Journal of Exposure Science & Environmental Epidemiology (2020) 30:962–970
https://doi.org/10.1038/s41370-019-0198-2

ARTICLE

Dynamic classification of personal microenvironments using a suite of wearable, low-cost sensors

Casey Quinn[1] · G. Brooke Anderson[1] · Sheryl Magzamen[1] · Charles S. Henry[2] · John Volckens[1,3]

Received: 2 April 2019 / Revised: 4 September 2019 / Accepted: 29 October 2019 / Published online: 14 January 2020

Abstract

Human exposure to air pollution is associated with increased risk of morbidity and mortality. However, personal air pollution exposures can vary substantially depending on an individual's daily activity patterns and air quality within their residence and workplace. This work developed and validated an adaptive buffer size (ABS) algorithm capable of dynamically classifying an individual's time spent in predefined microenvironments using data from global positioning systems (GPS), motion sensors, temperature sensors, and light sensors. Twenty-two participants in Fort Collins, CO were recruited to carry a personal air sampler for a 48-h period. The personal sampler was retrofitted with a GPS and a pushbutton to complement the existing sensor measurements (temperature, motion, light). The pushbutton was used in conjunction with a traditional time-activity diary to note when the participant was located at "home", "work", or within an "other" microenvironment. The ABS algorithm predicted the amount of time spent in each microenvironment with a median accuracy of 99.1%, 98.9%, and 97.5% for the "home", "work", and "other" microenvironments. The ability to classify microenvironments dynamically in real time can enable the development of new sampling and measurement technologies that classify personal exposure by microenvironment.

Introduction

Exposure to air pollution has been shown to have deleterious effects on human health [1–5]. Epidemiologic studies of air pollution have traditionally relied on exposure data reported by a central-site monitor. However, air pollution exposures are known to be spatially and temporally heterogeneous; thus, ambient air quality measures from a central-site monitor are not necessarily representative of an individual's exposure, particularly while indoors [6–10]. Recent research has focused on characterizing air pollution exposures with greater spatial and temporal resolution, often with the intent to quantify (post-hoc) air pollution levels within distinct microenvironments (e.g. "home", "work", "transit", etc.) [11–14]. These post-hoc evaluations have helped determine the timing, location (and source) of elevated exposures [15].

To estimate microenvironment exposures, studies have traditionally used surveys and diary methods. These methods are burdensome for study participants (particularly for children) and are subject to reporting bias and/or missing data [13, 16]. To alleviate the burden of self-reporting, several studies have used handheld [11, 12, 14, 17] or smartphone-based [18–21] global positioning systems (GPS) to estimate where participants spend their time. These studies have demonstrated that GPS sensors tend to provide accurate location data when outdoors. However, the positional accuracy is compromised by signal loss and drift in and around buildings. These GPS limitations create uncertainty and may also lead to exposure misclassification.

To account for GPS signal loss/drift, some studies define a spatially-explicit "buffer" around microenvironments such as "home" and "work" (rather than relying solely on geographic information system [GIS] geocoded building boundaries) [11, 12]. Other studies have utilized GPS-enabled buffers in conjunction with temperature

Supplementary information The online version of this article (https://doi.org/10.1038/s41370-019-0198-2) contains supplementary material, which is available to authorized users.

✉ John Volckens
john.volckens@colostate.edu

1 Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO 80523, USA

2 Department of Chemistry, Colorado State University, Fort Collins, CO 80523, USA

3 Department of Mechanical Engineering, Colorado State University, Fort Collins, CO 80523, USA

[11, 22, 23], light [24–27], or motion [28, 29] data to detect a transition between microenvironments. The algorithms developed to date have relied on data collection from multiple sources to collect the exposure (e.g. central-site monitors or continuous personal sampling devices) and microenvironment detection data (e.g. GPS, temperature, motion, light, etc.). These approaches are appropriate for analysis of microenvironment exposures in a post-hoc fashion (i.e., via post processing of data after collection); however, there are potential advantages to developing technologies and/or algorithms that can assist in dynamic (i.e., in real time) sampling and speciation of personal air pollution exposures as a function of microenvironment [30–33]. For example, the automated microenvironment aerosol sampler (AMAS) [34] is a wearable device containing four separate filter channels for collecting fine particulate matter ($PM_{2.5}$) with integrated GPS, motion, light, and temperature sensors. Devices such as the AMAS can use microenvironment-detection algorithms to dynamically adjust the flow of air through individual filters and collect $PM_{2.5}$ samples specific to a microenvironment.

The goal of this work was to develop an adaptive buffer size (ABS) algorithm capable of dynamically classifying an individual's location into one of three microenvironments: "home", "work", and "other" (a miscellaneous category encompassing all non-home and non-workplace activities). The ABS algorithm (Fig. 1, Fig. S1) uses GPS, motion, temperature, and light sensor data to determine a transition between microenvironments. Rather than relying on buffers with fixed radii around each location, the algorithm employs a dynamic buffer that is adjusted, in real time, based on GPS signal quality and environmental sensor data. The algorithm identifies transitions between microenvironments by detecting changes in motion, temperature, light, and GPS signal strength. Performance of the ABS algorithm was evaluated using data collected by participants in Fort Collins, CO and verified using a traditional diary method combined with a user pushbutton method for digital logging. Results from this work demonstrate that the ABS algorithm can successfully classify microenvironments to further advance the science of personal exposure monitoring.

Methods

The development and evaluation of the ABS algorithm involved four steps: (1) collection of a "reference" dataset of spatially and temporally resolved microenvironment transitions using a panel of volunteers, (2) data curation to address data entry errors (made by the participants) or gaps in the sensor data, (3) formulation of the ABS algorithm, and (4) evaluation of the ABS algorithm (and other, simpler algorithms) relative to the reference dataset.

Data collection

Adult participants were recruited from Colorado State University in Fort Collins, CO over the course of a

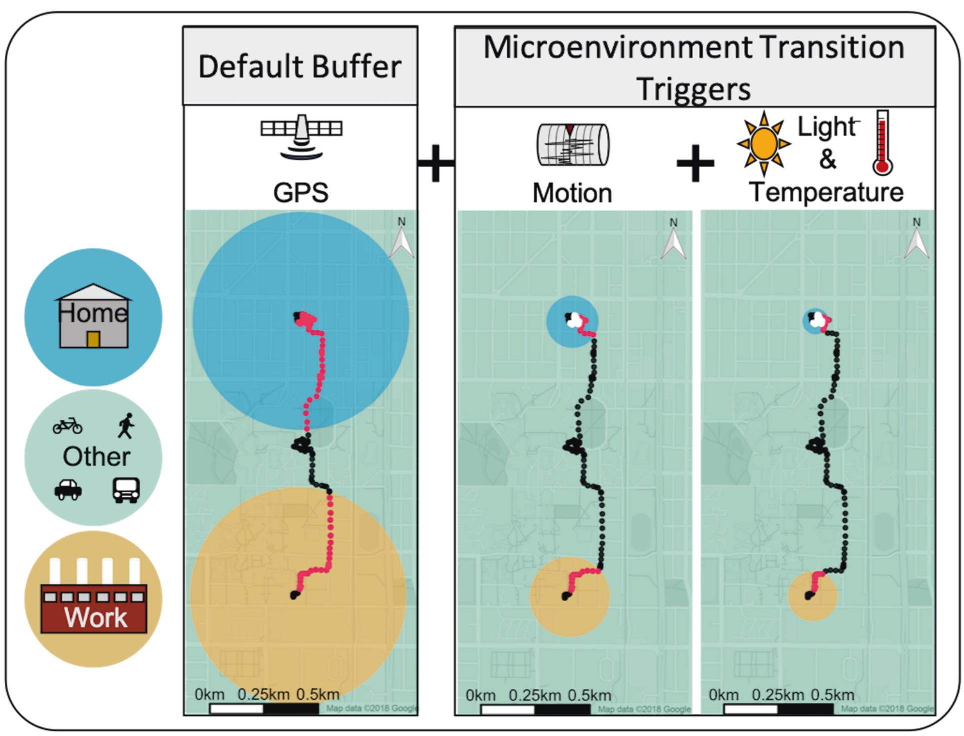

Fig. 1 Overview of the adaptive buffer size (ABS) algorithm. When only GPS location data are available, a large (default) buffer is used. Buffer radii are reduced when motion/environmental sensor data trigger an increased probability that a microenvironment transition has occurred. The black dots represent when an algorithm correctly classified the microenvironment. The red dots demonstrate points that were misclassified. The white dots represent points when the GPS accuracy had drifted, but since there was no motion detected the points were correctly classified within the microenvironment.

four-month period (Dec. 2015– Mar. 2016). The only requirements were that participants were working adults that did not live in the same residence as any of the other participants. Each participant provided their "home" and "work" addresses, which were used to determine a coordinate centroid (latitude, longitude) for these microenvironments using Google Maps. A Google Maps disclaimer states a maximum uncertainty of 15 m worldwide, and Goadarzi et al. [35] reported that Google maps has ~3 m accuracy in a major city, thus this method was deemed appropriate for the development and use of the microenvironment algorithm. Participants were asked to note the timing of each microenvironment transition (i.e., the times when they left or entered their "home" and "work" microenvironments) using a written diary. Participants were also asked to carry an ultrasonic personal aerosol sampler (UPAS) [36] for ~48 h. The UPAS measured and recorded ultraviolet light intensity (Silicon Labs, SI1145), temperature, (Bosch Sensortech, BME280), and motion (ST Microelectronics, LSM303). The UPAS was further modified (Fig. S2, Table S1) by adding a GPS sensor (Adafuit, 746) and a pushbutton recorder (Switchcraft Inc., ED913). The modified UPAS will be referred to as the personal monitor throughout the remainder of the manuscript. Due to the known limitations of the traditional diary method, participants were asked to depress the pushbutton during each microenvironment transition to serve as an additional source of reference data. The personal monitor logged all sensor data, including the pushbutton status, to non-volatile memory at a frequency of 0.2 Hz (every 5 s). The number of satellites tracked by the GPS was logged, in addition to the latitude and longitude, in degrees decimal minutes (DDM). Linear acceleration was measured on a ±2 g full scale for all three primary axes and the total acceleration was calculated using Eqn. S1. Light intensity data were collected as 100× the ultraviolet (UV) index. Temperature was logged in Celsius. Battery voltage was logged in volts. All study procedures were approved by the Institutional Review Board at the Colorado State University and informed consent was obtained from all participants.

Data curating

Microenvironment transition data were curated to establish a reference dataset for evaluation of the ABS algorithm. The GPS latitude and longitude data were converted to decimal degree (DD) formatting and used to calculate the distances from the "home" and "work" centroid coordinates using the Haversine formula [37]. The pushbutton and diary data were used to determine each participant's duration within each microenvironment and the precise timing of when they transitioned between microenvironments (e.g., when the participant left "home" and entered the "other" microenvironment). If the transition time recorded by the pushbutton and diary methods coincided, the pushbutton event was used unless there was a diary entry that specifically noted an error in the pushbutton transition time. The pushbutton data were also used to estimate time spent in each microenvironment as participants tended to record rounded times in their diaries or had discrepancies between the clock on the personal monitor and the clock or watch they were using. Some participants pressed the pushbutton on the personal monitor multiple times upon entering or exiting a microenvironment. In these instances, only one (the first) button press was counted as a microenvironment transition. When the personal monitor logged a button press without a corresponding diary entry (occasionally or for the entire sample), the data were manually evaluated to determine the GPS distance from both microenvironment centroids. If the GPS distance was less than 500 m from a "home" or "work" microenvironment when the button was pressed, the event was included and coded appropriately; otherwise, it was assumed that the button press was accidental and discarded. If the participant's diary included events that didn't have a corresponding pushbutton event, it was assumed that the participant made the microenvironment transition but forgot to press the button. Infrequently, participants failed to note a transition that caused an impossible result (i.e. a direct "home" to "work" transition). To estimate the time spent in each microenvironment for those datasets, the data were manually evaluated in an attempt to code the data based on GPS distances. Once the diary and pushbutton time-location data were compiled they were manually validated against the GPS data. For the reminder of the manuscript, the combined diary and pushbutton data are referred to as the reference dataset.

Adaptive buffer size (ABS) algorithm

The ABS algorithm dynamically adjusts the buffer size around both the "home" and "work" microenvironments to account for uncertainty in GPS positional accuracy. The ABS algorithm was formulated from the hypothesis that acute changes in GPS signal quality, acceleration, UV light intensity, and temperature are indicative of a transition between microenvironments (e.g., moving from indoors to outdoors). The ABS algorithm uses these triggers to adjust the buffer sizes in an attempt to optimize microenvironment classification (i.e., to maximize accuracy, sensitivity, and specificity as defined below). A detailed flow chart depicting the ABS algorithm operation is shown in Fig. S1.

The ABS algorithm defines three circular buffers with varying radial distances from the centroid of the "work" and "home" microenvironments; (1) large buffer: 500 m for both the "home" and "work" microenvironment, (2) medium buffer: 185 m and 120 m for the "work" and "home" microenvironment respectively, and (3) small buffer: 115 m and 60 m for the “work” and “home” microenvironment respectively. Wu et al. demonstrated that variations in GPS signal accuracy are on the order of 500 m or greater for a number of different GPS models when located in indoor environments, largely due to signal bounce/degradation [38]. Evaluation of our GPS signal showed similar performance indoors at several locations. For this reason, we set our large buffer to 500 m for periods when GPS signal quality was low. The medium and small buffers were set using the cumulative distributions of participant distance from “home” and “work” locations from the reference dataset (Fig. S3). The medium buffer represents the 85th percentile and the small buffer represents the 75th percentile.

The ABS algorithm uses a combination of categorical and continuous variables as triggers for detecting microenvironment transitions: (1) GPS signal status, (2) motion and environmental triggers, and (3) personal monitor charging status. The first step (see Fig. S1 for a process diagram) was to determine if the personal monitor was plugged into a wall outlet and charging by monitoring the battery voltage. The personal monitor had to be charged overnight and thus while charging, no microenvironment transition could occur. The next step was to determine if there was a GPS signal available. If no GPS signal was available, the last known location and large buffer were used; otherwise, motion and environmental triggers were used to check for microenvironment transitions. Motion and environmental trigger events are characterized as acute changes in acceleration, temperature, or light data that may be indicative of a transition between two microenvironments. To invoke a motion trigger, the absolute value of the change in total acceleration between each time step had to be greater than 0.37 g (95th percentile of cumulative distributions). A temperature trigger was defined as when the absolute value temperature change for a sixty-second period was greater than 1 °C and the current change in temperature had a sign change (negative, positive, or zero) as compared with the prior temperature change measurement. A UV light trigger is activated when there is a UV index [39] greater than 0.03 (95th percentile of cumulative distributions). Once a motion, UV light, or temperature trigger is detected, an individual timer for each trigger is reset and continues to increment once every second until the next trigger of that type is detected. Finally, in order for a transition to occur, the microenvironment must be confirmed for 30 s (six consecutive timestamps). This requirement helps reduce excessive “bouncing” in and out of microenvironments.

The ABS algorithm uses the status of the triggers to select one of the three radial buffers. If no transition triggers are active or the GPS signal quality is poor or unavailable, the default buffer with a radius of 500 m is used around each microenvironment. If motion and one of the environmental triggers is activated, the medium buffer is used. If triggers for motion, light, and temperature are all detected then the small buffer is used. This method allows the buffer size to be evaluated every 5 s and the radii to be reduced if the probability of a microenvironment transition increases.

ABS algorithm evaluation

Our hypothesis (that environmental and motion triggers can aid microenvironment classification) was evaluated by comparing output from the ABS and static buffer algorithms to the curated reference data. To aid the comparison, the size of the static buffers were set to match the three buffer sets used in the ABS algorithm. These static buffer algorithms are referred to as “GPS-S” (115 m and 60 m, “work” and “home” buffers), “GPS-M” (185 m and 120 m, “work” and “home” buffers), and “GPS-L” (500 m and 500 m, “work” and “home” buffers). A fourth algorithm, referred to as the “Motion” algorithm, was identical to the ABS algorithm except only motion triggers were used (light and temperature data were excluded) to determine if the use of environmental sensor data enhances microenvironment classification.

The raw GPS data from the reference dataset was feed into all five of the algorithms. Performance of each algorithm was calculated by comparing the classification output (at five-second intervals) to the participant’s known location (as determined from the curated reference data) to estimate the true positive, false positive, true negative, and false negative classifications for each microenvironment. For example, for the “home” microenvironment, a true positive indicates that an algorithm predicted the “home” microenvironment when the person was actually at “home”. A false negative indicates that an algorithm incorrectly predicted the “work” or “other” microenvironment when the person was actually at “home”. With these data, the sensitivity (Eq. S2), specificity (Eq. S3), and accuracy (Eq. S4) of each algorithm were calculated by microenvironment. The sensitivity characterizes the degree to which an algorithm underestimates the amount time spent in each microenvironment. The specificity indicates how much an algorithm overestimated time in each microenvironment. Finally, the total algorithm accuracy was also evaluated to determine the effectiveness of each algorithm across all microenvironments. The Friedman and pairwise post-hoc Nemenyi tests (PMCMR package [40]) were used to test for significant differences between the ABS and reference algorithms. All analyses were performed using R 3.5.0 (R Core Team, Vienna, Austria) and code developed and used for these analyses are available upon request.

Results

A total of 22 participants were recruited into the study; these participants provided a total of 25 microenvironment transition datasets (three participants collected data twice) from December 2015 through March 2016, with median, standard deviation, minimum and maximum data collection times of 47.8, 7.7, 30.9, and 71.9 h, respectively. These 25 datasets provided 274 microenvironment transitions (Table S2; 90 defined only by the traditional diary method, 5 required manual assignment using GPS coordinate data, and the remaining 186 were defined by the pushbutton or both the pushbutton and traditional diary method) into or out of 22 unique "home" and 4 "work" microenvironments (campus buildings at Colorado State University). Four of the participants failed to use the traditional diary method, but those participants did use the pushbutton to indicate their microenvironment transitions. The median (25th, 75th percentile) amount of time that participants spent in the "home", "work", and "other" microenvironments based on the reference dataset were 27.8 (24.3, 30.8), 11.5 (8.4, 14.9), and 4.6 (3.1, 9.2) hours. These results are shown in Fig. S4 and Table S3.

The number of predicted microenvironment transitions (Table S2) and median (25th, 75th percentile) amounts of time spent in the "home", "work", and "other" microenvironments (Fig. S4, Table S3) were determined for each of the five algorithms. The ABS algorithm predicted 350 microenvironment transitions and the median (25th, 75th percentile) amounts of time spent in the "home", "work", and "other" microenvironments were 28.0 (24.8, 30.3), 12.1 (9.2, 14.4), and 4.0 (2.8, 9.2) hours respectively.

The accuracy, sensitivity, and specificity were calculated for all five algorithms by microenvironment using all 25 datasets (See SI. XLSX data file). The accuracy comparisons by microenvironment are shown in Fig. 2 and the total accuracy for the microenvironments combined are shown in Fig. S5. The median (25th, 75th percentile) accuracy values are summarized in Table S4 and the pairwise post-hoc Nemenyi results are shown in Tables S5 and S6. The ABS had the highest median total accuracy of 0.975 (0.946, 0.980) of all the algorithms. The sensitivity and specificity by microenvironment are shown in Fig. 3. A summary of the median (25th, 75th percentile) data shown in Table S7, and the pairwise post-hoc Nemenyi results are shown in Tables S8 (Sensitivity) and S9 (Specificity).

Discussion

For microenvironments that are defined with a fixed buffer size (GPS-S, GPS-M, and GPS-L), the GPS-only algorithms show a positive trend between the size of the buffer and the classification accuracy for a given microenvironment (Fig. 2). However, there is a distinct tradeoff that occurs for the GPS-only algorithms; as the buffer size increases the increased accuracy for the "home" and "work" is gained through a loss of accuracy for the "other" microenvironment (i.e., the "home" and "work" buffers infringe upon the "other" microenvironment). The same tradeoff is also apparent with sensitivity: as the sensitivity for the "home" and "work" classifications increases, the sensitivity for the "other" classifications is negatively impacted leading to an underestimation of the amount of time spent outside of the "home" and "work" microenvironments. Inverse tradeoffs are also apparent for the specificity of the three GPS-only algorithms; the specificity for the "other" classifications increases as the buffer size increases, however, these increases are offset with decreases in the "home" and "work" classifications (Fig. 3). These tradeoffs can be seen in Fig. 1 where the larger buffer encompasses the entire area around the "home" and "work" microenvironment. However, the large boundary also encompasses the majority of points that are logged during the transition between the "home" and "work" microenvironments leading to an underestimate of the time spent in the "other" microenvironment. Proper classification of the "other" microenvironment is critical; while the amount of time spent in the "other" microenvironment is typically lower, air pollution concentrations in the "other" microenvironment tend to be elevated (e.g., while in transit on roadways). Thus, exposures in the "other" microenvironment could potentially account for a substantial portion of a person's daily cumulative exposure [41, 42]. However, other studies have shown that "home" and indoor environments can be the

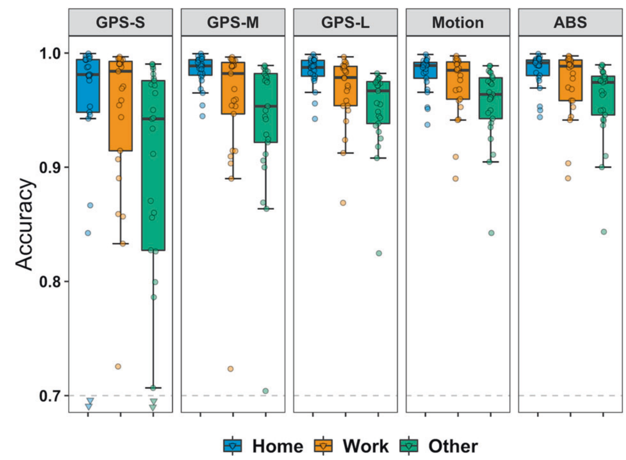

Fig. 2 Microenvironment determination accuracy for each of the five algorithms compared against the participant reference dataset. Data points represent single volunteers ($n = 25$); triangles represent outliers that fall below the 0.7 cutoff indicated by the dashed gray line. GPS-S location only with small buffer, GPS-M location only with medium buffer, GPS-L location only with large buffer.

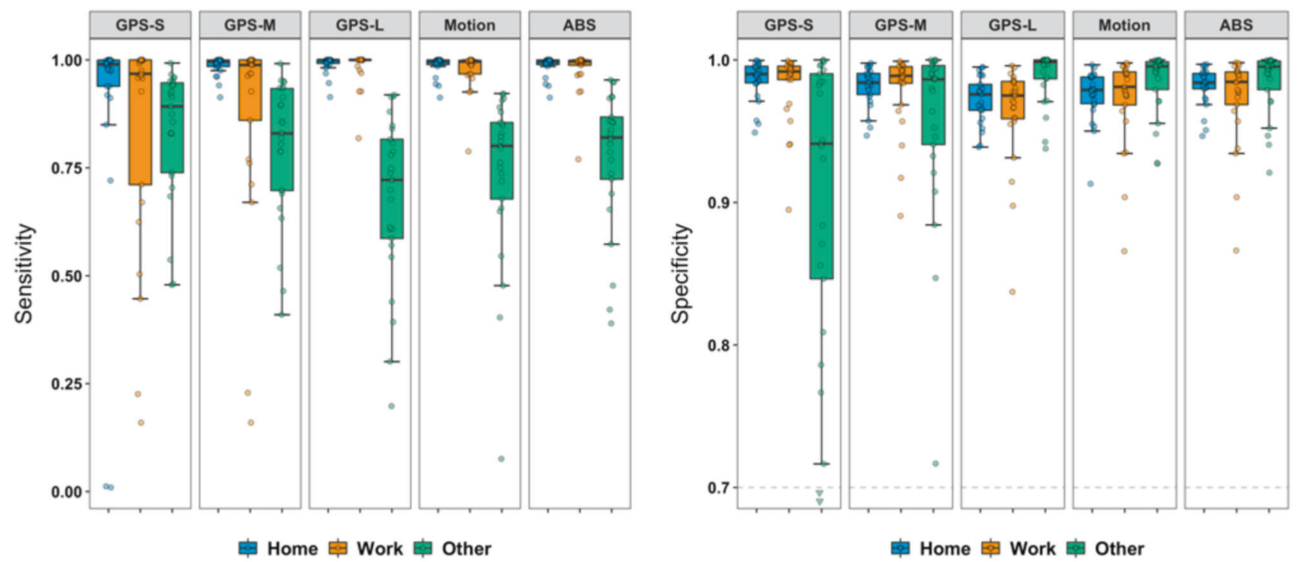

Fig. 3 The sensitivity and specificity for each of the five algorithms by microenvironment. Data points represent single volunteers ($n =$ 25); triangles represent outliers that fall below the 0.7 cutoff indicated by the dashed gray line. GPS-S location only with small buffer, GPS-M location only with medium buffer, GPS-L location only with large buffer.

majority source of exposure due to the amount of time one spends at "home", "work", or "school [43, 44].

Using trigger events to adjust the buffer size helps retain classification accuracy (Fig. 2) without loss of sensitivity and specificity (Fig. 3). The ABS algorithm showed greater total and microenvironment accuracies when compared with the three GPS-only and Motion algorithms; these improvements were statistically significant when the ABS algorithm were compared with the most accurate GPS-only (GPS-L) algorithm (Tables S5 and S6). Additionally, unlike the GPS-only algorithms, both the ABS and Motion algorithms were able to reduce the tradeoffs between the sensitivity and specificity; the sensitivity for the "other" classifications and the specificity for the "home" (with the exception of the Motion algorithm) and "work" classifications were improved and were statistically significant as compared with the GPS-L algorithm. While the ABS algorithm had determined more transitions (350) as compared with the Motion algorithm (332) the ABS algorithm had slightly better accuracies as seen in Table S4, and slight differences for the "other" sensitivity and "home" specificity (Table S7).

Previous work by Breen et al. [12] developed a microenvironment algorithm that was used to classify eight different microenvironments (inside and outside of the home, work, school; inside vehicles; other locations). They evaluated their algorithm and determined the sensitivity and accuracy for the Work (99.9% & 99.6%), Work-Out (60.4% & 99.7%), School (93.1% & 99.9%), School-Out (73.5% & 99.9%), Home (98.8% & 98.9%), and Home-out (81.4% & 98.9%) microenvironments. The "other" microenvironment described herein would encompass all of the microenvironments aside from the "work", "home", and "school". Thus, a direct comparison is not easily made to the sensitivity and accuracy results published by Breen et al., but we can see the ABS algorithm provides similar results. Both algorithms provide a high accuracy and sensitivity for detecting the "home" and "work" microenvironments, but both have reduced sensitivity values with detecting when an individual is outside the "home" and "work" microenvironment. This comparison is notable as the performance of the ABS algorithm in detecting microenvironments in real time is comparable to the performance of a post-hoc analysis.

Strengths, limitations, and future work

This work has several strengths worth noting. First, the reference dataset was collected by 22 volunteers who were not asked to change their daily routines and allowed for the evaluation of microenvironment transitions that occurred at random throughout the day. Second, this work demonstrates tradeoffs (between accuracy, sensitivity, and specificity) of various algorithms used to classify microenvironments. The primary concern is that a high accuracy for an algorithm can be achieved but at the cost of underestimating time spent in certain microenvironments. This type of tradeoff can impact health assessments if large exposures occur in microenvironments where people only spend a short amount of time during the day. Third, the ability to detect microenvironments accurately and in real time enables the use of emerging technologies that collect physical samples by microenvironment (such as the filter-based AMAS). Fourth, this study had 18% of the participants fail to use the diary method at all to record their daily habits. This further supports the challenges of using diary methods that have been previously published [13, 16].

There are several limitations worth noting. Circular buffers were used to define the “home” and “work” microenvironments rather than polygons or the actual building boundaries. This simplified approach was used [1] to accommodate memory limitations of the microcontroller within the personal monitor and [2] to simplify the implementation of the algorithm (i.e., requiring no prior knowledge of each microenvironment except the coordinate location). Future work could allow for adjustable radial buffer sizes for each microenvironment or could use a rectangular rather than circular bounding box, as most building footprints tend to be rectangular. Defining a microenvironment with a circular buffer and a centroid point rather than the building boundary can result in defining a portion of a “work” or “home” microenvironment as “other”. The use of a rectangular buffer, however, would require additional coordinate data.

To better account for issues caused by poor GPS signal quality or complete signal loss, future work should monitor GPS signal quality parameters such as horizontal dilution of precision [45]. Another option that is becoming more cost-effective would be to incorporate advanced spatial tracking techniques such as dead reckoning that integrate motion and GPS sensors to account for GPS signal degradation [21]. While leveraging light detection and temperature data didn’t provide a substantial benefit to the performance of ABS algorithm (beyond the inclusion of motion data), future work that combines these environmental data with enhanced motion data could lead to the development of algorithms that could potentially “learn” microenvironment boundaries. The light sensor used for this work could only measure UV light and thus could only help detect transitions during daylight hours. Also, UV light can be detected indoors, that is why the algorithm developed in this work required UV motion to be coupled with either temperature or acceleration motion. Future work should either replace the UV sensor with or add an additional light sensor capable of measuring both indoor low-level light and outdoor light.

This work included only four “work” locations within a single city (Fort Collins, CO) and during two seasons (Winter and Spring). Future improvements to the algorithm, such as those described above, should be validated with more extensive testing that spans a variety of locations and climates. In addition, we hope to continue the development of the algorithm and one of the first items we would like to include is further classification of the “other” microenvironment. We would initially start with the detection of transit, including vehicular transit, walking, and biking. We would begin by leveraging the number of studies that have already evaluated transit classification using GPS and/or motion sensors (e.g., Brunauer et al. [46], Ellis et al. [47], Feng and Timmermans [48], etc.) and potentially including hardware with pre-installed firmware with this capability.

Finally, the accuracy, sensitivity, and specificity metrics can be biased particularly when comparing “home” and “work” microenvironment classifications to the “other” microenvironment. This bias occurs because people tend spend the majority of their time at “home” or “work” as compared with the “other” microenvironments. Future work should develop metrics that can provide an improved assessment of the accuracy of microenvironment transitions.

Conclusions

This work describes an algorithm that is capable of detecting pre-determined microenvironments in real time using minimal user input. The ability to dynamically classify microenvironments in real time can aid researchers by reducing the amount of postprocessing work required to classify exposure by microenvironment, ease the integration of exposure data into spatial models, and potentially provide air quality notifications in real time. Improvements to all of these areas will help researchers better understand the links between air pollution exposure and health.

Acknowledgements This work was supported by grant ES24719 from the National Institute of Environmental Health Sciences and by grant OH010662 from the National Institute for Occupational Safety and Health. The authors wish to thank all of the participants who volunteered to help collect the samples. Finally, the authors would like to acknowledge those who helped with the development of the UPAS and AMAS: Daniel David Miller-Lionberg, Eric Wendt, Joshua Smith, Nathan Henry, and Nicholas Good.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Brook RD, Rajagopalan S, Pope CA, Brook JR, Bhatnagar A, Diez-Roux AV, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation 2010;121:2331–78.
2. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet 2017; 389:1907–18.
3. Mann JK, Balmes JR, Bruckner TA, Mortimer KM, Margolis HG, Pratt B, et al. Short-term effects of air pollution on wheeze in asthmatic children in Fresno, California. Environ Health Perspect. 2010;118:1497.

4. Peel JL, Tolbert PE, Klein M, Metzger KB, Flanders WD, Todd K, et al. Ambient air pollution and respiratory emergency department visits. Epidemiology 2005;16:164–74.
5. Pope CA III, Dockery DW. Health effects of fine particulate air pollution: lines that connect. J Air Waste Manag Assoc. 2006;56: 709–42.
6. Branco P, Alvim-Ferraz M, Martins F, Sousa S. The microenvironmental modelling approach to assess children's exposure to air pollution–a review. Environ Res. 2014;135:317–32.
7. Lim S, Kim J, Kim T, Lee K, Yang W, Jun S, et al. Personal exposures to PM 2.5 and their relationships with microenvironmental concentrations. Atmos Environ. 2012;47:407–12.
8. Van Ryswyk K, Wheeler AJ, Wallace L, Kearney J, You H, Kulka R, et al. Impact of microenvironments and personal activities on personal PM 2.5 exposures among asthmatic children. J Exposure Sci Environ Epidemiol. 2014;24:260.
9. Baxter LK, Dionisio KL, Burke J, Sarnat SE, Sarnat JA, Hodas N, et al. Exposure prediction approaches used in air pollution epidemiology studies: key findings and future recommendations. J Expo Sci Environ Epidemiol. 2013;23:654–9.
10. Özkaynak H, Baxter LK, Dionisio KL, Burke J. Air pollution exposure prediction approaches used in air pollution epidemiology studies. J Expo Sci Environ Epidemiol 2013;23:566.
11. Adams C, Riggs P, Volckens J. Development of a method for personal, spatiotemporal exposure assessment. J Environ Monit. 2009;11:1331–9.
12. Breen MS, Long TC, Schultz BD, Crooks J, Breen M, Langstaff JE, et al. GPS-based microenvironment tracker (MicroTrac) model to estimate time–location of individuals for air pollution exposure assessments: Model evaluation in central North Carolina. J Exposure Sci Environ Epidemiol. 2014;24:412.
13. Elgethun K, Yost MG, Fitzpatrick CTE, Nyerges TL, Fenske RA. Comparison of global positioning system (GPS) tracking and parent-report diaries to characterize children's time–location patterns. J Exposure Sci Environ Epidemiol. 2006;17:196.
14. Kim T, Lee K, Yang W, Yu SD. A new analytical method for the classification of time–location data obtained from the global positioning system (GPS). J Environ Monit. 2012;14:2270–4.
15. Steinle S, Reis S, Sabel CE. Quantifying human exposure to air pollution—moving from static monitoring to spatio-temporally resolved personal exposure assessment. Sci Total Environ. 2013;443:184–93.
16. Kelly P, Krenn P, Titze S, Stopher P, Foster C. Quantifying the difference between self-reported and global positioning systems-measured journey durations: a systematic review. Transp Rev. 2013;33:443–59.
17. Wu J, Jiang C, Houston D, Baker D, Delfino R. Automated time activity classification based on global positioning system (GPS) tracking data. Environ Health 2011;10:101.
18. Asimina S, Chapizanis D, Karakitsios S, Kontoroupis P, Asimakopoulos D, Maggos T, et al. Assessing and enhancing the utility of low-cost activity and location sensors for exposure studies. Environ Monit Assess. 2018;190:155.
19. Donaire-Gonzalez D, Valentín A, de Nazelle A, Ambros A, Carrasco-Turigas G, Seto E, et al. Benefits of mobile phone technology for personal environmental monitoring. JMIR mHealth and uHealth. 2016;4:356–64.
20. Glasgow ML, Rudra CB, Yoo E-H, Demirbas M, Merriman J, Nayak P, et al. Using smartphones to collect time–activity data for long-term personal-level air pollution exposure assessment. J Exposure Sci Environ Epidemiol. 2016;26:356.
21. Langlois C, Tiku S, Pasricha S. Indoor localization with smartphones. IEEE Consum Electron: Harnessing the sensor suite in your pocket. IEEE Consumer Electronics Magazine. 2017;6:70–80.
22. Nethery E, Mallach G, Rainham D, Goldberg MS, Wheeler AJ. Using Global Positioning Systems (GPS) and temperature data to generate time-activity classifications for estimating personal exposure in air monitoring studies: an automated method. Environ Health 2014;13:33.
23. Lee B, Lim C, Lee K. Classification of indoor-outdoor location using combined global positioning system (GPS) and temperature data for personal exposure assessment. Environ Health Prev Med. 2017;22:29.
24. Flynn JI, Coe DP, Larsen CA, Rider BC, Conger SA, Bassett JD. Detecting indoor and outdoor environments using the ActiGraph GT3X+ light sensor in children. Med Sci Sports Exerc. 2014;46: 201–6.
25. Hu M, Li W, Li L, Houston D, Wu J. Refining time-activity classification of human subjects using the global positioning system. PloS ONE 2016;11:e0148875.
26. Tandon PS, Saelens BE, Zhou C, Kerr J, Christakis DA. Indoor versus outdoor time in preschoolers at child care. Am J Prev Med. 2013;44:85–8.
27. Pagels P, Raustorp A, Guban P, Fröberg A, Boldemann C. Compulsory school in-and outdoors—implications for school children's physical activity and health during one academic year. Int J Environ Res Public Health 2016;13:699.
28. Kestens Y, Thierry B, Chaix B. Re-creating daily mobility histories for health research from raw GPS tracks: validation of a kernel-based algorithm using real-life data. Health Place. 2016;40: 29–33.
29. Dewulf B, Neutens T, Van Dyck D, De Bourdeaudhuij I, Panis LI, Beckx C, et al. Dynamic assessment of inhaled air pollution using GPS and accelerometer data. J Transp Health. 2016;3:114–23.
30. Natonal Research Council. Research priorities for airborne particulate matter: IV. Continuing research progress: National Academies Press; 2004.
31. Bell ML, Dominici F, Ebisu K, Zeger SL. Samet JM. Spatial and temporal variation in PM2. 5 chemical composition in the United States for healtheffects studies. Environ Health Perspect. 2007; 115:989–95.
32. Kelly FJ, Fussell J. Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. Atmos Environ. 2012;60:504–26.
33. Valavanidis A, Fiotakis K, Vlachogianni T. Airborne particulate matter and human health: toxicological assessment and importance of size and composition of particles for oxidative damage and carcinogenic mechanisms. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev. 2008;26:339–62.
34. Quinn C, Miller-Lionberg DD, Klunder KJ, Kwon J, Noth EM, Mehaffy J, et al. Personal exposure to PM2.5 black carbon and aerosol oxidative potential using an automated microenvironmental aerosol sampler (AMAS). Environ Sci Technol. 2018;52: 11267–75.
35. Goudarzi MA, Landry RJ. Assessing horizontal positional accuracy of Google Earth imagery in the city of Montreal, Canada. Geod Cartogr. 2017;43:56–65.
36. Volckens J, Quinn C, Leith D, Mehaffy J, Henry CS, Miller-Lionberg D. Development and evaluation of an ultrasonic personal aerosol sampler. Indoor Air 2017;27:409–16.
37. Sinnott RW. Virtues of the haversine. Sky Telesc. 1984;68:159.
38. Wu J, Jiang C, Liu Z, Houston D, Jaimes G, McConnell R. Performances of different global positioning system devices for time-location tracking in air pollution epidemiological studies. Environ Health Insights. 2010;4:S6246. https://doi.org/10.4137/EHI.S6246.
39. World Health Organization. WHO & International Commission on Non-Ionizing Radiation Protection. Global solar UV index: a practical guide. World Health Organization; 1995.

40. Pohlert T. The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). R package version 2014. http://CRAN.R-project.org/package=PMCMR.
41. HEI. Exposure factors handbook 2011 edition (final report). Washington D.C.: United States Environmental Protection Agency; 2011. Report no.: EPA/600/R-09/052F.
42. Wilson JG, Kingham S, Pearce J, Sturman AP. A review of intraurban variations in particulate air pollution: Implications for epidemiological research. Atmos Environ. 2005;39:6444–62.
43. Bekö G, Kjeldsen BU, Olsen Y, Schipperijn J, Wierzbicka A, Karottki DG, et al. Contribution of various microenvironments to the daily personal exposure to ultrafine particles: personal monitoring coupled with GPS tracking. Atmos Environ. 2015;110:122–9.
44. Koehler K, Good N, Wilson A, Mölter A, Moore BF, Carpenter T, et al. The Fort Collins commuter study: variability in personal exposure to air pollutants by microenvironment. Indoor Air 2019;29:231–41.
45. Langley RB. Dilution of precision. GPS World 1999;10:52–9.
46. Brunauer R, Hufnagl M, Rehrl K, Wagner A, editors. Motion pattern analysis enabling accurate travel mode detection from GPS data only. Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013). The Hague, Netherlands: IEEE; 2013.
47. Ellis K, Godbole S, Marshall S, Lanckriet G, Staudenmayer J, Kerr J. Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms. Front Public Health 2014;2:36.
48. Feng T, Timmermans HJ. Transportation mode recognition using GPS and accelerometer data. Transp Res Part C Emerg Technol. 2013;37:118–30.