Core Banking Solution (CBS) Project", *International Journal of Disaster Recovery and Business Continuity*, NADIA, ISSN: 2005-4289 (Print); 2207-6425 (Online), vol.8, November (2017), pp. 29-44, <http://dx.doi.org/10.14257/ijdrbc.2017.8.03>.
16. Ekbal Rashid, "Disease Detection on the Basis of Multiple Symptoms by Expert System", *International Journal of Disaster Recovery and Business Continuity*, NADIA, ISSN: 2005-4289 (Print); 2207-6425 (Online), vol.8, November (2017), pp. 1-10, <http://dx.doi.org/10.14257/ijdrbc.2017.8.01>.
17. Subhanil Guha and Anindita Dey, "Study on Morphological Changes in and around Nayachara Tail Using Remote Sensing Techniques", *International Journal of Disaster Recovery and Business Continuity*, NADIA, ISSN: 2005-4289 (Print); 2207-6425 (Online), vol.7, November (2016), pp. 13-26, <http://dx.doi.org/10.14257/ijdrbc.2016.7.02>.
18. Jun-hee Choi, Hyun-Sug Cho, "A Study on the Improvement of the Accuracy of Low-Cost Light Scattering Method Particulate Matter Sensors," *Asia-pacific Society of Convergent Research Interchange*, Jun., 2019.
19. J. B. Wedding and M. A. Weigand, "An Automatic Particle Sampler with Beta Gauging," *Jour. of the Air Waste Manage. Assoc.*, vol. 43, no. 4, pp. 475-479, 1993.
20. source: [ipac.caltech.edu](http://ipac.caltech.edu)
21. <http://www.horiba.com/fileadmin/uploads/Scientific/Documents/PSA/BC002.pdf>, 2007
22. C. Wu, "Implementing the Radix-4 Decimation in Frequency (DIF) Fast Fourier Transform (FFT) Algorithm Using a TMS320C80 D0 DSP," Texas Instruments, Taiwan, Application Report: SPRA152, 1998.
23. H. Gao, Y. Duan, L. Shao, X. Sun. Transformation-based processing of typed resources for multimedia sources in the IoT environment. *Wireless Networks*, 2019. DOI: 10.1007/s11276-019-02200-6
24. Y. Zhu, W. Zhang, Y. Chen, H. Gao. A novel approach to workload prediction using attention-based LSTM encoder-decoder network in cloud environment, *EURASIP Journal on Wireless Communications and Networking*, 2019, 2019(247). DOI:10.1186/s13638-019-1605-z
25. H. Gao, C. Liu, Y. Li, X. Yang. V2VR: Reliable Hybrid-Network-Oriented V2V Data Transmission and Routing Considering RSUs and Connectivity Probability. *IEEE Transactions on Intelligent Transportation Systems*, 2020, DOI: 10.1109/TITS.2020.2983835
26. H. Gao, W. Huang, Y. Duan. The Cloud-Edge Based Dynamic Reconfiguration to Service Workflow for Mobile Ecommerce Environments: A QoS Prediction Perspective. *ACM Transactions on Internet Technology*, 2020, DOI: 10.1145/3391198

**Jun-hee Choi** received the B.S. degree in Computer Engineering from the Daejeon University, Daejeon, Korea, in 2016 and the M.S. degrees from the Chung-Ang University, Seoul, Korea, in 2019. Since March 2019, he has been with the Daejeon University, where he is Ph.D. student with the Department of Fire and Disaster Prevention College of Engineering. His research interests are Fire-fighting ICT, Computer Science and bigdata.

**Hyun-Sug Cho** received the B.S. degree in Mathematics from the Daejeon University, Daejeon, Korea, in 1996 and the M.S. and Ph.D. degrees from the Daejeon University, Daejeon, Korea, in 2001 and 2008, respectively. Since March 2006, she has been with the Daejeon University, where she is Professor with the Department of Fire and Disaster Prevention College of Engineering. Her research interests are network security, fire-fighting ICT and bigdata.

*Received: September 05, 2019; Accepted: June 14, 2020*