

DYNAMIC ASSESSMENT OF PERSONAL EXPOSURE TO AIR POLLUTION FOR EVERYONE: A SMARTPHONE-BASED APPROACH

V. Casella^{1, *}, M. Franzini¹, R. Bellazzi², C. Larizza², D. Pala³

¹ Dept. of Civil Engineering and Architecture, University of Pavia, Italy - (vittorio.casella, marica.franzini)@unipv.it

² Dept. of Electrical, Computer and Biomedical Engineering, University of Pavia, Italy - (riccardo.bellazzi, cristiana.larizza)@unipv.it

³ Dept. of Electrical, Computer and Biomedical Engineering, University of Pavia, Italy – daniele.pala02@universitadipavia.it

Commission IV, WG IV/10

KEY WORDS: Personal exposure, smart city, exposomics, smartphone

ABSTRACT:

In Epidemiology, exposure assessment is the process of measuring or estimating the intensity of human exposures to an environmental agent such as air pollution. Healthcare agencies typically take into consideration yearly averaged pollution values and apply them to all citizens, in risk models. However distinct parts of cities can have significantly different levels of pollution and individual habits can influence exposure, too. Consequently, in epidemiology and public health, there is an increasing interest for personal exposure assessment, i.e. the capability of measuring the exposure of individuals. Within the EU H2020 PULSE project, an innovative mechanism for the individual and dynamic assessment of exposure to air pollution has been implemented. The present paper illustrates its technological and scientific components. The system has already been deployed to several pilot cities of the project and Pavia, Italy, has been the first one. In that city several hundreds of tracks have already been acquired and processed. Therefore, the paper thoroughly illustrates the assessment procedure with examples.

1. INTRODUCTION

1.1 Motivation

In Epidemiology, exposure assessment is the process of measuring or estimating the intensity of human exposures to an environmental agent. In the case of air pollution, namely exposure to particulate matter, yearly averaged values are typically considered. More precisely, healthcare agencies typically take into consideration yearly averaged pollution values and apply them to all citizens, in risk models.

The WHO (World Health Organization) globally monitors air quality and offers an online map (WHO website) of yearly-averaged pollution levels. As an example, the city of Pavia is reported to have had an average value of $45 \mu\text{g}/\text{m}^3$ of PM 10 (Particulate Matter 10) for the year 2015 (Figure 1). Data are not up-to-date and are aggregated at the level of the city and averaged on a one-year basis. So, there is no distinction between the seasons and the parts of the city are considered as a whole. This could be reasonable for Pavia, Northern Italy, being a small-size city (63 km^2 and 70000 inhabitants) but is not justified for larger ones. However distinct parts of cities can have significantly different levels of pollution and individual habits can influence exposure, too (Moon, 2001).

In epidemiology and public health, there is an increasing interest for personal exposure assessment, i.e. the capability of measuring the exposure of individuals. Personal exposure assessment is when exposure is evaluated for individuals: people living in one part of a city could have a lower exposure than others. People having different habits could also have different exposure values. Following most recent scientific papers (Steinle et al., 2016; Sanchez et al., 2019), personal exposure can be performed by equipping small groups of voluntaries with dedicated devices. The volunteers will carry the monitors for a time and then results will be summarized. Such implementations only allow to monitor

a limited number of people, as devices cost, and for a limited time, as devices typically are bulky and have memory and battery limitations.

Moreover, most of the available personal monitors can only perform integral monitoring (Sanchez et al., 2019): they can quantify the total amount of pollution the carrier inhaled in the considered time frame. In other words, they will supply information such as: in the previous 24h the carrier inhaled a certain amount of PM 10. They can't support the user in understanding where, when and what quantity of pollution he ingested.

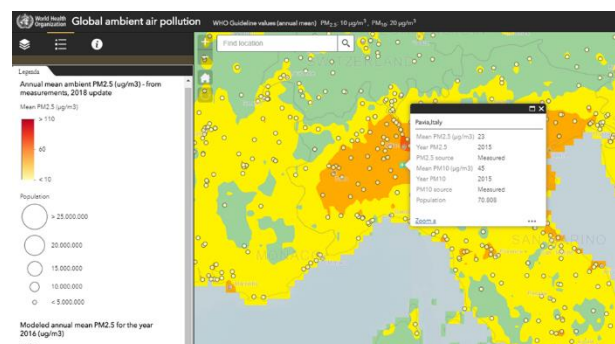


Figure 1. The WHO online map on air pollution and people's exposure for Pavia, Northern Italy

In summary, a limited number of people can perform, for a limited time, integral monitoring. On the contrary, it would be recommendable that all the citizens can perform lifelong exposure assessment in a differential or dynamic way, i.e. having the capability of distinguishing the contribution given by single activities or narrow time spans to the daily exposure amount.

* Corresponding author

With this approach the citizens could organize their life in order to minimize the inlet of pollution: they could appropriately choose the park to go jogging and the time.

Indeed, an advanced personal exposure functionality should have the following distinctive characteristics:

- to be open: being available to all the citizens and not requiring any special device;
- to be dynamic: having the capability to assess the instantaneous inhaled pollution and thus being able to sum up for any time frame chosen by the user (one minute or one hour) and for any time;
- being continuously available: it is performed routinely rather than for a special period when special arrangements are performed;
- being upgradable at no-cost for the user: in the case that more advanced methodologies become available for air quality monitoring, changes will be performed in the monitoring stations and in the processing equipment, supposed to be owned by the municipalities; users will only have to upgrade the App running in their mobile phone.

1.2 The PULSE project

Since the beginning of the 21st century, an important increase in the percentage of population living in urban areas has been observed (Department of Economic and Social Affairs, 2014), and this trend is projected to increase at least up to the year 2050. Cities are economic drivers for countries, but also the best labs for innovation aiming at managing public health challenges deriving from demographic and epidemiological transitions (Innovating Cities, Environment, Research and Innovation website).

Big cities are complex and heterogeneous environments in which socioeconomic and environmental conditions can vary considerably within small distances, and it has been widely observed that the technological progress and a progressive change in lifestyle are leading to increasing challenges for public health operators, deriving from increasing prevalence of respiratory and cardiovascular diseases (Anandan et al., 2010; Guarnieri and Balmes, 2014). In this context, the PULSE project (Participatory Urban Living for Sustainable Environments) has been funded by the European Commission under the Horizon 2020 framework to undertake research and innovation in cities in Europe, the United States and Asia. The project started in late 2016 and partners with municipality leaders of seven major cities - Paris, Singapore, Birmingham, Barcelona, Pavia, Keelung and New York - to collect information from the public health system, remote and fixed environmental sensors, and citizen-operated mobile devices, to develop a system for the management of public health policies in the urban environment. It focuses mainly on two pathologies, typical in the urban environments: asthma, usually linked to air pollution, and type 2 diabetes, linked to lifestyle and physical inactivity. Within PULSE, a multi-technological integrated platform that connects citizens and Public Health Authorities has been constructed. The background idea is that health risk is the consequence of a complex combination of exposures and human behaviour, and big data technology can be used to predict and mitigate public health problems, analysing the situation in the cities with a high level of spatial and temporal granularity in order to take proper actions quickly, assisting citizens personally and promoting healthy habits and well-being. PULSE features several state-of-the-art technologies: through a personal App and after signing a consent form, the users can send their own data and GPS/GNSS positions to a backend system that contains Big Data analytics and risk models, and they receive feedbacks containing personalized advice in return in order to apply the best behaviour to reduce

their risk of asthma and diabetes depending on their personal parameters and the environment they live in.

Furthermore, an innovative WebGIS allows to visualize, explore and analyse environmental, social and health data with a geographic description and a dashboard featuring several analysis and simulation tools allows the public health authorities to inspect aggregated data and quickly organize proper interventions in the areas of the city that mostly need them.

2. THE EXPERIMENTAL TESTSITE

Within the PULSE project, an advanced solution to perform dynamic personal exposure assessment has been deployed.

One of its main technological items is a dense network of low-cost sensors. The city of Pavia started deploying its own network on September 2018, constituted by Purple Air sensors (Purple Air website). The sensors record numerous parameters related to both air quality and environmental factors such as PM 1.0, PM 2.5, PM 10, temperature, humidity and pressure. In addition, thanks to the presence of a Wi-Fi transmitter, they provide measured data in real time which can be downloaded for subsequent analysis and viewed via a proprietary WebGIS interface (Figure 3). The sensors are installed on the balconies of some volunteers, specifically recruited, and on some structure belonging to the University of Pavia (Figure 2), or to the Municipality. Now, 48 devices are installed (Figure 3) but those used to perform personal exposure assessment are 37, because some of them are installed indoor for calibration purposes and other projects.



Figure 2. Example of a sensor installation at the University of Pavia

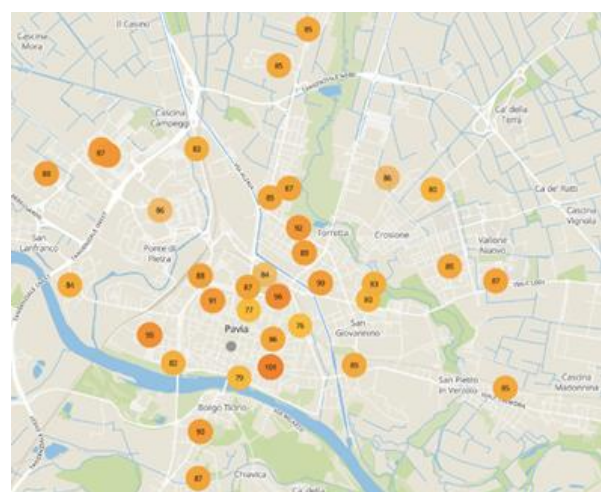


Figure 3. Maps of the sensors deployed in Pavia

An example of data acquired by the sensors is shown in Figure 4 in which the behaviour of PM 1.0, PM 2.5 and PM 10 for one sensor and for the first week of the year 2019 is reported. The separation between the curves is well understandable as it is, by

definition, $PM_{10} \geq PM_{2.5} \geq PM_{1.0}$. Indeed $PM_{1.0}$ quantifies the mass of the particle matter having size $\leq 1 \mu m$ being contained in a cubic meter of air. $PM_{2.5}$ quantifies the mass of particles having size $\leq 2.5 \mu m$ so $PM_{2.5}$ includes $PM_{1.0}$ plus the mass of particles having size between 1.0 and 2.5 microns. The same holds for PM_{10} .

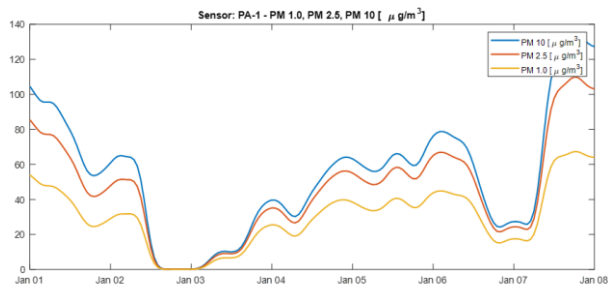


Figure 4. Joint graphics of $PM_{1.0}$, $PM_{2.5}$ and PM_{10} for sensor PA-1 and one week

The capability of surveying both air quality and climate variants gives the possibility to analyse the pollution phenomenon in a combined approach evaluating interrelations and cause-effect connections. Figure 5 reports, on the left y-axis, humidity and pressure while, on the right y-axis, the hourly averaged wind speed and its maximum value. Wind data shown are collected on an hourly basis. As the wind is characterized by gusts, average values might report the increase of wind or not. Maximum values are even more interesting for some analysis, represented by the red continuous line. It comes out that on January 2nd, 2019, the wind suddenly increased, and PM suddenly decreased. Figure 5 clearly demonstrates a connection between climatic variables (wind in our case) and air pollution and confirms that, when the wind arises and is strong, pollution almost disappears. Wind data are not recorded by the stations deployed by PULSE but were obtained from the repository of the local environmental agency (ARPA Lombardy).

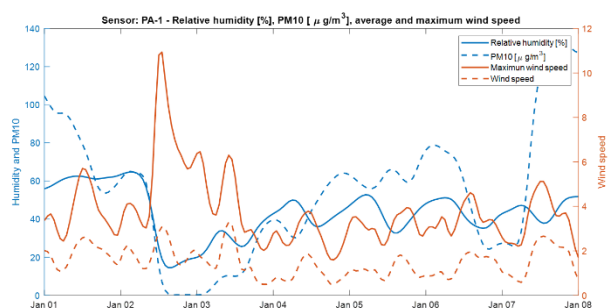


Figure 5. Joint graphic of relative humidity and PM_{10} in blue, referred to the left y-axis and average and maximum wind speed, in red and referred to the right y-axis. Data are related to one sensor and the first week of the year 2019

Since September 2019 the citizens of Pavia are constantly informed about the real time city's air quality, thanks to the Pulse@PV App. It is a free App purposely developed to show the most recent measurements of each active sensor of the network in Pavia. The data shown by the App are updated every 10 minutes. Figure 6 contains some screenshots of Pulse@PV. Its functionalities are:

- Display of the PM_{10} ($\mu g/m^3$) measured in real-time by the nearest sensor in the home page (Figure 6a). From this page the user can select one of the three main sections of the App: Sensors List, Map and About Pulse@PV;

- Display of the list of the sensors active in the last 30 minutes, sorted by distance from the current position. Each sensor is associated to a coloured icon, indicating the PM_{10} level (Figure 6b);
- Display of the latest measurements of $PM_{1.0}$, $PM_{2.5}$, PM_{10} , temperature, humidity and pressure of each sensor (Figure 6c), coloured according the US EPA colour codes for each pollutant (Figure 6d shows the PM_{10} colour codes);
- Display of the map of the sensors network in the city. Each sensor is coloured according to the PM_{10} level (Figure 6e);
- Display of different sections that report information about the PULSE project, the App, air pollution, the different types of particulates and the danger of being exposed to different levels of pollution.

Currently, 150 Android users and 126 IOS users have downloaded the App.

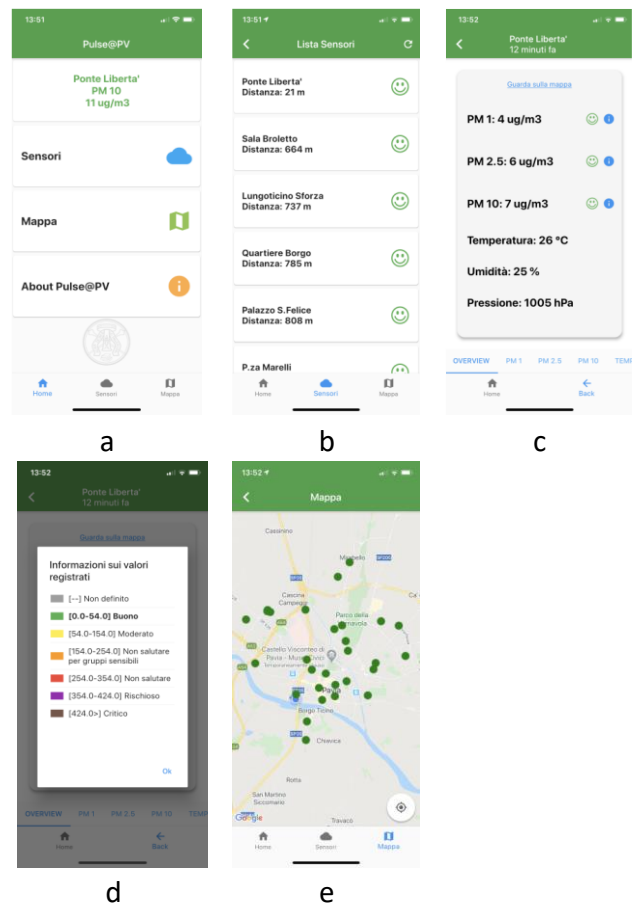


Figure 6. Screenshots of the Pulse@PV App. a. home page; b. list of all the active sensors; c. measurements provided by a selected sensor; d. colour codes for PM_{10} ; e. map of the active sensors.

3. METHODOLOGY

Personal exposure assessment implemented in PULSE needs some technological layers: (1) a methodology to determine dense air pollution models; (2) a methodology to track movements of people; (3) a breathing model, including the number of breaths per second and the volume of the inhaled air. In PULSE a dense network of small sensors is chosen as described in Section 2, smartphones equipped with inner GPS/GNSS sensor are used for trajectories survey and a basic model was adopted, which is described Section 3.4.

The section will illustrate the methodology used to construct the dense air pollution models and to calculate personal exposure.

3.1 Generation of dense and dynamic air quality maps

The construction of dense and dynamic air quality maps is mandatory for personal exposure assessment. By dense, we mean that the considered variants (PM 10, temperature, etc.) are known almost on each location; dynamic means instead that such dense maps must be known almost at each time.

Indeed, air pollution has a strong variability in time and a limited but significant variability on space (on the scale of the city of Pavia; space variability can be huge at the regional or national scale). Figure 7 reports the PM 10 time series for the three locations (shown in Figure 9) related to the first 20 days of January 2020. The figure highlights significant variations against time and position, thus highlighting the need of dense monitoring, in space and time. The two very low peaks are due to the wind. Figure 8 is related to one day, thus highlighting strong intra-day variability.

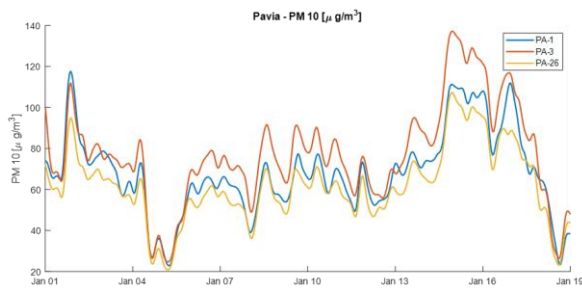


Figure 7. PM 10 for the first 20 days of January 2020 for the sensors PA-1, PA-3, PA-26

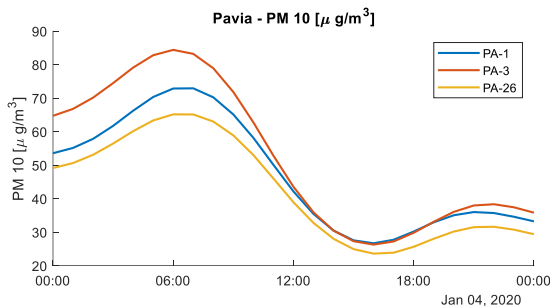


Figure 8. PM 10 for the three mentioned stations and for January 4th, 2019; graphic highlights strong intra-day variability

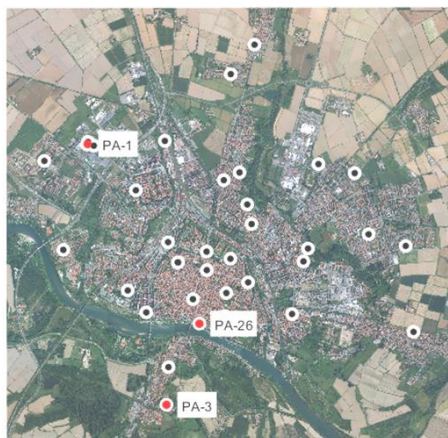


Figure 9. Map of the 37 outdoor sensors; the three ones selected are shown in red; their name is also displayed

It was decided to show only three sensors out of 37 regularly running in order not to clutter the figures: their position and name are shown in Figure 9.

Starting from dense but discrete measurements, continuous maps have been calculated by adopting the gaussian kernel interpolation (Section 3.2). Indeed, raster are generated having a spacing of 100 m in space and 1 hour in time: see Figure 10 in which red dots represent where air pollution is measured (the 37 Purple Air sensors), and black dots represent where it is estimated. Also, we underline that the developed methodology is capable to estimate the pollution level at any location and, the actual problem solved involves, together with spatial location, time: our methodology can estimate pollution level at any location and at any time.

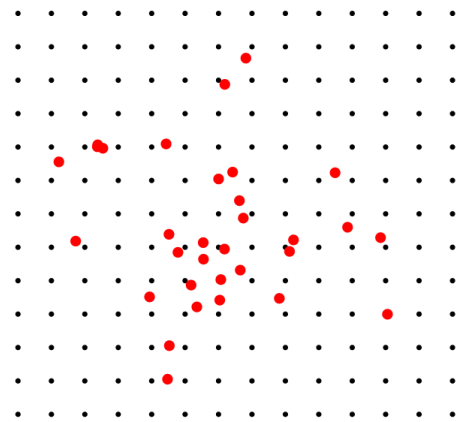


Figure 10. Interpolation of air quality data; red dots: where pollution is measured; black dots: where it is estimated

3.2 Gaussian kernel interpolation

Gaussian kernel interpolation (Wilson and Nickisch, 2015) is based on weighted average: the unknown pollution value z located at the location (x, y, z, t) , a point belonging to the 4D space, as time is added to the three spatial coordinates, is calculated by:

$$z = \frac{\sum_j^m \sum_{i=1}^n z_{ij} w_i^s w_j^t}{\sum_j^m \sum_{i=1}^n w_i^s w_j^t} \quad (1)$$

where z_{ij} is the pollution level measured by the i -th monitor at the j -th epoch; indeed a certain number of epochs (times when a measurement is acquired) are considered around the selected t time: in our case a time window having a semi-width of 2 hours was selected. m is the number of the considered epochs and n is the number of the monitors (red dots). As the formula highlights, the weight is the product of the factors w_i^s and w_j^t . Both weight functions are based on a gaussian kernel. The first one is related to space-distance:

$$w_i^s = e^{-\frac{(ds)_i^2}{2\sigma_s^2}} \quad (2)$$

where ds is the spatial distance between the estimation point and the location of the i -th monitor; the function decreases when ds increases; the σ_s parameter controls how quickly the weight decays. The second weight function is related to time-distance and has the form:

$$w_j^t = e^{-\frac{(dt)_j^2}{2\sigma_t^2}} \quad (3)$$

where dt is the time span between the time of the estimation point (x, y, z, t) and the time of the j -th measurement considered. See

Figures 11 and 13 for an example of the raw measurements used to perform the interpolation and of the continuous surface obtained.

3.3 Tracking of users

Smartphone's onboard GPS/GNSS sensor is used to record users' trajectories. The PulsAir App, which was developed within PULSE, has the tracking capability, among other features. The App measures the phone position at regular intervals, typically 1 or 5 seconds, and produces sort of tabular data as shown in Figure 15, including date, time and 3D position, for each recorded dot.

3.4 Personal exposure calculation

Once tracks are recorded, they must be processed in order to assess exposure. Velocity is determined first of all, by forming the quotient between space and time variations. Suitable smoothing methodologies are applied in order to limit propagation of significant measurement errors, as the App use the basic navigation solution. Figure 17 reports velocities for the considered track.

Next step is to associate each dot with local pollution i.e. the concentration of the considered pollutant at that location and at that time. This is performed by space-time interpolation techniques previously illustrated. Results are shown in Figure 18. Finally, a link must be established between the motion of the user, local pollution and breathing activity (Zhou et al., 2001; Adams et al., 2009). Within the PULSE tools, the model (Breathe, 2016) is currently implemented which is described in Table 1. We only consider four motion statuses, for the moment, and for each of them we know from literature the average breathing pace and volume on the inhaled air.

Each dot belonging to a recorded track is classified into the states listed in the Status column, according to the speed value, following the criteria reported in the Speed column. Each dot also has a time span, so the air volume inhaled at each dot is determined. Knowing pollution intensity (which is a concentration typically measured in micrograms per cubic meter) the mass is calculated of the pollutant inhaled at each dot. An example of such a calculation is shown in Figure 20.

Speed [km/h]	Status	# breaths per minute	Air volume per breath [litre]
< 2	At rest	15	0.6
2 - 6	Walking	28	1.8
6 - 15	Running	40	2.5
> 15	Driving a car	15	0.8

Table 1. The breathing model adopted so far

4. RESULTS

The section will present main results obtained for dense and dynamic air quality monitoring and for personal exposure assessment.

4.1 Continuous air quality maps construction

In the following, an example of a continuous map of air pollution and climate variants will be shown for the day January 4th, 2020. Figure 11 shows the color-coded values of PM 10 recorded by all the monitors at the time 10:00 AM. The colour bar shown on the right side expresses the PM 10 concentration in μg per m^3 . To further highlight that climate variants are monitored as well, Figure 12 is analogous to the previous one but are related to

temperature; colour bar on the right side express the temperatures in Celsius degrees.

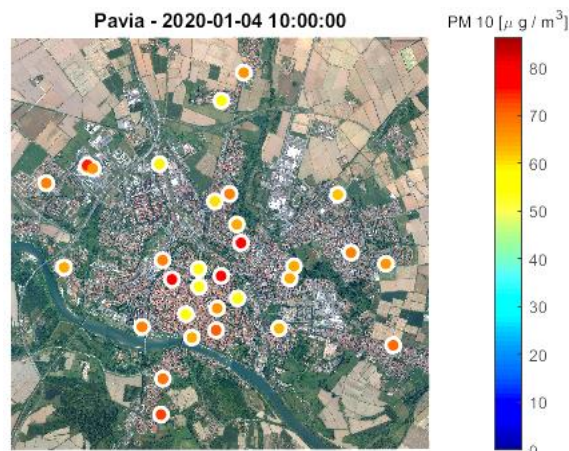


Figure 11. Color-coded measurements for PM 10 and all the monitors; day selected in January 4th, 2020 and time shown is 10:00 AM

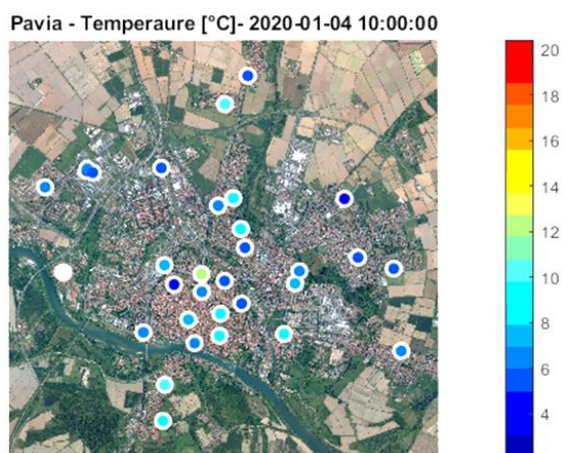


Figure 12. Color-coded measurements for temperature and all the monitors; day selected in January 4th, 2020 and time shown is 10:00 AM

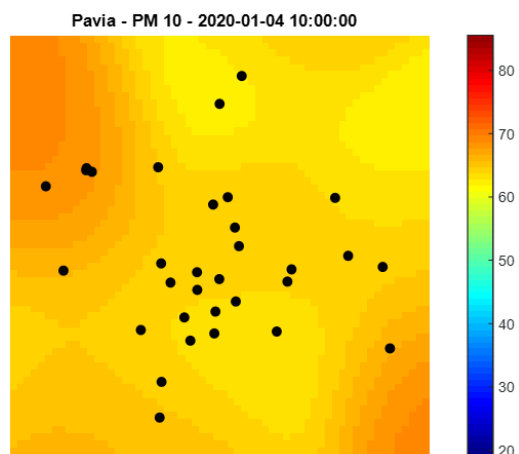


Figure 13. The interpolated continuous model for PM10; day selected in January 4th, 2020 and time shown is 10:00 AM

Continuous maps are obtained by gaussian-kernel interpolation applied to discrete measurements shown in Figure 10. It's worth

reminding that the map calculated for the time 10:00 is not only obtained by measurements acquired at the same time (Figure 11) but all the samples measured in a suitable time span: they are inserted in the calculation with a suitable weight.

An example of the so-obtained continuous map is reported in Figure 13; it shows the interpolated model for PM 10 referred once again to the 10.00 AM of the selected day. The colour bar shown on the right side expresses the PM 10 concentration in $\mu\text{g per m}^3$ whereas the black dots represent the position of the original input data corresponding to the air quality sensors network. Such maps are produced every hour and for all the surveyed variables: for air pollution PM 1.0, PM 2.5 and PM 10; for climate variants temperature, humidity and pressure.

4.2 Personal exposure calculation

Several volunteer users have been monitored in Pavia since December 2019 for personal exposure calculation and more than 300 tracks have already been acquired and processed. To illustrate the proposed methodology a track acquired by a single user at the beginning of the 2020 year will be used as example. The user can acquire a whole track related to all his day or monitor significant time frames only. For the sake of readability, reasonably shorter track is used in the following as example of the implemented functionality. The track is shown in Figure 14 and it is related to January 14th, 2020.

Pavia - User: 1214 - Serial: 2 - 2020-01-14 - Location

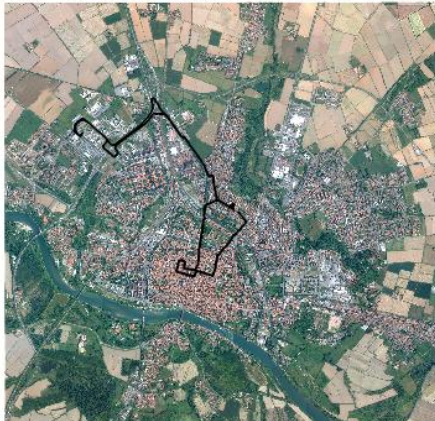


Figure 14. Example of the tracks acquired by a user in Pavia on January 14th, 2020

After the user has recorded the path, data is stored in tabular format like the fragment shown in Figure 15: each line is referred to a measure and reports time and the 3D coordinates expressed in East, North and ellipsoidal height, in the WGS84 datum and according to the UTM projection; date and time are stored, too. Indeed, the PULSE backend services stores WGS84 geographic coordinates, being the project international and involving cities belonging to three continents. When exposure is calculated, projected coordinates are used instead, for the sake of simplicity.

2020-01-14	12:01:27	510638.804	5005501.567	161.090
2020-01-14	12:01:28	510640.216	5005502.236	160.220
2020-01-14	12:01:29	510642.808	5005502.129	157.210
2020-01-14	12:01:30	510649.640	5005502.696	153.160
2020-01-14	12:01:31	510653.724	5005502.703	148.520
2020-01-14	12:01:32	510656.081	5005502.262	145.760
2020-01-14	12:01:33	510657.181	5005502.042	143.910
2020-01-14	12:01:37	510655.690	5005501.151	143.380
2020-01-14	12:01:38	510654.749	5005499.927	143.430
2020-01-14	12:01:39	510654.201	5005498.926	143.760

Figure 15. Excerpt from the stored tracks data

Being the points located in space and time, it is possible to calculate velocity. This step requires some cautions because some smoothing is required, in order to avoid that differentiation introduces too high noise in the produced data. The last three columns of Figure 16 represent derived quantities: velocity (in m/sec, shown in red), time duration (the time span between one dot and the following, which it is not always regular, shown in blue) and distance between the *i*-th dot and the next one (in m, shown in green). They are all used for further calculations.

2020-01-14	12:01:27	510638.804	5005501.567	161.090	1.19	1	1.56
2020-01-14	12:01:28	510640.216	5005502.236	160.220	1.19	1	2.59
2020-01-14	12:01:29	510642.808	5005502.129	157.210	1.19	1	6.86
2020-01-14	12:01:30	510649.640	5005502.696	153.160	1.19	1	4.08
2020-01-14	12:01:31	510653.724	5005502.703	148.520	1.19	1	2.40
2020-01-14	12:01:32	510656.081	5005502.262	145.760	1.19	1	1.12
2020-01-14	12:01:33	510657.181	5005502.042	143.910	1.08	4	1.74
2020-01-14	12:01:37	510655.690	5005501.151	143.380	0.89	1	1.54
2020-01-14	12:01:38	510654.749	5005499.927	143.430	0.67	1	1.14
2020-01-14	12:01:39	510654.201	5005498.926	143.760	0.74	1	1.07

Figure 16. Excerpt from the stored tracks data

Pavia - User: 1214 - Serial: 2 - 2020-01-14 - Speed [m/sec]

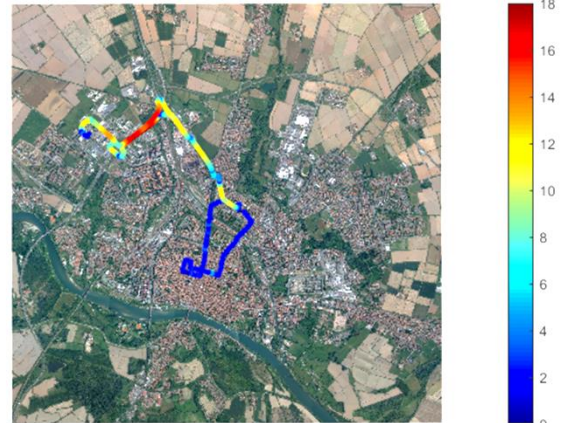


Figure 17. Velocity in m/sec for the path shown in Figure 14

Figure 17 shows velocity for the path shown in Figure 14. It highlights that a part of the track was travelled by car, as velocity is in the 5-18 m/sec range, represented by light blue, yellow, orange and red. Another part was walked instead and is shown in blue in the city centre. Next step is to associate each 4D dot with local pollution and this is done by interpolation described in Section 3. Figure 18 shows the considered track coloured according to the local and instantaneous values of PM 10.

User: 1214 - Serial: 2 - 2020-01-14 - Local PM100_ATM_A

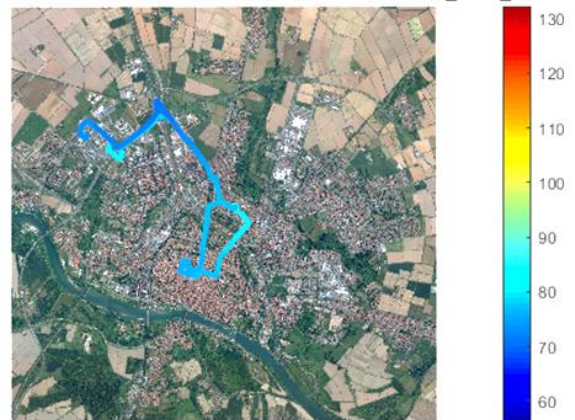


Figure 18. The track coloured according to the local and instantaneous values of pollution

Finally, the amount of pollutant inhaled at each dot is estimated. In Figure 19 the red columns report the local pollutant concentration and the blue one the ingested amount. Figure 20 shows point coloured according to exposure: noticeably, it is lower in the part where the car was used and is higher when the user walked.

2020-01-14	12:01:27	510638.804	5005501.567	161.090	1.19	1	1.56	81.58	0.0685
2020-01-14	12:01:28	510640.216	5005502.236	160.220	1.19	1	2.59	81.58	0.0685
2020-01-14	12:01:29	510642.808	5005502.129	157.210	1.19	1	6.86	81.57	0.0685
2020-01-14	12:01:30	510649.648	5005502.696	153.160	1.19	1	4.08	81.56	0.0685
2020-01-14	12:01:31	510653.724	5005502.703	148.520	1.19	1	2.40	81.55	0.0685
2020-01-14	12:01:32	510656.081	5005502.262	145.760	1.19	1	1.12	81.55	0.0685
2020-01-14	12:01:33	510657.181	5005502.042	143.910	1.08	4	1.74	81.54	0.2740
2020-01-14	12:01:37	510655.690	5005501.151	143.380	0.89	1	1.54	81.53	0.0685
2020-01-14	12:01:38	510654.749	5005499.927	143.430	0.67	1	1.14	81.53	0.0685
2020-01-14	12:01:39	510654.201	5005498.926	143.760	0.74	1	1.07	81.53	0.0685

Figure 19. The same example-data above with two more columns added: air pollution (in our case PM 10; unit is in $\mu\text{g}/\text{m}^3$; it is shown in red) and personal exposure (PM 10, units are absolute μg and are shown in blue)

User: 1214 - Serial: 2 - 2020-01-14 - Exposure PM100_ATM_A

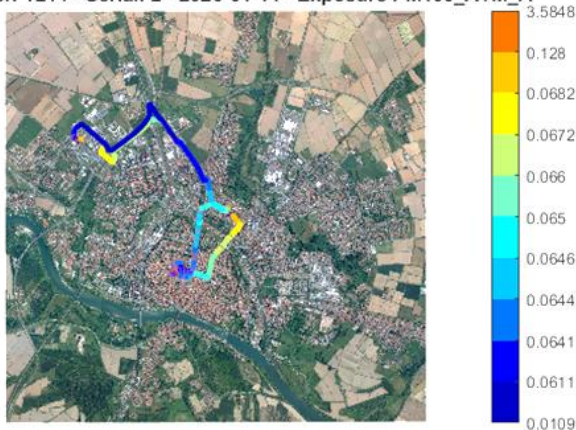


Figure 20. Exposure calculated at any location, measured in micrograms of PM 10

The developed toolbox also shows a summary display for each processed track, as shown in Figure 21. It highlights that the total length of the processed track is around 12 km and have time span of 1 hour and 10 minutes; recorded dots are 2900 roughly and the total amount of the inhaled PM 10 is 227 μg .

City: pavia
 User: 1214
 Instance: 2
 Start time: 2020-01-14 12:01:27
 End time: 2020-01-14 13:10:28
 Pollutant assessed: PM 10
 Number of dots: 2885
 Total length: 11.655 [km]
 Time span: 1.15 [hours]
 Total exposure: 227.4404 [micrograms]

Figure 21. Summary statistics

Personal exposure algorithms have already been integrated into the PULSE backend services and indeed the related data can be visualized through the PulsAir App and the project's WebGIS, as Figure 22 and Figure 23 demonstrate. It could be observed that the one shown by the WebGIS doesn't exactly agree with those shown in Figure 20. The reason is that, to speed up display, a hierarchical description of the paths is adopted, so that, according to the zoom level, a simplified version is shown. One last remark

concerns privacy issue; while it is perfectly acceptable that a citizen can see on his own mobile phone the trajectories he followed, that could be a problem for the WebGIS. We answer, first of all, that data are anonymized, and the data management plan is defined in detail and is described in one report of the project. Moreover, the WebGIS has two configurations, one for the whole public and one for qualified users, such as the personnel of local healthcare agencies. The presented screenshot was obtained from such a mode, while the public version of the site is not enabled to show individuals' trajectories.

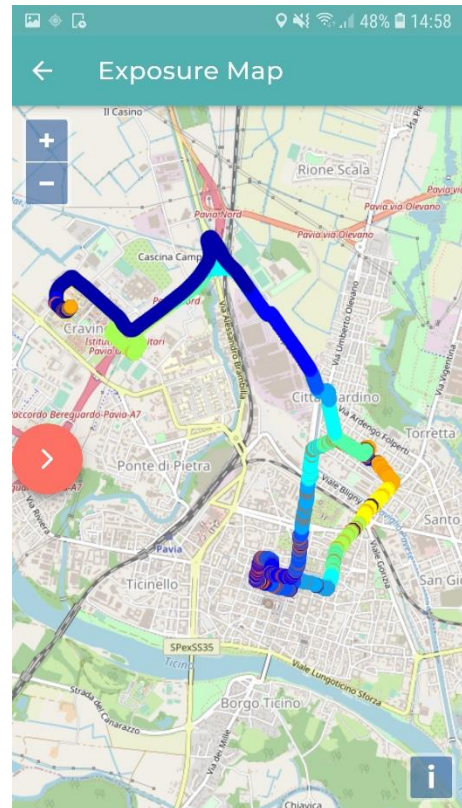


Figure 22. PULSEAIR App screenshot where personal expositor trajectory is visible

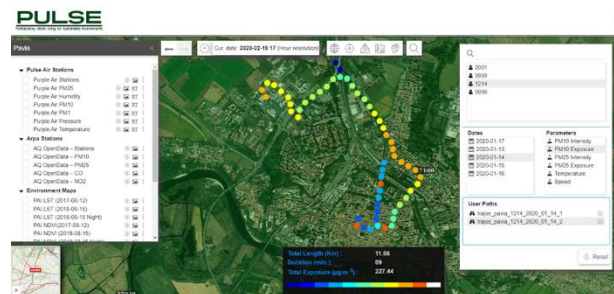


Figure 23. PULSE WebGIS screenshot where personal expositor trajectory is visible

5. CONCLUSIONS AND FURTHER ACTIVITIES

The paper illustrates a dense air quality monitoring network deployed in Pavia-Italy and a novel implementation of personal exposure assessment.

Data acquired by the monitoring network highlight that air pollution has a significant spatial variability, even at the scale of a medium city. Mostly, they have strong high-frequency time

variability, as Figure 8 confirms showing substantial intra-day variations. A first conclusion is that, to perform detailed assessment of personal exposure to air pollution, detailed air quality models are needed, which are dense in 4D (space plus time). Also, quasi-real time data availability is required, in order to make the service capable to notify people in real time on just after the end of the day; in PULSE, the second solution was adopted, but the first one is more attractive and could be really interesting for people.

The depicted personal exposure implementation is independent from the sensing system supplying the data. In Pavia we adopted a network of low-cost sensors, but this is not mandatory. Alternative approaches are based on satellite data (some specific EU missions are already operational, and others are planned) and physical diffusion models. In future, integrated models will probably become the gold standard. The capability of providing 4D-dense measurements in quasi-real time is compulsory anyway.

The implemented personal exposure solution meets all the requirements mentioned at the beginning. It is open because it can be accessible by an unlimited number of users (all the population, one day), once the monitoring infrastructure is deployed, simply by installing a dedicated App on their smartphone. It's dynamic or differential as it is able to quantify the inhaled pollution for each time frame chosen by the user: 1 second, 5 seconds, 1 minute... It's continuously available, meaning that it can be performed on a routinely basis and not for the limited time span of an experiment. Finally, it is freely upgradable when advances in the sensing technology will be available. This is true under the condition, which we assumed, that the development of the App installed by users is funded by local governments or governmental agencies. Finally, the App could become a tool for a more advanced interaction between citizens and local governments, healthcare agencies and environmental protection organizations.

The benefits for the citizens and the environment are manifold. It will lead to awareness increase about air quality and the amount of the pollution inhaled. This will induce, very likely, behavioural changes and the adoption of more sustainable mobility styles: this will positively impact the environment. Even if the pollution levels remained the same, our proposed methodology will allow the users to optimize their mobility strategies and thus to mitigate pollution's effects: "I'll go jogging in the evening, rather than in the morning, because the inhaled pollution will be lower".

Several developments are already envisaged and partially under development. Line simplification algorithms are under development: there are many in literature which are fully satisfactory from the geometric point of view, while we need that simplification does not impact the exposure assessment.

Line segmentation would be interesting too, i.e. the capability of separating a whole trajectory into homogeneous parts corresponding to different activities: at rest, walking, running, driving a car, biking, staying on a train... This is not yet under development and would be benefitted by considering, together with GPS/GNSS data, measurements coming from other sensors of the smart phone, such as accelerometers. We must say that the apps developed so far are not able to access such additional data. The breathing model adopted is quite simple. There is a running collaboration with the two major hospitals of the city focusing on the refinement of such a model, among several other topics. This part could be benefitted by the integration of data coming from smart watches and mobility trackers. Indeed, such devices can record additional data such as heart rate and several models (Greenwald et al., 2019; Klass et al., 2019; Schantz et al., 2019) have been developed to relate breathing activity to heart rate. Nevertheless, our developed methodologies will be scalable: basic functionalities are available through the phone-only

configuration while additional ones are activated when supplementary devices are available. Data coming from dedicated trackers would also be beneficial for the previously mentioned segmentation.

Finally, we aim at recruiting an adequate number of volunteers and at performing and initial, statistically significant sampling of personal exposure. This will be done after that the sensing network has been fully calibrated and tested. Such validation activity is running but is out of the scope of present paper and therefore not illustrated here.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the municipality of Pavia, ARPA Lombardy and all the PULSE consortium for their cooperation and contributions to this study.

Dr. Manuel Ottaviano of UPM (Universidad Politecnica de Madrid) is kindly acknowledged here for providing Figure 22. This work has been funded by the European Commission Horizon 2020 Framework programme under grant agreement GA-727816.

REFERENCES

- Adams, C., Riggs, P., Volckens, J., 2009. Development of a method for personal, spatiotemporal exposure assessment. *Journal of Environmental Monitoring*, 11(7), 1331-1339. doi.org/10.1039/B903841H
- Anandan, C., Nurmatov, U., Van Schayck, O.C.P., Sheikh, A., 2010. Is the prevalence of asthma declining? Systematic review of epidemiological studies. *Allergy*, 65(2), 152-167. doi.org/10.1111/j.1398-9995.2009.02244.x
- Breathe, 2016. Your lungs and exercise, vol. 12, no. 1, 97-100. doi.org/10.1183/20734735.ELF121.
- de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M.J., Badaloni, C., Beelen, R., Bolignano, A., Cesaroni, G., Pradas, M.C., 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environment international*, 73, 382-392. doi.org/10.1016/j.envint.2014.08.011
- Department of Economic and Social Affairs, United Nations. World urbanization prospects, the 2014 revision: highlights. 2014. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=857993> (Accessed April 2020)
- Greenwald, R., Hayat, M.J., Dons, E., Giles, L., Villar, R., Jakovljevic, D.G., Good, N., 2019. Estimating minute ventilation and air pollution inhaled dose using heart rate, breath frequency, age, sex and forced vital capacity: A pooled-data analysis. *PLoS One*, 14(7): e0218673. doi.org/10.1371/journal.pone.0218673.
- Guarnieri, M., Balmes, J.R., 2014. Outdoor air pollution and asthma. *The Lancet*, 383(9928), 1581-1592. doi.org/10.1016/S0140-6736(14)60617-6
- Innovating Cities, Environment, Research and Innovation. European Commission https://ec.europa.eu/research/environment/index.cfm?pg=future_cities (Accessed April 2020)

Johnson, M., Isakov, V., Touma, J.S., Mukerjee, S., Özkaynak, H., 2010. Evaluation of land-use regression models used to predict air quality concentrations in an urban area. *Atmospheric Environment*, 44(30), 3660-3668. doi.org/10.1016/j.atmosenv.2010.06.041

Klass, M., Faoro, V., Carpentier, A., 2019. Assessment of energy expenditure during high intensity cycling and running using a heart rate and activity monitor in young active adults. *PLoS One*, 14(11): e0224948. doi.org/10.1371/journal.pone.0224948.

McMillen, D.P., 2004. Geographically weighted regression: the analysis of spatially varying relationships.

Monn, C., 2001. Exposure assessment of air pollutants: a review on spatial heterogeneity and indoor/outdoor/personal exposure to suspended particulate matter, nitrogen dioxide and ozone. *Atmospheric environment*, 35(1), 1-32. doi.org/10.1016/S1352-2310(00)00330-7

Purple Air website, <https://www.purpleair.com/sensors> (Accessed April 2020)

Sanchez, M., Milà, C., Sreekanth, V., Balakrishnan, K., Sambandam, S., Nieuwenhuijsen, M., Kinra, S., Marshall, J.D., Tonne, C., 2019. Personal exposure to particulate matter in peri-urban India: predictors and association with ambient concentration at residence. *Journal of exposure science & environmental epidemiology*, 1-10. doi.org/10.1038/s41370-019-0150-5

Schantz, P., Salier Eriksson, J., Rosdahl, H., 2019. The heart rate method for estimating oxygen uptake: analyses of reproducibility using a range of heart rates from commuter walking. *Eur J Appl Physiol*, 119(11-12):2655-2671. doi.org/10.1007/s00421-019-04236-0.

Steinle, S., Reis, S., Sabel, C.E., 2013. Quantifying human exposure to air pollution—Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Science of the Total Environment*, 443, 184-193. doi.org/10.1016/j.scitotenv.2012.10.098

WHO website, <http://maps.who.int/airpollution/> (Accessed April 2020)

Wilson, A. Nickisch, H., 2015. Kernel interpolation for scalable structured Gaussian processes (KISS-GP). In *International Conference on Machine Learning*, 1775-1784.

Zhou, J., You, Y., Bai, Z., Hu, Y., Zhang, J., Zhang, N., 2011. Health risk assessment of personal inhalation exposure to volatile organic compounds in Tianjin, China. *Science of the Total Environment*, 409(3), 452-459. doi.org/10.1016/j.scitotenv.2010.10.022