Methodology for the investigation of IAQ inside French dwellings: the PAMELA study

Jerome Chesneau, Suzanne Crumeyrolle, Luc Dauchet, Benjamin Hanoune

Methodology for the investigation of IAQ inside French dwellings: the PAMELA study

Jerome Chesneau [1,*] , Suzanne Crumeyrolle [2] , Luc Dauchet [3] , Benjamin Hanoune [1,*]

[1] Univ. Lille, CNRS, UMR 8522 – PC2A – Physicochimie des Processus de Combustion et de l'Atmosphère, F-59000 Lille, France
[2] Univ. Lille, CNRS, UMR 8518 – LOA – Laboratoire d'Optique Atmosphérique, F-59000 Lille, France
[3] Univ. Lille, Inserm, CHU Lille, Institut Pasteur de Lille, U1167 - RID-AGE - Facteurs de risque et déterminants moléculaires des maladies liées au vieillissement, F-59000 Lille, France

[*] *Corresponding email: benjamin.hanoune@univ-lille.fr*

SUMMARY
The PAMELA (Particules Atmosphériques : Mesure de l'Exposition à Lille et aux Alentours) project aims to quantify the personal exposure to particles in the Lille area (Northern France), to differentiate the exposure according to the various environments, and to explore the drivers of exposure in terms of environment and activity. To that purpose we developed a compact particle sensor (0.3-10 µm, 6 size channels), also measuring additional parameters (T, RH, P, GPS positioning), that is worn during one week by participants recruited for the study. We present here the first results obtained, as well as general methodological considerations that emerged after a few weeks, concerning the design of the sensor units, the analysis of the data, and the issues related to the behavior of participants in such a study.

KEYWORDS
Individual exposure, sensors, particles, citizen science, microenvironments and activities

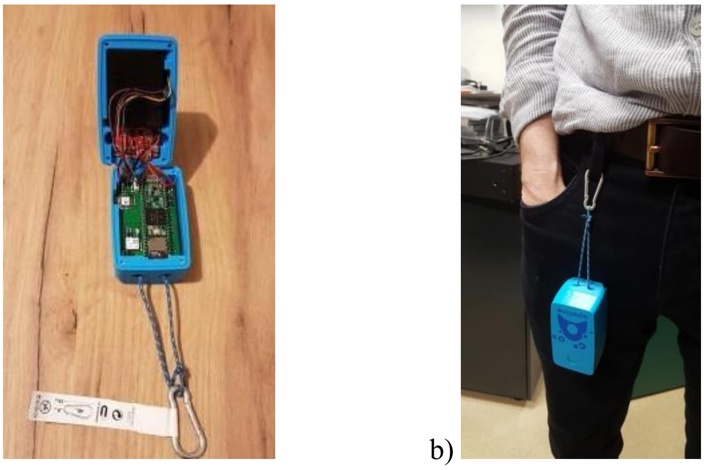
1 INTRODUCTION

The World Health Organization considers air pollution as the greatest environmental risk in the world, and estimates that seven million people die annually from exposure to air pollutants. The health impacts of acute and chronic exposure to particles have been the subject of many studies, with the most common effects on the cardiovascular and respiratory systems. Epidemiological studies have shown that particles even impact the health of healthy people (Dauchet et al, 2018). However, in these epidemiological studies, the individual exposure is not measured but only evaluated using the concentrations at the home address of the subjects of the cohort, calculated using chemistry-transport models and outdoors concentrations as input ("front door exposure"). While this approach benefits from its simplicity and availability of data from the national regulatory outdoor air monitoring networks, it does not consider the many environments in which we are exposed every day, or the diversity of activities and behaviors. In particular, this approach completely overlooks the indoor exposure of people (80 to 90 % of the time), starting with the home environment (50-60 % of the time).
It is therefore necessary, to better assess the health impacts of air pollution, to improve the assessment of the exposure, and if possible by measuring it instead of inferring it indirectly. Time-resolved personal measurements will also, in the first place, provide a well-needed description of the pollution levels, and of the environmental parameters, in many indoors and outdoors microenvironments, as well as the concentration levels of pollutants as a function of activities performed by people or close to people.

The objectives of the PAMELA project (Particules Atmosphériques : Mesure de l'Exposition à Lille et aux Alentours – Atmospheric Particles : measuring individual exposure in and around Lille), are to develop wearable real-time particle sensors, to progressively enroll at least 500 persons over the next two years in the metropolitan area of Lille (Northern France), who will carry the sensors during one week, to measure their individual exposure, to discriminate between indoor and outdoor exposure, to relate their exposure to the microenvironments and activities, and ultimately to set the ground for future full-size epidemiological studies.

2 MATERIALS/METHODS

Wearable particle sensors (50 units so far) were developed in our laboratories and are presented in Fig. 1. They measure 8*5*5 cm, and include an optical particle sensing element (HK-A5 model; output: particle number concentration in 6 size channels from 0.3 to 10 µm, mass concentration PM_1, $PM_{2.5}$ and PM_{10}), a temperature and relative humidity sensor (HDC 180), a pressure sensor (DPS 310), a GPS unit, a battery providing autonomy of about 10 hours, an SD storage card and a BLE communication system. Measurements are taken every second. Data are retrieved either by physically connecting the sensor unit to a computer after the measuring campaign or through a mobile phone app developed in our lab. This app allows quasi real-time transfer of the data to a server, and real-time visualization, but is not used in the present study, because its use by the wearers could influence their behaviour.



a) b)

Figure 1: a) inside view of the sensor unit b) sensor carried by a participant

The calibration (particle, T, RH) of the sensors is performed by installing them on the ATOLL atmospheric observation facility (http://www.loa.univ-lille1.fr/observations/plateformes.html) located at the University of Lille, which provides reference measurements for temperature, humidity and size-resolved number concentration up to 1 µm as well as PM_1 mass concentration (ACSM and SMPS instruments). $PM_{2.5}$ and PM_{10} reference measurements are taken from a reference monitoring station of the French air quality network, located 3 km away. Long-term measurements have shown that there are only rare moments with local pollution events that would perturb such a calibration.

Participants (40 so far) to the study are recruited among university members and staff, students, and local associations. For the moment, we do not consider the representativity of the participants with respect to age, sex, home address…, but these factors will be considered eventually as the number of volunteers increase. Each participant is given a sensor with instructions for use, and is requested to wear it or keep it close to himself or herself during one week, after which time they return it to the lab for data analysis and technical assessment of the sensor. The participants also complete questionnaires about themselves and their dwellings, and are asked to keep a space/time/activity log, under free format. The participants have no way of visualizing the measurements, so that they will not modify their behavior. After analysis of the data, a report is provided, containing a summary of their exposure, with comparison to French health guidelines and to the average results from all the previous participants of the study. Privacy issues are taken into account at all steps of data, questionnaires, logs and results handling.

Data analysis is performed using R language routines developed in the lab. They are used for calibration of the sensors, data preprocessing, and data analysis.
Data preprocessing includes (i) removal of incomplete dataframes that can happen when turning off the sensor when data are being written or from data transfer errors, (ii) change of units for the particle number concentrations because of a specific format delivered by the sensor, (iii) addition of flags corresponding to the state of the battery (full but still on charge, see below) and to a null number of satellites detected by GPS, (iv) calculation of the absolute humidity.
Data analysis aims to differentiate the environments and activities based on all measured parameters except the particles number and mass concentrations. For this we rely on segmentation methods, applied either to temperature, humidity (relative and absolute), number of visible satellites and speed as determined by the GPS, considering the mean values, slopes and fluctuations of each of these parameters, using the ‘changepoint’ R-package (Killick et al. 2012). Sensitivity of the segmentation methods still needs to be manually adjusted to optimize the number of significant environment/activity changes. Simultaneous use of all the parameters for segmentation should perform more efficient and robust segmentation (El Hafyani et al., 2020) but has not been implemented yet. The space/time/activity log is not used for the segmentation process, but only afterwards to explain and “put a name” on the different environments.

3 RESULTS

An example of data acquired by one participant is given in Figure 2.

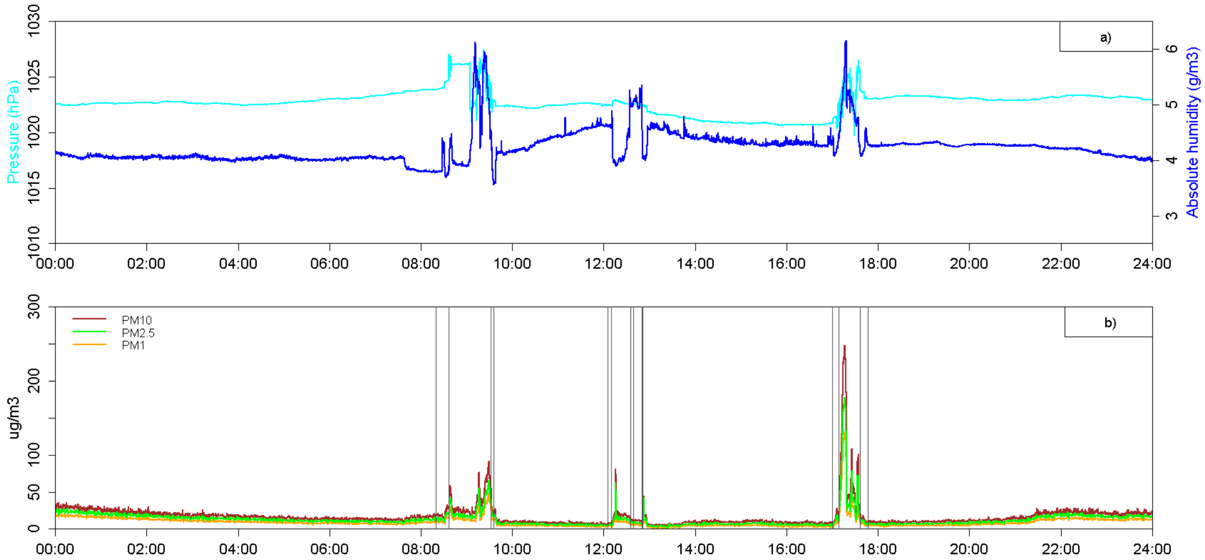


Figure 2: Time series acquired on Nov. 9, 2021 (volunteer #2, sensor unit APO_010). a) pressure and absolute humidity, b) PM concentration.

Pressure and absolute humidity were found, in this example as in others, to be the best indicators of change of microenvironment. The segmentation process identifies changes, most of them clearly visible on the pressure and humidity data, but not so on the PM_{10} mass concentration, with only three high concentration episodes sticking out from the background concentration. The main episodes were identified by the participant in its activity log as subway ride in the morning, lunch break at the university cafeteria, and subway ride in the evening. While concentrations of PM_{10} in the subway in the mornings tend to be quite stable from day to day, concentrations in the evenings show unexplained much larger variations. From this segmentation, and by comparison with the activity log, we can derive the duration that the participant spent in each environment. For instance, for the data shown on Figure 2, the participant spent 60% of his time at home, 28% in the office, 6 % in transportation (4% in the morning and 2% in the evening), 2% for lunch at the university cafeteria, and 4% outdoors.

Fig. 3 presents the statistical analysis of the PM concentrations during these events. The most polluted environments are the subway cars, which are well known confined environments often presenting high PM concentrations (Moreno et al. 2018). PM_{10} concentrations measured at the home of the participant as well as outdoors are around 30 µg.m^{-3}, and lower in his workplace (university laboratory) and in the university cafeteria.

This figure also presents the average exposure levels calculated from our measurements (color lines), as well as the “front door” exposure levels. For PM_{10} and $PM_{2.5}$ the measured exposure levels are 17.2 and 13.0 µg.m^{-3} respectively, lower than the 36.0 and 16.4 µg.m^{-3} front door exposure levels. For PM_1, the exposure level is 9.2 µg.m^{-3}. This illustrates further the need to to improve epidemiological studies using individual measurements instead of extrapolated front door values.

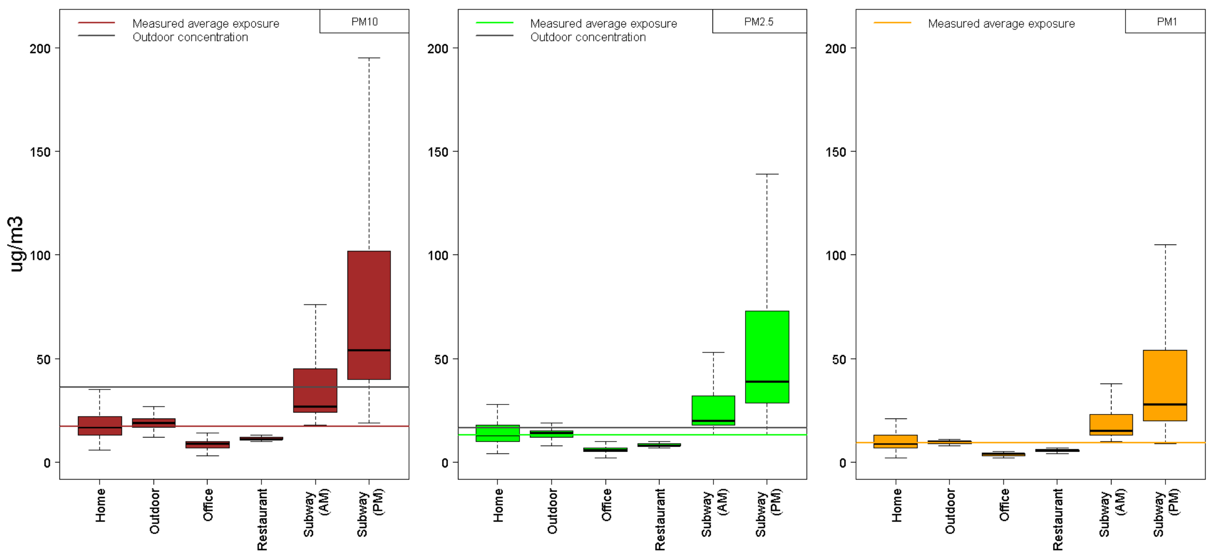


Figure 3: Statistical analysis of PM in each environment identified in Fig 2.

No other specific microenvironments have been identified so far, but activities (not shown here) such as smoking (outdoors), use of fireplace, and cooking activities are also clearly distinguished in the data from some participants.

4 DISCUSSION

Sensor design issues

In spite of careful design planning and testing while developing the sensor unit prototypes, several technical issues appeared when starting to analyze the first results. The integration of the temperature and humidity sensor inside the sensor unit, close to the battery and electronic board, leads to biased temperature measurements that cannot be completely corrected even with careful calibration, because of the thermal inertia of the unit (~20 minutes). In addition, when the sensor is plugged on the sector, which we recommended when measuring indoors to extend the duration of the measurements, the charging cycle of the unit leads to power overshoots that also perturb the temperature and humidity sensor. This makes the temperature and relative humidity readings for the moment unreliable when the battery level is over 4.5 V, but does not prevent their use for the segmentation procedure, because we test for breakpoints in the time profiles and do not search specific values. Testing is under way with a modified position of the temperature sensor inside the unit and a modified charging board that should fix these problems. No such effects have been noticed on the other sensors and outputs of the unit (P, GPS, particles), though the influence of temperature and humidity on the particle number and mass concentrations may not be correctly taken into account. However, even though the sensors have been calibrated against reference instruments, we expect that the uncertainty associated with the measurements may not be significantly changed with a better correction for environmental parameters.

User related issues

Many issues, some expected, some not, are connected with the incorrect use of the sensors by the participants of the study, and with the poor observance of the protocol. Some of the participants reported not being able to tell whether the sensor was on or off, in spite of the red and blue LED lights on the shield, and as a consequence they did not recharge the sensor. Some participants said that they simply forgot to recharge the sensor. In some cases, the participants only turned the sensor on when they suspected unsafe particle levels, instead of leaving the sensor active all the time. Some of the participants did not keep the sensor close to them, either they left it in one room when they were home, instead of carrying it with them, or they merely forgot to take it during their daily activities, which tended to happen more often after a three to four days, when the novelty of the experiment wore off. Some of them lent the sensor to a member of their family or a friend. While the questionnaires were in general correctly filled by the participants, this was not the case with the space/time/activity logs. In addition to being sometime illegible because of really bad handwriting, some logs contained irrelevant and unusable information, and were generally incomplete, even with participants educated in air pollution, in particular when the logs were filled either retrospectively, at the end of the day, leading to missing activities, but also when they were filled more often, because of rapid succession of environments and activities, such as leaving a building and entering another one, with a short smoking break outdoors in between, making it difficult to keep a written track of all these changes.

As a result, the amount of data usable is far from 100%, and it is difficult for the moment to ascertain for each participant the correct observance of the protocol, and to always compare exposure, environments or activities. It will be therefore necessary to improve the initial information package and follow-up that we give them. However, the first results obtained during this work, can still be exploited, with meaningful analysis and results, and we are starting to question the necessity of asking all the participants to keep a log.

5 CONCLUSIONS

The sensors and analysis methodology developed within the PAMELA project framework are now operational. Some adjustments need to be made to the sensors, so to have a better and more significant measurement of the temperature. The protocols for analysis are only at their early stages of development, with improvement expected in the segmentation process, and in the

automatic assignment of the microenvironment and/or activities based on environmental, GPS, and particles concentration and size, by way of unsupervised and supervised learning.
Further work is underway to improve the information given to participants to improve the quality of the space/time/activity logs, in order to better document the physical measurements, and to improve their knowledge and concern about air pollution and its effects on health.
A database with information on environmental parameters and particle load in all the microenvironments, and in particular indoors, including workplace, dwellings, and public transportation systems, is progressively growing, which will provide eventually statistically meaningful information useful for all aspects of IAQ studies, and for epidemiological studies as well.

ACKNOWLEDGEMENT
The work was supported by the French government through the Programme Investissement d'Avenir (I-SITE ULNE / ANR-16-IDEX-0004 ULNE), managed by the Agence Nationale de la Recherche as part of the project "Santé Environnement: du risque territorial au risque individual" (also funded by Lille European Metropole). This work is a contribution to the CaPPA project (Chemical and Physical Properties of the Atmosphere), funded by the French National Research Agency (ANR) through the PIA (Programme Investissement d'Avenir) under contract ANR-11-LABX-005-01, to the Institut de Recherches Pluridisciplinaires en Sciences de l'Environnement (IRePSE Fed 4129), and a contribution to the CPER research project CLIMIBIO, with financial support from the French Ministère de l'Enseignement Supérieur et de la Recherche, the Hauts de France Region and the European Funds for Regional Economic Development. One of the authors (JC) thankfully acknowledges the CaPPA project and the University of Lille for PhD support.

6 REFERENCES

Dauchet L., Hulo S., Cherot-Kornobis N., Matran R., Amouyel P., Edmé J.L. and Giovannelli J. 2018. Short-term exposure to air pollution: Associations with lung function and inflammatory markers in non-smoking, healthy adults. *Environ. Int.* 121(1):610-619

El Hafyani H., Zeitouni K. and Taher Y., Leveraging Change Point Detection for Activity Transition Mining in the Context of Environmental Crowdsensing. In: *Proceedings of the SIGKDD International Workshop on Urban Computing, UrbComp 2020*, San Diego

Killick R., Fearnhead P. and Eckley I.A. 2012. Optimal Detection of Changepoints With a Linear Computational Cost, *Journal of the American Statistical Association*, 107(500), 1590-1598

Moreno T., Martins V., Reche C., Minguillon M.C., de Miguel E. and Querol X. 2018. Chapter 13 – Air quality in subway systems. In Non-exhaust emissions – An urban air quality problem for public health; Impact and mitigation measures (Ed. F. Amato). Academic Press.