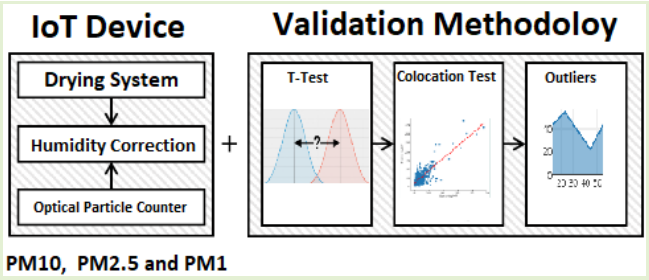
# Design and Evaluation of a Dryer System for IoT Hyperlocal Particulate Matter Monitoring Device

Eduardo Illueca Fernández, Iris Cuevas Martínez, Jesualdo Tomás Fernández Breis, *Senior Member, IEEE*, and Antonio Jesús Jara Valera, *Senior Member, IEEE*

***Abstract*—Particulate matter (PM) monitoring and climate change mitigation actions have been promoted due to the Paris agreement because of their impact on health and high mortality rate. High-resolution networks based on hyperlocal Internet of Things (IoT) sensors can play a fundamental role in improving data quality. In this context, hyperlocal refers to air quality IoT systems that allow collecting data in real time and in the cheapest way in comparison with local reference stations. Despite these methods are powerful and widely used by the scientific community, the signal is highly affected by relative humidity (RH). In this article, we present a system for measuring nanoparticles based on drying the air sampled and avoiding the hygroscopic growing of the particles. To the best of our knowledge, this is the first dryer system approach developed for IoT hyperlocal sensors. In addition, the relevance of our solution is supported by the following points: 1) we propose a new dryer system that has been patented; 2) our solution can be integrated into an IoT infrastructure that allows it to interact with other services; and 3) our solution has been validated in a real scenario in the city of Madrid. We have observed that the integration of a dryer system improves the performance of the OPC-N3 sensor and that we can measure the PM10 and PM2.5 fractions with high precision, $R^2 = 0.83$. In addition, our solution can measure small particles, such as PM1, with a good correlation against the reference air quality stations. Thus, our work contributes by improving high-spatial-resolution nanoparticle monitoring in correlation to official measurements to mitigate climate change.**




***Index Terms*— Particulate matter (PM), sensor system integration, sensor testing and evaluation, smart sensor system.**

## I. Introduction

ONE of the key aspects in current digital transformation, and into the challenges for the next years, is particulate matter (PM) monitoring as one of the main pollutants related to traffic emissions in urban areas [1]. In 2015, PM became one of the top five mortality risk factors in the world, with an estimation of 4.2 million people premature deaths [2]. In Europe, there are about 300 000 deaths per year due to PM2.5 and 245 000 due to inorganic particles [3]. In addition, particles with a high content of heavy metals, such as vanadium and nickel, correlate with cardiovascular mortality and morbidity [4]. The maximum levels of exposure to these particles are regulated in Directive 2008/50/EC of the European parliament and of the council on air quality and cleaner air for Europe, in addition to other pollutants, such as CO, $SO_2$, and $NO_x$. The categorization of PM mostly depends on the aerodynamic diameter size of particles. It is divided into three subclasses PM10 ($d \leq 10$ $\mu$m), PM2.5 ($d \leq 2.5$ $\mu$m), and PM1 ($d \leq 1$ $\mu$m), where the smaller particles generate a higher risk to health [5]. However, the PM monitoring methods proposed in Directive 2008/50/EC are too expensive to be implemented in a network with enough nodes to provide a high spatial resolution.

Hyperlocal air quality monitoring refers to high spatiotemporal measurements taken by combining Internet of Things (IoT) devices that allow the collection of real-time data and are easy to deploy in several locations, allowing public authorities to make policy changes [6]. Therefore, cost-effective high-resolution sensor networks monitoring pollution at fine spatial

Manuscript received 31 August 2023; revised 28 November 2023; accepted 6 February 2024. Date of publication 15 February 2024; date of current version 2 April 2024. This work was supported in part by Fundación Séneca under Grant 21300/FPI/19; in part by HOP Ubiquitous S. L., Región de Murcia, Spain, under Grant 21681/EFPI/21; in part by the Madrid City Council under Grant 145/2018/03144; and in part by the European Commission under Grant 101037648, Grant 10108653, and Grant 101100728. The associate editor coordinating the review of this article and approving it for publication was Prof. Xingwang Li. *(Corresponding author: Eduardo Illueca Fernández.)*

Eduardo Illueca Fernández and Jesualdo Tomás Fernández Breis are with the Department of Informatics and Systems, University of Murcia, 30100 Murcia, Spain (e-mail: eduardo.illueca@um.es).

Iris Cuevas Martínez and Antonio Jesús Jara Valera are with Research and Development, Libelium LAB, 30562 Ceutí, Spain.

Digital Object Identifier 10.1109/JSEN.2024.3364537

and temporal resolutions [7] are essential to regulating and performing smart management and making decisions in smart cities [8]. The importance of hyperlocal air quality sensors is that an extensive network of several compliant reference stations and a much larger number of hyperlocal sensors can deliver reliable, high-temporal-resolution data at neighborhood scales [9]. In this context, the use of IoT architectures and smart cities strategies allows for integrating the sensors into a bigger ecosystem and interacting with other components and services that improve its performance [10].

In this sense, the use of and research on so-called hyperlocal IoT sensors has increased in recent years. The need of the IoT on nanoparticles is highlighted by the fact that real-time data can ease the process of decision-making, and the IoT should address this [11]. The IoT devices are based on optical particle counters (OPCs) that base the measurement on the assumption that the number of particles is proportional to the light scattering. In this method, a light source illuminates the particles, and then, the scattered light from the particles is measured by a photometer. For particles with diameters greater than 0.3 $\mu$m, the scattered light is roughly proportional to their number concentration [12]. It has been shown that such sensors can have an acceptable linear response in the laboratory. However, there are deviations under real conditions due to ambient factors, such as temperature and, especially, relative humidity (RH) [13], [14]. Regarding this last factor, the correction for humidity growth reduces the bias of the particulate monitors. [15].

Related to this last issue, water vapor can condense on aerosol particles, making them grow hygroscopically under high RH conditions and changing the light scattering coefficient. So, the light scattering method quantifies a larger diameter and changes, and this error is propagated to calculate mass concentration [16]. Thus, accuracy and error influenced by RH are nonlinear with time and depend on the chemical composition of the particles and season, factors that affect the growth of the parcel under humidity conditions [17]. To correct this effect, reference instruments usually have dryer systems that remove water from particles before measurement. In contrast, hyperlocal air quality sensors do not include such drying processes, resulting in particle sizes being overestimated at high RH, resulting in PM values being then enhanced relative to reference measurements, supposing a limitation of state-of-the-art approaches.

Therefore, we hypothesize that adding a dryer system to hyperlocal air quality sensors will improve the quality of the measurements. In this work, we describe the development and validation of a humidity correction dryer system based on silica gel and infrared radiation. The dryer system has been developed and integrated into three IoT hyperlocal PM devices—based on the Alphasense OPC-N3 sensor technology—and validated by collocation tests in three reference stations in the city of Madrid. The system is a circuit composed of two columns that dries the sample before it arrives at the sensor. The main contribution of the work is the validation of a dryer integrated into a hyperlocal air quality device, since, to the best of our knowledge, this is the first dryer system implemented in a hyperlocal nanoparticle sensor and the first time this dryer system is tested in a real-time environment.

TABLE I
PM DEVICES AND TECHNOLOGIES

| Device | Method | Reference |
|---|---|---|
| BAM Monitor | Betta Attenuation | [18] |
| TEOM Microballance | Gravimetry | [19] |
| GRIMM 11-D | Optical Particle Counter | [20] |
| *FIDA Palas* | Optical Particle Counter | [21] |
| Alphasense OPC-N2/N3 | Optical Particle Counter | [22], [23] |
| Shiney PPD42NS | Optical Particle Counter | [24], [25] |
| Samyoung DSM501A | Optical Particle Counter | [24], [26] |
| Sharp GP2Y1010AU0F | Optical Particle Counter | [24], [27] |
| Plantower Serie | Optical Particle Counter | [28], [29] |
| Purple Air | Optical Particle Counter | [15] |

The rest of this article is structured as follows. First, Section II presents all the work done on the use of hyperlocal air quality sensor and humidity correction approaches. Then, Section III describes the physical design of our solution as well as the resources used. From these tests, Section IV is obtained, which are analyzed in Section V. Finally, the main milestones and future challenges are presented in Section VI.

## II. State of the Art

According to the European legislation, the reference method for PM determination is manual gravimetry, defined in the norm UNE-EN 12341:2015. However, the same directive states that methods showing a good correlation with manual gravimetry can be used for monitoring. This last group includes methods, such as automatic gravimetry, betta attenuation, and OPCs, which are quite important as they can be miniaturized in hyperlocal air quality sensors. For this reason, other methodologies have been explored to be miniaturized, such as separation methods. Table I states the technology behind the devices discussed in this article.

OPCs are based on light scattering, and each particle's diameter is computed. Then, the particle's mass is calculated using the density of the standard used for factory calibration and assuming that the particle is spheric. The mass of all particles belonging to one bin is summed to obtain the mass concentration, and all the bins belonging to one fraction are summed to compute the total mass concentration.

### A. Hyperlocal Air Quality Sensors

Numerous studies have been conducted to evaluate hyperlocal air quality devices against medium—and high-cost—reference techniques. In laboratory conditions, the hyperlocal air quality sensors perform well, normally $R^2 > 0.85$ [30]. However, this is not a realistic result, because these devices tend to be less precise under real conditions—named field evaluation in contrast to laboratory evaluation studied the performance of a low-cost sensor in California, finding a moderate agreement with beta attenuation monitoring (BAM) ($R^2 = 0.35$–$0.81$) [31]. On the other hand, evaluation against the tapered element oscillating microbalance (TEOM) monitor has a lower performance ($R^2 \leq 0.30$) with a low concentration range. However, a higher correlation with the BAM monitor

($R^2 > 0.80$) was observed in places with high concentrations in the same city [32], showing a great difference between concentration ranges and reference techniques used for validation. Consequently, it is necessary to understand which factors affect IoT hyperlocal sensors' performance.

The first study to make a significant impact on this field was conducted the study by Wang et al. [24] where three low-cost sensors Shiney PPD42NS, Samyoung DSM501A, and Sharp GP2Y1010AU0F were evaluated with the SidePark reference technique. As a result, very high linearity was obtained with all sensors ($R^2$ values greater than 0.89) in laboratory conditions. This study was the first to highlight the critical dependence on humidity, particle size, and particle chemical composition. Another problem typically exhibited by these sensors is scale bias. For example, in this study, Speck sensors for PM2.5 were used and found to overestimate the concentration by 200% in indoor conditions and 500% in outdoor conditions, compared to the GRIMM Reference Dust Monitor [21]. On the other hand, the Plantower PMS1003 sensor had a bias of +46% when measuring PM10 [28].

Another important feature is intrasensor variability, the variation between the measures taken from different devices for the same manufacturer. In [22], 12 low-cost sensors were evaluated in their measurement of the PM2.5 fraction, testing three sensors of the same brand under the same conditions. For the OPC-N2, a wide range of $R^2$ values were obtained—between 0.38 and 0.67. This variability was also related to the treatment done to report data. For example, there were differences between daily means and hourly means. For the PMS3003 sensor, the evaluation was conducted for hourly and daily averages, obtaining $R^2$ values ranging from 0.40 to 0.90. However, the slopes were close to one when the averages were hourly, and they were much larger in the case of daily averages [29].

Finally, a deeper analysis is done based on the performance of the Alphasense OPC-N3, the sensor that will be integrated into our solution. The South Coast AQMD1 conducted a field evaluation. This report showed the best performance for the PM1 fraction and a poor performance for the PM10 fraction. In this sense, the results for hourly means were $R^2 =$ 0.78–0.82 for PM1, $R^2 =$ 0.61–0.69 for PM2.5, and $R^2 =$ 0.48–0.53 for PM10. However, those results got improved for small fractions when daily mean are applied, $R^2 =$ 0.88–0.90 for PM1 and $R^2 =$ 0.69–0.76 for PM2.5, except for PM10, $R^2 =$ 0.22–0.26.

Regarding other technologies, capacitance-based PM sensors are becoming a hot topic in air quality monitoring, because they can be easily integrated into small devices and can be combined with particle discrimination methods [33], [34]. These separation techniques are based on the movement of phoretic particles in a fluid. The movement of the particles is based on a gradient that could be electrical (electrophoresis), chemical (diffusiophoresis), or thermal (thermophoresis) [35]. The motion of particles across gradients, whether thermal, electric, or concentration, is a well-established phenomenon, which already has applications in several fields, such as the study of proteins in molecular biology [36]. More concretely, thermophoresis has been studied to design real-time PM sensors, and a sensor has been modeled using the finite-elements method, with promising results [37], [38]. Thermophoresis is a transport force that occurs due to a temperature gradient, moving particles less than 2.5 $\mu$m in heater regions than bigger particles, allowing separation [39].

### B. Humidity Correction Systems

The effect of humidity on particle size and its application to particle measurements through light scattering has been widely studied. The physical explanation is that water vapor condenses on the surface of aerosol particles, making them grow hygroscopically—increasing their diameter—under high RH conditions [40]. To correct for this effect, reference instruments are usually equipped with dryer systems that remove water from particles by applying heat before measurement.. For instance, Palas Fidas 200 S is an EN 16450 certified static light scattering PM measurement instrument that incorporates a dryer system, the intelligent aerosol dryer system (IADS), used to remove water from particles before measurements [41]. On the other hand, the TEOM microbalance sample passes through a pair of Nafion dryers, with a specific geometry to control particle loss, remove water vapor, and reduce humidity levels. From the exit of the sample equilibration system (SES), the sample flow is routed through the TEOM mass sensor, and the bypass flow is routed to the control unit in the usual fashion. Because of the pressure drop created at the mass flow controllers, the zone between the mass flow controllers and the vacuum pump contains air with reduced water vapor pressure. This dried, low-pressure merged system flow then becomes the purge flow for the sample and bypasses Nafion dryers [42]. Other approaches consist of implementing technologies that are not temperature-dependent. For instance, a real-time PM2.5 sensor module with the high mass resolution was developed with the module consisting of a two-stage aerosol impactor, a thin-film piezoelectric on silicon (TPoS) MEMS oscillator, and a micropump, achieving a collection efficiency of 51% for 2.54 $\mu$m and 50% for 1.03 $\mu$m [43].

In contrast, hyperlocal sensors cannot implement these systems due to their low size and the complexity of the previous systems, so there is a need for coupling these devices with dryer systems. There are some approaches in this sense in the state of the art, such as the design of a sample inlet containing Perma Pure dryers [44], allowing the control of sample RH and temperature. Then, the system also contains a nephelometer and particle-sizing instrument. In Perma Pure dryers, the sample aerosol entered the inlet through a protective cover, which eliminated rain or insects from the sample train [45]. Another work has developed dryer systems for reducing the error by humidity on measuring PM in the ambient air with optical particle measuring instruments. Two types of dryer systems were designed: dryer systems using heating and dilution methods, showing a high correlation with reference systems [46].

Parallel to the physical dryer systems, there are many state-of-the-art approaches to the moisture problem from an algorithmic point of view. For example, a mathematical method was developed to assess the PM levels based on the resolution of the inverse problem in aerosol tomography [47].

Beyond this, other approaches seek to remove the signal due to the humidity. One of these solutions has proposed an RH correction factor that could be applied to PM data in high RH conditions. This factor was calculated from historical PM data measured with an Alphasense OPC-N2 compared against a TEOM reference instrument based on gravimetry, which is not affected by humidity measurements [20]. However, this strategy is not coherent with the physical phenomenon. On dehydration, particles would reduce in size, not in number, thus affecting the derived PM in ways that now would depend on the detailed particle size spectrum.

To overcome this, exhaustive work has been done to calculate the correction factor from the measured particle size distribution rather than mass values. The results not only showed significant improvement in sensor performance but also retained fundamental information related to particle composition. Thus, a particle size distribution-based correction algorithm was developed to correct the influence of RH on sensor measurements. The application of the correction algorithm, which assumed a physically reasonable correction factor, with the overestimation of PM measurements reduced from a factor of 5 before correction to 1.05 after correction [40]. The main drawback of this study is that this correction factor is highly dependent on the chemical composition, and it is difficult to extrapolate these results to other locations or to explain the seasonal intrinsic variation associated with PM chemical composition.

### C. Hyperlocal Air Quality Networks Across Europe

Beyond a research level, the implementation of hyperlocal air quality networks has been done by public authorities in cities under the guidelines of the European Union, aiming to be climate neutral by 2050, and it needs to reduce at least 55% of emissions by 2030. In Spain, cities, such as Valencia (https://geoportal.valencia.es/apps/GeoportalHome/es/inicio/), Santander (https://www.smartsantander.eu/), or Molina de Segura (https://ciudadinteligente.molinadesegura.es/), have developed their own smart city platforms. At the European level, the French AIR Parif platform is of special relevance, which integrates air quality data, emissions, and forecasting over an extensive network of sensors (https://www.airparif.asso.fr/). Helsinki has a promising real-time air quality monitoring platform (https://hri.fi/data/en/dataset/reaaliaikainen-ilmanlaatu-hsyn-ilmanlaadun-mittausasemilla), and Amsterdam provides hyperlocal heat maps with hyperlocal data at street level by combining observations with models (https://maps.amsterdam.nl/klimaatadaptatie/?LANG=nl). Finally, implementing policies through these local initiatives allows for reducing emissions. For instance, research suggests that a combination of direct regulation of diesel vehicles and a targeted rebate to encourage substitution between diesel vehicles and BEVs can reduce PM emissions by 13% [48].

## III. Materials and Methods

Three units have been developed and installed in three reference stations in Madrid. This section describes the different methodologies and resources used for this work. First, a description of the hardware system proposed is done, followed by a summary of the IoT-cloud architecture in which these sensors are integrated. This architecture allows us to perform a validation based on colocation tests described in Section III-C. Finally, information about the software and computational resources used is provided.

### A. Hardware of the Dryer System

The designed hyperlocal PM sensors are based on light scattering, where particles move through an airflow and are illuminated by a laser beam. The light is then scattered in all directions according to the particle properties—size, shape, absorption, and refractive index—and the beam's wavelength. This does not affect the measurement process, because it remains constant [47]. Our solution integrates the Alphasense OPC-N3 sensor, selected for its measurement capacity and price competitiveness. However, it is important to highlight that OPC-N3 is more expensive than the rest of the sensors presented in the state of the art, because its hardware is more complex and achieves higher accuracy [23], reaching a range of 0.35 a 40 $\mu$m—based on a refraction index of $1.5 + 0i$—and a maximum mass concentration of 2000 $\mu$m/m$^3$. It samples particles with a typical flow rate of 280 mL/min and a velocity of 1000 particles per second. This sensor is known as an OPC, because it calculates the diameter of the particle and classifies it into one bin, a size range in which all particles with a diameter within this size range are contained. Then, the particle's mass is calculated as PM1, PM2.5, and PM10 according to the following formula:

$$m = \sum_{i=0}^{n} \frac{4}{3}\pi r^2 \rho. \tag{1}$$

It is easy to note that particle mass scales with particle size according to a cubic relationship, so small errors in particle classification lead to large errors in calculated mass. For this reason, a patented system [49] has been implemented to dry the air sample and avoid hygroscopic error, as shown in Fig. 1. The dryer system is based on a Teflon column filled by silica gel and an infrared radiation lamp. When the lamp is on, the silica gel absorbs the moisture from the passing air, drying it, and when it is off, desorption takes place, and the moisture returns to the air. Therefore, the airflow is as follows: 1) it enters through the first column and dries out as the lamp is switched on; 2) it leaves the column and passes through the optical sensor; and 3) it leaves the sensor and passes through the exit column, taking up the moisture and regenerating the silica gel. In order to optimize this process, two optical sensors are placed, so when air enters one column, it is measured by one sensor, and regeneration of the second column is allowed. In the next cycle, the air enters the regenerated column, and the second sensor measures the particles.

As can be deduced from Fig. 1, the air sample is dried before passing the sensor. Thus, this affects the signal, because the hygroscopic growth is removed before the sensor takes the measurement. In this context, the computed size by the sensor is closer to the real size, implying a lower value of $r$ in (1). Consequently, computing mass concentration in a cubic way generates a more accurate value. This approach presents

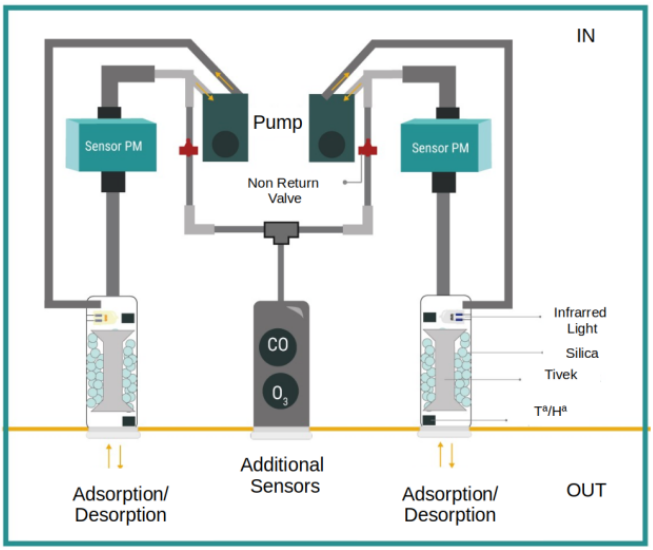



Fig. 1. Scheme of the proposed dryer system, based on a circuit with two silica columns and two OPC-N3.

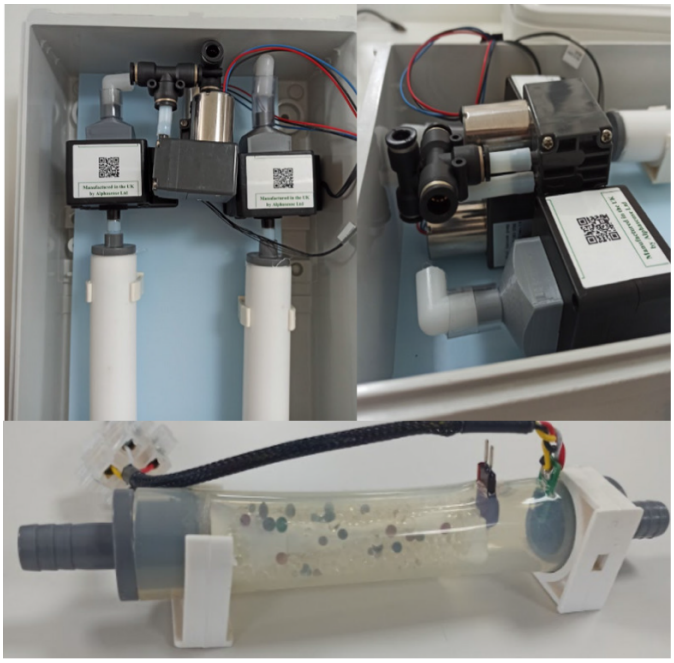


Fig. 2. On the left top, the circuit is presented with the two OPC-N3 sensors, the dryer columns, the pump, and the valve that changes the cycles. On the right top, a delta of the OPC-N3 and the connection tube to the valve are shown. On the bottom, a detail of what is inside the dryer column is shown.


Fig. 3. Scheme of dryer column, composed of Teflon column of 80 mm and two paper filters with a radio of 36 mm, and 14 mm the radio of the cavity.

the following advantages in comparison with the techniques exposed in state of the art: 1) it can be implemented inside an IoT device, which is not possible for the size selectors in TEOM microbalances [42]; 2) this approach does not need to generate heat to correct humidity, as occurs in dryer systems implemented in reference devices or thermophoretic approaches [37], which implies an energy consumption not feasible for IoT devices [41]; 3) it is more exhaustive than the approach proposed by Perma Pure dryers; and 4) it acts directly over the measurements, which has not done the algorithmic approaches [47].

Fig. 2 details how the circuit is organized inside the box. Both dryer columns are connected to one OPC-N3, and each one is connected to a valve connected to the pump. This component maintains the flux across the circuit and allows particles to pass through the sensors. At the bottom of Fig. 2, a detail of what is inside the column is filled with the silica gel. It is important to note that the infrared lamp is placed at the end of the tube, which is in charge of heating the gel and dry particles when it is on.

The prototype schematized in Fig. 1 is based on two dryer columns that are in charge of drying the sample and removing the hygroscopic growth of the particles. For this purpose, we use two Teflon tubes of 14 and 80 mm in radius, coupled with two paper filters at the input and output to remove particles—a surface of 26 239 $mm^2$ of Teflon (Fig. 3). The radius of the filter is 36 mm, and the radius of the overture that allows the air to pass is 14 mm—meaning an area of 6912 $mm^2$ of filter paper. The reason for using Teflon as the main material is that other options, such as silicone, could induce error, as the particles were prone to adhere to the walls. However, it is true that for some connections, silicone tubing must be used to achieve a fixed connection between the component (filter, solenoid valve, and so on) and the Teflon tubing. In other words, the aim is to achieve zero contact between the silicone and the airflow to obtain a lower error. The tubes are filled with silica gel that is in charge of absorbing water from particles. This process is accelerated by an infrared lamp controlled by a temperature sensor inside the column.

Despite the key components described previously, the new IoT system comprises several components. Table II lists all the elements needed to replicate the prototype with the required quantity for a single unit. It is important to note that here, we refer to the dryer column as the whole system described in the previous paragraph, with the Teflon cylinder, the silica gel, and the filters.

According to the Alphasense web page (https://www.alphasense.com/faqs/), OPC-N3 units are calibrated for sizing using controlled aerosols of monodisperse polystyrene latex microspheres of specific sizes. Aerosol number concentration is compared to an OPC *gold standard*, previously calibrated against a certified TSI 3330 OPS instrument. After this factory calibration, the whole unit, after integrating OPC-N3 with the dryer, is calibrated in the field with the GRIMM 11-D OPC reference system, which has a high correlation with manual gravimetry [20].

### B. Architecture

The data collection process is realized through an IoT-cloud architecture that is organized in different layers, as shown

TABLE II
COMPONENTS PER DEVICE

| Component | Units |
|---|---|
| Libelium Smart Spot Electronic Board | x1 |
| 12v chargers | x1 |
| Cable 12v to Double Enclosure | x1 |
| Double Enclosure - Feeding | x1 |
| Double Enclosure - Control | x1 |
| Umbilical Cord Cables 5v | x1 |
| Motor control plates H-bridge I2C | x1 |
| Standard OPCNx SPI cable 5v | x1 |
| Special OPCNx SPI cable 5v | x1 |
| Alphasense OPC-N3 | x2 |
| Pump 12v | x2 |
| Dryer column | x2 |
| TCA Temperature Sensor | x4 |
| Infrared Lamp G4 12W 10W | x2 |
| Lamp Holder | x2 |
| Red mosfet boards LP55 5v | x2 |
| Antenna | x1 |
| Silica Encapsulation S1-2016 | x1 |
| GT40 reductors | x2 |

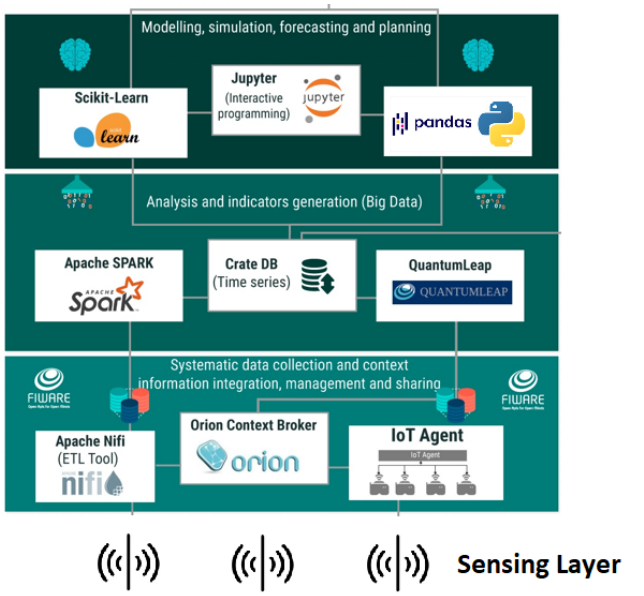



Fig. 4. IoT-cloud architecture for data collection, composed from bottom to top by a sensing layer, a contextual layer, a persistence layer, and a data analytics layer.

in Fig. 4. The *sensing layer* comprises the IoT systems and is the one that directly performs the measurements, using hyperlocal particle sensors as "things" according to the IoT paradigm. Then, the *contextual intelligence layer* integrates the collected data into an *Orion context broker* through an *IoT agent*. It is important to note that this layer integrates real-time contextual data. Next, these data are stored by the *persistence* layer in a CrateDB database. Finally, the analytics layer is the JupyterLab interface to perform the statistics through *pandas* and *sklearn*. In this layer, the reference data are injected from the European Air Quality Portal via the respective application programming interfaces (APIs) and merged into a single *csv* per location. In addition, the *csv* provided by the Madrid City Council is added to this layer, as the mobile station (see Section III-C) does not report to the European Air Quality Portal. The generated *csv* files are used to validate our solution, as explained in Section III-C.

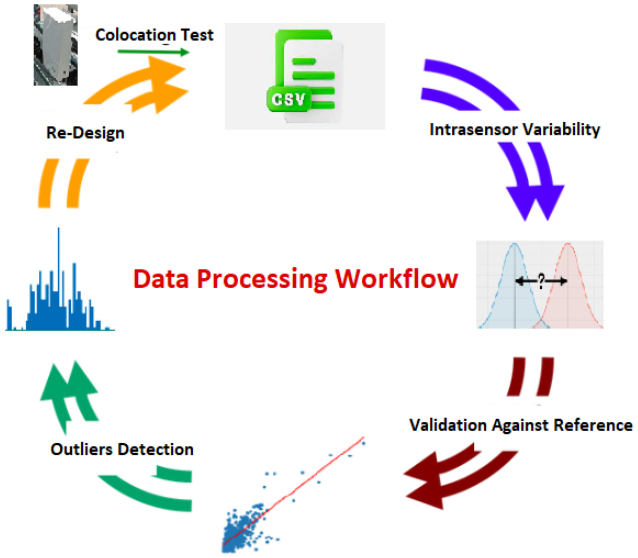



Fig. 5. Workflow for data processing and validation.

### C. System Validation

To validate our measurements, this work proposes a data mining approach based on statistics and inference to determine the quality of the data, according to the standardized workflow in data science, by placing our devices next to the reference stations of local air quality network (colocation test). The proposed workflow is summarized in Fig. 5, starting on the *csv* file generated in the data collection. It is composed of the following steps: 1) the prevalidation and intercomparison of the different PM sensors using the *student's t-test*, since we can assume the normality of the data because of the central limit theorem and the correlation matrix; 2) validation against an air quality station using the Pearson test and the linear regression model; 3) study of the noise with the $Z$-score algorithm for outliers detection; and 4) the iterative redesign of the solution and/or colocation tests to improve the performance of the system.

The evaluation corresponding to the second step of the workflow has been performed in three different locations in Madrid, and it was conducted at three granularity levels: on hourly averaged real data—to validate against the reference devices, on daily averaged real data—to assess the capacity for long-term monitoring, and on maximum daily—to study the effect of high concentrations. At these three levels, we have applied metrics to compare the different devices to the reference station concentrations quantitatively. These metrics included Pearson $r$ (2), which is a measure of the strength and direction of a linear relationship; mean absolute error (MAE) (3); and coefficient of determination $R^2$ (4). In these equations, $\hat{y}$ and $y$ are the predicted and real vectors, $m_{\hat{y}}$ is the mean of the vector $\hat{y}$, and $m_y$ is the mean of the vector $y$

$$r = \frac{\sum_{i=1}^{n} \left(y_i - m_y\right)\left(\hat{y}_i - m_{\hat{y}}\right)}{\sqrt{\sum_{i=1}^{n} \left(y_i - m_y\right)^2} \sqrt{\sum_{i=1}^{n} \left(\hat{y}_i - m_{\hat{y}}\right)^2}} \tag{2}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{3}$$


Fig. 6. Colocation tests in reference air quality stations.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n} \left(y_i - m_y\right)^2}. \tag{4}$$

The reference devices used to validate our solution were located on the official air quality stations. These devices are based on technologies that shows a high correlation with the manual normalized gravimetry defined in the norm UNE-EN 12341:2015. The official air quality stations report measurements every hour, as the mean of all values measured in an hour. Later, they are validated and stored in the European Air Quality Portal from where they can be accessed through an API. There are two kinds of reference devices. The first one is based on automatic microbalance (TEOM), and the physical principle beyond this technology is gravimetry. This device is installed in the Cuatro Caminos and Escuelas Aguirre air quality stations. On the other hand, optical methods show a high relationship with manual gravimetry. In this work, we will use data from two devices: the FIDA Palas, installed into Unidad Móvil and the GRIMM 11-D OPC [31], available in our laboratory. Fig. 6 shows the devices installed in the reference stations.

### D. Data Analysis

Data analysis has been done in a Python 3 environment, due to the wide availability of data science tools. The libraries used in this work and its references are summarized in Table III.

These libraries have been used to implement the workflow proposed in Fig. 2. The first step is the extraction of the csv file, which can be found in Zenodo (Zenodo Respository). Once the csv file is obtained, the intravariability test is performed with the ttest_ind() function from *scipy* by separating data from different columns. Then, the linear models are performed with the oms() function from *statmodel* and the $Z$-score and outliers by implementing (5). The reference data were obtained from the European Air Quality Portal API

TABLE III
LIST OF USED LIBRARIES

| Name | Use | Reference |
|---|---|---|
| *numpy* | Scientific computing | https://numpy.org/ |
| *pandas* | Data analysis | https://pandas.pydata.org |
| *scikit-learn* | Linear regressions models | https://scikit-learn.org |
| *scipy* | Mathematics and statistics | https://www.scipy.org |
| *joblib* | Running models | https://joblib.readthedocs.io/ |
| *matplotlib* | Data visualization | https://matplotlib.org |

(https://fme.discomap.eea.europa.eu), by setting as parameters CountryCode = ES, Pollutant = 10 (PM10) or 6001 (PM2.5), Year_from = 2021, Year_to = 2022, Source = E1a, and EoICode = ES0118A (Escuelas Aguirre) or ES1525A (Cuatro Caminos). The meteorology has been extracted from the Madrid Open Data Portal (https://datos.madrid.es). The data from the mobile unit have been provided by the Madrid city council, as it is not available and it is necessary to ask for them

$$Z = \frac{x - \mu}{\sigma}. \tag{5}$$

The system is hosted in an OVHcloud (OVH) server with an Intel[1] Xeon[1] CPU E3-1245 V2 @ 3.40 GHz—eight cores operating system with 31 GB of RAM and 1 TB of storage. The operating system is Ubuntu 18.04 × 86_64.

## IV. RESULTS

The following results were obtained from three collocation tests. A total of two devices were deployed in the Escuelas Aguirre reference station and Cuatro Caminos reference station, and one device was deployed in the Unit Movil station. Just to clarify, the three deployed devices' locations were the same during all the tests. Thus, three locations were tested, and the devices did not move during the testing period. All experiments were performed simultaneously, with an hourly sampling rate from July 10, 2021 02:00:00 to August 31, 2021 23:00:00. We obtained data from 1205 observations. Regarding weather conditions, collocations tests were taken in summer conditions in Madrid, with humidity values ranging from 11% to 85% and temperature values ranging from 17 °C to 42 °C. The RH median is equal to 32%, and the temperature median is equal to 26.6 °C.

### A. Data Analytics and Data Mining

Fig. 7 shows the correlations between the different variables that our device can measure against our reference OPC, GRIMM 11-D. This is important, because a high correlation could help to implement better calibration models. It is possible to observe a high correlation against the largest bins—particles with a diameter between 0.35 and 10 $\mu$m—and the PM1, PM2.5, and PM10 fractions. Also, the correlation with humidity and temperature is also significant.

As explained in materials and methods, the solution is based on a two-component system with two OPC-N3 sensors fed by two different columns, so it is important to test if there is high variability between sensors and columns. Using an independent $t$-test to compare the measured means between

[1]Registered trademark.

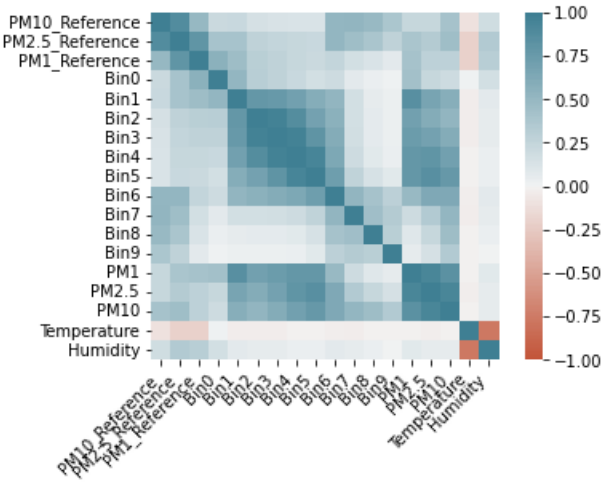



Fig. 7. Correlations between reference variables and device features.

columns, we obtain significant $p$-values for PM1, PM2.5, and PM1 ($p = 2.75 \times 10^{-14}$, $p = 2.75 \times 10^{-39}$, and $p = 2.75 \times 10^{-34}$). Fig. 8 shows the probability distribution for both columns and the comparison of the mean and standard deviation values. It is observed that the distribution represented in orange is wider, counting particles for a bigger size. On the other hand, the blue distribution is narrower, meaning lower variability in the measurements.

### B. Humidity Effect and Drying

Regarding the humidity monitoring, the results are summarized in Fig. 9. It can be noted that the humidity values were low in general. Still, higher values were indeed obtained at night. The same occurred with the temperature, with high values during all the monitoring periods—even greater than 40 °C. The comparison between the ambient humidity and the OPC-N3 internal humidity provided interesting insights, allowing us to quantify the effect of the dryer system. This analysis was done in Escuelas Aguirre and Cuatro Caminos, since the mobile unit did not provide humidity data. Fig. 9 shows that internal humidity (orange) was lower than ambient humidity (blue), especially at night—for Escuelas Aguirre, the mean and maximum humidity reduction were 43% and 56%, respectively. In contrast, for Cuatro Caminos, they were 46% and 60%, respectively. However, the correlation between the internal and external humidities is quite high. This pattern occurs in both stations, with Pearson coefficient 0.82 for Escuelas Aguirre and 0.98 for Cuatro Caminos.

### C. Hourly Mean Validation

The hourly mean validation was performed against the hourly values of the reference air quality data (Fig. 10). For this, hourly means were computed for the data of each device. This treatment is unnecessary for the reference measurements, because air quality stations report hourly means by default. The metrics used were the Pearson coefficient and the MAE. The best results were obtained for the mobile station (mobile unit) with $r = 0.84$, $r = 0.88$, and $r = 0.68$ for PM10, PM2.5, and PM1, respectively. In the other two stations, the results were quite similar for PM10 (0.86 for Cuatro Caminos

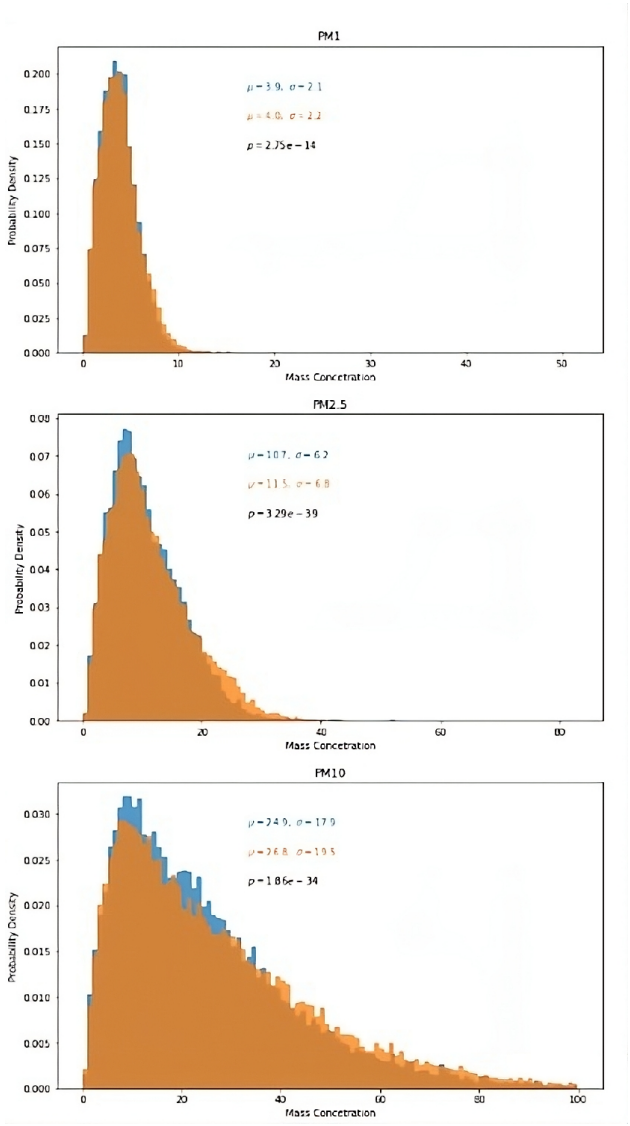



Fig. 8. Comparison of probability distributions between columns per fraction (PM1, PM2.5, and PM10). The text in each histogram shows the results of the independent $t$-test.

and 0.79 for Escuelas Aguirre) but worse for PM2.5 (0.60 and 0.59, respectively).

Regarding MAE, the patterns differed for PM10 fraction, where the higher value was for the Unidad Móvil (18.12 $\mu$g/m$^3$). In contrast, the PM2.5 fraction showed a higher error for Cuatro Caminos and Escuelas Aguirre (5.89 and 5.30 $\mu$g/m$^3$) in comparison with Unidad Móvil (4.65 $\mu$g/m$^3$). For PM1, the error observed was larger than PM2.5 (4.11 $\mu$g/m$^3$).

The next step in the validation of the measures is to use linear regression to test how close the measured values and the real ones were. Thus, the relationship between the two variables was a linear function of slope equal to 1 and intercept equal to 0. The slope represents the bias error, and the intercept represents the stochastic error. The graphic outputs of the model are presented in Fig. 11, from where a high linearity could be deduced, and error bars have been calculated from the tests at the three locations and coincide with the ranges of $R^2$ and MAE for all the tests.

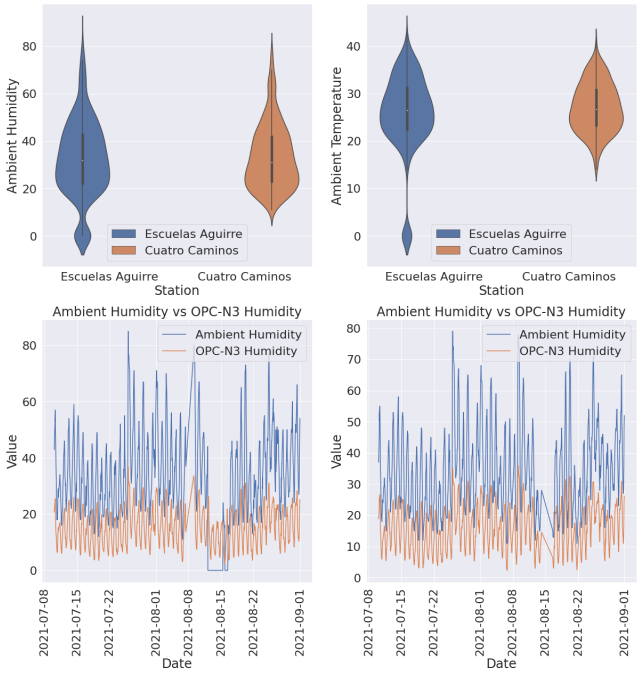



Fig. 9. Hourly mean validation of our device against air quality reference station. The metrics used are Pearson coefficient and MAE.

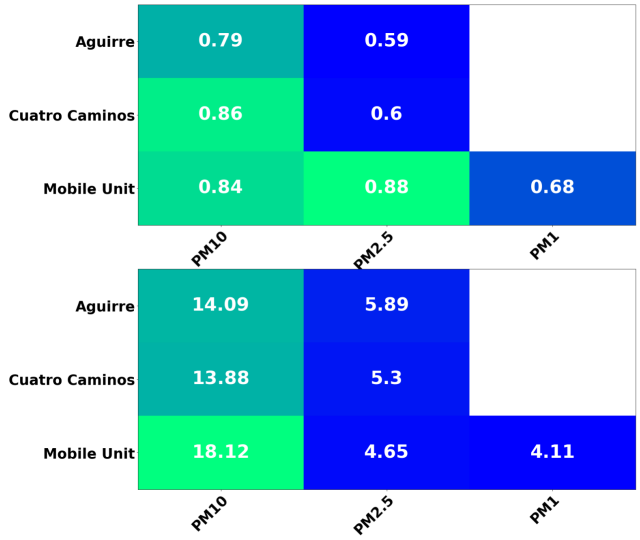



Fig. 10. Hourly mean validation of our device against air quality reference station. The metrics used are the Pearson coefficient and MAE. The 1264 hourly samples were taken.

Studying this hypothesis requires to analyze the statistics related to the regression model: the intercept, the slope, the $F$-test, and the associated $P$-value. In addition, the most important metric to evaluate the model is the coefficient of determination. All these results are shown in Table IV. The results were quite similar for all the fractions. For PM10, $R^2$ was equal to 0.74, 0.62, and 0.83 for Cuatro Caminos, Escuelas Aguirre, and Unidad Móvil, respectively. In contrast, for PM2.5, $R^2$ was equal to 0.356, 0.344, and 0.828 for Cuatro Caminos, Escuelas Aguirre, and Unidad Móvil. Finally, for PM1, the result was $R^2 = 0.50$.

## D. Daily Mean Validation

The daily mean validation was performed against the hourly values of the reference air quality data (Fig. 12). For this, daily means were computed for the data of each device and each air quality station. The best results were obtained for Unidad Móvil with $r = 0.97$, $r = 0.93$, and $r = 0.78$ for PM10, PM2.5, and PM1, respectively. In the other two stations, the results were quite similar for PM10 (0.96 for Cuatro Caminos and 0.95 for Escuelas Aguirre), but worse for PM2.5 (0.80 in both devices). It is important to note that the correlations improved when applying daily means.

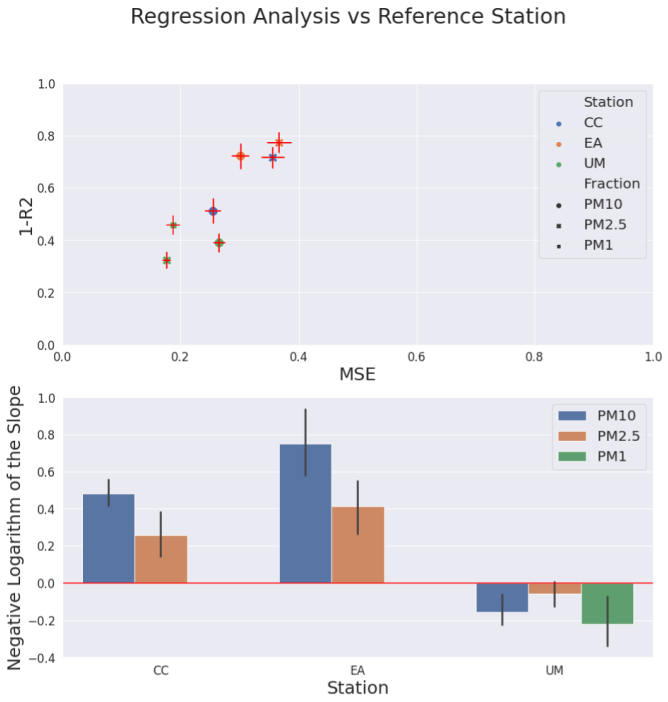



Fig. 11. Linear regression model for each fraction at each air quality station.

TABLE IV
LIST OF USED LIBRARIES

| Model | Intercept | Slope | F-Value | P-Value | $R^2$ |
|---|---|---|---|---|---|
| EA PM10 | 10.1 | 0.35 | 1950 | 0.00 | 0.62 |
| CC PM10 | 9.55 | 0.34 | 3317 | 0.00 | 074 |
| UM PM10 | 6.52 | 1.68 | 6046 | 0.00 | 0.83 |
| EA PM2.5 | 7.03 | 0.33 | 631 | 0.00 | 0.34 |
| CC PM25.5 | 6.82 | 0.41 | 658 | 0.00 | 0.36 |
| UM PM2.5 | 4.85 | 0.92 | 5974 | 0.00 | 0.83 |
| UM PM1 | 3.51 | 1.22 | 1241 | 0.00 | 0.50 |

About MAE, the patterns differed for the PM10 fraction, where the highest value was for the Unidad Móvil (16.95 $\mu$g/m$^3$). However, the PM2.5 fraction presented a lower error for Cuatro Caminos and Escuelas Aguirre (3.79 and 3.94 $\mu$g/m$^3$) in comparison with Unidad Móvil (4.12 $\mu$g/m$^3$). For PM1, the error observed was 4.11 $\mu$g/m$^3$. An improvement in comparison with hourly means was also observed, but it was less notorious than for the Pearson coefficient.

## E. Daily Maximum Validation

The daily maximum value validation was performed against the hourly values of the reference air quality data (Fig. 13). For this, maximum values were computed for each device's data and air quality station. The best results were obtained for the mobile station (Unidad Móvil) with $r = 0.94$, $r = 0.94$, and $r = 0.58$ for PM10, PM2.5, and PM1, respectively. The results

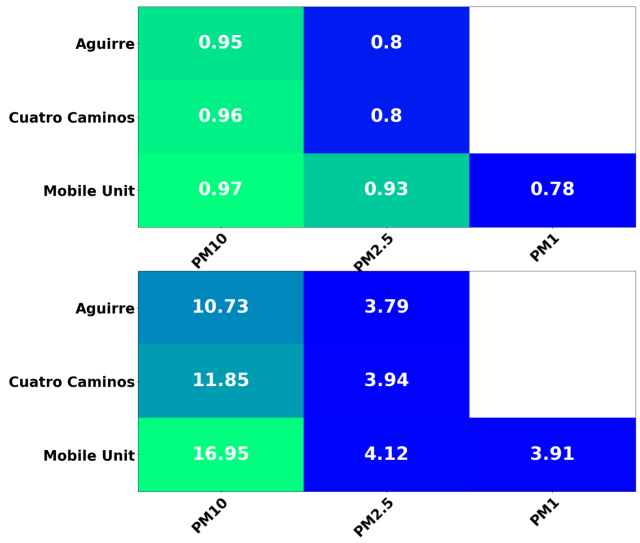



Fig. 12. Daily mean validation of our device against air quality reference station. The metrics used are Pearson coefficient and MAE. The 53 daily samples were taken.

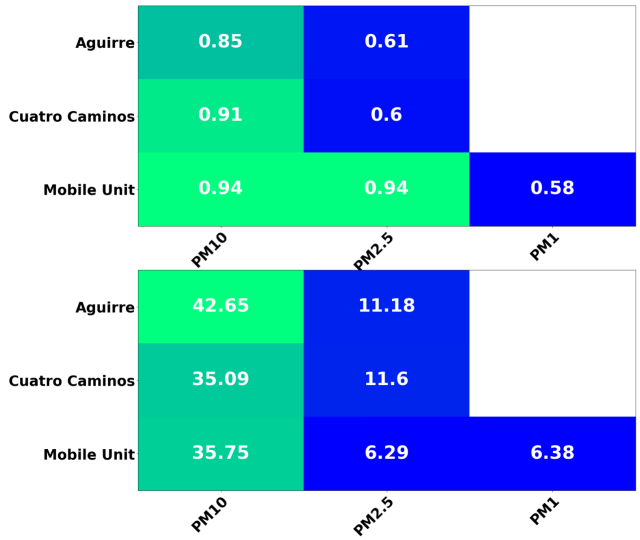



Fig. 13. Daily maximum validation of our device against air quality reference station. The metrics used are Pearson coefficient and MAE. The 53 daily samples were taken.

in the other two stations were $r = 0.91$ for Cuatro Caminos and $r = 0.85$ for Escuelas Aguirre for PM10. On the other hand, the results for PM2.5 were $r = 0.60$ and $r = 0.61$ for Cuatro Caminos and Escuelas Aguirre, respectively.

Regarding MAE, the values were quite higher than daily and hourly means. For PM10 fraction, the highest value was for Escuelas Aguirre (42.65 $\mu$g/m$^3$). In addition, the PM2.5 fraction showed a higher error for Cuatro Caminos and Escuelas Aguirre (11.18 and 11.60 $\mu$g/m$^3$) in comparison with Unidad Móvil (6.29 $\mu$g/m$^3$). For PM1, the error observed was larger in comparison with PM2.5 (6.36 $\mu$g/m$^3$).

### F. Outliers Detection

The outliers analysis was done for the PM10 fraction and in the Cuatro Caminos device, because it was the noisiest device, and the PM10 fraction had a high variability and a large error due to the size of the particles belonging to this

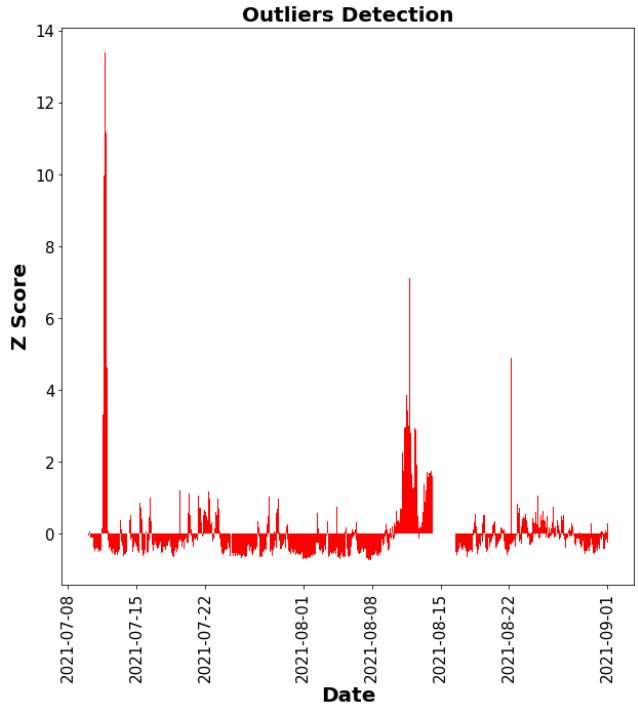



Fig. 14. Temporal distribution of $Z$-score for PM10 in Cuatro Caminos air quality station.

fraction. The first step was to compute the $Z$-score according to the definition exposed in statistical methods. The results are presented in Fig. 14. Therefore, we show the temporal evolution of the $Z$-score across time. Positive values mean the measurements were higher than expected, and negative values mean the values were lower than expected. In general, values over 3 and under $-3$ are usually considered outliers. Here, measurements with a high $Z$-score in the first week and some peaks in the second week of August are possible. The highest value was a peak close to $Z = 14$.

Once the outliers were identified, we applied a filter and an imputation method (mean imputation) to substitute the abnormal values. This step is critical, because a high $Z$-score could be an outlier or an event with extreme values—in the context of particulate, a dust Saharan intrusion may be an example. To analyze if these high $Z$-score values were outliers, we computed the Pearson coefficient after applying a filter of $z = 10$, $z = 3$, and $z = 1$ (Fig. 15). The initial correlation ($r = 0.86$) became lower when the filter was stricter ($r = 0.78$, $r = 0.53$, and $r = 0.45$ for $z = 10$, $z = 3$, and $z = 1$). In contrast, the filters could improve the initial error (13.9 $\mu$g/m$^3$) when using a narrow window for the $Z$-score (13.5, 12.5, and 10.4 $\mu$g/m$^3$).

## V. Discussion

Emerging hyperlocal IoT nanoparticle sensors are devices based on light scattering that allow to implementation of high-resolution networks and assess the effectiveness of climate change mitigation policies. This technology is currently limited by external factors, such as humidity. For this reason, in this article, we have proposed a new hyperlocal IoT system that helps in sample drying and removes the effect of humidity in the measurement.

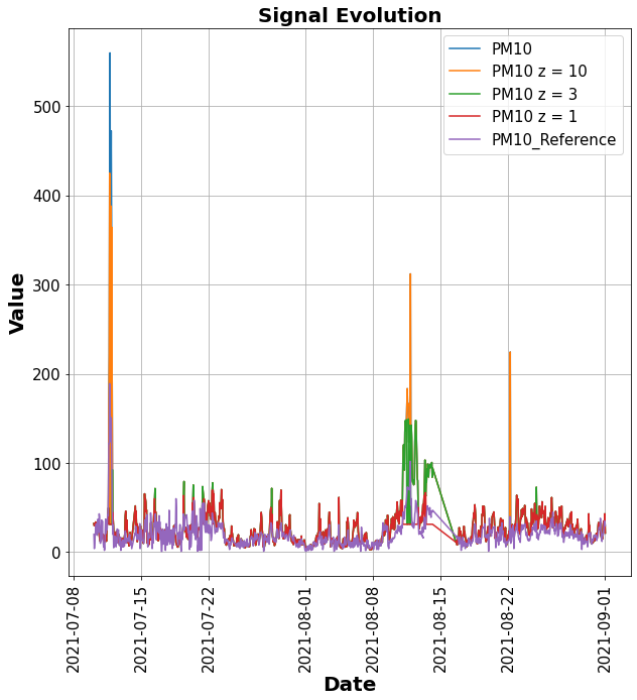



Fig. 15. Comparison between signals after applying a $Z$-score filter ($z = 10$, $z = 3$, and $z = 1$).

The first important point that can be extracted from the results presented is that a high variability exists between the two sensors integrated into the device. The $p$-values for the $t$-test showed a significant difference between the means in all fractions, and Fig. 8 also showed a significant difference between the probability distributions. In this sense, this variability could affect the results and should be considered. This may be due to load loss in different points of the circuit that cause asymmetries or differences in the sensor coming from the factory, especially related to the wear of the laser of the optical sensor. A different hypothesis could be that the variability of the sensors is high [22].

The proposed solution considerably improved the performance of the Alphasense OPC-N3 sensor in field conditions, showing high linearity with the reference equipment, as seen in Fig. 10. In general terms, our solution outperformed OPC-N3 according to the evaluation carried out by the South Coast AQMD. For PM10, we obtained $R^2$ ranging from 0.62 to 0.83 for hourly measurements, contrasting with $R^2 =$ 0.48–0.53 for the single sensor. A similar pattern occurred for the PM2.5 fraction, achieving values ranging from 0.34 to 0.82 compared with $R^2 = 0.61$–$0.69$. In this case, we achieved a greater score in the best scenario but a lower score in the worst one. However, the PM2.5 results were more coherent with the 0.38–0.67 proposed for OPC-N2 [29]. Finally, $R^2 =$ 0.50 for PM1, which was lower than the obtained in the validation done by the South Coast AQMD (0.78–0.82). It is important to notice that we could only validate PM1 against one station, so there is a lack of statistical evidence to compare with the proposed ranges. On the other hand, we achieved similar performances to those obtained in the state of the art for PM2.5 using machine learning approaches, with $R^2$ ranging from 0.78 to 0.83, compared with 0.78 and 0.88 [50].

Regarding the dryer's effectiveness, the analysis presented in Fig. 9 shows a significant reduction in the humidity levels after passing the columns. All internal humidity values collected were below the ammonium sulfate efflorescence point (35% RH) [41], which can be considered representative of urban environments [51]. That means that the effect of the hygroscopic growth is not significant on the measurement. However, to confirm this hypothesis, some limitations of our approach should be addressed. The first one is that the correlation between internal and ambient humidity is quite high, meaning that the dryer is reducing the humidity but not removing it. Thus, the humidity changes still affect the system, but with a lower impact. The second one is related to locations, as we have only tested our system in the city of Madrid. The weather conditions could be quite different in coastal cities, such as Valencia or Santander, more continental locations, such as Berlin, or Nordic places, such as Helsinki. It is true that the experiments' humidity and temperature ranges are wide—covering from 11% to 85% and from 17 °C to 42 °C. Still, it is necessary to test the system in different cities to confirm its effectiveness, because the humidity and temperature distributions are centered in dry and hot values.

Another important topic is how the temperature and the chemical composition affect the system and the measurements. The results show a low correlation between the temperature and the mass concentration signal, suggesting that there is no significant dependency between them. This is in line with (1), as temperature is not used for the computation of mass concentration. It could indeed affect density, but this effect is not significant in the study ranges. In addition, we demonstrate that our system is robust to extreme high temperature conditions, as it happens in Madrid during July. However, future tests should evaluate the system for extreme cold conditions. Finally, the effect of the chemical composition cannot be deduced from this study, as it is necessary to test the system in colocation tests in stations with different aerosol compositions. Urban environments indeed present similar compositions, with a $\kappa = 0.61$ [52]. Thus, to test this effect, comparing urban with rural locations is necessary.

Despite the linearity of sensor response not being assumed, it is the proper way to check the quality of the air quality system according to new standards, such as CEN/TS 17660 [53]. In an ideal scenario, the relation between the response of the sensor and the reference should be linear, with a slope = 1 and an intercept = 0. The slope stands for the bias error, whereas the intercept represents the nonsystematic error. Our results show that the slope is close to one, reducing the systematic error. There is indeed a considerable nonsystematic error. The same occurs with the MAE, which has high values in the three fractions. These facts show a high noise and unbiased error that should be addressed in the following iterations. The hypothesis is that we are dealing with two sensors in the devices, and the uncertainties related to each one are added. New designs could try to use only one device and check if the noise problems are reduced. In addition, more robustness checks are needed in more cities and different reference stations.

Regarding daily means, the devices were colocated to the reference station for 52 days. The results improved for all the fractions and all the locations in terms of correlations and, in consequence, in terms of $R^2$, as occurred when we compare against the reference technique. In contrast, as concluded in previous studies, the MAE does not get large when we apply daily mean measurements [40]. Compared with the evaluation of the South Coast AQMD, the solution improved the performance of the OPC-N3. For PM10, we obtained $R^2 = 0.90$–$0.94$ in contrast to the $R^2 = 0.22$–$0.26$ for the OPC-N3, and for PM2.5, we achieved $R^2 = 0.80$–$0.86$. Finally, for PM1, we reached $R^2 = 0.68$, lower than the values obtained by AQMD.

Finally, the results obtained by the daily maximum showed a good correlation with the reference, but the values of MAE were so high, especially for PM10, with values higher than 40 $\mu$g/m$^3$. This is because the signal is especially noisy in the higher values, so the computation is inaccurate. For this reason, it is necessary to analyze the effect of this noise and the presence of outliers. In this sense, we observed that, for the Cuatro Caminos station, there were some measurements with a $Z$-score over than $Z = 10$—close to $Z = 14$—that supposes a high deviation of the means. When we imputed these values with the mean, we obtained a reduction of the error, but also a reduction of the correlation. That could mean that these extreme peaks are associated with real events but, due to the intrinsic noise to the signal, are also amplified and obtain an extreme value and reduce the accuracy, especially for the daily maximum values. This increment in the noise linked to maximum values is related to optical methods, since the error in the measurement increases cubically when we compute mass concentration from number concentration, thus being higher for large values.

An important discussion point is the effect of the reference technology on the PM2.5 fraction. This study used FIDAS Palas as a reference instrument in the Unidad Móvil, based on light scattering, and TEOM in Cuatro Caminos and Escuelas Aguirre, based on oscillating microbalance. In this sense, we observed better results against optical methods than against TEOM, and this aspect should be studied. Other factors could cause the difference observed, but the similar results for PM10 suggest that this is not a problem for the stations.

However, the proposed system presents some limitations, which are detailed as follows: 1) the use of two sensors in the circuit could add error due to the intrasensor variability; 2) the error is still important when measuring higher concentrations; and 3) the proposed system does not take into account the effect of particle composition and size [13]. To better understand these issues, it is necessary to implement more exhaustive networks. Thus, the main challenge is to validate the system in more reference points across Europe, implementing a dense monitoring network covering all the factors affecting performance. The next steps will be to implement and perform this validation, also trying different approaches to optimize the intrasensor variability. In addition, some algorithmic approaches should be implemented to consider size and chemical composition. Thus, coupling optical particle methods with spectrometric techniques to assess chemical composition in real time can help in assessing particle composition impact on measurements [54].

The technology proposed in this article highlights several challenges and future research steps in applying hyperlocal devices. The first one is the evaluation of low emission zones (LEZs), because these devices allow monitoring the evolution of nanoparticles and their correlation with the policies taken by public administrations. In this sense, it is possible to compute key performance indicators (KPIs) in a function of the concentration of nanoparticles. Those indicators could be related to several measures implemented in some cities in state of the art. For instance, several studies highlight the importance of implementing electric buses as urban transport and correctly managing schedules and routes [55]. In this sense, our solution could help to optimize the parameters seeking to reduce nanoparticle emissions. On the other hand, different approaches look for the security of different users, such as cyclists [56]. In this sense, our solution could improve the recommendation of healthy routes and bike stations.

Thus, the quality of the hyperlocal devices allows us to go beyond and use this information as input for more complex models. Recent studies have shown that hyperlocal air quality sensors could identify the sources of the particles by coupling the sensor to a mobile vehicle [57]. Our solution meets the requirements to be linked to a mobile device well. Therefore, it could improve the accuracy in the state of the art. Other approaches deal with exhaustively quantifying the impact of different scenarios in terms of environmental benefits (i.e., reduction of $CO_2$ emissions, noise pollution, and traffic congestion) and quality of service for the users [58]. In this context, monitoring nanoparticles could be a key strategy, because it provides a wide variety of real-time data that could be used to feed these models. So, the research and improvement on hyperlocal air quality sensors will play an essential role in the solutions related to LEZs and impact estimations.

## VI. Conclusion

In this work, we have described and evaluated a nanoparticle measurement system based on two silicon columns, demonstrating that the drying of the air samples substantially improves the measurements made with respect to the OPC-N3 hyperlocal sensor. Moreover, the results obtained with respect to the reference measurements are quite good, with the values of $R^2 = 0.83$ for PM10 and PM2.5 in the best scenario for hourly averages and Pearson's coefficients above 0.80 for daily averages and maxima. In addition, we have shown that these devices can be integrated into an IoT architecture and interact with a real-time calibration service to improve data quality.

The results obtained open new research possibilities, which can be classified along two lines. On the one hand, improving the quality of the data in three different aspects: 1) design of the dryer system, seeking to optimize certain aspects, such as airflow pressure losses, dryer efficiency, or sensor wear; 2) eliminating intrinsic noise from the signal, by means of methods to reduce electronic noise or the application of moving averages or the elimination of outliers; and 3) using real-time calibration models. On the other hand, the generation

of high-quality hyperlocal data allows for progress in the following lines that have been limited so far in state of the art: 1) the generation of high-resolution air quality networks that allow estimating the impact of measures, such as LEZ; 2) the improvement and calibration of chemical transport models used to model the spatiotemporal distribution of pollutants; and 3) the calculation and estimation of the origins of pollution.

## REFERENCES


[1] R. N. Colvile, E. J. Hutchinson, J. S. Mindell, and R. F. Warren, "The transport sector as a source of air pollution," *Atmos. Environ.*, vol. 35, no. 9, pp. 1537–1565, Mar. 2001.

[2] A. J. Cohen et al., "Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the global burden of diseases study 2015," *Lancet*, vol. 389, pp. 1907–1918, May 2017.

[3] J. Lelieveld, C. Barlas, D. Giannadaki, and A. Pozzer, "Model calculated global, regional and megacity premature mortality due to air pollution," *Atmos. Chem. Phys.*, vol. 13, no. 14, pp. 7023–7037, Jul. 2013, doi: 10.5194/acp-13-7023-2013.

[4] B. Ostro, W.-Y. Feng, R. Broadwin, S. Green, and M. Lipsett, "The effects of components of fine particulate air pollution on mortality in California: Results from CALFINE," *Environ. Health Perspect.*, vol. 115, no. 1, pp. 13–19, Jan. 2007.

[5] P. Patel and S. G. Aggarwal, "On the techniques and standards of particulate matter sampling," *J. Air Waste Manage. Assoc.*, vol. 72, no. 8, pp. 791–814, Aug. 2022.

[6] F. Qamar, A. L. Pierce, and G. Dobler, "Covariance in policy diffusion: Evidence from the adoption of hyperlocal air quality monitoring programs by U.S. cities," *Cities*, vol. 138, Jul. 2023, Art. no. 104363, doi: 10.1016/j.cities.2023.104363.

[7] M. I. G. Daepp et al., "Eclipse: An end-to-end platform for low-cost, hyperlocal environmental sensing in cities," in *Proc. 21st ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, May 2022, pp. 28–40, doi: 10.1109/IPSN54338.2022.00010.

[8] I. A. T. Hashem et al., "Urban computing for sustainable smart cities: Recent advances, taxonomy, and open research challenges," *Sustainability*, vol. 15, no. 5, p. 3916, Feb. 2023, doi: 10.3390/su15053916.

[9] R. Sserunjogi et al., "Seeing the air in detail: Hyperlocal air quality dataset collected from spatially distributed AirQo network," *Data Brief*, vol. 44, Oct. 2022, Art. no. 108512, doi: 10.1016/j.dib.2022.108512.

[10] Y. Sun, H. Song, A. J. Jara, and R. Bie, "Internet of Things and big data analytics for smart and connected communities," *IEEE Access*, vol. 4, pp. 766–773, 2016.

[11] H. Kim, S. Tae, P. Zheng, G. Kang, and H. Lee, "Development of IoT-based particulate matter monitoring system for construction sites," *Int. J. Environ. Res. Public Health*, vol. 18, no. 21, p. 11510, Nov. 2021, doi: 10.3390/ijerph182111510.

[12] A. C. Rai et al., "End-user perspective of low-cost sensors for outdoor air pollution monitoring," *Sci. Total Environ.*, vols. 607–608, pp. 691–705, Dec. 2017.

[13] M.-L. Aix, S. Schmitz, and D. J. Bicout, "Calibration methodology of low-cost sensors for high-quality monitoring of fine particulate matter," *Sci. Total Environ.*, vol. 889, Sep. 2023, Art. no. 164063, doi: 10.1016/j.scitotenv.2023.164063.

[14] S. K. Jha et al., "Domain adaptation-based deep calibration of low-cost $PM_{2.5}$ sensors," *IEEE Sensors J.*, vol. 21, no. 22, pp. 25941–25949, Nov. 2021, doi: 10.1109/JSEN.2021.3118454.

[15] J. Tryner et al., "Laboratory evaluation of low-cost PurpleAir PM monitors and in-field correction using co-located portable filter samplers," *Atmos. Environ.*, vol. 220, Jan. 2020, Art. no. 117067, doi: 10.1016/j.atmosenv.2019.117067.

[16] P. Zieger, R. Fierz-Schmidhauser, E. Weingartner, and U. Baltensperger, "Effects of relative humidity on aerosol light scattering: Results from different European sites," *Atmos. Chem. Phys.*, vol. 13, no. 21, pp. 10609–10631, Nov. 2013.

[17] J. Li et al., "Effects of chemical compositions in fine particles and their identified sources on hygroscopic growth factor during dry season in urban Guangzhou of South China," *Sci. Total Environ.*, vol. 801, Dec. 2021, Art. no. 149749, doi: 10.1016/j.scitotenv.2021.149749.

[18] B. I. Magi, C. Cupini, J. Francis, M. Green, and C. Hauser, "Evaluation of $PM_{2.5}$ measured in an urban setting using a low-cost optical particle counter and a federal equivalent method beta attenuation monitor," *Aerosol Sci. Technol.*, vol. 54, no. 2, pp. 147–159, Feb. 2020, doi: 10.1080/02786826.2019.1619915.

[19] J. J. Caubel, R. Trojanowski, T. Butcher, and V. H. Rapp, "A review of regulatory standard test methods for residential wood heaters and recommendations for their advancement," *Renew. Sustain. Energy Rev.*, vol. 184, Sep. 2023, Art. no. 113501, doi: 10.1016/j.rser.2023.113501.

[20] P. Görner, X. Simon, D. Bémer, and G. Lidén, "Workplace aerosol mass concentration measurement using optical particle counters," *J. Environ. Monit.*, vol. 14, no. 2, pp. 420–428, 2012.

[21] N. Zikova, P. K. Hopke, and A. R. Ferro, "Evaluation of new low-cost particle monitors for $PM_{2.5}$ concentrations measurements," *J. Aerosol Sci.*, vol. 105, pp. 24–34, Mar. 2017, doi: 10.1016/j.jaerosci.2016.11.010.

[22] B. Feenstra et al., "Performance evaluation of twelve low-cost $PM_{2.5}$ sensors at an ambient air monitoring site," *Atmos. Environ.*, vol. 216, Nov. 2019, Art. no. 116946.

[23] S. Sousan, S. Regmi, and Y. M. Park, "Laboratory evaluation of low-cost optical particle counters for environmental and occupational exposures," *Sensors*, vol. 21, no. 12, p. 4146, Jun. 2021, doi: 10.3390/s21124146.

[24] Y. Wang, J. Li, H. Jing, Q. Zhang, J. Jiang, and P. Biswas, "Laboratory evaluation and calibration of three low-cost particle sensors for particulate matter measurement," *Aerosol Sci. Technol.*, vol. 49, no. 11, pp. 1063–1077, Nov. 2015, doi: 10.1080/02786826.2015.1100710.

[25] K. M. Krishna, Y. D. Borole, S. Sharma, S. P. Singh, R. S. Kumar, and A. Shrivastava, "Smart robotic system for remote health monitoring of the environment using IoT and global positioning system," in *Proc. Int. Conf. Technological Advancements Innov. (ICTAI)*, Tashkent, Uzbekistan, Nov. 2021, pp. 460–465, doi: 10.1109/ICTAI53825.2021.9673276.

[26] H. Nandanwar and A. Chauhan, "IoT based smart environment monitoring systems: A key to smart and clean urban living spaces," in *Proc. Asian Conf. Innov. Technol. (ASIANCON)*, Pune, India, Aug. 2021, pp. 1–9, doi: 10.1109/ASIANCON51346.2021.9544596.

[27] N. L. Datta, A. Tabassum, J. Bhaavani, K. Y. R. Kumar, R. V. H. Prasad, and A. R. Raja, "Design and development of dust detection and filtering system," in *Proc. 9th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Coimbatore, India, Mar. 2023, pp. 739–743, doi: 10.1109/ICACCS57279.2023.10112892.

[28] R. Jayaratne, X. Liu, P. Thai, M. Dunbabin, and L. Morawska, "The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog," *Atmos. Meas. Techn.*, vol. 11, no. 8, pp. 4883–4890, Aug. 2018.

[29] T. Zheng et al., "Field evaluation of low-cost particulate matter sensors in high- and low-concentration environments," *Atmos. Meas. Techn.*, vol. 11, no. 8, pp. 4823–4846, Aug. 2018.

[30] L. Bai et al., "Long-term field evaluation of low-cost particulate matter sensors in Nanjing," *Aerosol Air Quality Res.*, vol. 20, no. 2, pp. 242–253, 2020.

[31] G. N. Carvlin et al., "Development and field validation of a community-engaged particulate matter air quality monitoring network in Imperial, California, USA," *J. Air Waste Manage. Assoc.*, vol. 67, no. 12, pp. 1342–1352, Dec. 2017.

[32] K. K. Johnson, M. H. Bergin, A. G. Russell, and G. S. W. Hagler, "Field test of several low-cost particulate matter sensors in high and low concentration urban environments," *Aerosol Air Quality Res.*, vol. 18, no. 3, pp. 565–578, 2018.

[33] J. Sun, Z. Liu, K. Yang, and Y. Lu, "A miniature system for particulate matter (PM) measurement," in *Proc. IEEE SENSORS*, Nov. 2015, pp. 1–4.

[34] M. Carminati et al., "Capacitive detection of micrometric airborne particulate matter for solid-state personal air quality monitors," *Sens. Actuators A, Phys.*, vol. 219, pp. 80–87, Nov. 2014.

[35] A. Domínguez and M. N. Popescu, "A fresh view on phoresis and self-phoresis," *Current Opinion Colloid Interface Sci.*, vol. 61, Oct. 2022, Art. no. 101610, doi: 10.1016/j.cocis.2022.101610.

[36] C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun, and S. Duhr, "Protein-binding assays in biological liquids using microscale thermophoresis," *Nature Commun.*, vol. 1, no. 1, pp. 1–7, Oct. 2010.

[37] P. W. Oluwasanya, G. Rughoobur, and L. G. Occhipinti, "Design, modeling and simulation of a capacitive size-discriminating particulate matter sensor for personal air quality monitoring," *IEEE Sensors J.*, vol. 20, no. 4, pp. 1971–1979, Feb. 2020, doi: 10.1109/JSEN.2019.2950775.

[38] J. P. Specht et al., "AlN FBAR particle sensor with a thermophoretic sampling mechanism," *IEEE Sensors J.*, vol. 21, no. 17, pp. 19427–19435, Sep. 2021, doi: 10.1109/JSEN.2021.3086528.
[39] D. Zhang, "Ash fouling, deposition and slagging in ultra-supercritical coal power plants," in *Ultra-Supercritical Coal Power Plants*. Amsterdam, The Netherlands: Elsevier, 2013, pp. 133–183.
[40] D. A. Lundgren and D. W. Cooper, "Effect of humidify on light-scattering methods of measuring particle concentration," *J. Air Pollut. Control Assoc.*, vol. 19, no. 4, pp. 243–247, Apr. 1969, doi: 10.1080/00022470.1969.10466482.
[41] A. Di Antonio, O. Popoola, B. Ouyang, J. Saffell, and R. Jones, "Developing a relative humidity correction for low-cost sensors measuring ambient particulate matter," *Sensors*, vol. 18, no. 9, p. 2790, Aug. 2018.
[42] J. J. Schwab, O. Hogrefe, K. L. Demerjian, and J. L. Ambs, "Laboratory characterization of modified tapered element oscillating microbalance samplers," *J. Air Waste Manage. Assoc.*, vol. 54, no. 10, pp. 1254–1263, Oct. 2004.
[43] C.-H. Weng, G. Pillai, and S.-S. Li, "A $PM_{2.5}$ sensor module based on a TPoS MEMS oscillator and an aerosol impactor," *IEEE Sensors J.*, vol. 20, no. 24, pp. 14722–14731, Dec. 2020.
[44] D. E. Day, W. C. Malm, and S. M. Kreidenweis, "Aerosol light scattering measurements as a function of relative humidity," *J. Air Waste Manage. Assoc.*, vol. 50, no. 5, pp. 710–716, May 2000.
[45] D. S. Kang et al., "Development of drying systems for accurate measurement of particulate matter by means of optical particle measuring instruments," *Part. Aerosol Res.*, vol. 14, no. 4, pp. 191–203, 2018.
[46] O. Vernik, A. Degani, and B. Fishbain, "Mathematical estimation of particulate air pollution levels by aerosols tomography," *IEEE Sensors J.*, vol. 22, no. 9, pp. 8274–8281, May 2022.
[47] L. R. Crilley et al., "Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring," *Atmos. Meas. Techn.*, vol. 11, no. 2, pp. 709–720, Feb. 2018.
[48] H. Choi and Y. Koo, "Effectiveness of battery electric vehicle promotion on particulate matter emissions reduction," *Transp. Res. D, Transp. Environ.*, vol. 93, Apr. 2021, Art. no. 102758, doi: 10.1016/j.trd.2021.102758.
[49] *Continuous Device for Air Quality Measurement and Corresponding Air Quality Measurement Procedure*, document PCT/ES2022/070076, HOP Ubiquitous S.L, Murcia, Spain, Feb. 15, 2022.
[50] B. G. Loh and G. H. Choi, "Calibration of portable particulate matter–monitoring device using Web query and machine learning," *Saf. Health at Work*, vol. 10, no. 4, pp. 452–460, Dec. 2019.
[51] Y. Liu, Z. Yang, Y. Desyaterik, P. L. Gassman, H. Wang, and A. Laskin, "Hygroscopic behavior of substrate-deposited particles studied by micro-FT-IR spectroscopy and complementary methods of particle analysis," *Anal. Chem.*, vol. 80, no. 3, pp. 633–642, Feb. 2008, doi: 10.1021/ac701638r.
[52] B. Svenningsson et al., "Hygroscopic growth and critical supersaturations for mixed aerosol particles of inorganic and organic compounds of atmospheric relevance," *Atmos. Chem. Phys.*, vol. 6, no. 7, pp. 1937–1952, Jun. 2006, doi: 10.5194/acp-6-1937-2006.
[53] E. I. Fernández, N. B. Mulero, A. P. Pérez, J. M. M. Garcia, I. C. Martínez, and A. J. J. Valera, "CEN/TS 17660 in air quality systems for data quality validation and certification over smart spot air quality systems," in *Proc. Int. Conf. Ubiquitous Comput. Ambient Intell. (UCAmI)*, Nov. 2022, pp. 642–653, doi: https://doi.org/10.1007/978-3-031-21333-5_65.
[54] M. Allers et al., "Real-time remote detection of airborne chemical hazards—An unmanned aerial vehicle (UAV) carrying an ion mobility spectrometer," *IEEE Sensors J.*, vol. 23, no. 15, pp. 16562–16570, Aug. 2023, doi: 10.1109/JSEN.2023.3287448.
[55] R. Deng, Y. Liu, W. Chen, and H. Liang, "A survey on electric buses—Energy storage, power management, and charging scheduling," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 1, pp. 9–22, Jan. 2021.
[56] A. Herrmann, M. Liu, F. Pilla, and R. Shorten, "A new take on protecting cyclists in smart cities," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 12, pp. 3992–3999, Dec. 2018.
[57] X. Chen, A. Marjovi, J. Huang, and A. Martinoli, "Particle source localization with a low-cost robotic sensor system: Algorithmic design and performance evaluation," *IEEE Sensors J.*, vol. 20, no. 21, pp. 13074–13085, Nov. 2020.
[58] F. Bistaffa, C. Blum, J. Cerquides, A. Farinelli, and J. A. Rodríguez-Aguilar, "A computational approach to quantify the benefits of ridesharing for policy makers and travellers," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 1, pp. 119–130, Jan. 2021.

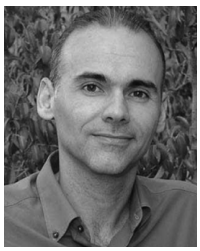
**Eduardo Illueca Fernández** was born in Cieza in 1996. He received the bachelor's degree in biochemistry and the M.Sc. degree in bioinformatics from the University of Murcia, Murcia, Spain, in 2018 and 2019, respectively.

He is the author of the Bioconductor package org.Mxanthus.db: Genome-wide annotation for *Myxococcus xanthus* DK 1622, and he has presented his work in several congress. His research interests focus on machine learning and artificial intelligence solutions to air quality monitoring, as well as air quality forecasting.

Mr. Illueca Fernández was awarded for his academic performance and the Research Initiation Fellowship from the University of Murcia from November 2019 to December 2019, participating in the project development of semantic bioinformatics tools for the bacterium *Myxococcus xanthus*. In 2020, he obtained the Fellowship from the Seneca Foundation to pursue an industrial Ph.D. at the University of Murcia and HOP Ubiquitous S. L.

**Iris Cuevas Martínez** received the bachelor's degree in chemical engineering from the University of Murcia, Murcia, Spain, in 2016, and the MBA degree in business analysis from UCAM, Murcia, in 2020.

She is the Laboratory Manager at Libelium Murcia, Ceutí, Spain. She is an expert in air quality, and she manages the Gases and Particles Laboratory in the process of ISO17025 certification as a National Calibration Laboratory, having several lines of research open in suspended particles (nanoparticles) and electrochemical gas sensors, and coordinates the Data Sciences Group that develops algorithms in the field of pollutant dispersion and automatic identification in air pollution sources and zoning studies.



**Jesualdo Tomás Fernández Breis** (Senior Member, IEEE) received the bachelor's degree in computer engineering and the Ph.D. degree in computer science from the University of Murcia, Murcia, Spain, in 1999 and 2003, respectively.

He is currently working as a Full Professor with the Faculty of Computer Science, University of Murcia. He has been leading research projects related to semantic web technologies since 2004. He is the co-founder of the spin-off longseq applications. He is also a member of the IMIB-Arrixaca Bio-Health Research Institute, Murcia. His current research interests include the application of semantic technologies for the development of learning health systems and the development of quality assurance methods for ontologies and terminologies.

**Antonio Jesús Jara Valera** (Senior Member, IEEE) received the Ph.D. (cum laude) degree from the University of Murcia (UMU), Murcia, Spain, in 2013, and the MBA degree from the ENAE Business School, UCAM, in 2012, where he did his entrepreneurship formation.

The Ph.D. results present a novel way to connect objects to Internet-enabled platforms in an easy, secure, and scalable way. He is the Director of the Research and Development Department, Libelium Murcia, Ceutí, Spain. As part of Libelium, he focuses on the smart cities market with solutions for citizens engagements and environmental monitoring (air quality sensors). He has also participated in over 100 international events about the IoT as a speaker, and he holds over 100 international publications/articles (5000 citations and H-index = 37). He holds several patents in the IoT domain, and he has also advised in the IoT domain to companies, such as Microsoft and Fujitsu.

Dr. Jara Valera received Entrepreneurship Awards from the ENAE (sponsored by SabadellCAM Financial Services) and emprende Go (sponsored by Spanish Government), and the IPSO Alliance Award (Sponsored by Google) for his disruptive innovation in the IoT. He was selected and mentored by the acceleration program FIWARE. He is the Chair of data quality and the Internet of Things (IoT) in IEEE.