Improved Fine Particles Monitoring in Smart Cities by means of Advanced Data Concentrator

Paolo Castello, *Member, IEEE*, Carlo Muscas, *Senior Member, IEEE,* Paolo Attilio Pegoraro, *Senior Member, IEEE*, Sara Sulis, *Senior Member, IEEE*

***Abstract*— Traffic reduction and air-quality improvement are among the main goals of several projects worldwide. This paper presents a fine particle monitoring based on heterogeneous air quality mobile sensors and an advanced data concentrator, AdDC, so that the level of pollution in the urban area, where few accurate fixed measurement stations are present, can be assessed with better accuracy. Some urban buses are used to carry low-cost sensors, thus implementing a mobile sensor network and increasing the time and space resolution of air quality information. The data obtained by these low-cost sensors is significantly affected by uncertainties, also due to atmospheric factors, such as humidity. The proposed AdDC processes all the obtained measurements and exploits the information obtained by the accurate fixed stations to improve the accuracy of the low-cost mobile sensors. In particular, a new compensation methodology, specifically targeted to the fine particles monitoring, is proposed. The monitoring of relative humidity is added, with the relevant on-the-fly calibration, so that the measured values can be used to correct the effects of humidity on $PM_{2.5}$ sensors. The validity of the proposed system is proved by means of simulations performed on an appropriate set-up.**

Index terms— Air pollution sensors, Air quality, Data processing, Public transportation, Measurement uncertainty, Smart city, Fine particles monitoring

I. INTRODUCTION

The diffusion of information and communication technology has driven various smart city projects in different areas as public transportation, vehicular traffic management, and smart building, in order to improve quality of life of residents. These projects often mainly aim at decreasing the traffic level and, as a consequence, the air pollution, especially in urban regions. Air pollution emerged in high-density urban areas as a result of industrial growth and has become one of the most important problems in modern cities [1], [2]. The levels of sulfur dioxide and carbon monoxide are usually kept within the quality standard by suitably controlling emissions (see, for example, [3] referring to the city of Tokyo). Nevertheless, nitrogen dioxide (NO_2) and ozone (O_3) often exceed the required quality levels in major cities around the world [4]. In addition, fine outdoor particulate matter ($PM_{2.5}$) is considered one of the risk factors with the highest impact on health in the world [5]-[7].

Generally, specialized staff monitor air pollution, through fixed monitoring stations equipped with certified reference instruments, with a high cost per device [8]. The high cost of this kind of instrumentation does not allow a high spatial density in the monitoring process to be obtained, since only a few locations characterized by a high-density population or by the presence of sensitive sites (e.g., factories, high-density highways, etc.) can be equipped with these monitoring stations. Owing to this limited spatial resolution, combined with the geographical and climatological conditions, the description of pollutants in the cities can be not well representative of the actual situation [3]. Thus, the idea of using mobile stations to help air pollution monitoring is spreading [1], [4], [9]. Currently, several projects worldwide for air quality monitoring are based on different low-cost sensors, which can be either found in the form of a commercial device for specific gas detection or integrated into an array that includes various sensors, [10]-[14]. The behaviour of these low-cost sensors is affected by chemical interference and environmental conditions. Moreover, in the same platform, the performance can vary significantly from sensor to sensor [8]. For this reason, it is a key point to take into account appropriately the uncertainties of all the quantities to monitor and to verify the performance of the sensors [15].

This paper, starting from the study presented in [16], suggests an efficient solution for fine particulate matter monitoring, based on an improved monitoring architecture that exploits the availability of buses and vehicles in the urban area. The centre of the proposal is an advanced data concentrator (AdDC), which collects information from several players and then shares refined measured data with different city authorities. In particular, the concentrator is necessary to provide a trustworthy correlation between the measures collected on the field, which are appropriately processed, and the data provided by the operators of the public utilities.

The starting point of the proposed procedure is the data concentrator (DC) presented in [16], capable of managing a variable reporting rate from the low-cost remote sensors and applying an enhanced *on-the-fly calibration* based on [17]. The DC compensates the errors of the sensors installed in the buses with respect to the fixed stations. Moreover, a multi-hop compensation facilitates the error compensation process of sensors which are never in the proximity of a reference air quality station. PM_{10} sensors were considered as examples in [16].

The improved fine particle monitoring here proposed is based on an AdDC that presents new functionalities and applies suitable uncertainty models. The AdDC collects heterogeneous information, such as the position of the buses and the measured values of pollutants and environmental parameters, along with the time-tag of the all the available measurements, and uses all this information to improve the overall knowledge of the air quality in the city.

As an example of the possible refining process, considering that the uncertainty of the low-cost sensors is very often significantly affected by atmospheric factors, in this paper the AdDC considers and compensates the impact of the relative humidity (RH) on the $PM_{2.5}$ measurements. The accuracy of $PM_{2.5}$ mobile sensors is improved by also exploiting the low-cost RH measurements of the mobile stations. These RH measurements are in turn refined with an

The authors are with the Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, Italy. Contact email: paolo.castello@unica.it.

 DOI:10.1109/TIM.2020.2977744 Publisher version: https://ieeexplore.ieee.org/document/9229105

appropriate multi-hop process relying on the high-accuracy data obtained from a fixed monitoring station.

The impact of this procedure on the performance and trustworthiness of the overall air quality monitoring system is evaluated through simulations and discussed.

The paper is organized as follows: Section II describes the distributed sensor network and the proposed uncertainty models; Section III defines the functionalities of the proposed AdDC; Section IV describes in detail the compensation process; Section V shows and discusses test results, followed by the conclusions.

II. Distributed Sensor Network

The distributed sensor network considered in this research activity is part of an industrial research initiative developed with the primary goals of decreasing road travel time and improving air quality in urban areas. The core of the project is the application of the netcentric paradigm through both fixed and mobile nodes enabling the sensor integration of the distributed equipment and the transformation of public transport in "mobile platforms" to monitor the quantities of interest, through the continuous acquisition of data. The project is based on two main types of sensor networks:

1. A network of fixed stations for the real-time monitoring of the number of vehicles entering and exiting the city, to improve the efficiency of the traffic information signals and allow optimizing the flux of cars during the rush hours.
2. A distributed sensor network composed of one or more fixed stations and a mobile sensor network installed in the public transportation system (buses, car-sharing vehicles) for the acquisition of air quality measurements.

The public transportation system is directly related to the mobility of citizens living in the city area. The city buses have been chosen for the distributed sensor network so that an extended area outside the range of the fixed air quality stations can be covered. Commonly, bus routes cover the city territory quite extensively. This system allows creating a widespread and pervasive sensor network for the monitoring of different pollution sources. The main idea is to improve the time and space monitoring resolution compared to conventional monitoring systems based on fixed stations.

Two types of mobile communication systems have been taken into account: the mobile system based on LTE and UMTS technologies, and a low-power wide-area network (LPWAN) with a multi-hop architecture [18]. For the purpose of this paper, a monodirectional communication from the mobile stations to the concentrator is assumed because the compensation process is performed at the concentrator level.

In the following, the main characteristics of the low-cost monitoring stations installed in the buses of the public transportation system will be described.

A. Low-cost sensors

The low-cost air quality sensors are identified as emerging measuring devices for "indicative measurements" regulated in the Air Quality Directive [19]. They can evaluate several quantities, but their accuracy is often affected by atmospheric factors. Focusing on PM measurements, gas sensors would allow air pollution monitoring at a lower cost than reference measurements [17]. Nevertheless, $PM_{2.5}$ measurements are significantly affected by RH. In this regard, it is worth noticing that different factors can influence the distribution of urban humidity during the day, for example, the reduced evapotranspiration due to limited vegetation or the emission of water vapour from industrial sources [20]. RH values could also be affected by mixing influences of surface roughness, moisture sources and thermal fields [21]. Furthermore, the RH changes inside the urban area from downtown to suburban areas, and its value also depends on the trees in a specific area [20]. These considerations drive to the need of evaluating suitably the RH, and its effects on the measures of the polluting quantities, not only at the fixed stations but all over the city by means of the distributed sensor network associated with the buses.

For the scope of the project, three low-cost commercial sensors for PM measurements have been considered. They are shown in Table I, where the metrological characteristics, as provided by their datasheets, are reported. All these sensors are based on the laser scattering principle.

TABLE I. Metrological characteristics of low-cost sensors for PM2.5 measurements.

Parameter	Sensors		
	Nova Fitness SDS 021	**Plantower PMS 7003**	**Alphasense OPC-N2**
Range [μg/m³]	0.0 ÷ 999.9	0 ÷ 500	0.1 ÷ 1500
Minimum Resolution [μm]	< 0.3	0.3 @ 1 PPM 1.0 @ 2.5 PPM 2.5 @ 10 PPM	0.38 ÷ 17
Temperature Range [°C]	-10 ÷ 50	-10 ÷ 60	-20 ÷ 50
Relative Humidity Range [%]	0 ÷ 70	0 ÷ 99	0 ÷ 95
Accuracy	Max ±15 % and ±10 μg/m³ [25°, 50 % RH]	±10 μg/m³ [$PM_{2.5}$, range: 0 ÷ 100 μg/m³]	See [22]
Digital Output	UART (TTL) @3.3V	Serial (TTL) @3.3V	SPI or USB protocol
Cost [$]	30	20	450

Digital output with different protocols characterizes all three sensors. In particular, the SDS021 sends the data through the universal asynchronous receiver-transmitter, UART, with serial communication protocol. The PMS 7003 is characterized by a serial communication and, finally, the OPC-N2 can be interfaced with the Serial Peripheral Interface (SPI) or USB protocol.

Considering the reporting rate, these sensors have quite high data reporting rates with respect to other types of sensors, because they can provide one measurement per second. In particular, with the PMS7003 it is possible to choose a fast mode with reporting intervals in the range 200 - 800 ms, depending on the concentration of PM. Generally, the sensors can send the obtained measurements through serial communication and can be easily interfaced with a device as a low-cost single board computer to collect and transmit the acquired data.

The typical power consumption is in the order of a few watts, also considering the fan used to push the air on the detecting area. This characteristic makes these sensors suitable for mobile applications where the power supply could be an issue. It is worth recalling that the low-cost sensors based on scattering principles differ from higher cost sensors for the heating system required to reduce the uncertainty introduced by the hygroscopic effect due to the humidity conditions [22]. In order to take into account this unwanted phenomenon, the sensor OPC-N2 is equipped with onboard temperature and humidity sensors, which leads to the highest cost shown in Table I.

The RH range for the SDS 021 is the lowest. The accuracy reported by the manufacturer for SDS 021 and PMS 7003 considers different operating condition. For SDS 021, the accuracy is indicated for specific temperature and relative humidity values (25° and 50% of RH, respectively). For PMS 7003, the accuracy refers to a range of concentration of $PM_{2.5}$. The datasheet of OPC-N2 does not report clearly the accuracy values; thus, the information can be obtained from [22], where the level of accuracy is obtained in different experimental conditions of temperature and relative humidity.

It is worth recalling that these sensors have been studied in the scientific literature to evaluate their performance in realistic scenarios (see, for instance [23]-[24]). In these studies, variations of the actual performance with respect to the expected one are described.

Among the three options analyzed in the project and summarized in Table I, the sensor SDS 021 has been considered in this study. Therefore, the performance of this device, in terms of possible outputs, has been reproduced in the simulations, based on both its nominal specifications and the variations described in [23]-[24], in order to build a more realistic uncertainty model for the measurements provided by this sensor.

B. *Mobile and fixed stations*

The mobile stations acquire measurements of pollutants and relative humidity, and include a GPS receiver for correlating these data with the time of the measurements, position, and speed of the buses (see Table II) [16]. The time synchronization accuracy provided by commercial GPS receivers with respect to the coordinated universal time (UTC), generally in the order of microseconds, is largely sufficient, for the aim of this project, to correctly time-align all the measurements provided by the mobile sensors and collected in the concentrator.

TABLE II. INFORMATION PROVIDED BY THE MOBILE STATIONS.

Code	Values
Sensor ID	Integer number to recognize the station
Time	UTC-time
Position	GPS coordinates
Speed	Instantaneous speed provided by the GPS receiver
Measurements	CO, PM_{10}, $PM_{2.5}$, Temperature, Humidity,…

An ID identifies each mobile station. Measurements are sent in real-time using the mobile communication channel. Thus, the AdDC receives the data from different stations with low latency. In the proposed architecture, the buses do not talk to each other. This is a technical choice that allows reducing the requirements of the mobile stations. The compensation process is entirely handled by the AdDC by exploiting the geolocalization of the measurements.

The GPS receiver is also necessary for triggering the desired reporting rate of the data. To increase the quality of the information provided by the distributed measurement system, the data reporting rate of the sensors can change in proximity of a fixed monitoring station. The coordinates of the fixed station are pre-stored in the mobile monitoring station, so that, in the area in proximity, the mobile station increases the measurement rate autonomously. By suitably exploiting this feature, the accuracy of the overall process can be improved as described in the following section.

The fixed air pollution monitoring station features higher accuracy compared to the low-cost sensors [8], and is an essential element in the *on-the-fly calibration* applied in this paper. The data from this high-accuracy station are downloaded by the AdDC from a remote web server and are considered as a reference for all the measurements obtained by the low-cost sensors.

In the considered scenario, the station is close to the route of a bus and between two bus stops. This position allows correlating the data collected by the bus at specific instants with those of the fixed station.

III. ADVANCED DATA CONCENTRATOR FUNCTIONALITIES

Fig. 1 depicts the base elements of the monitoring system, with the AdDC, two air-quality stations (1 and 2) installed on the buses, and the fixed high-quality monitoring station. The advanced concentrator can implement different types of policies, and, in particular, the information enhancement strategy described in the following. The description focuses on $PM_{2.5}$ concentration measurements.

A. *On-the-fly calibration based on variable reporting rate*

The *on-the-fly calibration* is a technique designed to increase the quality of the information provided by low-cost gas sensors. In this paper, the technique is applied by the AdDC preferably when the bus is close to the high-quality air monitoring fixed station and, if necessary, when the bus is close to a previously calibrated bus. Let us consider Bus 1 in Fig. 2, whose route passes near the fixed station. When the bus is in a given proximity area of the fixed station, the reporting rate *RR* of its measurements increases. For instance, between t_3 and t_4 the *RR* of Bus 1 can be changed from one measurement every ten seconds to one measurement per second (mps). It is worth noticing that the number of

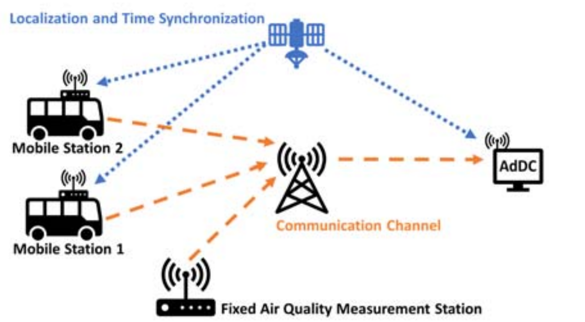

Fig. 1. Base elements of the monitoring system.

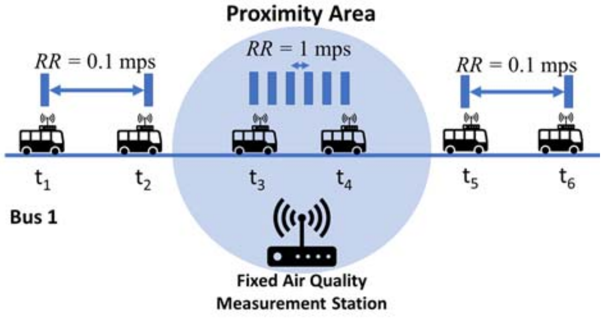

Fig. 2. Variable reporting rate in the proximity area of the fixed measurement station.

measurements sent by the mobile station in this stage is not set a priori but depends on the time spent to cross the proximity area and can change due to the traffic condition or packet loss in the communication network [16].

The AdDC acquires all the available measurements with the corresponding positions and allows the *on-the-fly calibration* when the bus is inside the proximity area [17]. The average difference between bus measurements and the corresponding measurements gauged in the reference station is evaluated to compensate for the systematic error affecting the sensor in the mobile station. Once the measurements of this bus are considered as refined, the “compensated” bus can act as the available master towards other mobile stations, on buses with routes far from the fixed station, in a “second hop” *on-the-fly calibration* (see Bus 2 in Fig. 3).

The low-cost sensors on the buses can be indeed programmed to change the reporting rate also in memorized common bus stops (e.g. at the final stop of the programmed routes), so that the number of comparable measurements increases. Again, this operation does not require a specific command from the AdDC or a two-way communication with it. The proposed AdDC is then able to check when two buses are close to each other so that, for instance, Bus 1 can serve as a “reference” for compensating the measurements provided by Bus 2 in a multi-hop process.

B. $PM_{2.5}$ compensation considering RH influence

To refine the quality of the particulate matter monitoring, it is necessary to consider the specificity of the monitored area, characterized by different green fields, ponds, high-density commercial and industrial sites. In this scenario, the impact of the relative humidity on the accuracy of the particular matter estimations, especially for smaller sizes, can be critical. Proposal of correction factor has been presented in the literature to compensate this effect [22].

In this paper, a compensation procedure taking into account residual uncertainty sources has been applied. Together with different types of data and particulate matter measurements, the AdDC receives the relative humidity values measured in the buses and uses these data, along with the information concerning their uncertainty, to evaluate a correction factor $C(RH)$ that will be applied to improve the accuracy of $PM_{2.5}$ estimation.

Equation (1) shows how the correction factor $C(RH)$ depends on the RH value through a statistically derived value *k*, which falls in the range 0.38 ÷ 0.4 [22].

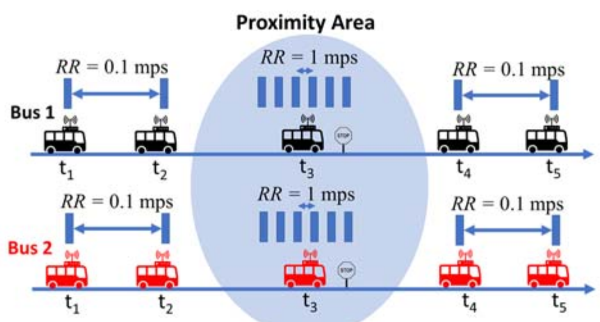

Fig. 3. Variable reporting rate in the proximity area (bus stop) between Bus 1 and Bus 2.

$$C(RH) = 1 + \frac{\frac{k}{1.65}}{-1 + \frac{1}{RH \cdot 100}} \quad (1)$$

Fig. 4 shows the values of $C(RH)$ as a function of the RH for *k* = 0.4, which is the value considered in this paper.

Following this approach, the correct value of particulate matter can be obtained dividing the measurement uncompensated by the correction factor $C(RH)$.

In this paper, a refined procedure is applied to define the correction factor for the mobile station measurements, using the multi-hop process described in the following section. The AdDC computes such factors for all the buses in the network and continuously updates them.

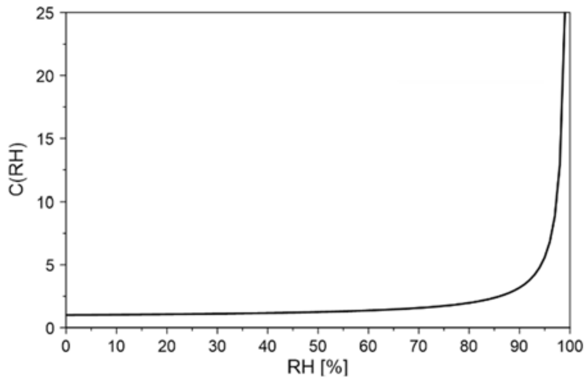

Fig. 4. Correction factor as a function of the relative humidity.

IV. Compensation Process

The compensation process could be summarized with three different stages.

1. In the initial stage (*waiting for the proximity*), the mobile station is still far from the proximity area and its measurements cannot be compensated due to the distance.
2. In the second stage (*proximity area*), the mobile station is close to the reference station (either the fixed station or a previously “compensated” bus, according to the multi-hop compensation process described above) until the bus exits the proximity area. In this stage, the reporting rate rises and the AdDC evaluates the compensation factor.
3. In the third stage (*running*), the mobile station is outside the proximity area, but the AdDC continues to compensate for the systematic error calculated in the previous stage.

In the following, each stage is described in detail for the scenario that includes the process between mobile and fixed station.

1. Waiting for the proximity

The first stage is related to the waiting time until the first rendezvous with the reference station. By assuming that the uncertainty associated with the measurements obtained in the reference station is negligible, so that the “reference” values provided by that station correspond to the “true” values of the measurands, the model to describe the behaviour of the mobile station is represented by equations (2) and (3):

$$RH_i = RH_i^r + es_{RH} + er_{RH_i} \quad (2)$$

$$PM_i = PM_{2.5i}^r \cdot C(RH_i^r) + es_{PM} + er_{PM_i} \quad (3)$$

In (2), the *i-th* measurement of relative humidity, RH_i, differs from the corresponding reference value, RH_i^r, due to a systematic error, es_{RH}, and a random error, er_{RH_i}.

In (3), the *i-th* measurement of particulate matter PM_i differs from the reference value $PM_{2.5i}^r$, due to the effect of humidity and again to the presence of a bias (es_{PM}) and a random error (er_{PM_i}).

These errors summarize, in a general way, the behaviour of the mobile sensors as described by manufacturers (see Table I).

2. *Proximity area*

The AdDC can evaluate the impact of the systematic error by processing the measurements obtained from the mobile stations while travelling in the proximity area. The stage starts when the AdDC is able to evaluate the differences among the measurements (RH_j and PM_j) and the reference values (RH^r and $PM_{2.5}^r$) obtained from the fixed station:

$$\Delta RH_j = RH_j - RH^r = es_{RH} + er_{RH_j} \quad (4)$$

$$\Delta PM_j = PM_j - PM_{2.5}^r \cdot C(RH^r) = es_{PM} + er_{PM_j} \quad (5)$$

where j indicates the measurements obtained in the proximity area (for the sake of simplicity it will be considered $j = i_1, \ldots, i_1 + M - 1$). $PM_{2.5}^r$ and RH^r, due to the short time spent (commonly less than one minute) to cross the proximity area, can be considered constant and for this reason their subscripts are dropped in the equations. The AdDC obtains the compensation values (k_{RH}, k_{PM}) by averaging these differences over the M measurements taken during the rendezvous:

$$k_{RH} = \frac{1}{M}\sum_{j=i_1}^{i_1+M-1} \Delta RH_j = es_{RH} + \frac{1}{M}\sum_{j=i_1}^{i_1+M-1} er_{RH_j} \quad (6)$$

$$k_{PM} = \frac{1}{M}\sum_{j=i_1}^{i_1+M-1} \Delta PM_j = es_{PM} + \frac{1}{M}\sum_{j=i_1}^{i_1+M-1} er_{PM_j} \quad (7)$$

From (6) and (7) it emerges clearly that the random contributions represent possible residual errors after compensation processes. Thus, the possibility of increasing the reporting rate, as described in Section III.A, allows processing more values, so that this residual error is reduced, as it will be shown in Section V.B.

3. *Running*

During the running stage, the AdDC defines the values $\widehat{RH}_i$ and $\widehat{PM}_i$ ($i \geq i_1 + M$), representative of the actual values of humidity and the particular matter obtained from the mobile station, by compensating the values received from the remote sensors with the coefficients obtained during the last engagement with the fixed station, as follows:

$$\widehat{RH}_i = RH_i - k_{RH} \quad (8)$$

$$\widehat{PM}_i = \frac{PM_i - k_{PM}}{C(\widehat{RH}_i)} \quad (9)$$

The same procedure can be applied in the second hop, the *multi-hop,* between two mobile stations. In this stage, the AdDC uses the best available data to compensate the measurements obtained from the buses with routes far from the fixed station. In this case, the model must be properly adapted to take into account the lower accuracy of the mobile "compensated" reference value. In particular the above equations (4)-(9) still hold replacing the reference values with the compensated ones. For instance, equation (4) becomes:

$$\Delta RH_{2j} = RH_{2j} - \widehat{RH}_{1j} \quad (10)$$

where subscripts 1 and 2 have been used to distinguish data from the two involved buses, see Fig. 3.

V. Tests and Results

A. *Scenario*

To validate the proposed solution, the analyses and results are summarised in two parts. First, the multi-hop process with the variable reporting rate is applied in order to study the impact of the number of measurements in the proximity areas on the estimation of the quantity of interest. Then, the accuracy of the $PM_{2.5}$ concentration measurements is evaluated considering the whole compensation process carried out also taking into account the relative humidity measurements.

The two parts share the same test architecture and simulated scenario which, according to the description in Section III, considers two buses, each one equipped with a mobile monitoring station, with different routes. The analysis covers 24 hours, which means at least 8640 measurements/day for each station, the actual number depending on the reporting rates. In this period, Bus 1 passes 24 times close to the fixed station (see Fig. 2), and 24 crossings are also considered between bus 1 and bus 2 (taking place at the common bus stops, as in Fig. 3).

For the sake of simplicity, the uncertainty introduced by the high-accuracy fixed station is neglected, as in [8] and [17], and the geographical concentration profiles of $PM_{2.5}$ and RH are considered constant (there is no loss of generality since the proposed algorithm perfectly fits also a variable scenario). This means that both values have been assumed to be constant over the entire bus route when a specific time instant is considered. In order to test the methodology with possible realistic concentrations of pollutants, the reference values of $PM_{2.5}$ and RH are obtained by interpolation of 24-hourly values, following a typical profile, measured by an air pollution monitoring station in Italy during an autumn day and reported in Fig. 5.

The sensor SDS 021 has been assumed in this study and its metrological behavior, modeled according to Section II.A, has been reproduced in the simulations. Due to the characteristics of the sensor, the method presented in Section IV has been applied directly on the digital output values.

To properly consider the effects of the random errors in the compensation process, the results are presented in terms of root mean square error (RMSE) evaluated over 1000 repetitions of the 24-hours scenario shown in Fig. 5.

B. *Multi-hop process*

First, preliminary tests are considered to define the most suitable number of repeated measurements M to be used in each compensation stage of the multi-hop process. In this

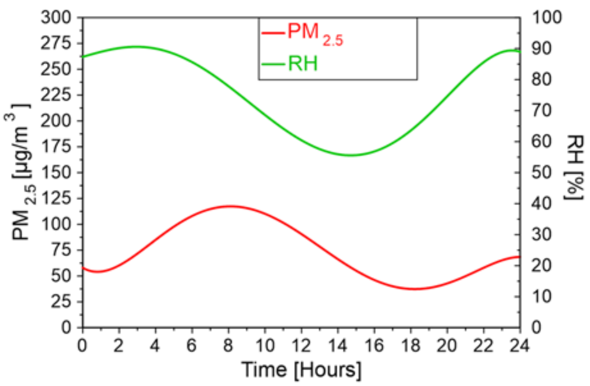

Fig. 5. Reference values of $PM_{2.5}$ and RH for 24 hours.

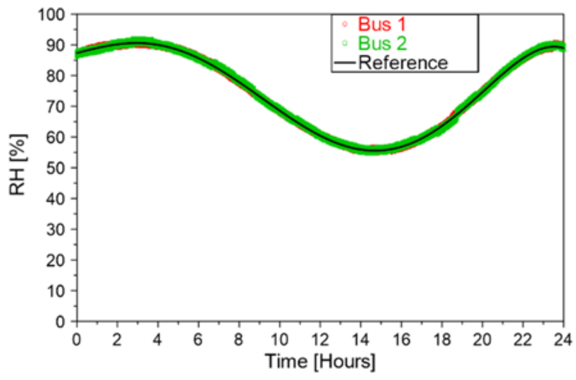

Fig. 6. Reference and estimated values of RH for 24 hours considering the case $(M_1, M_2) = (10{,}10)$.

scenario, the influence of RH is assumed negligible ($C(RH) = 1$).

To evaluate different engagement scenarios taking into consideration different durations, different numbers of $PM_{2.5}$ concentration measurements are considered during the engagement intervals between bus 1 and the fixed station (M_1) and between bus 1 and bus 2 (M_2). In particular, cases from (M_1, M_2) = (1,1) up to (M_1, M_2) = (30,10) are considered.

The following accuracy values are assumed for both mobile stations in the simulations, performed in LabVIEW environment: constant errors es_{PM} equal to 10 µg/m^3 and uniform random errors er_{PM} in the range ± 10 µg/m^3.

Table III reports the results obtained by the two mobile stations in terms of RMSE. The first result, with only one value measured in the proximity areas, i.e. (M_1, M_2) = (1,1), could be considered as a comparison baseline without variable reporting rate. When a higher number of measurements is considered during the engagement among the stations, the RMSE values for the estimated concentrations from buses 1 and 2 clearly decrease. Due to the limited time in the proximity area, the higher reporting rate allows collecting more measured values, so that the random effects can be reduced and a better estimation of the offset error of the low-cost sensors is achieved, compared to what could be obtained with a lower reporting rate. A further rise in the reporting rate, from 10 to 30 measurements per second, has a more limited impact on the overall quality of the compensation.

TABLE III. RMSE OF ESTIMATED VALUES OF $PM_{2.5}$ WITH DIFFERENT NUMBER OF MEASUREMENTS.

	$PM_{2.5}$ RMSE [µg/m³]					
(M_1, M_2)	(1, 1)	(2,2)	(5, 5)	(10, 5)	(10, 10)	(30, 10)
Bus 1	8.14	7.07	6.29	5.98	5.98	5.63
Bus 2	11.56	9.14	7.31	7.07	6.56	6.40

Table IV reports the RMSEs obtained with a different value of random error: er_{PM} equal to ± 15 µg/m^3. As expected, the RMSE values are larger but, again, the same decrease pattern can be identified. This is not surprising, since it is strictly related to the averaging effect on random errors and its propagation in the multi-hop scenario. A number of concentration measurements higher than 10, as in (M_1, M_2) = (30,10), only leads to a slight decrease of the errors and thus will not be further investigated in the following.

TABLE IV. RMSE OF ESTIMATED VALUES OF $PM_{2.5}$ - FIXED NUMBER OF CONCENTRATION MEASUREMENTS WITH er_{PM} UP TO 15 MG/M^3.

	$PM_{2.5}$ RMSE [µg/m³]					
(M_1, M_2)	(1, 1)	(2, 2)	(5, 5)	(10, 5)	(10, 10)	(30, 10)
Bus 1	12.25	10.57	9.43	8.96	8.96	8.41
Bus 2	17.34	13.66	10.92	10.58	9.84	9.63

C. Estimation of RH Values

As previously discussed, knowing the RH is crucial to increase the overall quality of the monitoring. For this reason, measurements of RH values from the mobile stations are here considered to implement the complete compensation process. The multi-hop method with variable reporting rate is thus here applied to the RH estimation. In the following, er_{RH} within the range ± 1 % has been considered as a reasonable and typical value that can be found in the datasheet of low-cost sensors for RH evaluation [25], and constant errors es_{RH} equal to 4 % have been assumed for both buses. The fixed high-quality station is assumed to be able to measure high accuracy reference values also for RH.

Table V reports the RMSEs of the relative RH estimation achieved, starting from the measurements given by the humidity sensors on the buses. Considering the multi-hop technique with a variable reporting rate, the RMSE decreases promptly, and Fig. 6 shows the results of the multi-hop process evaluated for 24 hours and applied to the RH measurements considering the case $(M_1, M_2) = (10{,}10)$.

TABLE V. RMSE OF ESTIMATED VALUES OF RH %.

	RH RMSE [%]	
(M_1, M_2)	(1, 1)	(10, 10)
Bus 1	0.81	0.60
Bus 2	1.15	0.66

D. Estimation of PM2.5 with RH evaluation

The final part of the validation process refers to the influence of the RH measurement and its uncertainty on the evaluation of the $PM_{2.5}$.

Fig. 7 shows the reference value of particulate matter with ($PM_{2.5}$) and without ($PM_{2.5}$ (RH)) exact RH compensation. The figure gives an idea of the impact of the humidity on $PM_{2.5}$, when the estimation does not take into account the high levels of RH and does not consider any sort of compensation. In this case, the obtained concentrations are clearly useless for any monitoring and, in particular, for possible limit verification activity.

Table VI shows instead the RMSE of $PM_{2.5}$ estimation when only a multi-hop compensation is used, without considering the impact of RH in the compensation process.

TABLE VI. RMSE OF ESTIMATED VALUES OF $PM_{2.5}$ WITHOUT CONTINUOUS RH COMPENSATION

	$PM_{2.5}$ RMSE [µg/m³]	
(M_1, M_2)	**(1, 1)**	**(10, 10)**
Bus 1	16.62	16.37
Bus 2	21.96	20.37

Constant errors es_{PM} equal to 10 µg/m³ and a uniformly distributed random error er_{PM} = ± 10 µg/m³ for both mobile stations are assumed in the simulations (as in Table III). The mobile stations follow the uncompensated values of $PM_{2.5}$ influenced by RH. Before the first rendezvous, this effect could be modelled as a bias in the measurements: this preliminary stage is not considered in the RMSE values shown in Table VI. With the multi-hop compensation, during the engagement, the bias between the reference $PM_{2.5}$ and the influenced measurements of $PM_{2.5}$(RH) obtained from bus 1 and 2 is evaluated and removed, as shown in Fig. 8 for the case $(M_1, M_2) = (10, 10)$. However, if RH is not measured in the mobile stations, the influence of humidity outside the rendezvous area cannot be fully compensated, and the measurements of $PM_{2.5}$ are affected by the variability of RH. It is important to recall that, due to the variability of humidity during the day, but also, more in general, to the passage of the buses within different humidity areas depending on the geographical position, the measurements cannot be compensated only with the RH obtained from the reference stations.

Then, other tests have been carried out by continuously compensating the effects of RH on the measurements of $PM_{2.5}$ from the mobile stations.

Table VII shows the RMSE of $PM_{2.5}$ estimations when the whole compensation process described in Section IV is

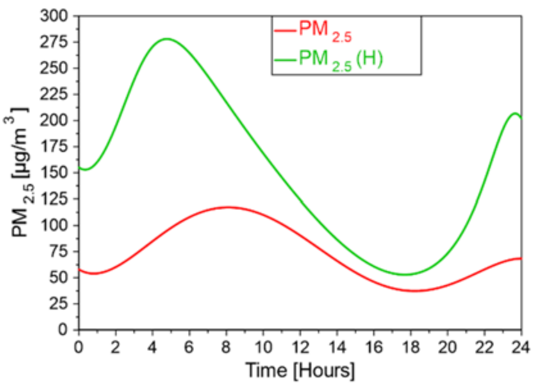

Fig. 7. Reference values of PM with ($PM_{2.5}$, red line) and without ($PM_{2.5}$(RH), green line) compensating the influence of RH during 24 hours.

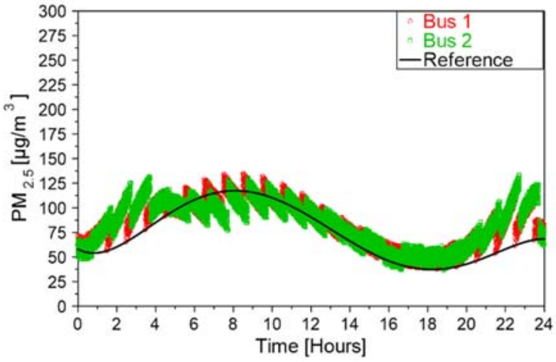

Fig. 8. Reference and measured values of $PM_{2.5}$ considering the case $(M_1, M_2) = (10,10)$ and without applying the compensation of RH during 24 hours.

applied. Different numbers of measurements in the proximity areas have been considered, and the random error in RH sensors is assumed be in the range ± 1 %, while constant errors equal to 4 % are considered (as in Table VI).

TABLE VII. RMSE OF ESTIMATED VALUES OF $PM_{2.5}$ CONSIDERING RH COMPENSATION

	$PM_{2.5}$ RMSE [µg/m³]	
(M_1, M_2)	**(1, 1)**	**(10, 10)**
Bus 1	5.35	3.92
Bus 2	7.58	4.32

Fig. 9 shows the trends during 24 hours of the obtained refined $PM_{2.5}$ values in the case $(M_1, M_2) = (10, 10)$. As expected, the quality of information improves with respect to the case where only the impact of the systematic error was removed, without any specific RH compensation (Fig. 8) and the multi-hop procedure itself becomes more effective.

VI. CONCLUSIONS

This paper has presented an efficient solution for fine particulate matter monitoring. Mobile monitoring stations with low-cost air quality sensors installed in the public transportation system are used to increase the time and space resolution of air quality monitoring in an urban area. These mobile monitoring stations feature low accuracy and are affected by the variability of environmental conditions. To improve the accuracy of the measurement process, a fine particle monitoring based on an advanced data concentrator

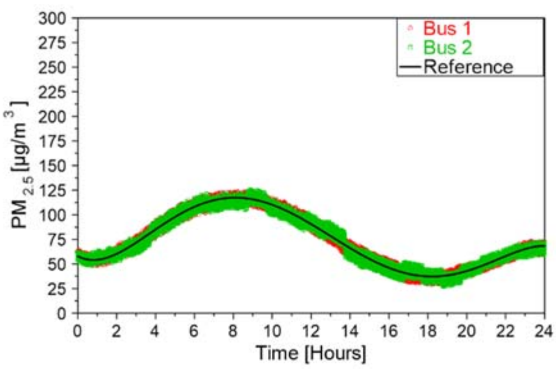

Fig. 9. Reference and measured values of $PM_{2.5}$ (case $(M_1, M_2) =$ (10,10)) considering the compensation of RH during 24 hours.

has been presented. The advanced concentrator, exploiting the geolocalization of the mobile distributed stations, enables the *on-the-fly calibration* among the mobile monitoring stations and the high-quality fixed station.

Moreover, the relative humidity is measured by the stations on the buses and its effects on the behavior of the mobile sensors are properly taken into account, so that the overall quality of the $PM_{2.5}$ monitoring is further improved.

Acknowledgement

The research activities described in this paper have been conducted within the R&D project "Cagliari2020" partially funded by the Italian University and Research Ministry (grant# MIUR_PON04a2_00381).

References

[1] P. W. Oluwasanya, A. Alzahrani, V. Kumar, Y. A. Samad, and L. G. Occhipinti, "Portable multi-sensor air quality monitoring platform for personal exposure studies," *IEEE Instrum. Meas. Mag.*, vol. 22, no. 5, pp. 36–44, Oct. 2019.

[2] M. Carratù, M. Ferro, A. Pietrosanto, P. Sommella, and V. Paciello, "A Smart Wireless Sensor Network For PM10 Measurement," in *2019 IEEE International Symposium on Measurements Networking (M&N)*, Jul. 2019.

[3] W. Tsujita, A. Yoshino, H. Ishida, and T. Moriizumi, "Gas sensor network for air-pollution monitoring," *Sens. Actuators B Chem.*, vol. 110, no. 2, pp. 304–311, Oct. 2005.

[4] L. Spinelle, M. Gerboles, M. G. Villani, M. Aleixandre, and F. Bonavitacola, "Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide," *Sens. Actuators B Chem.*, vol. 215, pp. 249–257, Aug. 2015.

[5] G. Yue, K. Gu, and J. Qiao, "Effective and Efficient Photo-Based PM2.5 Concentration Estimation," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 3962–3971, Oct. 2019.

[6] T. Savu, B. A. Jugravu, and D. Dunea, "On the Development of a PM2.5 Monitoring Network for Real-time Measurements in Urban Environments," *Rev. Chim.*, vol. 68, no. 4, pp. 796–801, May 2017.

[7] W. Hernandez, A. Mendez, R. Zalakeviciute, and A. M. Diaz-Marquez, "Analysis of the information obtained from PM2.5 concentration measurements in an urban park," *IEEE Trans. Instrum. Meas.*, 2020.

[8] N. Castell *et al.*, "Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?," *Environ. Int.*, vol. 99, pp. 293–302, Feb. 2017.

[9] D. Hasenfratz, O. Saukh, S. Sturzenegger, and L. Thiele, "Participatory air pollution monitoring using smartphones". *Mobile Sensing*, 2012.

[10] F. Lamonaca, M. Vasile, and A. Nastro, "Non-invasive monitoring of the relation between building materials and indoor radon pollution," *IET Sci. Meas. Amp Technol.*, vol. 8, no. 3, pp. 150–154, Feb. 2014.

[11] R. Lu, R. P. Turco, and M. Z. Jacobson, "An integrated air pollution modeling system for urban and regional scales: 2. Simulations for SCAQS 1987," *J. Geophys. Res. Atmospheres*, vol. 102, no. D5, pp. 6081–6098, 1997.

[12] L. Morawska *et al.*, "Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?," *Environ. Int.*, vol. 116, pp. 286–299, Jul. 2018.

[13] 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 Qual. Res.*, vol. 18, no. 3, pp. 565–578, 2018.

[14] J. Dutta, F. Gazi, S. Roy, and C. Chowdhury, "AirSense: Opportunistic crowd-sensing based air quality monitoring system for smart city," in *2016 IEEE SENSORS*, Oct. 2016.

[15] A. Carullo, S. Corbellini and S. Grassini, "A Remotely Controlled Calibrator for Chemical Pollutant Measuring-Units," in *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 4, pp. 1212-1218, Aug. 2007.

[16] P. Castello, C. Muscas, P. A. Pegoraro, and S. Sulis, "A Proposal for a Data Concentrator for Smart City Applications," in *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, May 2019.

[17] D. Hasenfratz, O. Saukh, and L. Thiele, "On-the-Fly Calibration of Low-Cost Gas Sensors," in *Wireless Sensor Networks*, Berlin, Heidelberg, 2012.

[18] M. Anedda, C. Desogus, M. Murroni, D. D. Giusto, and G.-M. Muntean, "An Energy-efficient Solution for Multi-Hop Communications in Low Power Wide Area Networks," in *2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Jun. 2018.

[19] Directive 2008/50/ec of the European Parliament and of the Council of 21 may 2008 on ambient air quality and cleaner air for Europe.. Available [Online, accessed on September 2020]:

https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32008L0050

[20] M. Unkašević, O. Jovanović, and T. Popović, "Urban-suburban/rural vapour pressure and relative humidity differences at fixed hours over the area of Belgrade city," *Theor. Appl. Climatol.*, vol. 68, no. 1, pp. 67–73, Mar. 2001.

[21] A. Wypych, "Twentieth century variability of surface humidity as the climate change indicator in Kraków (Southern Poland)," *Theor. Appl. Climatol.*, vol. 101, no. 3, pp. 475–482, Aug. 2010.

[22] L. R. Crilley *et al.*, "Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring," *Atmospheric Meas. Tech.*, pp. 709–720, Feb. 2018.

[23] 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," *Atmospheric Meas. Tech.*, vol. 11, no. 8, pp. 4883–4890, Aug. 2018.

[24] A. Di Antonio, O. A. M. Popoola, B. Ouyang, J. Saffell, and R. L. Jones, "Developing a Relative Humidity Correction for Low-Cost Sensors Measuring Ambient Particulate Matter," *Sensors*, vol. 18, no. 9, Art. no. 9, Sep. 2018.

[25] Temperature and Humidity Module DHT11 Product Manual [Online]. Available:https://components101.com/sites/default/files/component_datasheet/DHT11-Temperature-Sensor.pdf, Accessed on: Jun. 24, 2020