Low-cost coarse airborne particulate matter sensing for indoor occupancy detection

Kevin Weekly, Donghyun Rim, Lin Zhang,
Alexandre M. Bayen, William W. Nazaroff, and Costas J. Spanos

***Abstract*— In the energy-efficient smart building, occupancy detection and localization is an area of increasing interest, as services, such as lighting and ventilation, could be targeted towards individual occupants instead of an entire room or floor. Also, an increasing quantity and diversity of environmental sensors are being added to smart buildings to ensure the quality of services provided by the building. The need for particulate matter (PM) sensors in consumer devices such as air purifiers, is an example where manufacturing advances have made the sensors much less expensive than laboratory equipment. Beyond their original intended use, air quality, they can also be used for occupancy monitoring. The work presented in this article proposes to use a low-cost (< 8 USD) PM sensor to infer the local movement of occupants in a corridor by sensing the resuspension of coarse ($\geq 2.5\,\mu m$) particles. To obtain meaningful values from the inexpensive sensors, we have calibrated them against a laboratory-grade instrument. After calibration, we conducted a 7.8 hour experiment measuring coarse PM within a pedestrian corridor of a heavily-used office area. Comparing against ground truth data obtained by a camera, we show that the PM sensor readings are correlated with human activity, thus enabling statistical methods to infer one from the other.**

I. INTRODUCTION

As the price of global energy rises, an increasing number of researchers have adopted the challenge of reducing energy consumption of buildings and improving the comfort of its occupants. A major component of this effort is monitoring the building and observing it more comprehensively. For example, traditionally, an occupant is responsible for turning off the lights when she leaves a room, however, it would be more efficient for the building's control system to do this automatically by knowing where its occupants are and their needs. Acceptable indoor air quality is an occupant need and should be sensed and controlled for. Particulate matter (PM) sensors measure the particulate concentration, which, at high levels, correlates to health problems [1]. Thus, in this study, we have deployed several PM sensors in office spaces.

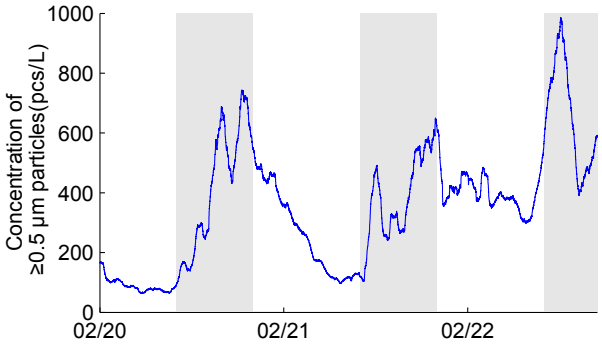


Fig. 1. Particulate matter detected inside office on UC Berkeley campus. Shaded area represents times between 10:00 PST and 20:00 PST each day.

While conducting continuous monitoring of PM concentration for this study, we noticed trends intuitively correlating to occupant activity. For example, Figure 1 shows a time-series of PM detected in an office space at the UC Berkeley campus over two full 24-hour periods. Clearly, there is a trend of a higher concentration of particles during the working hours of the day, and we can even see when there is a surge of activity inside of the workday, as indicated by the rise in PM. This trend is also consistent with previous work that showed significant increases of PM concentrations due to human activity such as walking [2].

To detect individual activity and movement in a building, the most straightforward methods are light-based, such as passive infrared (PIR) sensors, which detect movement of body heat, or emitter-detector pairs which trigger when a light beam, such as from a laser, is broken. Magnetic switches can also be added to doors to determine when they are open or closed. More recently, there has been a host of indoor positioning systems (IPS) [3] developed to obtain a person's location (i.e. x-y coordinates) continuously.

Our hypothesis is that PM concentration is an additional method that can be used to detect human activity. In the present study, we focus on the effect within a corridor, allowing us to easily obtain a ground-truth via camera. However, motivated by results such as those shown in Figure 1, human occupancy and activity for an entire room could be estimated. The advantage of using PM readings over installing a light-based system is that other motivations, such as monitoring air quality, drive the deployment of PM sensors.

*This research is funded by the Republic of Singapore's National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore.

K. Weekly, A.M. Bayen and C.J. Spanos are with the Department of Electrical Engineering and Computer Sciences at the University of California Berkeley, USA. kweekly at eecs.berkeley.edu, bayen at berkeley.edu, spanos at eecs.berkeley.edu

D. Rim and W.W. Nazaroff are with the Department of Civil and Environmental Engineering at the University of California Berkeley, USA. mcdhrim at gmail.com, nazaroff at ce.berkeley.edu

L. Zhang is with the Department of Electronic Engineering at Tsinghua University, China.linzhang at tsinghua.edu.cn

978-1-4799-1515-6/13/$31.00 ©
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6653970&isnumber=6653883

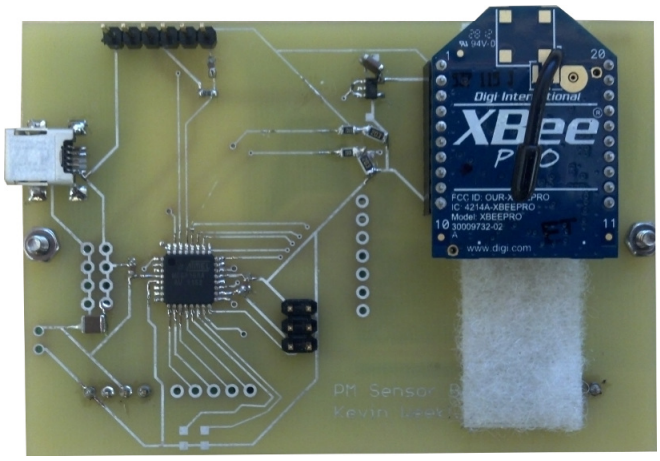


Fig. 2. Custom designed electronic board, which when mounted to the PPD-20V, provides wireless PM reporting capability.

The rest of the article is organized as follows: In Section II, we review some of the related studies of indoor PM. Section III describes the hardware developments enabling real-time PM monitoring. Section IV describes the PM sensors we tested and the test apparatus we built to collect PM readings. Section V describes how we interpret the raw data and ultimately, the correlation between human activity and PM concentration. Finally, we conclude the article in Section VI.

II. RELATED WORK

In the literature, the relevant physical phenomena is termed *resuspension*, which is when particles lift off of a surface and become airborne. Since human activity, such as walking or vacuuming, is a large contributor to resuspension in an indoor office space, we reason that it can act as a proxy for occupancy.

A mass conservation model given by [4] is commonly used to model the PM concentration of an entire room or building and includes a term accounting for resuspension. There have been several studies of experimentally investigating the effect of human activity on PM concentration in real [5], [6], [7], [8] and laboratory [9], [2] conditions. Although a mass conservation model is simple and intuitive, much of the complexity of an indoor space, such as space geometry and air mixing characteristics, is hidden in the parameters of the model, which are difficult to accurately determine. The model also only tracks a single variable, making it difficult to integrate sensor readings and locate emission sources, both which occur at a single point in 3D space. We need to understand the physics at a much finer spatial granularity to accomplish these goals. To this end, there has been some efforts to understand the deposition and resuspension of particles at a fine level [10], [11], [12], however, although these models identify more system parameters, it is again difficult to determine their values in a real setting.

Most studies of indoor air quality rely on expensive equipment such as Optical Particle Counters (OPCs) to obtain accurate PM readings of fine ($< 0.5\,\mu$m) particles, however, inexpensive sensors have been used to allow economical large-scale deployments [13], [14], [15]. Our hypothesis is that coarse and less accurate PM readings, such as those obtained by commodity and mass-produced optical sensors, can be made useful if their inaccuracies are mitigated through multiplicity and trend analysis.

TABLE I
SENSOR MODELS TESTED.

Model	Qty. Tested	Cost	PM Size detected
DSM501A[17]	5	< 8 USD	Two outputs: $\geq \{1, 2.5\}$ µm
PPD-20V[18]	3	~ 700 USD	$\geq 0.5\,\mu$m
GT-526S[19]	1 (reference)	2990 USD	Six outputs: $\geq \{0.3, 0.5, 1, 2, 5, 10\}$ µm

III. REAL-TIME MONITORING

To enable long-term experiments, we designed and built a wireless-enabled PM sensor based on the PPD-20V sensor, shown in Figure 2. The circuit board has the same dimensions as the PPD-20V sensor and is attached to the back by two bolts and two sets of header pins. An on-board ATMega128 microcontroller reads the digital output of the PM sensor and calculates the PM sensor ratio over 10 minutes. Every 10 minutes, the microcontroller uses an 802.15.4 transceiver (Digi XBee series) to send the reading to a sink node containing an embedded Linux computer (Beaglebone) and its own 802.15.4 transceiver. Python software communicates with the transceiver and forwards the raw sensor packet data to our Internet server for interpretation. Software running on our Internet server receives and interprets the data packet, then inserts the data into an sMAP [16] server instance running on itself. Thus, software for our future research projects can retrieve the data using the standardized sMAP protocol. Figure 1 shows one example of data that was collected with the device in Figure 2 and retrieved using sMAP. We can also use the sMAP web interface to quickly visualize the data as it arrives. Given the prohibitive expense of the PPD-20V sensor, our plan is to adapt the design to accommodate the DSM501A sensor and deploy many of these in the office for future coarse PM studies.

IV. PROCEDURE

A. Particulate Matter Sensing

Our experiments consisted of measuring particle concentration using 8 PM sensor modules. The specifications of the two models tested are given in Table I. Figure 3 shows the principle of operation of the PPD-20V sensor and the DSM501A sensors operate using a similar principle. Essentially, the sensor employs a Light-Emitting Diode (LED) aimed at a small point inside the device. When a particle of sufficient size passes by this point, a detector picks up the scattering of the LED light and outputs a digital signal on the output pin. A feature of the DSM510A device is that it provides two outputs with different sensitivities. A resistive heater causes convective airflow which ensures a small flow of air is passed through the device. We have

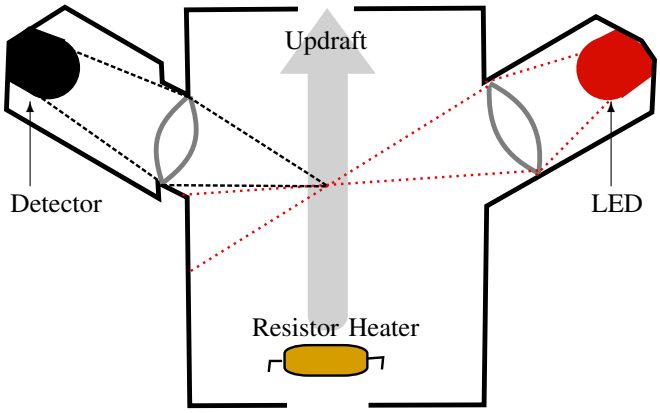


Fig. 3. Operation of low-cost PM sensor measuring scattering of LED light (Shown: PPD-20V).

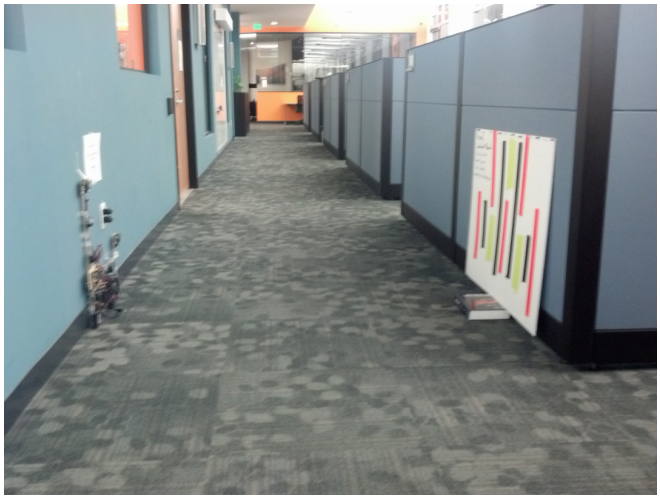


Fig. 4. Main corridor of 490 Cory Hall, prior to experiment being started. Rooms on the left are lab areas and cubicles on the right are occupied work spaces.

programmed an ATMega168 microcontroller to receive the 13 digital signals from the PPD-20V and DSM501A sensors and record the data onto a Secure Digital (SD) memory card. It is programmed to sample every signal at 2.5 kHz to see if it is 0 V (logic 0) or 5 V (logic 1). Every 50 ms, it logs the number of samples that were logic 1 during the last period. This data set can be processed into a ratio of the time that the signal is 0 V within a time window. After calibration against an accurate sensor, the GT-526S laser particle counter, this ratio value can be mapped to meaningful units (described in Section V-C).

B. Physical Configuration

The experiment was carried out in the main corridor of 490 Cory Hall (also known as the SWARM lab) on Berkeley campus, which is a heavily used and trafficked office location. Figure 4 gives a picture of the corridor prior to the experiment being run.

The camera and PM Sensors were mounted on an aluminum structure leaned against one side of the corridor and visual landmarks were placed on the other side of the corridor. The sensors and landmarks were arranged according to Figure 5. Figure 6 is a picture of the apparatus we constructed to achieve this configuration. Although the analysis in this article does not look into differences in PM readings due to spatial variations, the apparatus allows us to investigate this later. The most valuable contribution of the apparatus was allowing us to run a multiplicity of sensors with all of them relatively close and in the same orientation, and we found that averaging the readings from several sensors gave the smoothest results.

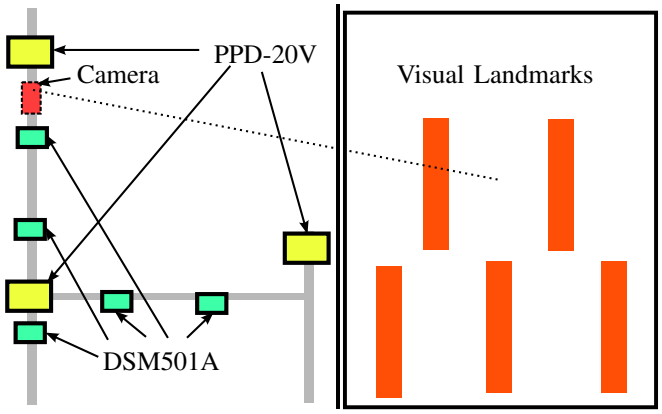
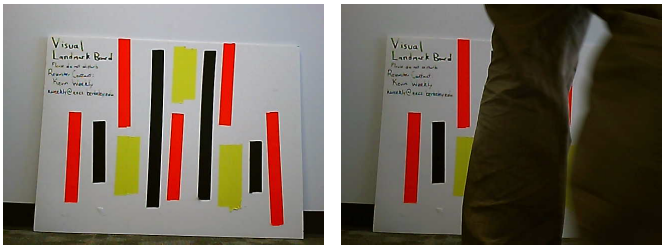


Fig. 5. Physical Configuration of the sensors and visual landmarks. Dotted line indicates center of view of camera.

Fig. 6. Physical implementation of the design in Figure 5.

C. Ground Truth Sensing

The goal of the ground truth observation was to determine when humans or other objects passed in proximity of the PM sensor. Ground truth data was collected using a Veho VCC-003-MUVI digital video recorder. After replacing the stock memory card with a 16 GB card, and the battery with a much larger one, the camcorder was capable of recording approximately 10 hours of video at a resolution of 640×480 pixels. We placed visual landmarks (bright orange tape strips) in the field-of-view of the camera, to make post-processing more straightforward. Examples of the images captured are shown in Figure 7.



(a) Clear View (b) Obscured View

Fig. 7. Example images showing detection of occupant in corridor from obscured visual landmarks (orange strips).

To obtain the ground truth, the video was later post-processed to detect when something obscured the visual landmarks. We made the determination based on whether the hue, saturation, and value of the color at the landmark locations fall outside a specified range. Since the video is also archived, we can validate the computer vision manually.

V. RESULTS

A. Filtering

Data recorded by the apparatus was taken at a high sample rate relative to the effects we intended to measure. By inspection, the individual PM samples are not easy to interpret as they are mostly either a ratio of 1 or 0. By doing several filtering steps we are able to extract a meaningful signal.

The raw data of the experiment comes in the form of a text file from the PM sensor apparatus as well as a video file from the digital video recorder. The video file was processed by a simple computer vision algorithm to give a set of times, $\mathcal{C} = \{t_1, t_2, \ldots, t_n\}$, when human activity occurred in front of the camera. There were a total of 312 such occurrences during the experiment. We calculated a *camera occurrence rate*, x_i, by the following:

$$x_i = \frac{|\{t | t \in \mathcal{C} \wedge t \geq W_s(i - 0.5) \wedge t < W_s(i + 0.5)\}|}{W_s}$$

where W_s is the "sample period" and is set to 10 s for these results. Thus, x_i is a discrete signal representing the number of camera occurrences per second, evaluated every 10 s.

The PM sensor signals, originally at a sample period of 50 ms, are down-sampled to a sample period of W_s (10 s), by averaging every 200 readings.

We then passed each signal, both PM readings and camera occurrence rate, through a sliding window average filter with a window size of 30 samples. Thus, the resulting samples still have a sample period of 10 s, but each sample represents the average of 300 s, or 5 min, worth of data.

B. Selection of variables

The text file provides timeseries of 13 variables, 2 per DSM501A sensor and 1 per PPM-20V sensor, so we sought to find which signals were relevant. We found that the coarse outputs correlated most with camera occurrences, an effect which is also supported by [8], which found that occupancy is a large factor in supermicron particle concentration, but an insignificant factor for finer particles. Intuitively, this is because larger particles are overcome by gravity and settle faster than finer particles which are kept afloat by air currents. Sensors which are sensitive to smaller particles are actually less effective for inferring occupancy since their readings will ultimately be dominated by submicron particles (which have a concentration almost 10 times as large as supermicron particles). To verify, we computed the Pearson's correlation coefficient between the camera data and PM readings, grouped by the size of particles sensed. The results, presented in Table II, show that the amount of linear correlation between camera occurrences and PM readings is the most when the particles are 2.5 µm or larger.

TABLE II

PEARSON'S CORRELATION COEFFICIENT FOR DIFFERENT OUTPUTS.

Output	r
≥ 0.5 µm from PPD-20V	-0.03
≥ 1 µm from DSM501A	0.28
≥ 2.5 µm from DSM501A	0.49

TABLE III

MSE OF LINEARLY CALIBRATED OUTPUTS COMPARED TO OPC

Output	MSE
≥ 0.5 µm from PPD-20V	4.7%
≥ 1 µm from DSM501A	19.4%
≥ 2.5 µm from DSM501A	9.5%

Ultimately, we decided to construct a signal, y_i, composed of the average of all 5 DSM501A ≥ 2.5 µm outputs. That is,

$$y_i = \frac{1}{5}\left(\sum_{k=1}^{5} \mathrm{PM}_i^{k,2.5\,\mu\mathrm{m}}\right)$$

where $PM_i^{k,2.5\,\mu\mathrm{m}}$ is the ith reading of the ≥ 2.5 µm output from the kth DSM501A sensor. This signal, by inspection, was smoother than any individual PM output and showed more correlation with the camera data.

C. Calibration

The calibration procedure consisted of a 29.5 hour experiment where data was collected by the test apparatus while the OPC measured accurate values of PM concentration. For both the ≥ 0.5 µm and ≥ 1 µm outputs, we used the corresponding channel from the OPC to compare against. For the ≥ 2.5 µm output of the DSM501A sensor, we used the ≥ 2 µm channel of the OPC to compare against.

For each output we performed a linear fit against the reference data. The scatter plot in Figure 8 illustrates the data collected and the fitted line for the ≥ 2 µm output. To determine the accuracy of the linearly fitted model, we also calculated the relative Mean Squared Error (MSE) of the modelled concentration against the real concentration, shown in Table III. Examining the time series, we found that there was a spike of airborne PM at around 00:00 hours during the experiment. The spike could be seen mostly in the micron or lower sized particles and was picked up by the PPD-20V unit. However, the DSM501A did not pick up the spike

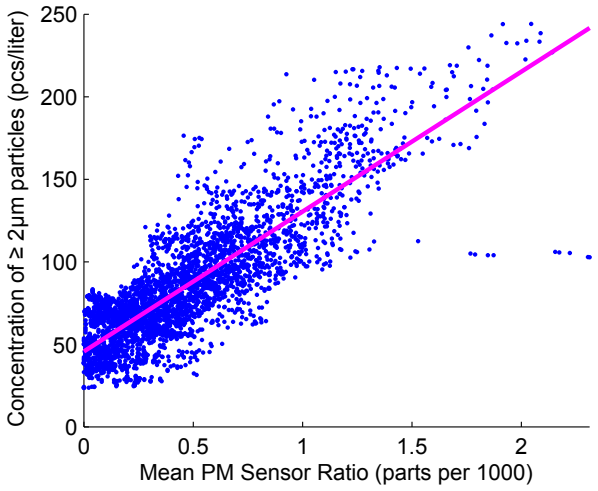


Fig. 8. Scatter plot of $\geq 2\,\mu$m particle concentration from OPC against $\geq 2.5\,\mu$m output from the DSM501A.

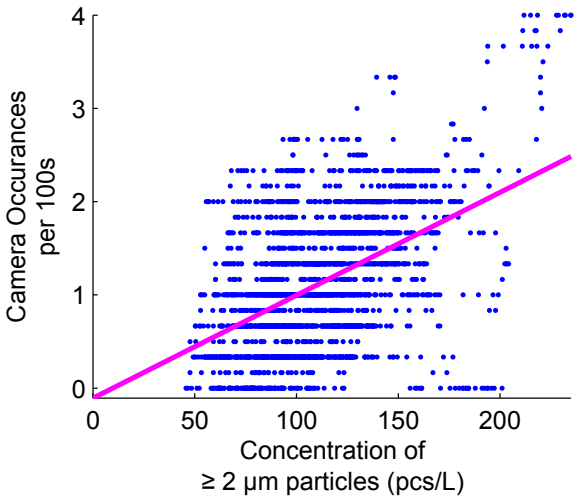


Fig. 9. Scatter plot of the data shown in Figure 10. Magenta line is the linear fit to the data.

in PM, perhaps because some characteristic of the particles such as surface (changing the refraction characteristics) or composition (causing them to not be caught in the heater updraft) caused them to not be detected by the DSM501A. This complicated the calibration results significantly as can be seen by the $\geq 1\,\mu$m results of Table III.

Fortunately, as described in the previous section, we opted to use the $\geq 2\,\mu$m outputs of the DSM501A for correlation with activity. This output could be calibrated reasonably accurately and thus we used the following linear model to map the PM sensor raw units to real-world values:

$$\text{PM}_{\geq 2\,\mu\text{m}} = \left(8.48 \times 10^4\right)(\text{PM Sensor Ratio}) + 45.7$$

D. Synchronization of PM sensor and camera data

We also shifted the PM sensor data earlier in time so that it more closely correlates to the camera data. We calculate the shift amount as the peak location of the cross-correlation of the two signals. For this experiment, the shift amount was found to be 30 s. Physically, this represents the time it takes for activities happening in front of the camera to affect the sensors via air circulation or diffusion. It could also account for experimental errors in synchronizing the time offset of the two signals.

E. 490 Cory corridor experimental results

Figure 10 is a plot of the two signals x (Top) and y (Bottom) after the filtering steps described above. The green lines are plotted at the times in the set $\mathcal{C}$. Visually, we can see a clear correlation of the two signals, particularly noting that many of the sharp increases in PM are correlated with increase in camera occurrences. There are some cases (e.g. at 5.5 hours) where there is a spike in PM concentration, but no

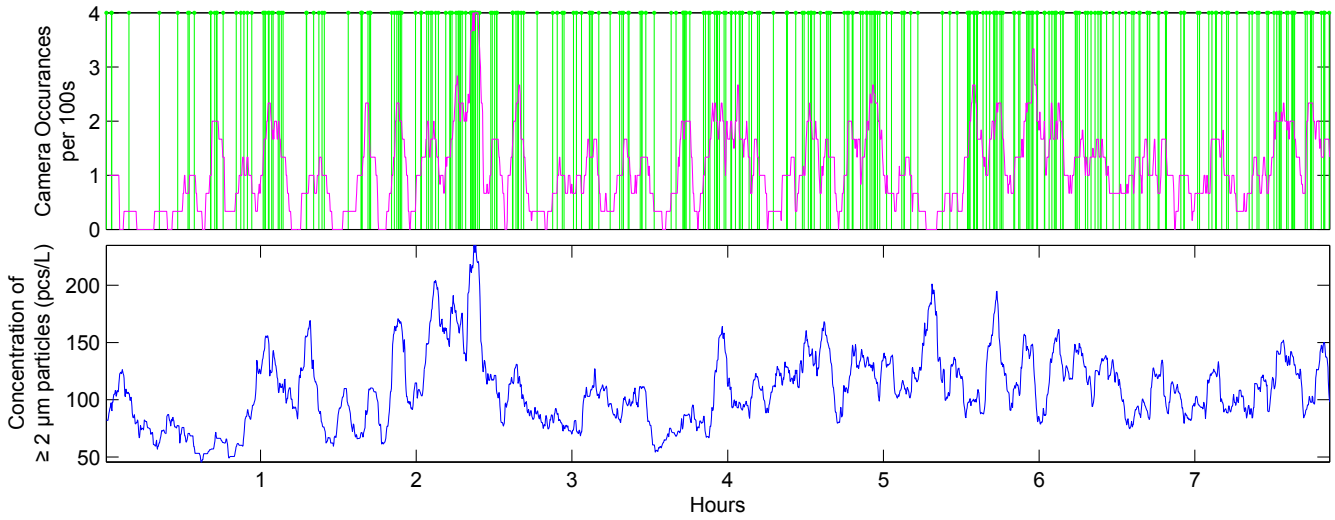


Fig. 10. Timeseries of filtered data over 7.8 hr experiment. Top: Green lines mark camera obstruction occurrences and magenta line is the filtered camera occurrence rate. Bottom: Filtered $\geq 2.5\,\mu$m outputs from DSM501A (average of 5).

corresponding spike in camera occurrences. One explanation is that there were significant resuspension events in the nearby cubicle without any persons moving in front of the camera.

The same data, is presented in a scatter plot in Figure 9 and we also plot the linear fit to the data. The linear fit is given by: $y_i = \left(2.6 \times 10^{-2}\right) x_i + 4.6 \times 10^{-4}$.

This data could be used to construct the statistical models: $\Pr(x_i, y_i)$ (i.e. joint probability), $\Pr(x_i|y_i)$ (e.g. for maximum likelihood detection), or $\Pr(y_i|x_i)$ (e.g. as a sensor noise model). In particular, the noise model, $\Pr(y_i|x_i)$, is an input to Baysian estimation algorithms such as the particle filter [20], [21].

An example which demonstrates using the correlated data is a simple binary detector. Consider a detector which outputs true if the camera occurrence rate is less than once per 100 s, using the information of whether the $\geq 2\,\mu$m concentration is less than 100 per liter. Used on this data set, the detector would be correct 66% of the time with a 42% false positive ratio and 28% false negative ratio. While this alone is far from ideal, it could be combined in a multiple agent decision framework containing other sensors such as passive infrared detectors.

VI. CONCLUSION

We have shown via experimentation that local human activity, measured visually by a camera, is correlated to the concentration of coarse particles, particularly those $\geq 2.5\,\mu$m. These types of particles can be easily sensed with low-cost PM sensors such as the DSM501A. Furthermore, if smoother data is needed, more low-cost sensors could be added, while still being economical.

We have also described the hardware developments which enable real-time reporting of PM sensor data to a central server using wireless technology. These hardware modules, while primarily intended for air quality monitoring, can also indicate occupancy as we have shown by this article. We believe it will prove to be yet another piece of valuable information for smart buildings.

An important next step will be to investigate the causality of the variables we are measuring. Of particular concern is describing the hidden variable which is measured by both the camera and PM sensor. That is, we have not determined whether how dependent the PM measurement is on the occupancy of the entire room versus the local occupancy directly in front of the camera. An experiment which could disambiguate the two would be to also place several PM sensors away from the test apparatus and measure the PM level of the entire room. For a permanent deployment, multiple sensors in the same room will be useful for estimating this background PM. We would then look at deviations of the local PM from the background level.

Also, more detailed calibration will be needed if we wish to compare our results against other published PM results. For instance, we can calibrate the sensors in varied locations which may have different average amounts of PM or different material composition of the PM. We should also fully study the sensor characteristics of the DSM501A and the amount of manufacturing variability.

REFERENCES

[1] A. Zanobetti and J. Schwartz, "The effect of fine and coarse particulate air pollution on mortality: a national analysis," *Environmental Health Perspectives*, vol. 117, no. 6, p. 898, 2009.

[2] J. Qian and A. R. Ferro, "Resuspension of dust particles in a chamber and associated environmental factors," *Aerosol Science and Technology*, vol. 42, no. 7, pp. 566–578, 2008.

[3] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1067–1080, 2007.

[4] W. W. Nazaroff, "Indoor particle dynamics," *Indoor air*, vol. 14, no. s7, pp. 175–183, 2004.

[5] J. Kildesø, P. Vinzents, T. Schneider, and N. P. Kloch, "A simple method for measuring the potential resuspension of dust from carpets in the indoor environment," *Textile research journal*, vol. 69, no. 3, pp. 169–175, 1999.

[6] T. L. Thatcher and D. W. Layton, "Deposition, resuspension, and penetration of particles within a residence," *Atmospheric Environment*, vol. 29, no. 13, pp. 1487–1497, 1995.

[7] H. Fromme, D. Twardella, S. Dietrich, D. Heitmann, R. Schierl, B. Liebl, and H. Rüden, "Particulate matter in the indoor air of classroomsÛexploratory results from munich and surrounding area," *Atmospheric Environment*, vol. 41, no. 4, pp. 854–866, 2007.

[8] J. Qian, D. Hospodsky, N. Yamamoto, W. W. Nazaroff, and J. Peccia, "Size-resolved emission rates of airborne bacteria and fungi in an occupied classroom," *Indoor air*, 2012.

[9] C. Gomes, J. Freihaut, and W. Bahnfleth, "Resuspension of allergen-containing particles under mechanical and aerodynamic disturbances from human walking," *Atmospheric Environment*, vol. 41, no. 25, pp. 5257–5270, 2007.

[10] B. Hu, J. D. Freihaut, W. Bahnfleth, C. A. Gomes, and B. Thran, "Literature review and parametric study: Indoor particle resuspension by human activity," *Proceedings of Indoor Air*, pp. 1541–1545, 2005.

[11] G. A. Loosmore, "Evaluation and development of models for re-suspension of aerosols at short times after deposition," *Atmospheric environment*, vol. 37, no. 5, pp. 639–647, 2003.

[12] A. R. Ferro, R. J. Kopperud, and L. M. Hildemann, "Source strengths for indoor human activities that resuspend particulate matter," *Environmental science & technology*, vol. 38, no. 6, pp. 1759–1764, 2004.

[13] M. Budde, M. Busse, and M. Beigl, "Investigating the use of commodity dust sensors for the embedded measurement of particulate matter," in *Ninth International Conference on Networked Sensing Systems (INSS)*. IEEE, 2012, pp. 1–4.

[14] C.-H. Wang, Y.-K. Huang, X.-Y. Zheng, T.-S. Lin, C.-L. Chuang, and J.-A. Jiang, "A self sustainable air quality monitoring system using wsn," in *5th IEEE International Conference on Service-Oriented Computing and Applications (SOCA)*. IEEE, 2012, pp. 1–6.

[15] Y. Fu and B. Hallberg, "A personal environment monitoring system for pulmonary disease management," in *4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE)*. IEEE, 2010, pp. 1–4.

[16] S. Dawson-Haggerty, X. Jiang, G. Tolle, J. Ortiz, and D. Culler, "smap: a simple measurement and actuation profile for physical information," in *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2010, pp. 197–210.

[17] "Dust sensor module P/N: DSM501 Specifications," 2013. [Online]. Available: http://www.samyoungsnc.com/products/3-1%20Specification%20DSM501.pdf

[18] "PPD 20V (Particle Sensor Unit)," 2013. [Online]. Available: http://www.shinyei.co.jp/stc/optical/main_ppd20v_e.html

[19] "GT-526S Handheld Particle Counter," 2013. [Online]. Available: http://www.metone.com/documents/GT-526S-Datasheet.pdf

[20] D. Fox, J. Hightower, H. Kauz, L. Liao, and D. Patterson, "Bayesian techniques for location estimation," in *Proceedings of The 2003 Workshop on Location-Aware Computing*, 2003, pp. 16–18.

[21] S. Maskell and N. Gordon, "A tutorial on particle filters for on-line nonlinear/non-gaussian bayesian tracking," in *Target Tracking: Algorithms and Applications (Ref. No. 2001/174), IEEE*. IET, 2001, pp. 2–1.