Decoding Physical and Cognitive Impacts of Particulate Matter Concentrations at Ultra-fine Scales

Shawhin Talebi[1*], David J. Lary[1], Lakitha O. H. Wijeratne[1†], Bharana Fernando[1†], Tatiana Lary[1†], Matthew D. Lary[1†], John Sadler[1†], Arjun Sridhar[1††], John Waczak[1††], Adam Aker[1††] and Yichao Zhang[1††]

[1*]Hanson Center for Space Sciences, The University of Texas at Dallas, 800 W Campbell Rd, Richardson, 75080, TX, USA.

*Corresponding author(s). E-mail(s): shawhin.talebi@utdallas.edu;
Contributing authors: david.lary@utdallas.edu;
†These authors contributed equally to this work.

Abstract

The human body is an incredible and complex sensing system. Environmental factors trigger a wide range of automatic neurophysiological responses. Biometric sensors can capture these responses in real time, providing clues to the underlying biophysical mechanisms. Here we show biometric variables can be used to accurately estimate ultra-local particulate matter concentrations in the ambient environment with high fidelity ($r^2 = 0.91$) and that smaller particles are better estimated than larger ones. Inferring environmental conditions solely from biometric measurements allows us to disentangle key interactions between the environment and the body. A deeper understanding of these interactions can have countless important applications in public health, preventative healthcare, city planning, human performance, and much more. By tapping into our body's 'built-in' sensing abilities, we can gain insights to how our environment influences our physical health and cognitive performance.

Keywords: Biometrics, Particulate Matter, Machine Learning

1 Introduction

Over 4 million premature deaths worldwide were attributed to outdoor air pollution in 2016 [1]. In 2019, 99% of the global population resided in areas that fell short of the World Health Organization (WHO) air quality guidelines [1]. There has been mounting evidence that poor air quality negatively impacts respiratory, cardiovascular, and cerebrovascular health [2–7]. Further, there is emerging evidence on the impact of poor air quality on neurological outcomes including chronic diseases (e.g. Alzheimer's disease and dementia) [2, 8, 9] and acute cognitive impairment [10–14].

Although several large-scale epidemiological studies show the negative effects of air pollution on physical and cognitive health [2–7], these studies largely focused on coarse spatial (∼10 miles) and temporal (∼1 day) scales. Much less research focuses on ultra-local spatial (∼1 m) and temporal (∼10 seconds) scales that make simultaneous environmental and holistic biometric observations of the human physiological responses.

Before an extreme result such as a disease occurs, poor air quality already negatively impacts human physical and cognitive performance [10–14]. Through this work, we investigate how air pollution impacts human performance by examining the relationship between environmental air quality measurements and automatic physiological responses at ultra-fine scales.

This study extends past works that examined interactions of cardiovascular variables such as heart rate (HR), heart rate variability (HRV), and blood pressure (BP) with air quality on fine scales [15–17]. The main contribution of this study is that we augment cardiovascular markers with other biometrics, including electroencephalography (EEG), pupillometry, galvanic skin response (GSR), body temperature, blood oxidation, and respiration rate. This extended set of variables provides insight into both the cardiovascular and cognitive status of the participant. A study of air quality and human physiology at the ultra-local level may shed light on the biophysical mechanisms that underlie their interactions.

2 Results

In this work we used a data-driven experimental paradigm to develop and explore several empirical machine learning models which describe the connection between ambient air particulate matter (PM) concentrations and the biometric variables of an individual breathing that air. Due to logistical constraints imposed by the COVID-19 pandemic, we were only able to collect data from one participant. Additional participants will be included in future research. Two factors, however, mitigate the limited population size in this study. First, the data collection took place over three days, which allowed for contextual variability. Furthermore, the participant repeatedly circled the same trail, allowing for multiple observations of identical spatial positions and 360-degree changes in wind direction angles.

The estimated PM values included: PM_1, $PM_{2.5}$, PM_4, PM_{10}, PM_{Total}, and 45 different PM size bins ranging of 0.18 – 10 μm measured in $\mu g/m^3$, as well as particle count density (dCn) measured in particles per m^3. For model development, 329 biometric predictor variables were available. Two subsets of 9 biometric predictor variables were used in training a set of empirical machine learning models. The first subset includes EEG variables, and the second subset does not. The cognitive effects of air quality can be identified by evaluating predictive models with and without EEG values.

Each machine learning model used was a trained ensemble of decision trees for multi-variate non-linear non-parametric regression with full hyperparameter optimization [18–23]. The empirical models are evaluated using two key metrics. First, the model accuracy assessed using the squared correlation coefficient (r^2) between the model prediction and the true PM values. Second, a ranking of predictor variable importance obtained as the weighted average importance of each predictor across the ensemble.

We first evaluated the six machine learning models for particulate matter (PM) values which estimated: the particle count density (dCn), PM_1, $PM_{2.5}$, PM_4, PM_{10}, and PM_{Total}. 329 biometric predictor variables were used as model inputs including: delta (1 – 3 Hz), theta (4 – 7 Hz), alpha (8 – 12 Hz), beta (13 – 25 Hz), and gamma (25 – 70 Hz) band power densities for each of the 64 EEG electrodes, body temperature, galvanic skin response (GSR), heart rate (HR), heart rate variability (HRV), respiration rate (RR), peripheral capillary blood oxygen saturation (SpO_2), average pupil diameter, the difference between the left and right pupil diameters (anisocoria), and the 3D spatial distance between the left and right pupil centers (vergence eye movement). Then, using an Occam's razor principle, the top 9 important biometric predictor variables were used to train an additional six models for the same PM variables.

The best performing model using the top 9 EEG and non-EEG biometric predictors was for PM_1. This model had the highest accuracy with a validation dataset $r^2 = 0.91$. Comparison plots between estimated and ground truth PM_1 values are given in Figure 1. In the top-left plot, the estimated and true PM_1 concentrations in both the training (blue circles) and validation (red pluses) datasets closely follow to the perfect fit (black) line. In the top-right plot, the quantile-quantile comparison shows the distribution of measured PM_1 values closely resembles the distribution of estimated PM_1 values. Finally, in the bottom plot, the time series of the estimated PM_1 values (dashed red line) tracks very closely to the true values (solid black line) over seven different trials spanning three separate days.

The performance of the PM_1 and five other PM models in this cohort are ranked in the left panel of Figure 2. The training and independent validation dataset performances are plotted in blue and orange, respectively, and sorted in descending order of independent validation performance. As previously discussed, PM_1 measured in $\mu g/m^3$ was best reproduced by the 9 biometric predictors (validation $r^2 = 0.91$). The empirical models based on

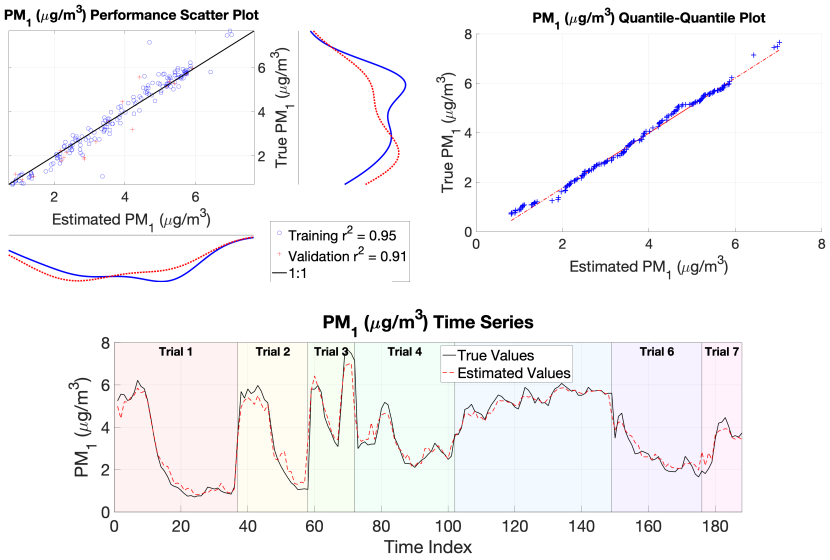

Fig. 1 Top performing model (PM_1) plots comparing predict and ground truth values. (**Top − Left**) scatter plot of true versus predicted PM_1 values. A perfect fit is indicated by the 1:1 line shown in black. Training data are shown as blue circles and validation data are plotted as red pluses. (**Top − Right**) quantile-quantile plot of true versus predict PM_1 values. Identical true and predicted distributions would results in a perfect y=x line. (**Bottom**) Time series plot of true PM_1 values (solid black line) and predicted PM_1 values (dashed red line).

the same biometric predictors were less able to accurately estimate the larger PM_{10} (validation $r^2 = 0.67$) values and PM_{Total} (validation $r^2 = 0.72$) which is dominated by PM_{10} due to the larger masses. The poor performance of these models could be explained the fact that there are significantly fewer large particles than small particles, and thus the larger particles are not as well mixed as the far more numerous and well mixed smaller particles. Because of their greater bulk, larger particles settle more quickly. As a result, the concentrations of large particles collected by the survey vehicle and those inhaled by the subject a few meters away are likely to differ more than for the smaller particles. Second, it's possible that the larger particles have less of an impact on the participant's physical and cognitive state because they are less likely to penetrate deeply into the respiratory and circulatory systems [26].

Each of the six empirical machine learning models has an associated predictor importance ranking, which quantifies the role of individual input predictor variables in estimating the respective PM target variable. The aggregated ranking of top predictors, shown in the right plot in Figure 2, elucidates which biometric variables are most helpful to the empirical models in discerning PM values. The most important predictor variable in estimating PM values was the body temperature measured at the participant's right temple. Surprisingly, the respiratory variable HRV played less of a role. Other important biometrics included GSR and the distance between the pupil centers of the eyes. GSR is

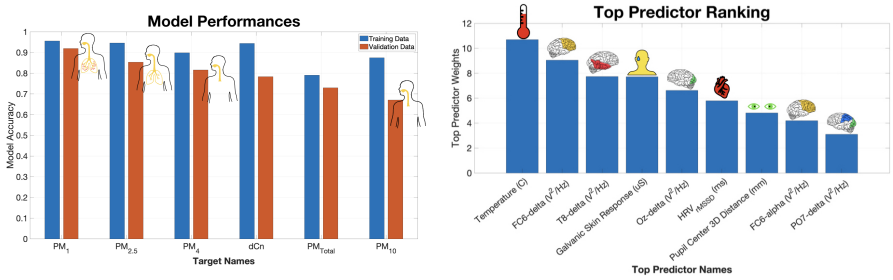

Fig. 2 Summary of empirical PM concentration models estimated from 9 EEG and non-EEG biometric predictor variables. (**Left**) Ranking of model performance defined as squared correlation coefficient between predicted and true PM values. Training and validation dataset performances for each model are shown in blue and orange, respectively. Sorting is based on validation dataset performance. Overlaid graphics indicate the deposition of the respective PM size bins in the airways [26]. (**Right**) Predictor importance ranking aggregated across all 6 models.

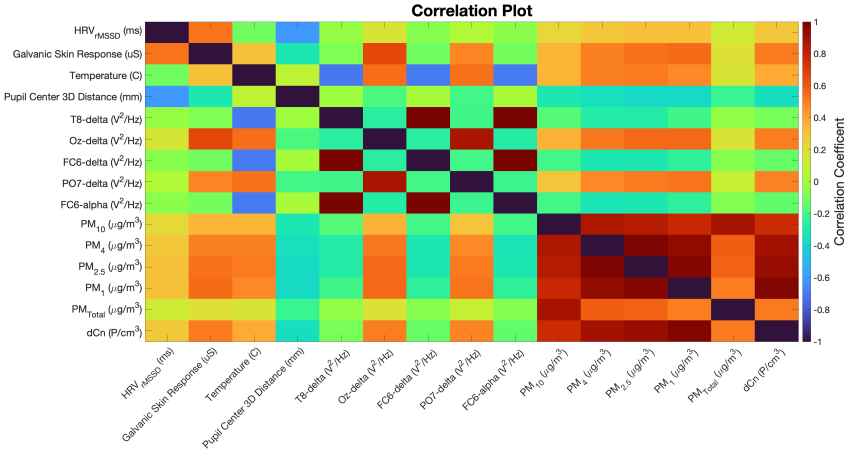

Fig. 3 Correlation plot of top 9 EEG and non-EEG biometric predictor variables, along with 6 target PM variables. Positively correlated variable pairs are indicated by a red box, negatively correlated pairs are shown by blue boxes, and non-correlated pairs have green boxes.

a strong correlate of body temperature. While the distance between the pupil centers is a proxy for vergence eye movements, which have been associated with attentional load and to be a strong predictor of cognitive status [24, 25]. EEG variables found to play an important role in estimating PM values were the delta band (1 – 3 Hz) power densities for the FC6, T8, and Oz electrodes. FC6 is above the frontal cortex on the right side of the head, T8 corresponds to the right temporal lobe, and Oz sits on top of the primary visual cortex.

Correlations between predictor and target variables are visualized as a color filled correlation plot in Figure 3. As seen by the red-orange streaks in the bottom-left and top-right of the correlation plot, HRV, GSR, temperature, and the delta power density of the Oz and PO7 electrode signals have strong

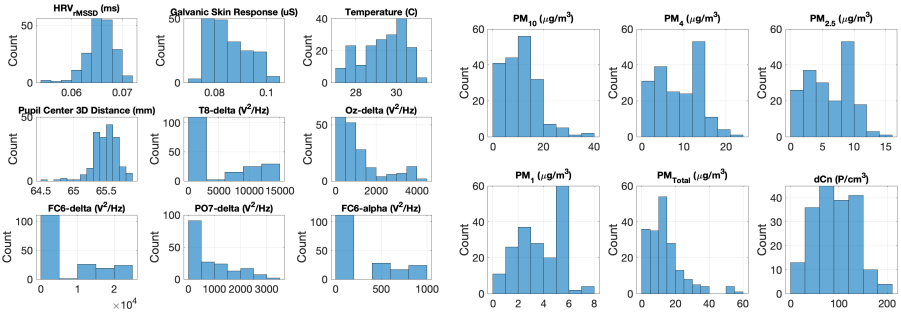

Fig. 4 (**Left**) Histograms of 9 EEG and non-EEG predictor variables. Plots are titled by variable name and its physical units. (**Right**) Histograms of 6 different PM target variables variables. Plots are titled by variable name and its physical units.

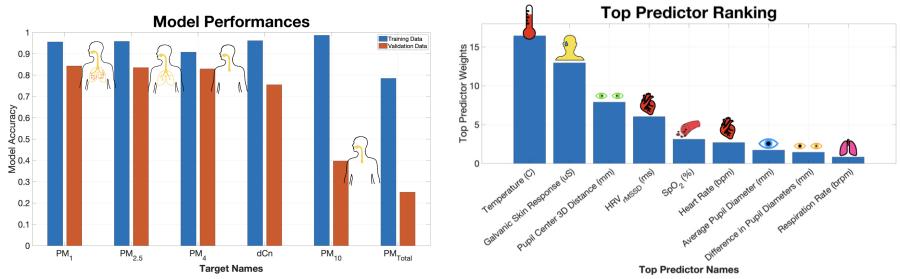

Fig. 5 Summary of empirical PM concentration models estimated from 9 non-EEG biometric predictor variables including eye tracking, respiratory, and other physiological variables. (**Left**) Ranking of model performance defined as squared correlation coefficient between predicted and true PM values. Training and validation dataset performances for each model are shown in blue and orange, respectively. Sorting is based on validation dataset performance. Overlaid graphics indicate the deposition of the respective PM size bins in the airways [26]. (**Right**) Predictor importance ranking aggregated across all 6 models.

positive correlations with all target variables except PM_{Total}. In other words, as these predictor variables increase, so do the corresponding PM target variables. PM target variables show the greatest negative correlation with the 3D spatial distance between left and right pupil centers. Suggesting that the pupils tend to converge with an increase in PM concentrations. Lastly, of all the target variables, PM_{Total} is most strongly correlated with PM_{10} values, which reflects the strong contribution of PM_{10} particles to PM_{Total}.

Histograms for both predictor and target variables are displayed in Figure 4. Plots are titled by the variable name and its respective physical units. From the target PM variable histograms in the right plot of Figure 4, the mass scales of different particle sizes are evident. Namely, the larger sized PM_{10} particles vary over a much larger range ($0 - 40\ \mu g/m^3$) than the smaller PM_4 ($0 - 20\ \mu g/m^3$), $PM_{2.5}$ ($0 - 15\ \mu g/m^3$), and PM_1 particles ($0 - 8\ \mu g/m^3$). This further explains the strong influence of PM_{10} values on PM_{Total}.

Next, an additional set of six empirical machine learning models for the same set of PM targets (dCn, PM_1, $PM_{2.5}$, PM_4, PM_{10}, and PM_{Total}) were

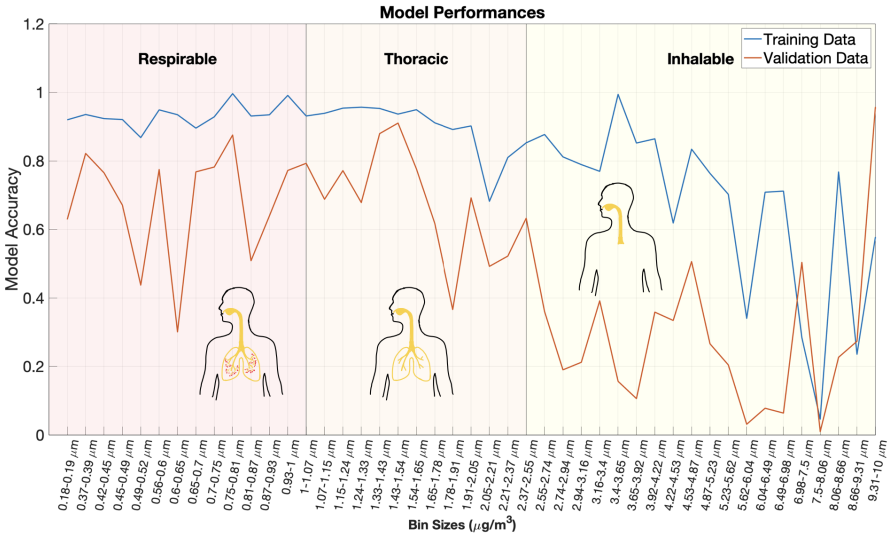

Fig. 6 Model accuracies plotted against bin size. 45 separate PM models were trained for size bins ranging from 0.18 to 10 micrometers. PM values were estimated solely from 9 non-EEG biometric variables. Training dataset performance is plotted as a blue line and validation dataset performance is plotted in orange. A clear drop in model performance is observed between 2 – 3 micrometers. Overlaid graphics indicate the deposition of the respective PM size bins in the airways [26, 27].

evaluated, except this time the PM targets were estimated from 9 non-EEG biometric predictor variables (body temperature, GSR, HR, HRV, RR, SpO_2, average pupil diameter, difference between left and right pupil diameters, and the 3D spatial distance between left and right pupil centers).

The model performance ranking for the six empirical PM models estimated from the 9 non-EEG biometric predictor variables is shown in the left panel of Figure 5. We see that the smaller particles are better estimated by the non-EEG biometrics. Again, this result may be due to better mixing of smaller particles or to deeper penetration of those particles into the respiratory system or both.

Comparing the performance rankings in Figure 2 and Figure 5, there are clear changes in model accuracies. All models with the exception of PM_4 exhibit a drop in performance. The largest drop occurs for the already poor performing PM_{Total} (drop in validation $r^2 = 0.47$) and PM_{10} (drop in validation $r^2 = 0.28$) models.

There is overlap between the importance rankings of Figure 2 and Figure 5. In both cases, body temperature is the most significant predictor of the PM values. Additionally, GSR maintains its order in the ranking as the 2nd most important non-EEG predictors. Although respiratory variables such as HRV and HR appear in the top six of the importance ranking, these variables trail behind temperature, GSR, and the distance between the eye pupil centers.

Fig. 7 Data collection images. (**Left**) Custom made backpack to house biometric devices and recording computer. (**Middle**) Participant and environmental survey vehicle riding in tandem during data collection. (**Right**) Environmental sensors organized in trunk of electric survey vehicle.

The observation that smaller particles are better estimated than larger sized particles, is explored further by evaluating model performances for finer scaled size bins. Here, 45 models were trained to estimate different PM size bins ranging from 0.18 to 10 micrometers using the 9 non-EEG biometrics listed above. Model accuracy is plotted against bin size in Figure 6. Training and validation accuracies are plotted as blue and orange lines, respectively. The regional depositions of each particle size bin is indicated by a label and background shading [26, 27]. The smallest particles (PM_1) are classified as respirable and can penetrate to the alveoli. The next smallest size bin is thoracic ($PM_{2.5}$) which consists of particle penetrating into the bronchioles. The largest size bin are the inhalable particles (PM_{10}) which can enter into the nose, mouth, and trachea.

There is a clear drop in both training and validation dataset accuracies for size bins between 2 to 3 micrometers, corresponding to thoracic and inhalable particles. For particle size bins above this drop, there is large degree of variation in model performances, however most have poor performance with a validation r^2 below 0.4. While the results may imply that smaller particles have a greater impact on physiological systems due to their deeper deposition, that conclusion cannot be reached based upon the present data. The drop in performance for larger particles may be explained in part or completely by the fact that smaller particles are more plentiful and better mixed. An evaluation of the relative contributions of each of these factors requires further investigation.

3 Materials & Methods

3.1 Holistic Sensing

The data in this study are a subset of a holistic biometric and environmental sensing paradigm. The aim of holistic sensing is to capture all relevant information about a system of interest. The full sensor array includes biometric monitors such as: electroencephalography (EEG), eye tracking glasses, electrocardiography (ECG), galvanic skin response (GSR), body temperature, blood oxygen saturation, and heart rate (Figure 8), in addition to environmental factors such as: particulate matter, chemical composition of air, temperature,

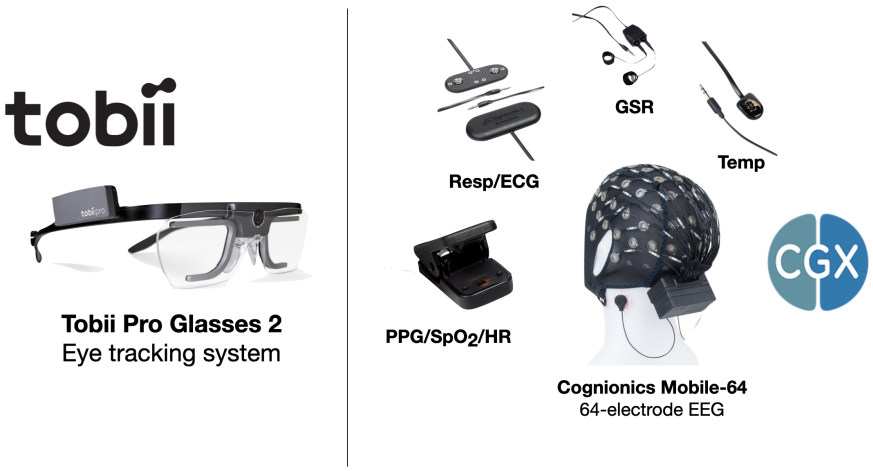

Fig. 8 Biometric sensing systems. (**Left**) Tobii Pro Glasses 2 eye tracking system. This instrument performs eye tracking data, pupillometry, and provides two videos streams of the participant's POV and eyes, respectively. (**Right**) Cognionics Mobile-64 and AIM2 systems. Sensing suite includes 64-electrode EEG, PPG which measures SpO_2 and HR, respiration/ECG sensors, GSR, and temperature probe.

pressure, humidity, visible light spectrum, and more (Figure 9). After processing raw sensor recordings, the full sensor array has a feature space approaching 20,000 variables (∼16,500 biometric and ∼2,000 environmental). In the present study we focus on a relatively small subset, consisting of 329 biometric and 51 environmental variables.

The biometric sensing suite used in this research aims to comprehensively capture the physiological and cognitive status of the participant, without restricting the participant's actions, movements, or decision making. The goal is to gather the maximum amount of information with minimal interruption of normal behaviors. Biometric sensors are placed on the participant in such a way to allow for unrestricted mobility (Figure 10). Sensor recording units and other devices are organized in a backpack worn by the participant that all together weighs less than 10 lbs (Left panel in Figure 7).

Over 100 biometric markers are measured at sampling rates of 500 Hz and 100 Hz. These quantities are processed to derive over 329 variables for the present analysis. This holistic biometric sensing suite integrates two independent sensing systems (Figure 8). Eye tracking is recorded 100 times a second using the Tobii Pro Glasses 2. Data from the glasses produced average pupil diameter, the difference in pupil diameter between left and right eyes, and the 3D spatial distance between pupil centers. All other biometric data are measured 500 times a second using the Cognionics Mobile-64 and AIM2 systems. These systems include a 64-electrode EEG, temperature sensor, respiration sensor, Photoplethysmogram (PPG), and galvanic skin response (GSR) measurement. Heart rate and SpO_2 values are automatically computed by the

Fig. 9 Images of environmental sensing systems. Fidas® Frog Fine Dust Monitoring System measures particulate matter concentrations at 100 different size bins. The AIRMAR 220WX WeatherStation® Instrument samples barometric pressure, wind speed and direction, ambient temperature, and more. The 2B Technologies Black Carbon Photometer measures atmospheric black carbon particulates using long-path photometry. The 2B Technologies Model 205 Dual Beam Ozone sensor is a UV-based ozone monitor. The Konica Minolta CL-500A Illuminance Spectrometer measures the spectral irradiance from 360 to 780 nm at every nanometer. The portable mass spectrometer was constructed by the UNT Laboratory of Imaging Mass Spectrometry and measures charge mass ratios ranging 1 - 300 amu. The 2B Technologies Model 405 nm $NO_2/NO/NO_x$ Monitor™ directly measures atmospheric Nitrogen Dioxide (NO_2) and Nitric Oxide (NO). The LI-COR LI-850 Gas Analyzer measured CO_2 and water vapor in the air.

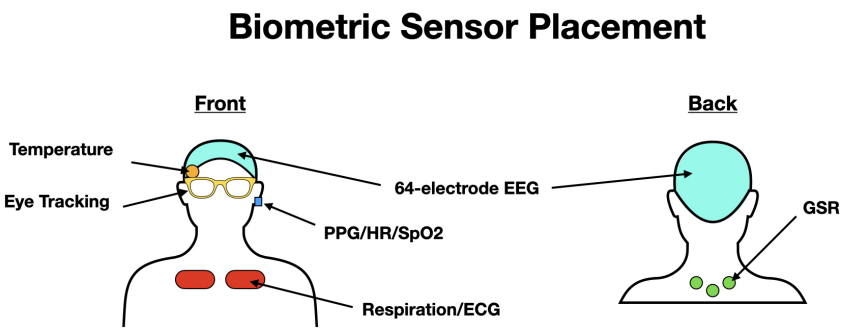

Fig. 10 Schematic of biometric sensor placement on participant. (**Left**) Cartoon of front participant view. The 64-electrode EEG sits on the participant's head. A temperature probe is placed under the EEG cap on the right temple. Eye tracking glasses are carefully placed on participant, avoiding EEG electrodes. PPG sensor is secured to left ear lobe. Respiration sensors are place near the top of the chest. (**Right**) Cartoon of back participant view. GSR sensors are placed below the back of the neck.

AIM2 system using the PPG. Heart Rate Variability (HRV) and Respiration Rate (RR) values are derived from respiration sensor data with a custom

MATLAB script. All biometric data were down-sampled to 1/30 Hz (every 30 seconds) to match particulate matter recordings.

A holistic evaluation of an environmental setting is the ultimate goal of the environmental sensing suite used in these studies. This suite brings together several sensing packages (Figure 9). However, due to its significant societal relevance, for this study we focus on particulate matter (PM) concentrations recorded using the Fidas® Frog fine dust monitoring system. This instrument simultaneously measures PM mass fractions of PM_1, $PM_{2.5}$, PM_4, PM_{10}, and a size distribution within a size range of 0.18 - 100 micrometers as well as the total particle count density (dCn). PM data was recorded at sampling rate of 1 Hz and down-sampled to 1/30 Hz (every 30 seconds).

3.2 Data Collection

Biometric data collection was restricted to a single participant due to logistical constraints arising from the COVID-19 pandemic. However, future works will include data from multiple participants. The small population size in the present study is mitigated by two factors. First, data was collected over three separate days, providing a range of contexts. Additionally, the participant circled the same trail multiple times, offering multiple observations of identical positions and 360-degree changes in wind direction angles.

Data were collected while the participant rode a bicycle in a dynamic outdoor setting. An electric survey vehicle equipped with a suite of environmental sensors followed safely behind the participant during all rides (Middle image in Figure 7. Although several dimensions of the environmental context were sampled (e.g. ambient light, temperature, pressure, mass spectra, etc.), here we focus on the relationship between particulate matter values and biometric variables. Additional relationship will be explored in future works.

Data collection took place in May and June of 2021 at Breckenridge Park located in Richardson, TX over three separate days which included four to five trials per day. The first two trials consisted of two minutes of eyes closed and eye open baseline biometric measurements, respectively. The third trial consisted of a "warm-up" ride, where the participant cycled to a public bike trail in tandem with the electric survey vehicle. Additional trials consisted of the participant repeatedly cycling a one-mile loop on a public bike trail. The participant was free to stop cycling at their discretion. Data collection was halted whenever cycling stopped. If the participant chose to continue, a new data collection trial was initiated.

The complete dataset consists of 188 data records collected every 30 seconds (total time of about 1.5 hours) with 329 biometric predictor variables and 51 PM target variables. Biometric predictor variables include: delta (1 – 3 Hz), theta (4 – 7 Hz), alpha (8 – 12 Hz), beta (13 – 25 Hz), and gamma (25 – 70 Hz) band power densities for each of the 64 EEG electrodes, body temperature, GSR, HR, HRV, RR, SpO_2, average pupil diameter, difference between left and right pupil diameters, and the 3D spatial distance between left and right pupil centers. Environmental PM target variables include: PM_1, $PM_{2.5}$, PM_4, PM_{10},

PM_{Total}, and 45 different PM size bins ranging of 0.18 – 100 μm measured in $\mu g/m^3$, as well as particle count density (dCn) measured in P/cm^3. The data is made publicly available at the Zenodo datastore: https://zenodo.org/record/6326357#.Yieu4RPMJb8.

Ethical approval declarations All experimental protocols were approved by The University of Texas at Dallas Institutional Review Board and informed consent was obtained from the study participant.

3.3 Model Development

All models of PM concentration are obtained by an ensemble of decision trees for regression with a hyperparameter optimization process [18–23]. 90% of the data is used for training, while 10% is help back as an independent validation dataset. Scripts for model training are freely available at the GitHub repository: https://github.com/mi3nts/DUEDARE.

4 Conclusion

The human body and environment form a complex ecosystem. A key aspect of this system is air quality and the effects it has on our bodies. Environmental factors trigger physiological responses that can be detected by holistic biometric sensing. Here we used an ultra-fine holistic sensing paradigm to show particulate matter concentrations in the ambient environment can be accurately estimated using only nine biometric variables. In addition, smaller particles were found to be more accurately estimated. Two potential causes may explain this result. First, smaller particles are much more abundant and well mixed in the ambient environment than larger ones, thus resulting in a greater similarity between particles inhaled by the participant and collected by the survey vehicle. Secondly, smaller particles can deposit into the respiratory system more deeply, and may have a greater impact on the body. Further investigation is needed to assess the relative contributions, if any, of these two factors, since they are not mutually exclusive.

Although the present work shows preliminary findings from a single participant over multiple days, future research will include data from multiple participants. Additionally, several other variables collected (e.g. ambient light, temperature, pressure, mass spectra, etc.) will be evaluated for their physiological interactions. By understanding the key interactions between the environment and the human body, health and performance can be improved across many different domains.

Supplementary information. The data and code has been made publicly available. The full data set is available at the Zenodo data store: https://zenodo.org/record/6326357#.Yieu4RPMJb8 and code is available at the GitHub: https://github.com/mi3nts/DUEDARE.

Declarations

- This research was funded by the following grants: The US Army (Dense Urban Environment Dosimetry for Actionable Information and Recording Exposure, U.S. Army Medical Research Acquisition Activity, BAA CDMRP Grant Log #BA170483). EPA 16th Annual P3 Awards Grant Number 83996501, entitled Machine Learning Calibrated Low-Cost Sensing. The Texas National Security Network Excellence Fund award for Environmental Sensing Security Sentinels. SOFWERX award for Machine Learning for Robotic Teams. Support from the University of Texas at Dallas Office of Sponsored Programs, Dean of Natural Sciences and Mathematics, and Chair of the Physics Department are gratefully acknowledged. The authors acknowledge the OIT-Cyberinfrastructure Research Computing group at the University of Texas at Dallas and the TRECIS CC* Cyberteam (NSF 2019135) for providing HPC resources that contributed to this research (https://utdallas.edu/oit/departments/circ/, accessed 02/22/2022.
- The authors declare no conflicts of interest
- **Ethical approval declarations** All experimental protocols were approved by The University of Texas at Dallas Institutional Review Board.
- Informed consent was obtained from the study participant.
- Data used in this study is publicly available at the Zenodo datastore: https://zenodo.org/record/6326357#.Yieu4RPMJb8
- Code to reproduce analysis and visualization in this work are available at the GitHub repository: https://github.com/mi3nts/DUEDARE
- Authors' contributions: Methodology, D.J.L., S.T., T.L.; software, S.T.; formal analysis, D.J.L., S.T., T.L., B.F.; data curation, S.T., D.J.L, L.O.H.W., B.F., T.L., M.D.L., J.S., A.S., A.A., Y.Z.; writing—original draft preparation, S.T., D.J.L; writing—review and editing, S.T., D.J.L, B.F., T.L., J.S.; visualization, S.T.; supervision, D.J.L.

References

[1] WHO. Ambient (outdoor) air pollution. Retrieved January 24, 2022, from https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

[2] Manisalidis I, Stavropoulou E, Stavropoulos A, Bezirtzoglou E. Environmental and Health Impacts of Air Pollution: A Review. Front Public Health. 2020;8:14. Published 2020 Feb 20. doi:10.3389/fpubh.2020.00014

[3] Orellano, P., Reynoso, J., Quaranta, N., Bardach, A., & Ciapponi, A. (2020). Short-term exposure to particulate matter (PM 10 and PM 2.5), nitrogen dioxide (NO 2), and ozone (O 3) and all-cause and cause-specific mortality: Systematic review and meta-analysis. Environment International, 142. https://doi.org/10.1016/J.ENVINT.2020.105876

[4] Daellenbach, K. R., Uzu, G., Jiang, J., Cassagnes, L. E., Leni, Z., Vlachou, A., Stefenelli, G., Canonaco, F., Weber, S., Segers, A., Kuenen, J. J. P., Schaap, M., Favez, O., Albinet, A., Aksoyoglu, S., Dommen, J., Baltensperger, U., Geiser, M., El Haddad, I., ... Prévôt, A. S. H. (2020). Sources of particulate-matter air pollution and its oxidative potential in Europe. Nature, 587(7834), 414–419. https://doi.org/10.1038/S41586-020-2902-8

[5] Pope, C. A., Brook, R. D., Burnett, R. T., & Dockery, D. W. (2010). How is cardiovascular disease mortality risk affected by duration and intensity of fine particulate matter exposure? An integration of the epidemiologic evidence. Air Quality, Atmosphere & Health 2010 4:1, 4(1), 5–14. https://doi.org/10.1007/S11869-010-0082-7

[6] Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Whitsel, L., & Kaufman, J. D. (2010). Particulate Matter Air Pollution and Cardiovascular Disease. Circulation, 121(21), 2331–2378. https://doi.org/10.1161/CIR.0B013E3181DBECE1

[7] Brook, R. D., Bard, R. L., Kaplan, M. J., Yalavarthi, S., Morishita, M., Dvonch, J. T., Wang, L., Yang, H. Y., Spino, C., Mukherjee, B., Oral, E. A., Sun, Q., Brook, J. R., Harkema, J., & Rajagopalan, S. (2013). The effect of acute exposure to coarse particulate matter air pollution in a rural location on circulating endothelial progenitor cells: results from a randomized controlled study. Inhalation Toxicology, 25(10), 587–592. https://doi.org/10.3109/08958378.2013.814733

[8] Schikowski T, Altuğ H. The role of air pollution in cognitive impairment and decline. Neurochem Int. 2020 Jun;136:104708. doi: 10.1016/j.neuint.2020.104708. Epub 2020 Feb 21. PMID: 32092328.

[9] Kelly FJ, Fussell JC. Air pollution and public health: emerging hazards and improved understanding of risk. Environ Geochem Health. 2015 Aug;37(4):631-49. doi: 10.1007/s10653-015-9720-1. Epub 2015 Jun 4. PMID: 26040976; PMCID: PMC4516868.

[10] Lavy, V., Ebenstein, A., & Roth, S. (2014). The Impact of Short Term Exposure to Ambient Air Pollution on Cognitive Performance and Human Capital Formation. https://doi.org/10.3386/W20648

[11] Chatzidiakou, L., Mumovic, D., & Dockrell, J. (2015). The Effects of Thermal Conditions and Indoor Air Quality on Health, Comfort and Cognitive Performance of Students. https://doi.org/978-0-9930137-3-7

[12] Zhang, X., Chen, X., & Zhang, X. (2018). The impact of exposure to air pollution on cognitive performance. PNAS, 115(37), 9193–9197. https://doi.org/10.1073/pnas.1809474115

[13] Shehab, M. A., & Pope, F. D. (2019). Effects of short-term exposure to particulate matter air pollution on cognitive performance. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-44561-0

[14] Künn, S., Palacios, J., & Pestel, N. (2021). Indoor Air Quality and Cognitive Performance. SSRN Electronic Journal. https://doi.org/10.2139/SSRN.3460848

[15] Buteau, S., & Goldberg, M. S. (2016). A structured review of panel studies used to investigate associations between ambient air pollution and heart rate variability. Environmental Research, 148, 207–247. https://doi.org/10.1016/J.ENVRES.2016.03.013

[16] Amoabeng Nti, A. A., Robins, T. G., Mensah, J. A., Dwomoh, D., Kwarteng, L., Takyi, S. A., Acquah, A., Basu, N., Batterman, S., & Fobil, J. N. (2021). Personal exposure to particulate matter and heart rate variability among informal electronic waste workers at Agbogbloshie: a longitudinal study. BMC Public Health 2021 21:1, 21(1), 1–14. https://doi.org/10.1186/S12889-021-12241-2

[17] de Paula Santos, U., Ferreira Braga, A. L., Artigas Giorgi, D. M., Amador Pereira, L. A., Grupi, C. J., Lin, C. A., Bussacos, M. A., Trevisan Zanetta, D. M., Hilário Do Nascimento Saldiva, P., &; Terra Filho, M. (2005). Effects of air pollution on blood pressure and heart rate variability: a panel study of vehicular traffic controllers in the city of São Paulo, Brazil. European Heart Journal, 26 (2), 193–200. https://doi.org/10.1093/EURHEARTJ/EHI035

[18] Breiman, L, Friedman, J H, Olshen, R A, and Stone, C J, 1984, Classification and regression trees: Wadsworth, Inc.

[19] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324

[20] Tin Kam Ho, The random subspace method for constructing decision forests. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 8, pp. 832-844, Aug. 1998, doi: 10.1109/34.709601.

[21] Friedman, Jerome H. Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics, vol. 29, no. 5, Institute of Mathematical Statistics, 2001, pp. 1189–232, http://www.jstor.org/stable/2699986.

[22] Freund, Y. and Schapire, R.E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. Journal of Computer and

System Sciences, 55, 119-139. http://dx.doi.org/10.1006/jcss.1997.1504

[23] Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(3), e1301. https://doi.org/10.1002/WIDM.1301

[24] Balaban, C. D., Kiderman, A., Szczupak, M., Ashmore, R. C., & Hoffer, M. E. (2018). Patterns of Pupillary Activity During Binocular Disparity Resolution. Frontiers in Neurology, 9(NOV), 990. https://doi.org/10.3389/fneur.2018.00990

[25] Huang, M. X., Li, J., Ngai, G., Leong, H. V., & Bulling, A. (2019). Moment-to-Moment Detection of Internal Thought from Eye Vergence Behaviour. http://arxiv.org/abs/1901.06572

[26] Poh, T. Y., Ali, N. A. T. B. M., Mac Aogáin, M., Kathawala, M. H., Setyawati, M. I., Ng, K. W., & Chotirmall, S. H. (2018). Inhaled nanomaterials and the respiratory microbiome: Clinical, immunological and toxicological perspectives. Particle and Fibre Toxicology, 15(1), 1–16. https://doi.org/10.1186/s12989-018-0282-0

[27] World Health Organization. Occupational and Environmental Health Team. (1999). Hazard prevention and control in the work environment: : airborne dust. World Health Organization. https://apps.who.int/iris/handle/10665/66147