Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

Article

Long-term indoor-outdoor $PM_{2.5}$ measurements using PurpleAir sensors: An improved method of calculating indoor-generated and outdoor-infiltrated contributions to potential indoor exposure

Lance Wallace [1,*] and Wayne Ott [2]

1 US EPA (retired); lwallace73@g-mail.com
428 Woodley Way, Santa Rosa, CA 95409 USA
2 Stanford University; wott1@stanford.edu
1008 Cardiff Lane, Redwood City, CA 94061 USA
* Correspondence: lwallace73@gmail.com

Abstract: Low-cost monitors make it possible now for the first time to collect long-term (months to years) measurements of potential indoor exposure to fine particles. Indoor exposure is due to two sources: particles infiltrating from outdoors and those generated by indoor activities. Calculating the relative contribution of each source requires identifying an infiltration factor. We develop a method of identifying periods when the infiltration factor is not constant, and searching for periods when it is relatively constant. From an initial regression of indoor on outdoor particle concentrations, a Forbidden Zone can be defined with an upper boundary below which no observations should appear. If many observations appear in the Forbidden Zone, they falsify the assumption of a single constant infiltration factor. This is a useful quality assurance feature, since investigators may then search for subsets of the data in which few observations appear in the Forbidden Zone. The usefulness of this approach is illustrated using examples drawn from the PurpleAir network of optical particle monitors. An improved algorithm is applied with reduced bias, improved precision, and a lower limit of detection than either of the two proprietary algorithms offered by the manufacturer of the sensors used in PurpleAir monitors.

Keywords: exposure; indoor particles; infiltration factor; $PM_{2.5}$; PurpleAir; Random Component Superposition (RCS); Plantower sensors; indoor-generated particles;

1. Introduction

Airborne fine particles are the single most important environmental cause of mortality. Outdoor and indoor sources are roughly equally important, with annual estimated deaths in the millions [1,2]. The London smog of the 1950s, which killed thousands of persons over a period of a week, was instrumental in activating many countries to monitor outdoor particle concentrations [3]. Indoor risks were perhaps slower to be recognized and still receive only a fraction of the attention and resources devoted to monitoring and controlling sources of outdoor particle pollution. However, most people spend most of their lives indoors [4]. Therefore, multiple studies have been carried out to measure either personal or indoor exposures to particles [5-12]. However these studies have been limited to short-term measurements (days or weeks), since the expense of maintaining research-grade instrumentation in homes, together with the noise and clutter of the instruments, has made long-term (months or years) studies with actual measurements of indoor concentrations nearly impossible.

Recently, however, the advent of low-cost monitors has made it possible to measure $PM_{2.5}$ continuously for months or even years inside and outside residences. Multiple manufacturers have produced optical particle counters employing tiny noiseless sensors to estimate fine particle concentrations (manufacturers). These monitors have been tested for response to particles produced by multiple

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

common indoor sources and many have performed well [13,14]. The unobtrusive size and lack of noise has made them attractive to homeowners interested in their exposure to fine particles.

An important goal in studies of human exposure to particles has been to differentiate particles produced by indoor sources from those that infiltrate homes from outdoors. Multiple methods have been adopted. In the pioneering Harvard 6-City Study, indoor air concentrations were regressed on outdoor concentrations [15.]. The slope of the regression was interpreted as an infiltration factor F_{inf} that could be multiplied by the outdoor concentration to determine the contribution of outdoor particles to total indoor concentration.

Alternative methods have also been applied, such as using a tracer (typically sulfur) with few indoor sources to reduce the indoor-generated particle concentrations [16-19]; using a "censoring" method on continuous indoor measurements to eliminate indoor sources [20-23]; using a recursive modeling approach based on dynamic modeling of near-continuous measurements [24,25]; measuring deposition and penetration directly [26]; applying mass balance analysis to estimate resuspension [27-29]; and using LOESS regression to detect minima in the indoor/outdoor ratio [30].

However, the regression method is still being widely used. A careful statistical analysis of the regression method, called the Random Component Superposition (RCS) method is provided in [31]. In this approach, the following model is used:

$$C_{in}(t) = Pa/(a+k)\, C_{out}(t) + G(t) \tag{1}$$

where C_{in} and C_{out} are the instantaneous indoor and outdoor concentrations (µg/m^3), P is the penetration factor (dimensionless number describing the fraction of outdoor particles able to cross the building envelope), a is the air exchange rate (h^{-1}), k is the deposition rate (h^{-1}), and G is the instantaneous concentration at time t caused by indoor activities.

The term $Pa/(a+k)$ is the infiltration factor F_{inf} and describes the fraction of outdoor PM that enters the home. In the case of a time-averaged monitoring period, the equation may be written

$$\langle Cin \rangle = F_{inf} \langle C_{out} \rangle + \langle G \rangle \tag{2}$$

where now $\langle Cin \rangle$ and $\langle Cout \rangle$ are the averaged values, $\langle G \rangle$ is the average contribution of indoor activities to the indoor concentration and F_{inf} is considered to be constant. This equation is in the form of a simple linear regression, where the intercept is the average contribution of indoor-generated particles to the total indoor concentration, and the slope is the infiltration factor F_{inf}. Multiplying the observed outdoor average concentration by F_{inf} determines the indoor average concentration of infiltrated particles, or C_{inf}:

$$\langle C_{inf} \rangle = F_{inf} \langle C_{out} \rangle \tag{3}$$

Health effects depend on exposure. Remembering that exposure is defined as the contact of a pollutant with a receptor (in this case fine particles in the breathing zone of a human), for a person over time T, a rough estimate of exposure would be given by

$$\langle Cin \rangle T = F_{inf} \langle C_{out} \rangle T + \langle G \rangle T \tag{4}$$

Of course, persons are not always indoors, so equation (4) is not a measure of their exposure. It is, however, a measure of the *potential exposure* of a person indoors over time T. In particular, the first term on the right-hand side of equation (4) is described by epidemiologists as the *exposure to particles of ambient origin*. Epidemiologists then seek a relationship between health effects and this exposure to particles of ambient origin. The second term is the potential exposure to indoor-generated particles. Epidemiologists do not generally include this second term, since G is usually completely unknown.

There are three sources of error in this formulation. First, indoor-generated concentrations are seldom known. These can sometimes be substantial, as in smoking tobacco or marijuana, cooking on gas or electric stoves, burning candles, or simply moving about and resuspending particles [32-38]. Secondly, little is known about the times the residents are not at home. Thirdly, the equation assumes a constant infiltration factor, despite a number of influences such as the indoor-outdoor temperature difference, wind speed or direction, and, in particular, occupant behaviors such as opening a window or running a kitchen fan [39-42].

The first source of error can be rectified by measuring indoor and outdoor concentrations. This is now being done for an ever-growing number of homes using low-cost particle sensors. The second source of error can only be dealt with statistically at present, due to the lack of low-cost personal monitors. The third source of error can be reduced by using the RCS model to identify periods when different infiltration factors may occur for a given location. This is the focus of our investigation.

2. Materials and Methods

2.1 RCS model

The RCS model assumes a constant infiltration factor over the time the house is monitored. The concentration due to particles of ambient origin is given by equation (3) above. Suppose there is a set of average indoor concentrations and corresponding average outdoor concentrations. A regression of the indoor on the outdoor concentrations results in a regression line with slope F_{inf}. Now a second line is drawn showing the contribution of outdoor particles to indoor $PM_{2.5}$. That line is the solution to equation (3) above for the outdoor-infiltrated particles only. The line is parallel to the slope of the regression and passes through the origin. Considering that the indoor contribution <G> cannot be negative, the line through the origin defines a "Forbidden Zone" where few observations are expected (Figure 1). For any given observation, the total indoor air concentration consists of two components: the outdoor-penetrated and indoor-generated particles. The outdoor-penetrated particles (vertical red line in Figure 1) are determined by the upper boundary of the Forbidden Zone. The indoor-generated particles (vertical blue line in Figure 1) are calculated by subtracting this value from the observed total indoor concentration.

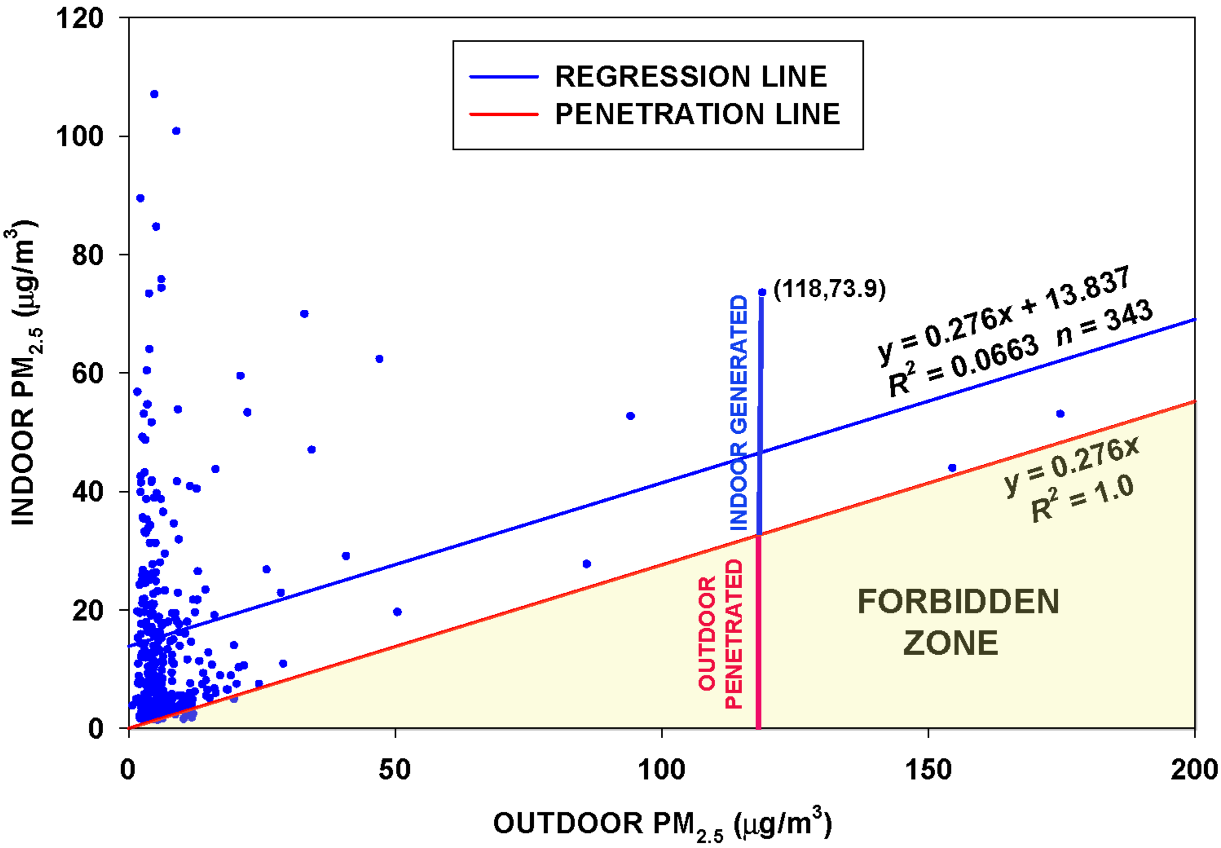

Figure 1. Regression of indoor on outdoor $PM_{2.5}$ measured over 343 days for a home. The slope of the regression line (blue) is the infiltration factor F_{inf}. The parallel line (red) is the product of the infiltration factor F_{inf} and the outdoor concentration. This line forms the upper boundary of a "Forbidden Zone". No observed point should appear in this zone if F_{inf} is constant and applies to all data, because it would indicate a negative indoor-generated concentration.

A major weakness of the regression approach (including the RCS method) is the assumption of a constant infiltration factor, despite the influences mentioned above that alter the infiltration factor. For example, a major EPA study measured 37 homes in the Research Triangle Park of North Carolina for personal, indoor, and outdoor concentrations of sulfur and fine particles [11, 17-19]. Each home was monitored for 7 days in each of the four seasons. The infiltration factor was considerably lower (0.49) in summer compared to the fall, winter, and spring values of 0.64, 0.62, and 0.59, respectively. The authors attributed this effect to increased use of air conditioning during the summer.

The existence of the Forbidden Zone is a feature that allows a check on the assumption of a constant infiltration factor. If many observations appear in the Forbidden Zone, the infiltration factor was not constant. However, it may then be possible to divide the dataset into two or more datasets, each with its own constant infiltration factor.

This study deals with the problem of a variable infiltration factor. First, the RCS approach is applied to an entire dataset of matched indoor and outdoor daily average $PM_{2.5}$ measurements. If there are multiple violations of the Forbidden Zone, subsets of the data are then sought that will have few or no violations. A rough estimate of the variability of the infiltration factor is the daily indoor/outdoor (I/O) ratio as a function of time. If there is a marked variation, it may suggest subsets of the data likely to have more stable ratios. For example, a seasonal variation (due to open windows in spring or use of air conditioning in summer) might appear. If the resulting RCS regressions on each subset (or at least one of the subsets) show no or few violations of the Forbidden Zone, the result can be accepted as a robust estimate of the infiltration factor for that subset and thus a good estimate of the relative contributions of indoor-generated and outdoor-penetrated particles to the total indoor $PM_{2.5}$.

The RCS approach is illustrated using matched indoor and outdoor long-term measurements (hundreds of days) at six California locations. The indoor sites are in Menlo Park, Santa Rosa, Redwood City, and San Francisco. The outdoor sites are in those four locations and also at Bennett Valley, near Santa Rosa, and Alexander Avenue, in Redwood City.

2.2 Algorithm for calculating $PM_{2.5}$

Although two algorithms (CF1 and CF_ATM)) for calculating $PM_{2.5}$ are offered by the Plantower (https://www.plantower.com/en/) manufacturer of the sensors used in the PurpleAir monitors, we reject both. There are several reasons for rejection:

1. Both algorithms are biased high, from 40-90% [43-45].
2. Both algorithms assign values of zero to some results below an arbitrary value. Yet there are no cases in which the reported number of particles is zero. This problem is particularly prevalent for indoor $PM_{2.5}$ measurements, typically below outdoor levels. In one study with >900,000 indoor measurements, more than 200,000 values were reported as zero by the Plantower algorithms (unpublished data, available from LW).
3. No information is provided by Plantower regarding the calibration aerosol properties, such as the chemical makeup and its density.
4. The precision of measurements is poor, in part because one or both sensors within the PA-II monitor often return zero values.
5. The limit of detection (LOD) is elevated, again in part because of poor precision caused by zero values.

The ALT-CF3 algorithm was developed to provide a transparent and reproducible alternative to the "black box" algorithms provided by Plantower. It is described fully in [46]. The algorithm uses the particle numbers reported for the three size fractions 0.3-0.5 μm, 0.5-1 μm, and 1-2.5 μm. By assigning an average diameter to each size fraction, the total particle volume is calculated. The total $PM_{2.5}$ mass is then determined in two steps: first a density is assigned (in this case the density of water) and then a calibration factor is determined by comparison to nearby regulatory monitors employing Federal Reference or Federal Equivalent Methods (FRM/FEM). The original calibration factor for 33 PurpleAir PA-II monitors within 0.5 km of 27 regulatory monitors was 3.0, leading to naming this alternative algorithm ALT-CF3 (48). The equation giving the ALT-CF3 estimate of $PM_{2.5}$ is

$$PM_{2.5} = 3(0.00030418\ N1 + 0.0018512\ N2 + 0.2069706\ N3) \tag{5}$$

where N1, N2, and N3 are the number of particles per deciliter in the three smallest size categories.

The ALT-CF3 method has been found to provide less bias, better precision, and a lower limit of detection than either the CF1 or CF_ATM algorithms [47-48]. This method is presently available on the PurpleAir mapping page at the PurpleAir Website (https://www2.purpleair.com/) as an alternative to the proprietary algorithms offered by Plantower. The method is also available on the PurpleAir API page (https://api.purpleair.com/) under a different name: "$PM_{2.5}$_alt". All $PM_{2.5}$ data in this paper were determined using the ALT-CF3 algorithm.

2.3 Calculation of the limit of detection (LOD)

A method for calculating the limit of detection for continuous monitors has been developed [49]. By analogy with integrated methods such as weighing particle masses on filters, it is required that the mean of multiple collocated instruments of one type all measuring the same environment at some low concentration exceeds 3 times the standard deviation to be considered as evidence at the 95% confidence level of a non-zero concentration. This approach can be applied to a single PurpleAir monitor with two sensors A and B. Every data point has an associated standard deviation based on the two sensors. The data can be ordered by the mean of the two sensors. At very low mean values, it is likely that the ratio of the mean to the standard deviation (μ/σ) will be <3 for many of the observations. For higher mean concentrations, the fraction of cases with μ/σ <3 will decline. When this fraction falls below 5%, the associated mean concentration will be the LOD. Operationally, this approach can be implemented by considering successive batches of, say, 100 measurements of the ordered data. When the number of ratios <3 falls to fewer than 5 of the 100 values in the batch, the associated mean of that batch may be considered a provisional LOD. However, a batch at a higher concentration may have 5 or more ratios <3, in which case the LOD will be greater than the provisional LOD and the search for the LOD must continue to higher concentrations. This problem can be solved by analyzing all possible batches of 100 (often in datasets of more than 100,000 entries) and picking the highest concentration with only 4 out of 100 values having μ/σ <3.

3. Results

3.1 Menlo Park

This site has one outdoor PA-II monitor (with two Plantower PMS 5003 sensors) and one indoor PA-I monitor (with one Plantower PMS 1003 sensor). The monitors operated from 11/13/2019 to 12/4/2021. The site is a private residence occupied by a renter. Between July 2, 2020 and Feb 14, 2021, the renter was out of the country and the residence was unoccupied. Daily means were calculated for all days using the ALT-CF3 algorithm. At least 540 of the 720 possible measurements per day were required (18 h/day). For every two-minute measurement made by the outdoor monitors with two identical sensors A and B, a measure of precision (abs(A-B)/(A+B)) was calculated. Precision was required to be less than 0.2 to be included in the analysis. This resulted in the loss of about 8% of the data.

The data for the time of occupancy (270 days) consist of 372,621 two-minute average $PM_{2.5}$ concentrations (Table 1). Note the loss of about 70,000 data points for the Plantower CF1 algorithm (bottom three rows) compared to the ALT-CF3 algorithm (top five rows). This is partly due to the Plantower decision to assign values of zero to concentrations below an arbitrary cutoff. No values of zero ever occur in the ALT-CF3 algorithm, since there are never occasions when the number of particles between 0.3 and 0.5 µm are zero. The CF1 algorithm overestimates the outdoor $PM_{2.5}$ concentration by about 85% (8.3 vs. 4.4 ug/m3) and the indoor concentration by 50% (3.3 vs. 2.2 ug/m3). The mean $PM_{2.5}$ concentration due to indoor activities (1.78 $\mu g/m^3$) is larger than the mean concentration (0.78 $\mu g/m^3$) attributable to penetration of ambient $PM_{2.5}$. The Spearman correlation coefficient was reasonably high at 0.67.

Table 1. $PM_{2.5}$ ($\mu g/m^3$) and associated precision at the Menlo Park site using the ALT-CF3 and Plantower CF1 algorithms.

	N obs	Mean	Std.Err.	Lower quartile	Median	Upper quartile	Max
ALT-CF3							
Mean outdoor $PM_{2.5}$	372621	4.4	9.2E-03	1.4	2.7	5.1	311
Indoor $PM_{2.5}$	369610	2.2	0.015	0.65	1.3	2.4	1050
Outdoor-infiltrated	372621	0.78	1.6E-03	0.25	0.5	0.9	55

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

Indoor-generated	369610	1.78	0.015	0.21	0.6	1.5	1049
Precision outdoor	372621	0.06	7.3E-05	0.02	0.1	0.1	0.2
Plantower CF1							
Mean outdoor $PM_{2.5}$	304345	8.3	0.019	2.5	5.1	9.6	552
Indoor $PM_{2.5}$	301625	3.3	0.031	0.61	1.8	3.8	1874
Precision outdoor	304345	0.07	9.1E-05	0.03	0.1	0.1	0.2

For the 270-day period when the residence was occupied, the RCS estimate of the infiltration factor is 0.0849 (Figure 2). The RCS-predicted Forbidden Zone created by the product of F_{inf} and the outdoor concentration is not violated by a single daily measurement. This suggests that the estimate of a single constant value for the infiltration factor throughout the entire period is robust.

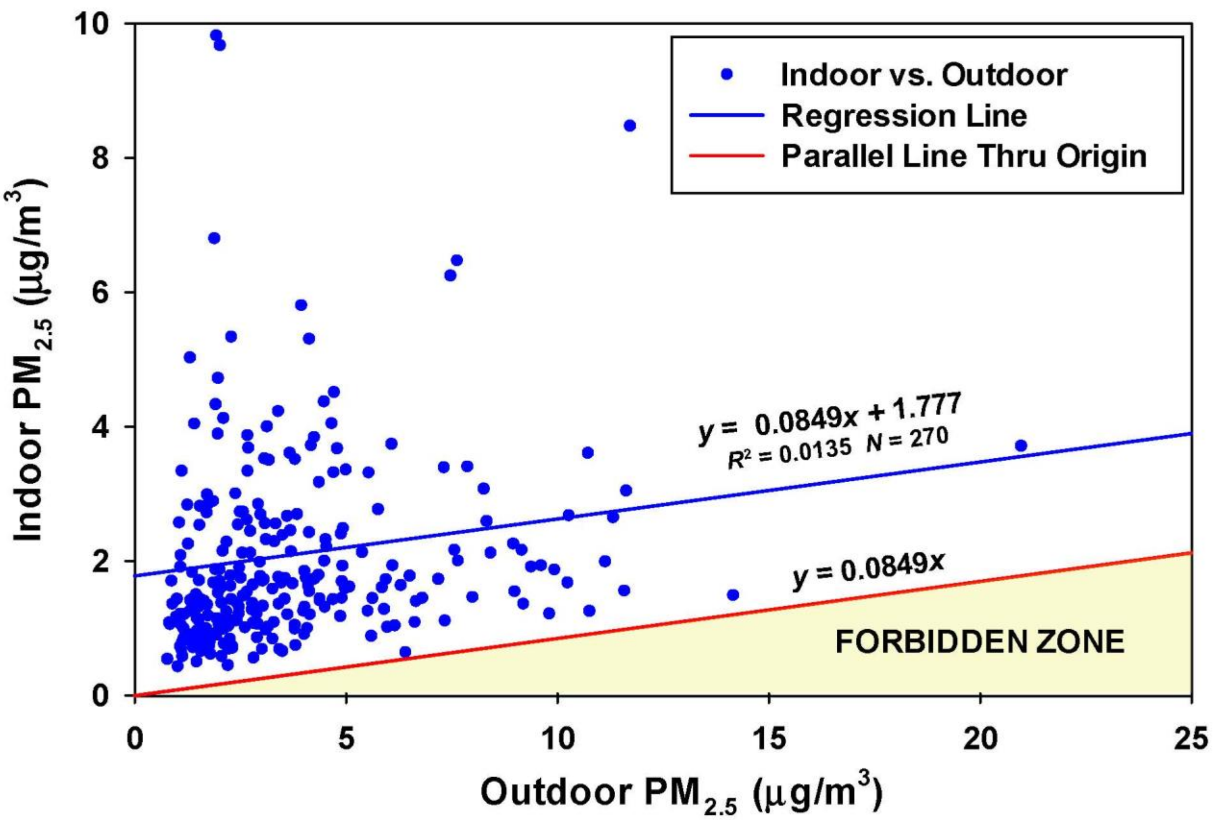

Figure 2. Regression of daily outdoor $PM_{2.5}$ on indoor $PM_{2.5}$ during the time the Menlo Park residence was occupied (11/13/2019 to 7/2/2020 and 2/14/21 to 12/4/2021).

3.1.1 Limit of detection (LOD) at Menlo Park

Using the approach to calculate the LOD described in [49], the ALT-CF3 algorithm resulted in an estimated LOD of 1.10 µg/m^3, with 73,458 (16.6%) $PM_{2.5}$ ALT-CF3 values below the LOD. For the Plantower CF1 algorithm, the LOD was almost 4 times higher at 4.15 µg/m^3 and nearly half (48.6%) of all values fell below the LOD.

3.2 Oakmont, Santa Rosa, CA

Oakmont is a community within Santa Rosa, CA. Two indoor PA-II monitors 1 & 2 were set up on 7/23/19 and operated until 6/18/20; at that point, two additional PA-II monitors (3 & 4) were deployed. Monitors 1 & 4 were operated exclusively indoors, and monitor 3 (except for one month) outdoors for the next 19 months to September 28, 2022. Daily means were calculated for all days using the ALT-CF3 algorithm. At least 360 of the 720 possible measurements per day were required. A table of results for the Plantower algorithms CF1 and CF_ATM for comparison with the ALT-CF3 algorithm is provided (Table 2). The Plantower algorithms overestimated $PM_{2.5}$ values by 50% (CF1) and 30% (CF_ATM) compared to the ALT-CF3 algorithm, which has been calibrated against multiple regulatory monitors [30]. The Plantower algorithms resulted in a major loss of data (18%) due to the Plantower practice of assigning values of zero to results below an arbitrary cutoff. (The ALT-CF3 algorithm resulted in no zero values.) The mean precision of 22% for both Plantower algorithms is nearly 3 times the mean of 8% shown by the ALT-CF3 algorithm.

Table 2. All $PM_{2.5}$ data from Monitor 1 at the Oakmont site between 7/23/19 and 9/28/22

	Valid N	Mean	Std. Err.	Min	Lower quartile	Median	Upper quartile	Max
MEAN 1 CF1	593793	4.41	0.03	0	0.105	1.1	3.28	1678
MEAN 1 CF ATM	593793	3.9	0.02	0	0.105	1.1	3.28	1118
MEAN 1 ALT-CF3	593793	2.97	0.02	0.0063	0.543	1.167	2.39	865
precision 1 CF1	488117*	0.22	0.0004	0	0.04	0.098	0.25	1
precision 1 CF ATM	488117*	0.22	0.0004	0	0.039	0.098	0.25	1
precision 1 ALT-CF3	593695	0.08	0.0001	3.02E-07	0.029	0.062	0.11	0.87
* Missing 105,676	observations assigned a value of zero by the Plantower CF1 or CF_ATM algorithms							

Quality assurance of the data was performed by setting a limit of 0.2 (20%) on the precision of both the indoor and outdoor measurements. This resulted in ~9% of the data being deleted. Also, several experiments resulting in elevated $PM_{2.5}$ for periods of several hours in one room of the house (which included at least one of the indoor monitors) were removed from consideration. This resulted in a further removal of ~0.6% of the data. The final database consisted of 804 days with matched indoor and outdoor daily means. At the Oakmont site, monitors 1 & 4 were always indoors, monitor 2 was mostly indoors, and monitor 3 was outdoors except for one month. The total number of zeros reported by the Plantower CF1 (and CF_ATM) algorithm for $PM_{2.5}$ were >100,000 (17-22%) for the two indoor monitors and less than about 40,000 (6-7%) for monitor 3, which was mostly outdoors (Table 3). This indicates how the lower concentrations found indoors lead to increased numbers of zeros reported by the Plantower algorithms compared to outdoors.

Table 3. Number of zeros reported by CF1 (and CF_ATM) and percentage of total measurements

Monitor/sensor	Location	Zeros	Total	%
1a	Indoors	127195	593804	21
1b	Indoors	129114	593804	22
4a	Indoors	114186	593688	19
4b	Indoors	102616	593688	17
2a	Some outdoors	52000	587309	9
2b	Some outdoors	121399	587309	21
3a	Mostly outdoors	40682	593813	7
3b	Mostly outdoors	33836	593813	6

3.2.1 Limit of detection (LOD) for Oakmont

The LOD was calculated for each monitor-location for both the Alt-CF3 and Plantower CF1 algorithms (Table 4). For the ALT-CF3 algorithm, the LOD ranged from 0.60 to 1.32 µg/m^3, and was always exceeded by a majority of the observations. For the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

Plantower CF1 algorithm, the LOD ranged from 2.9 to 9.9 µg/m^3, and fewer than half the observations exceeded the LOD except for a one-month period when monitor 3 was moved indoors.

Table 4. LODs and percent of observations exceeding the LOD for Oakmont comparing the ALT-CF3 algorithm to the Plantower CF1 algorithm

Monitor	Location	Valid N	ALT-CF3 LOD (µg/m^3)	ALT-CF3 N>LOD	ALT-CF3 (%>LOD)	CF1 LOD (µg/m^3)	CF1 N>LOD	CF1 (%>LOD)
1	Indoors	406108	0.99	233900	58	2.9	177908	44
2	Outdoors	253454	0.92	203384	80	9.9	39487	16
2	Indoors	146229	0.72	110674	76	3.2	44289	30
3	Outdoors	363797	0.6	334973	92	4.4	156850	43
3	Indoors	42304	0.52	42114	99.6	2.9	32150	76
4	Indoors	406092	1.32	215872	53	5.3	79371	20

The regression of the indoor on outdoor $PM_{2.5}$ using the ALT-CF3 algorithm is provided (Figure 3). However, there are a large number of values in the Forbidden Zone, suggesting that the estimate of the infiltration factor is not robust.

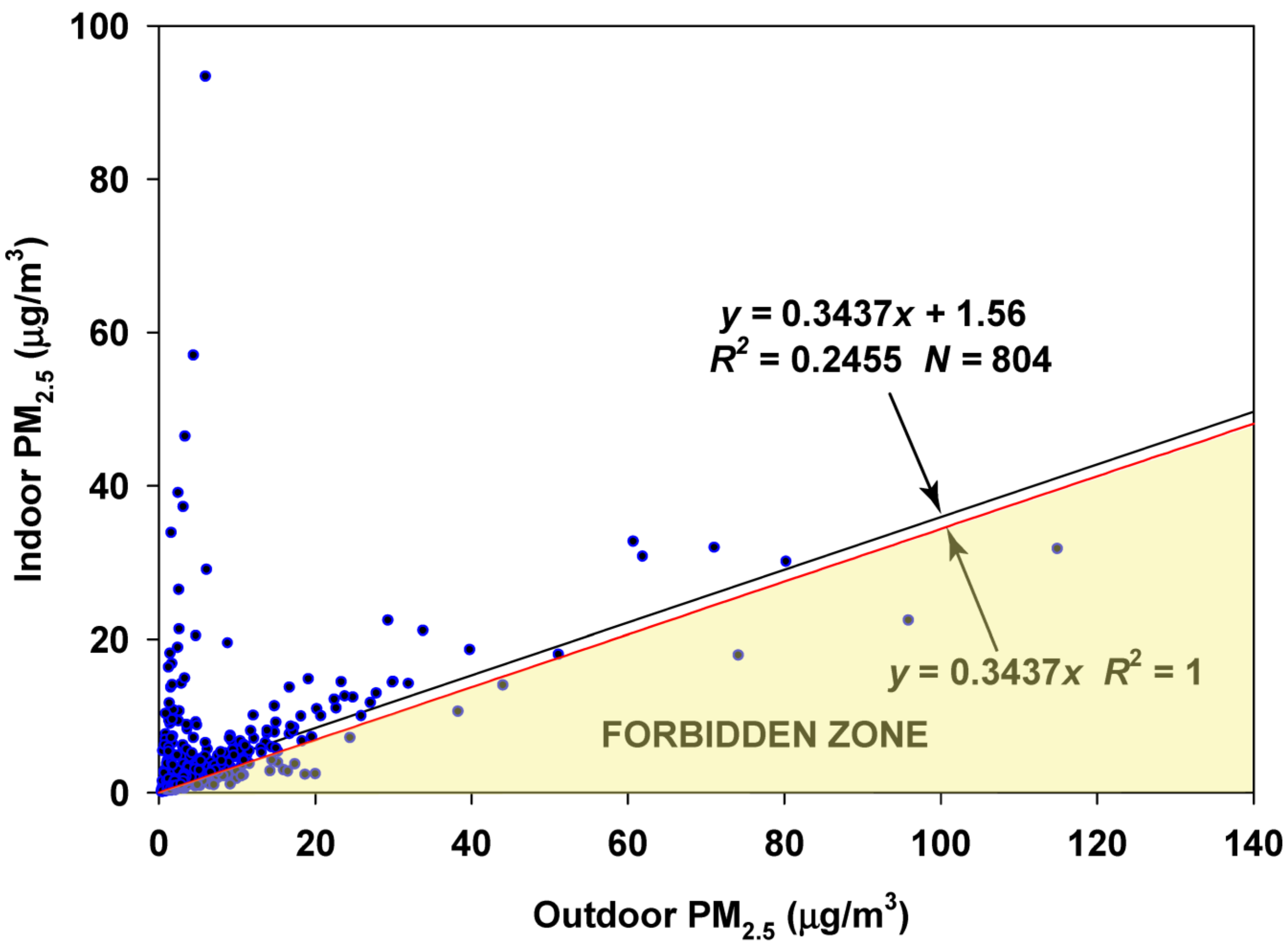

Figure 3. Random Component Superposition (RCS) regression for Oakmont.

Analysis of the monthly regressions showed a clear pattern of low infiltration factors (mean 0.11, SE 0.02) for the cool wet months of December through May (227 days), and high infiltration factors (mean 0.37, SE 0.04) for the warm dry months of June through November (285 days) (Figure 4). The 6 months (June-November) with high infiltration factors and 6 months (December through May) with low infiltration factors matches up very well both with temperature and rainfall. The 6 months with highest average temperatures in Santa Rosa (59-66 ºF) are May-October, while the 6 months with lowest temperatures (46-55 ºF) occur in November-April. Rainfall is also limited to the November-April time frame, with the six months from May to October seldom experiencing rain. Doors and windows are more likely to be open when the weather is warm and dry, resulting in a higher infiltration factor in the warm months. The I/O values appear to be a sine wave with a period of one year, and the peaks in the summer months and troughs in the winter months.

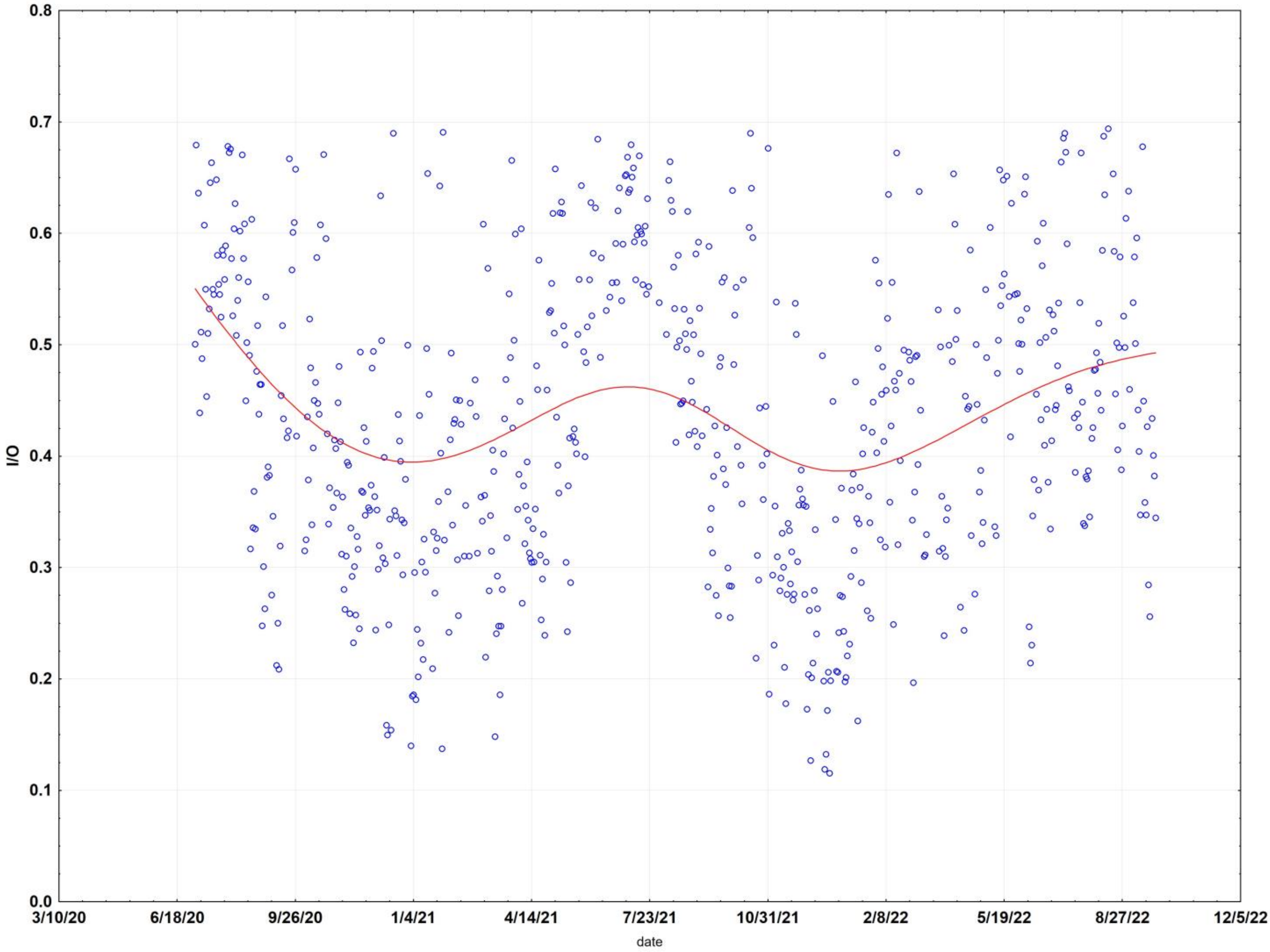

Figure 4. Indoor-outdoor (I/O) ratio of daily means at the Santa Rosa site.

Therefore the data are split into two 6-month periods. The 6-month period from December to May had few points falling into the Forbidden Zone (Figure 5). Therefore the estimated infiltration factor of 0.1355 is accepted, resulting in an average contribution of 1.54 μg/m^3 made by indoor-generated particles during this period.

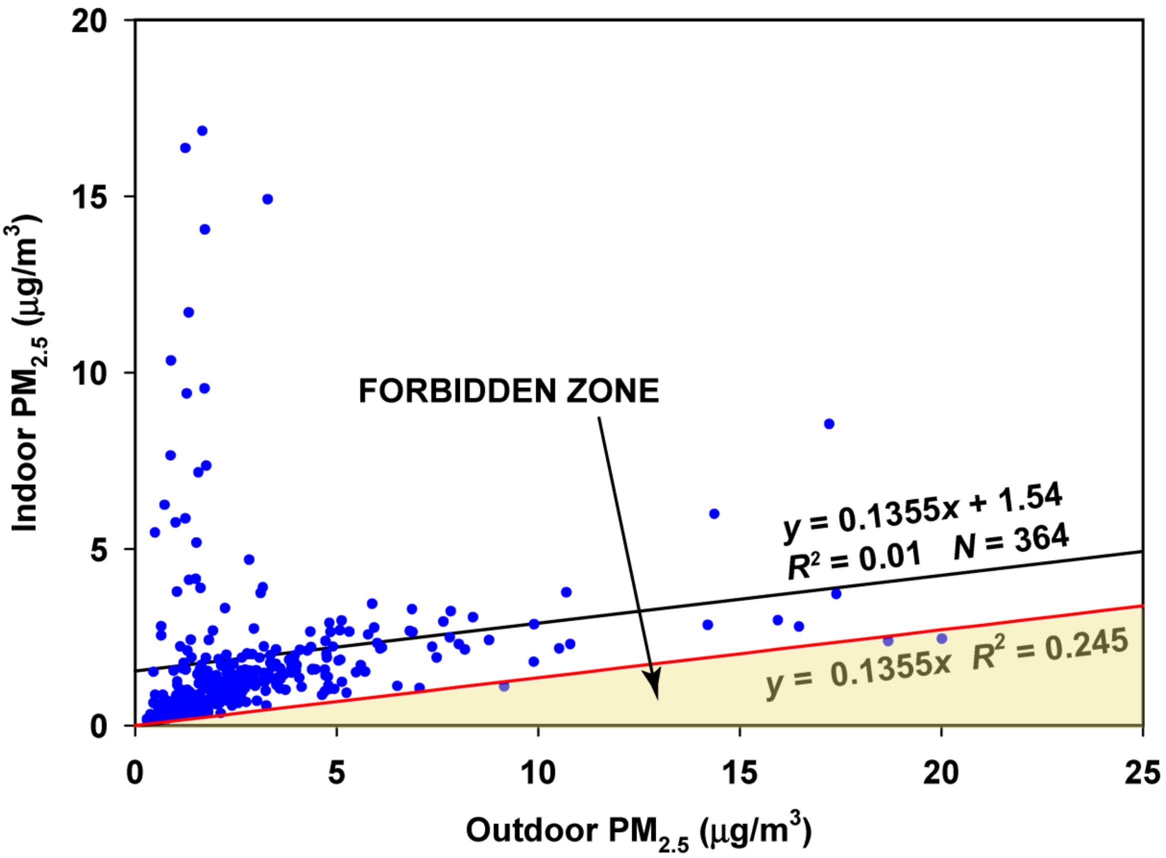

Figure 5. RCS regression of indoor on outdoor air for the December through May data in Oakmont.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

During the remaining 6 months of June-November, once again few points fall into the Forbidden Zone (Figure 6). The estimated infiltration factor was 0.3115, more than twice as high as the winter value of 0.1255, and the estimated contribution of indoor-generated particles is 1.90 µg/m^3.

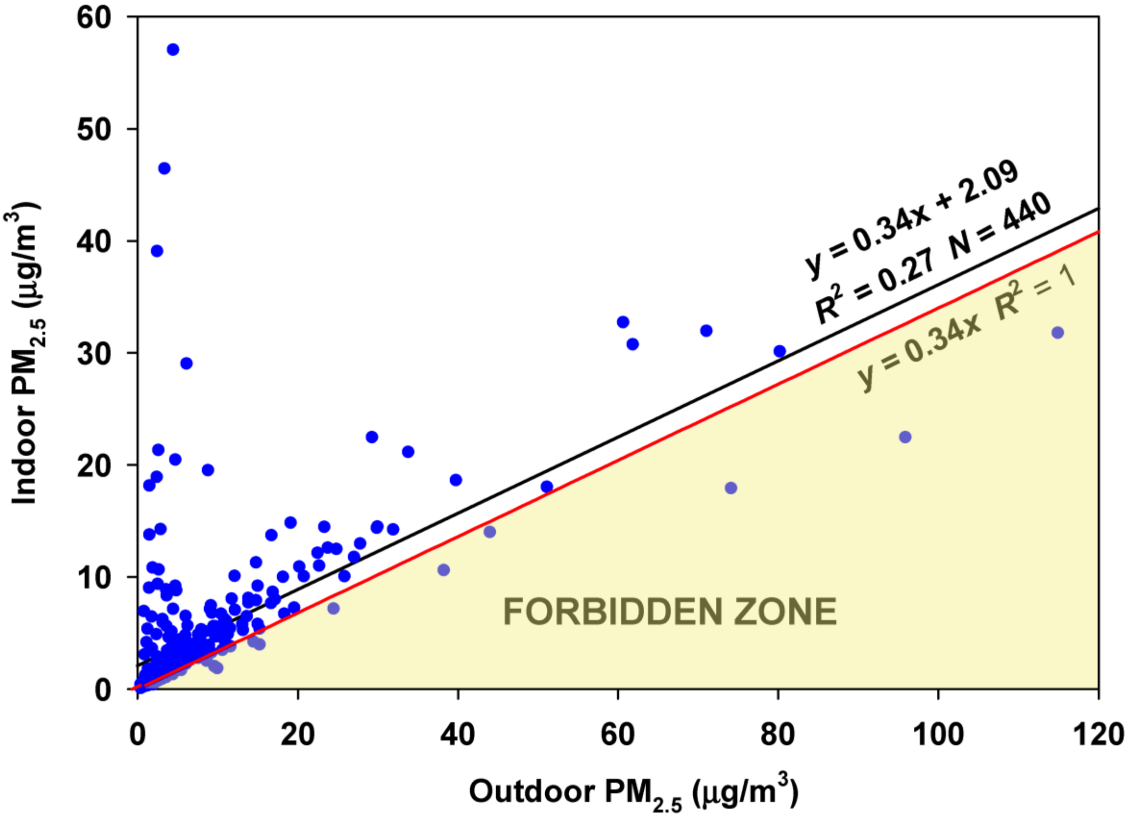

Figure 6. RCS regression of indoor on outdoor air for the June through November data in Oakmont.

 doi:10.20944/preprints202212.0301.v1

3.3 Redwood City, CA

A site in Redwood City, CA operated two PurpleAir PA-II monitors indoors and one PurpleAir PA-II monitor outdoors from May 1, 2021 until October 16, 2022. The 2-minute average indoor data for the full 19-month period are provided for both the Plantower CF1 and ALT-CF3 algorithms in Table 5. The Plantower CF1 algorithm overestimated $PM_{2.5}$ by about 50% (6.2 vs. 4.1 ug/m3). The mean precision of 17-21% in the CF1 algorithm was about twice the 9-10% observed for the ALT-CF3 algorithm. Applying an upper bound of 20% to the precision resulted in about 15% of the $PM_{2.5}$ CF3 data being removed from the analysis, compared to 40% of the Plantower CF1 data.

Table 5. All indoor $PM_{2.5}$ data (µg/m^3) and associated precision for monitors 1 & 2 compared for the ALT-CF3 and Plantower CF1 algorithms for the 19-month period 7/3/19 to 2/6/22.

	N obs.	Mean	Std.Err.	Lower quartile	Median	Upper quartile	Max
ALT-CF3							
Mean 1	662773	4.2	0.020	0.93	1.8	3.3	735
Mean 2	662635	4.0	0.017	0.92	1.8	3.4	539
precision 1	662719	0.09	9.2E-05	0.032	0.069	0.12	0.98
precision 2	662561	0.10	1.2E-04	0.043	0.080	0.13	0.99
Plantower CF1							
Mean 1	662773	6.3	0.033	0.74	2.3	5.1	1149
Mean 2	662633	6.1	0.029	0.70	2.3	5.1	883
precision 1	626667	0.17	3.1E-04	0.036	0.083	0.18	1
precision 2	617016	0.21	3.2E-04	0.066	0.12	0.22	1

The extreme loss of data from use of the Plantower CF1 algorithm is largely due to the practice of setting all $PM_{2.5}$ values falling below a cutoff to zero. The number of zeros reported by the Plantower CF1 algorithm ranged from 49,013 to 72,442 (7-11% of all data). No zeros are ever reported by the ALT-CF3 algorithm, since particles are always present in the 0.3-0.5 µm size category.

Daily means were calculated for all days with the requirement that at least half of the 720 possible measurements per day were available at an acceptable precision (abs (A-B)/(A+B)) of better than 20%. Out of 411 days meeting these requirements, 3 days had I/O values>1, indicating strong indoor sources, and were dropped from the database. One outlier on 8/8/21 of >35 ug/m3, probably due to a wildfire, was dropped, since the ALT-CF3 algorithm calibration factor (CF) of 3 may be inadequate to describe the response to wildfire smoke. The final database consisted of 407 days with matched indoor and outdoor daily means.

The RCS approach for the entire database results in a large number of values in the Forbidden Zone, indicating that the infiltration factor varies too much to be estimated using the entire database (Figure 7).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

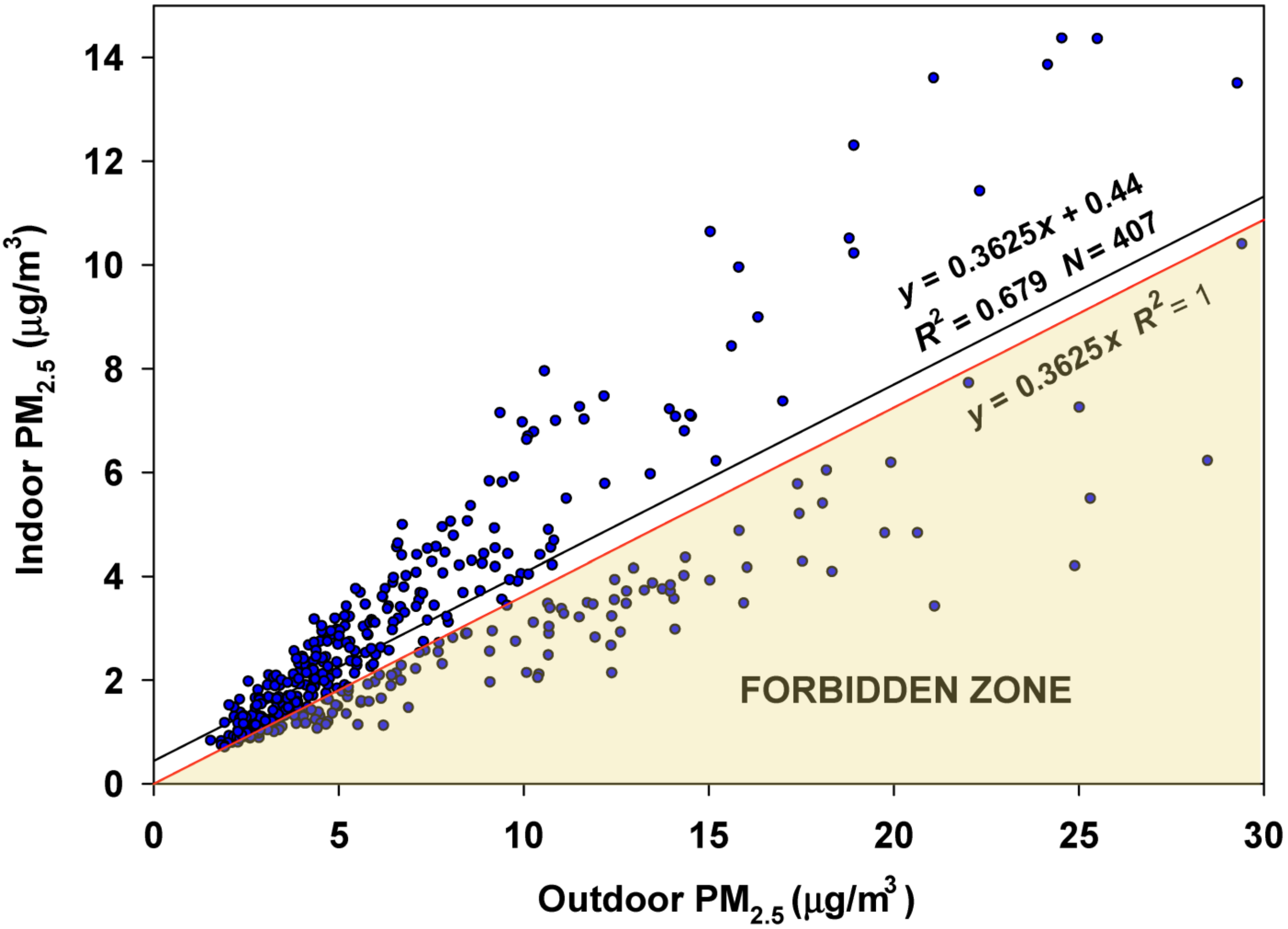

Figure 7. Random Component Superposition (RCS) regression for the Redwood City site. The slope (F_{inf}) of 0.3625 results in many datapoints falling into the Forbidden Zone (below the estimated outdoor-penetrated particles given by the red line).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

The I/O ratio is shown for the entire period (Figure 8). It suggests a sine wave with a period of one year, peaks occurring in the summer months and troughs in the winter months. A period of low infiltration factors appears between about November and March.

Figure 8. Indoor-outdoor (I/O) ratio of daily means at the Redwood City site.

The RCS regression on the November-March period results in only 15 points falling into the Forbidden Zone, and 12 of these seem to "hug" the line (Figure 9). Therefore the estimated infiltration factor of 0.2329 and mean contribution of 0.556 μg/m^3 made by indoor-generated particles during this period appears robust.

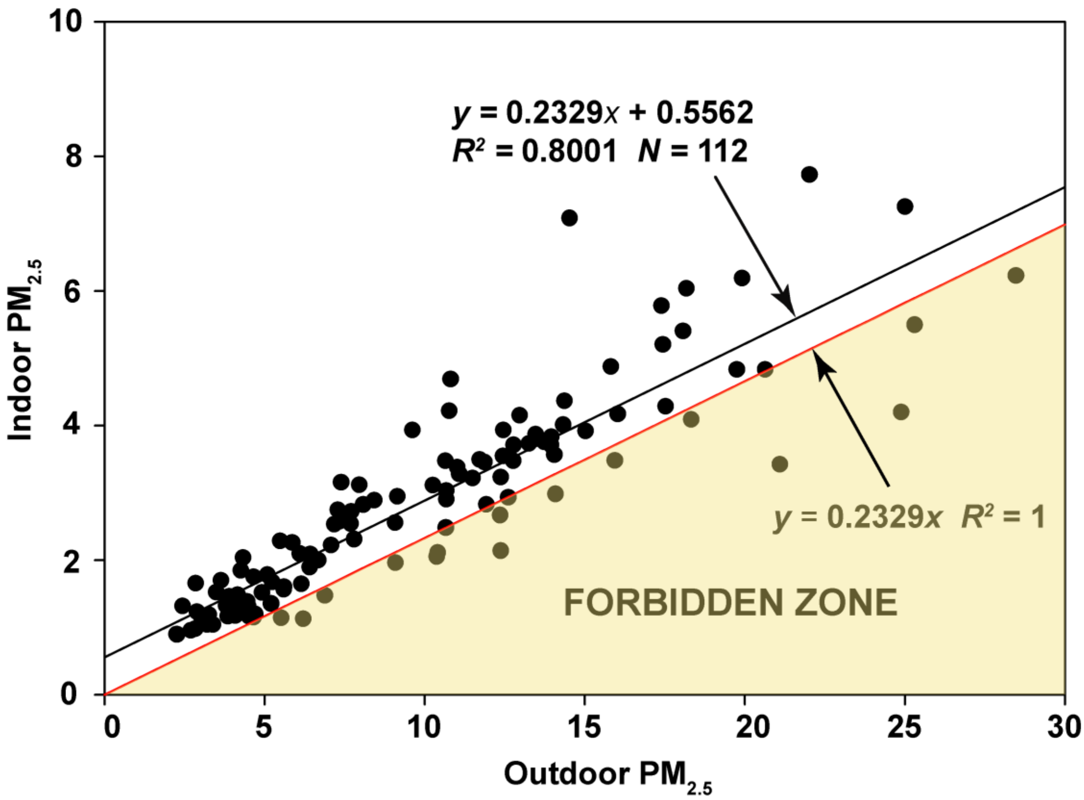

Figure 9. Random Component Superposition (RCS) regression for the November through March data at Redwood City.

An effort was made to estimate the infiltration factor for the summer months (June-July-August). However, the effort failed, suggesting too much variability of the infiltration factor. Two of these months were strongly affected by wildfires, which have been shown to affect the PurpleAir calibration by a factor of ~2 [50]. This may explain the failure of the RCS approach to establish a reasonable estimate of outdoor-penetrated and indoor-generated particles for these months.

3.3.1 Limit of detection (LOD)

For monitor 1 indoors, the LOD for the ALT-CF3 algorithm was about 1.4 μg/m^3, corresponding to about 97% of all observations exceeding the LOD. For the same monitor, the Plantower CF1 algorithm resulted in an estimated LOD of 4.75 μg/m^3, with only 42% of observations exceeding the LOD.

3.4 Alexander Avenue

Indoor and outdoor PurpleAir $PM_{2.5}$ measurements at a residence on Alexander Avenue in Redwood City were collected between 12/5/17 and 9/7/22 using the ALT-CF3 algorithm. Daily means were calculated for all days with the requirement that at least half of the 720 possible measurements per day were available at an acceptable precision (abs (A-B)/(A+B)) of better than 20%. The final database consisted of 1645 days with matched indoor and outdoor daily means.

The I/O ratio is shown for the entire period (Figure 10). Unlike our previous findings with homes in Santa Rosa and Redwood City, the I/O ratio shows little or no seasonal variation. There is however, an apparent sudden switch to less variable values occurring near January 2020.

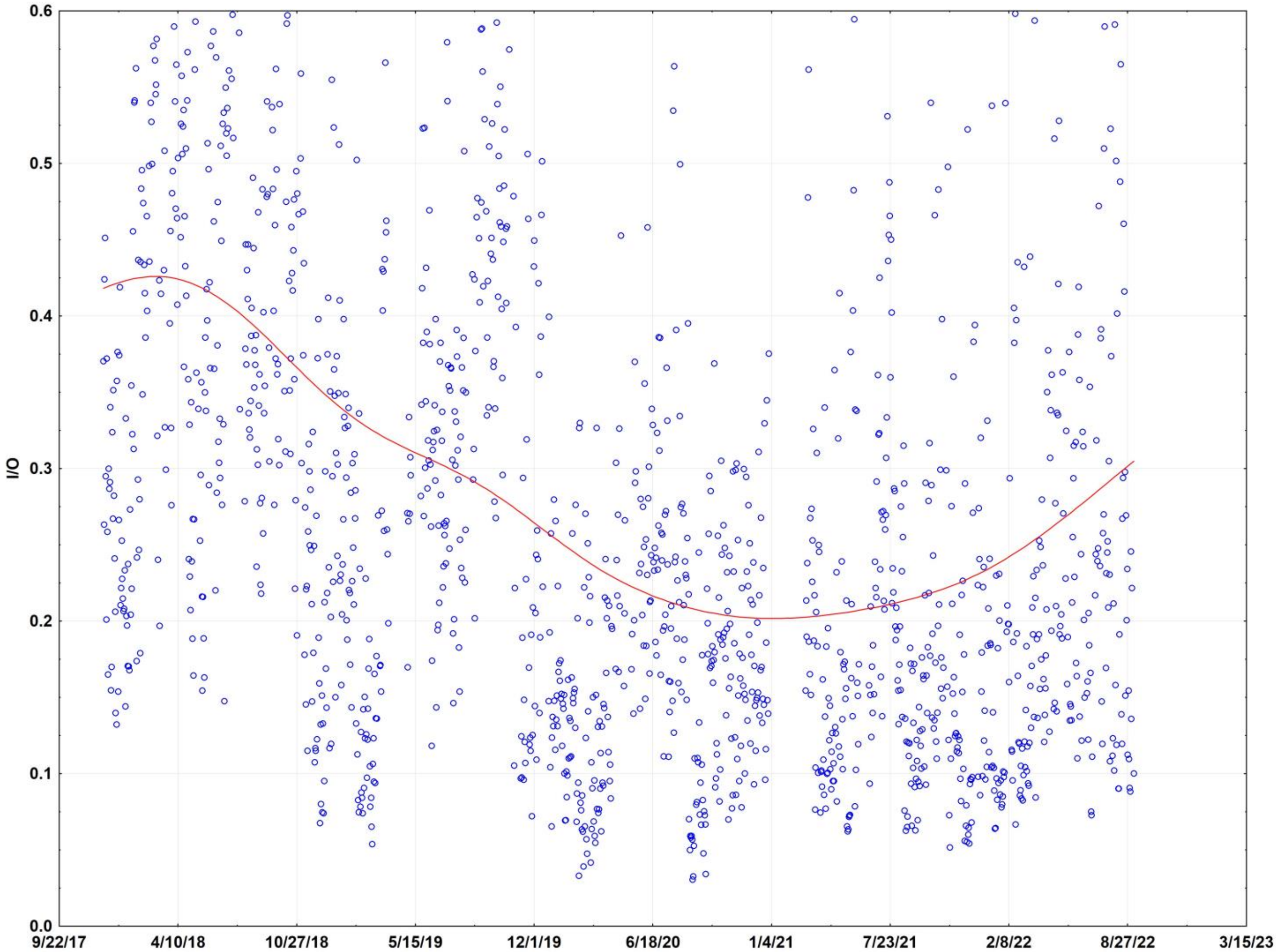

Figure 10. Indoor-outdoor (I/O) ratio of daily means at the Alexander Avenue site.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

The RCS approach to estimate the infiltration factor for the entire database resulted in a large number of values in the Forbidden Zone (Figure 11).

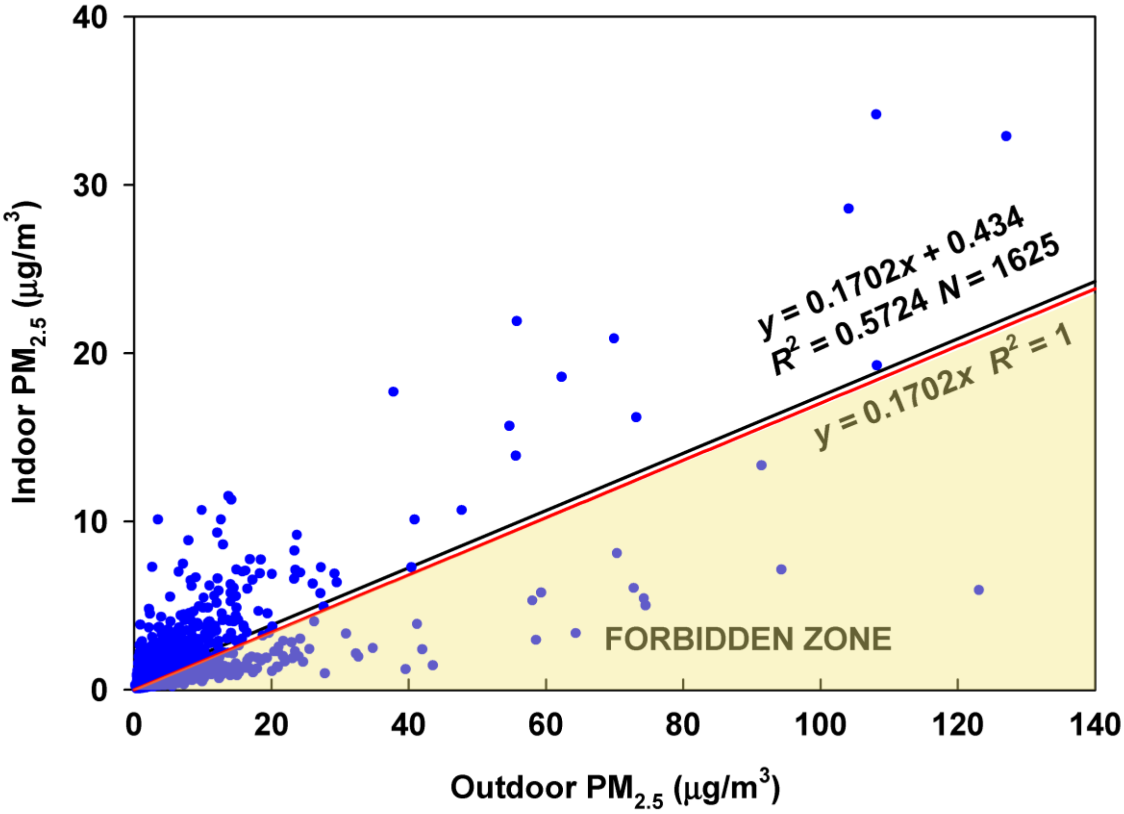

Figure 11. Regression of Alexander Avenue indoor $PM_{2.5}$ on Alexander Avenue outdoor $PM_{2.5}$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

Based on the apparently smaller variability of the I/O ratios in the latter part of the period from about January 1, 2020 to September 7, 2022 (938 days), the RCS approach is applied to that period (Figure 12). In this period, 75 (<8%) of the observations fall into the Forbidden Zone. This is somewhat marginal. However, if applicable, it would suggest an infiltration factor of 0.08 and an average contribution of about 0.51 μg/m³ from indoor-generated particles. The contribution of particles of ambient origin was 0.44 μg/m³, so the indoor-generated fraction of the total indoor potential exposure was about 54%.

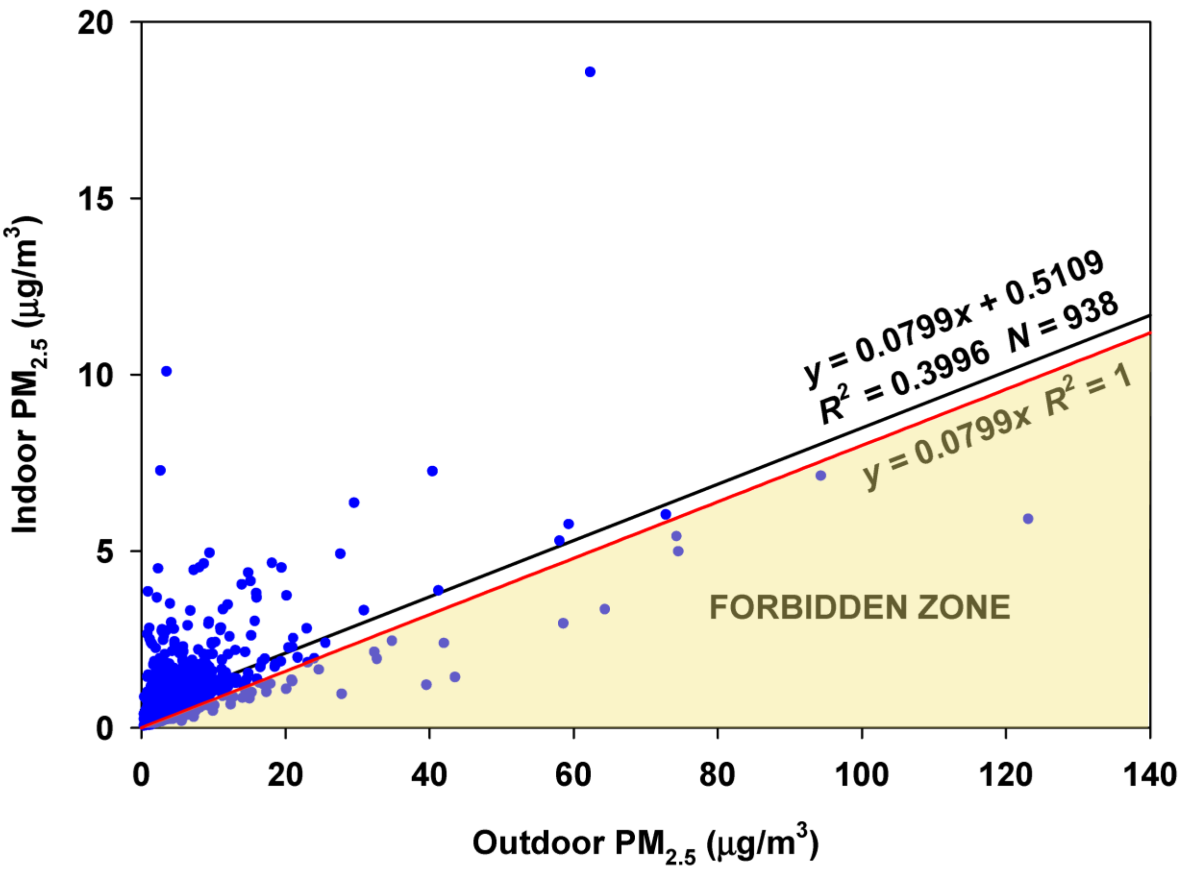

Figure 12. RCS regression for the period from 2020-2022.

An effort was made to estimate the infiltration factor for the period from 2017 through 2019. However, this effort was unsuccessful. For this location, there appears to have been a period before 2020 with a higher infiltration factor and a period from 2020 to the present with a lower infiltration factor. Only the latter period allowed a finding of a likely stable infiltration factor and therefore an estimate of the relative importance of indoor-generated and outdoor-penetrated particles.

3.5 Bennett Valley

Outdoor PurpleAir $PM_{2.5}$ measurements at the Bennett Valley PurpleAir site in Sonoma County. CA were collected using the ALT-CF3 algorithm and compared to the Oakmont site in Santa Rosa (about 4 km distant) covering the period from 6/19/2020 to 9/27/2022. Daily means were calculated for all days with the requirement that at least half of the 720 possible measurements per day were available at an acceptable precision (abs (A-B)/(A+B)) of better than 20%. 821 days met these requirements. However, an unusually high number (280) of the Bennett Valley measurements showed precision poorer than 20%. Only 21 of the 821 corresponding Oakmont measurements had precision poorer than 20%. The final database consisted of 522 days with matched indoor and outdoor daily means and precisions better than 20%.

The I/O ratio is shown for the entire period (Figure 13). There are two strong troughs in winter 2021 and 2022, with peaks in the summer of 2020 and 2021.

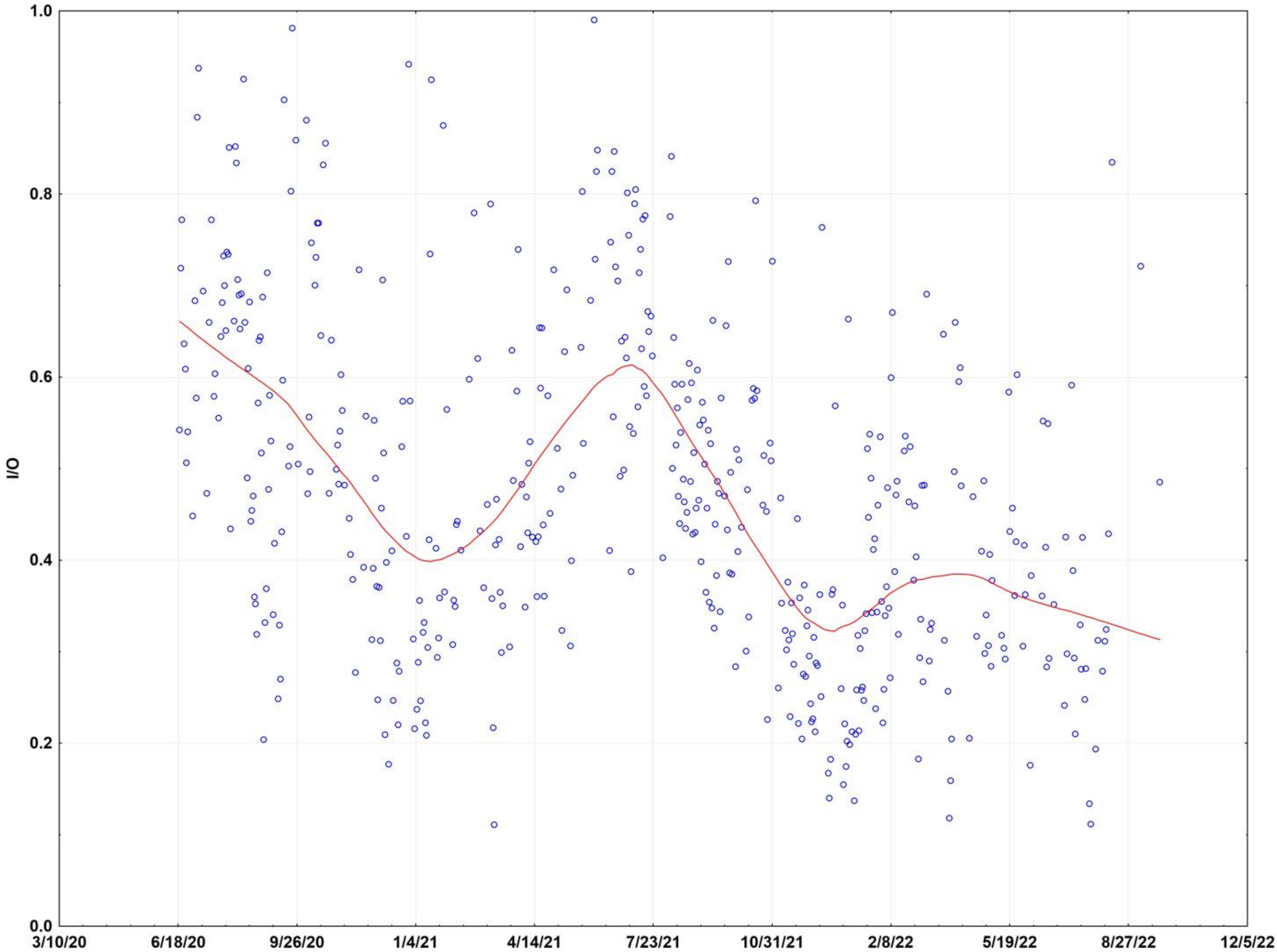

Figure 13. Indoor-outdoor (I/O) ratio of daily means at the Bennett Valley (outdoor) and Oakmont Santa Rosa (indoor) sites.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

The RCS approach to estimate the infiltration factor for the entire database results in a large number of values in the Forbidden Zone, indicating that the infiltration factor varies too much to be estimated using the entire database (Figure 14).

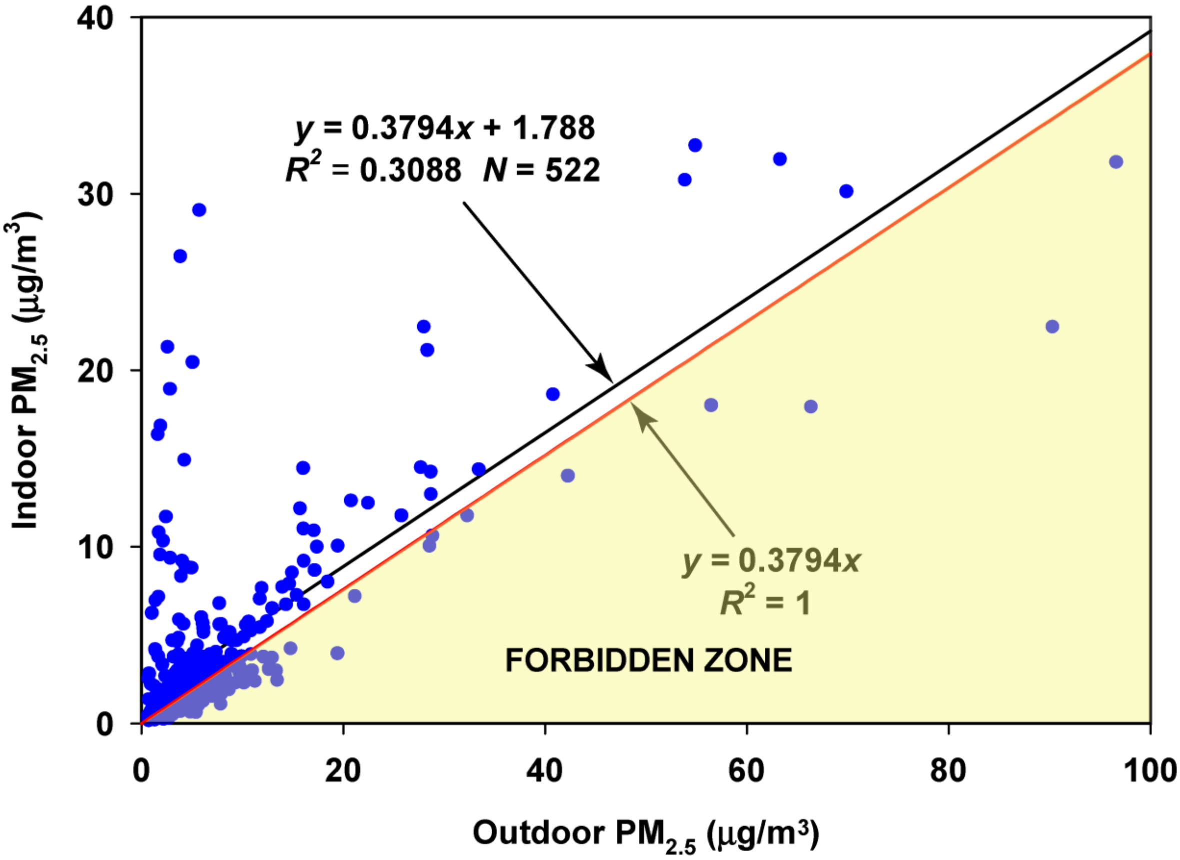

Figure 14. RCS regression of indoor values at Oakmont on outdoor values at Bennett Valley, about 4 km from Oakmont.

The variation of the I/O ratio suggests that a period of low infiltration factors might occur in the winter and perhaps the spring. Indeed, when the RCS regression is performed on the winter months (December through February) and the spring months (April-June), few points fall into the Forbidden Zone (Figure 15). Therefore the estimated infiltration factor of 0.2195 appears robust. An average contribution of 1.1 µg/m^3 is made by indoor-generated particles.

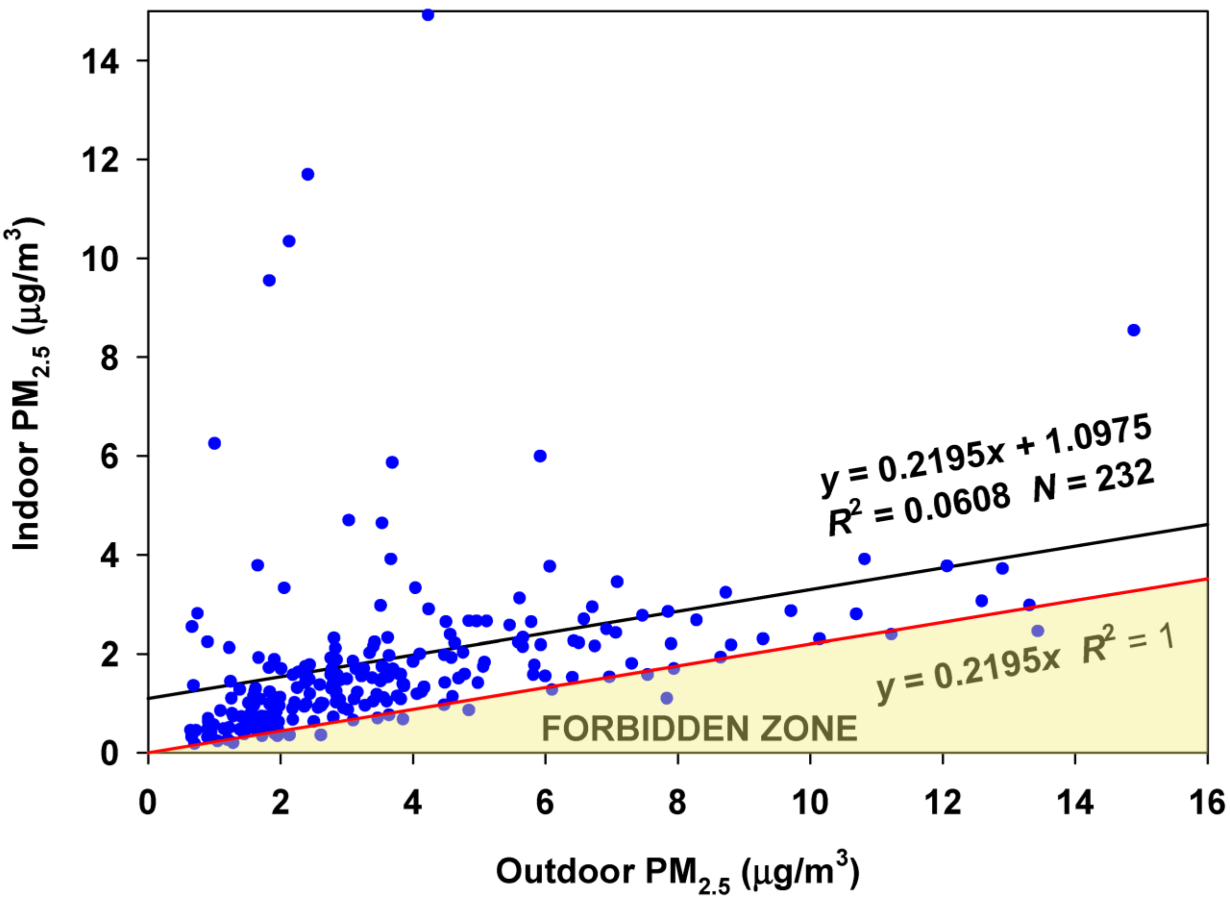

Figure 15. Regression of Oakmont indoor concentrations on Bennett Valley outdoor concentrations for the Winter and Spring months (December-June (March omitted)).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

The corresponding regression on the Summer-Fall months (August-November) again produced few observations violating the Forbidden Zone, with an estimated infiltration factor of 0.3637 and an estimated contribution of indoor-generated particles of 2.10 µg/m^3 (Figure 16).

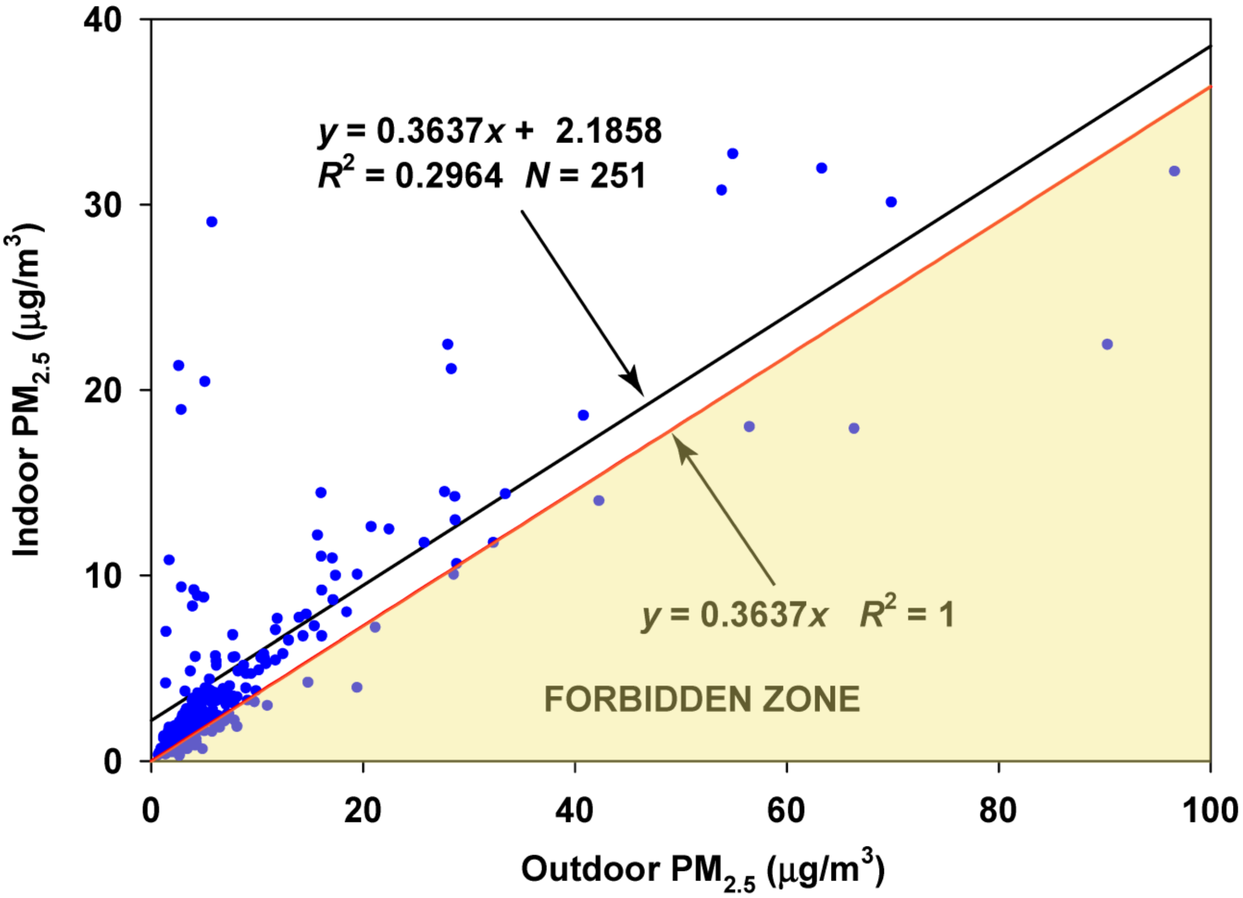

Figure 16. Regression of Oakmont indoor concentrations on Bennett Valley outdoor concentrations for the Summer and Fall months (August-November).

Plotting the indoor-outdoor ratio (I/O) vs. time was useful in showing the seasonal change in the infiltration factor. Two periods (Winter-Spring and Summer-Fall) were found to have a stable infiltration factor, allowing good estimates of both the infiltration factor and the relative contribution of indoor-generated and outdoor-penetrated particles to total indoor potential exposure. The good results for sites 4 km apart is notable, suggesting that outdoor sites may often have little spatial variation. This opens up the possibility of using distant outdoor sites (e.g., those with high-quality Federal Reference Method (FRM) $PM_{2.5}$ measurements) to calculate infiltration factors at many indoor sites.

3.6 Outer Sunset (San Francisco)

1. An outdoor and corresponding indoor site is located in the Outer Sunset area in northwest San Francisco. Data from 10/20/19 to 8/13/21 (409 days) were examined (Figure 17). This site in San Francisco had very few points falling in the Forbidden Zone. The intercept of 10.3 μg/m^3 is an unusually strong estimate of the dominance of the (unknown) indoor source(s).

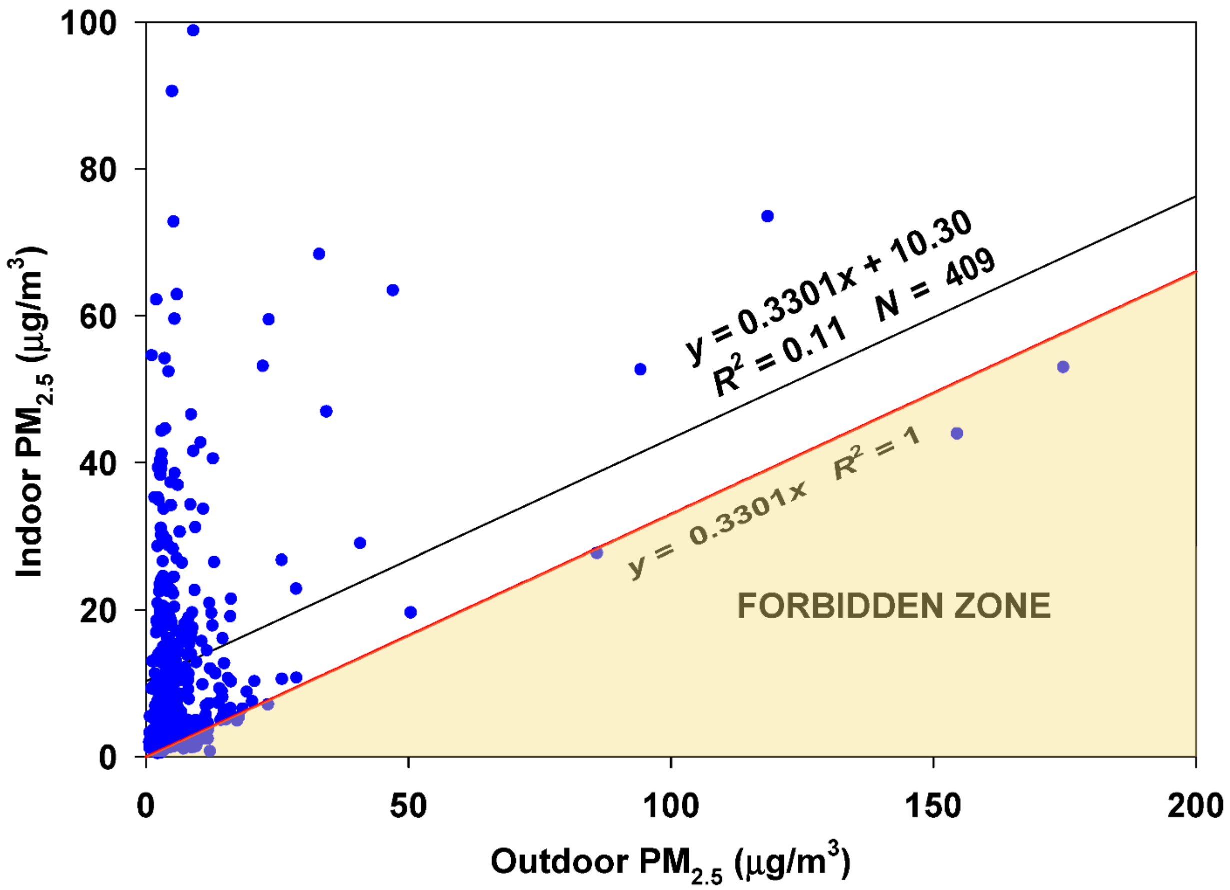

Figure 17. RCS regression of indoor on outdoor data at two nearby sites in the Outer Sunset area of San Francisco.

4. Discussion

The Alt-CF3 algorithm outperformed the Plantower CF1 algorithm in every aspect considered. Precision was better, the limit of detection was lower, the number of values > LOD was higher, and no zero values were recorded compared to hundreds of thousands of zeros for the Plantower CF1 algorithm. Although the Plantower CF_ATM algorithm was not directly considered in all analyses, it is identical to the Plantower CDF_1 algorithm for all values less than ~28 ug/m3, and since these constitute the majority of observations, and include all the zeros reported by Plantower CF_1, the findings regarding precision, LOD, and number of zeros for the CF_1 algorithm would be equally applicable to CF_ATM. Since all the alternative algorithms (or "conversion factors") other than ALT-CF3 offered by PurpleAir depend on one or another Plantower algorithm, they are all equally vulnerable to the Plantower shortcomings.

A summary of results for all six cases is offered (Table 6).

Table 6. Infiltration factors (F_{inf}) and mean $PM_{2.5}$ for outdoor-infiltrated (Out_{inf}) and indoor-generated (In_{gen}) particles, including the percent contribution to total indoor $PM_{2.5}$ from the indoor-generated particles (In_{gen} (%)). These variables are shown for the entire datasets and for the subsets (1 and 2) created when the Forbidden Zone was violated by too many observations.

Outdoor site	Oakmont	Bennett Valley	Redwood City	Alexander Ave	Menlo Park	Outer Sunset

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0301.v1

Indoor site	same	Oakmont	same	same	same	same
Finf	0.34	0.38	0.36	0.17	**0.085**	**0.33**
$PM_{2.5}$ Outinf	3.2	2.28	3.0	0.99	**0.31**	**2.6**
$PM_{2.5}$ Ingen	1.7	1.3	2.6	0.43	**1.8**	**10**
Ingen (%)	34	36	46	31	**85**	**80**
Finf 1	**0.34**	**0.36**	**0.23**	**0.080**	NA	NA
$PM_{2.5}$ Outinf 1	**2.2**	**4.10**	**2.9**	**0.44**	NA	NA
$PM_{2.5}$ Ingen 1	**2.1**	**2.2**	**2.3**	**0.51**	NA	NA
Ingen 1 (%)	**48**	**35**	**45**	**54**	NA	NA
Finf 2	**0.14**	**0.22**	NA	NA	NA	NA
$PM_{2.5}$ Outinf 2	**0.38**	**0.75**	NA	NA	NA	NA
$PM_{2.5}$ Ingen 2	**1.5**	**1.1**	NA	NA	NA	NA
Ingen 2 (%)	**80**	**59**	NA	NA	NA	NA

*Outinf = outdoor-infiltrated; Ingen = indoor-generated; 1,2=subsets 1,2; NA = Not Applicable

In Table 6, the values shown in red represent the four cases in which the Forbidden Zone was violated by numerous observations. These results have been discarded and replaced by the values shown in the bottom 8 rows for subsets of the data with many fewer violations of the Forbidden Zone.

In two cases, (Menlo Park and Outer Sunset), the number of observations violating the Forbidden Zone were near zero, and thus provided reassurance that the entire data set (last 2 columns, top four rows) could be used to determine robust estimates for the infiltration factor and also for calculating the contributions to total indoor $PM_{2.5}$ concentration of indoor-generated and outdoor-penetrated particles. In these two cases, the indoor-generated particles greatly outweighed (80-85%) particles of ambient origin. In the remaining four cases, the regression on the full dataset failed to identify a stable infiltration factor (red typeface, first four columns, top four rows). For two of these four cases (Oakmont and Bennett Valley), the use of the I/O ratio identified two 6-month periods (subsets 1 & 2, first two columns, bottom 8 rows) that could be used to calculate robust estimates of the infiltration factor for each period. In the remaining two cases (Redwood City and Alexander Avenue), it was possible to identify one subset with a stable infiltration factor (columns 3-4, rows 5-8), although the remaining subset could not be quantitatively analyzed.

The importance of determining the correct infiltration factors can be seen by considering the first column in Table 6 for Oakmont. Using the entire dataset (with the many violations of the Forbidden Zone), the contribution of indoor-generated particles was 34%, indicating that the outdoor-infiltrated particles (66%) were about twice as important. However, using the two subsets 1 & 2 with good estimates of the infiltration factor shows that the actual contribution of the indoor-generated particles was 48% in one case and 80% in the other, an average share of 64% or about twice as important as the particles of ambient origin (36%).

5. Conclusions

Epidemiological studies without indoor measurements are unable to arrive at estimates of indoor-generated particles, yet some of these particles can have health effects. The advent of low-cost particle sensors can improve our knowledge of total exposure to $PM_{2.5}$. These sensors make it possible for the first time to estimate long-term (months to years) $PM_{2.5}$ concentrations from indoor sources. One recent study has calculated such long-term exposures for more than 3000 PurpleAir indoor sites in three states over a 4.6-year period and found that indoor-generated particles make about the same contribution to total indoor exposures as particles

 doi:10.20944/preprints202212.0301.v1

of ambient origin [51]. The present study has shown how to use the Forbidden Zone to evaluate the quality of the estimate of the infiltration factor. In particular, the use of the simple I/O ratio can often suggest subsets of the data to analyze with a chance of identifying periods with a stable infiltration factor and thus providing a more accurate determination of both indoor-generated and outdoor-penetrated $PM_{2.5}$. Finally, a recent improved algorithm for calculating $PM_{2.5}$ from PurpleAir measurements was found to be superior to the proprietary algorithms offered by the manufacturer of the sensors used in PurpleAir monitors.

Author Contributions: Conceptualization, L.W and W.O.; Methodology, L.W. and W.O.; Software, L.W. and W.O.; Validation, L.W. and W.O.; Formal Analysis, L.W. and W.O.; Investigation L.W. and W.O.; Resources, L.W. and W.O.; Data Curation, L.W. and W.O.; Writing – Original Draft Preparation, L.W.; Writing – Review & Editing, L.W. and W.O.; Visualization, L.W. and W.O.; Supervision, L.W. and W.O.; Project Administration, L.W. and W.O..; Funding Acquisition, Not Applicable.

Funding: This research received no external funding.

Institutional Review Board Statement: "Not applicable" for studies not involving humans or animals.

Data Availability Statement: All raw data publicly available from PurpleAir websites (https://www2.purpleair.com/), (https://api.purpleair.com/). All RCS regression analyses available on request from the corresponding author.

Acknowledgments: We acknowledge the contribution to science made by the PurpleAir organization in making all public data available on the Internet.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Murray, C.J.L. The Global Burden of Disease Study at 30 years. *Nat Med* **2022,** *28*, 2019–2026. https://doi.org/10.1038/s41591-022-01990-1 https://www.nature.com/articles/s41591-022-01990-1

2. WHO, 2022. https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health

3. (Britannica 2022) https://www.britannica.com/event/Great-Smog-of-London

4. Klepeis, N.E.; Nelson, W.C.; Ott, W.R.; Robinson, J.; Tsang, A.M.; Switzer, P.; Behar, J.V.; Hern, S.; Engelmann, W. The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants. *J. Expos. Anal. Environ. Epidemiol* **2001**, *11*, 231-252. https://www.nature.com/articles/7500165

5. Koutrakis, P.; Briggs, S.L.K..; Leaderer, B.P. Source apportionment of indoor aerosols in Suffolk and Onondaga counties, New York. *Environ Sci Technol* **1992,** *26*, 521–527. doi:10.1021/es00027a012. https://pubs.acs.org/doi/10.1021/es00027a012

6. Özkaynak, H.; Xue, J.; Spengler, J.D.; Wallace, L.A.; Pellizzari, E.D.; Jenkins, P. Personal Exposure to Airborne Particles and Metals: Results from the Particle TEAM Study in Riverside, CA. *J Expos Anal Environ Epidemiol* **1996,** *6*, 57-78., https://pubmed.ncbi.nlm.nih.gov/8777374/

7. Wheeler, A.J.; Wallace, L.A.; Kearney, J.; Van Ryswyk, K.; You, H.; Kulka, R.; Brook, J.R.; Xu, X. Personal, indoor, and outdoor concentrations of fine and ultrafine particles using continuous monitors in multiple residences. *Aerosol Sci Technol* **2011,** *45*, 1078-1089. https://www.tandfonline.com/doi/full/10.1080/02786826.2011.580798

8. Allen, R.; Larson, T.; Sheppard, L.; Wallace, L.; Liu, L-J S. Use of real-time light scattering data to estimate the contribution of infiltrated and indoor-generated particles to indoor air. *Environ Sci Technol* **2003**, 3484-3492. https://pubs.acs.org/doi/abs/10.1021/es021007e

9. Allen, R.; Larson, T.; Sheppard, L.; Wallace, L.; Liu, L-J S. Estimated hourly personal exposures to ambient and non-ambient particulate matter among sensitive populations in Seattle, Washington. *J Air Waste Manage Assoc* **2004,** 54: 1197-1211. https://pubmed.ncbi.nlm.nih.gov/15819228/

10. Liu, L-J S.; Box, M.; Kalman, D.; Kaufman, J.; Koenig, J.; Larson, T.; Lumley, T.; Sheppard, L.; Wallace, L.A. Exposure assessment of particulate matter for susceptible populations in Seattle. *Environ Health Perspect* **2003,** *111*:909-918. https://pubmed.ncbi.nlm.nih.gov/12782491/

11. Wallace, L.; Williams, R.; Suggs, J. Estimating contributions of outdoor fine particles to indoor concentrations and personal exposures: effects of household characteristics and personal activities. **2006,** US EPA Report 600/R-06/023. https://nepis.epa.gov/Exe/ZyPDF.cgi/9100CCPM.PDF?Dockey=9100CCPM.PDF

12. Diapouli, E.; Chaloulakou, A.; Koutrakis, P. Estimating the concentration of indoor particles of outdoor origin: A review, *J Air Waste Manage Assoc* **2013,** *63*, 1113-1129. DOI: 10.1080/10962247.2013.791649. https://www.tandfonline.com/doi/full/10.1080/10962247.2013.791649

13. Wang, Z.; Delp, W.W.; Singer, B.C. Performance of low-cost indoor air quality monitors for $PM_{2.5}$ and PM_{10} from residential sources. *Build Environ* **2020,** *174*: 106654. https://doi.org/10.1016/ *Build Env*.2020.106654. https://www.sciencedirect.com/science/article/abs/pii/S0360132320300123

14. Singer, B.C.; Delp, W.W. Response of consumer and research grade indoor air quality monitors to residential sources of fine particles. *Indoor Air* **2018,** *28*:624–639. https://doi.org/10.1111/ina.12463 Accessed Jan 9, 2022.

15. Dockery, D.W.; Spengler, J.D.. Indoor-outdoor relationships of respirable sulfates and particles. *Atmos Environ* **1981,** *15*, 335–343. doi:10.1016/0004-6981(81)90036-6. https://www.sciencedirect.com/science/article/abs/pii/0004698181900366

16. Sarnat, J.A.; Long, C.M.; Koutrakis, P.; Coull, B.A.; Schwartz, J.; Suh, H.H. Using sulfur as a tracer of outdoor fine particulate matter. *Environ Sci Technol* **2002,** *36*:5305–5314. doi:10.1021/es025796b https://pubmed.ncbi.nlm.nih.gov/12521154/

17. Wallace, L.A.; Williams, R.; Rea, A.; Croghan, C. Continuous weeklong measurements of personal exposures and indoor concentrations of fine particles for 37 health-impaired North Carolina residents for up to four seasons. *Atmos Environ* **2006,** 40:399-414. https://www.sciencedirect.com/science/article/abs/pii/S1352231005009106

18. Wallace, L.; Williams, R. Use of personal–indoor–outdoor sulfur concentrations to estimate the infiltration factor and outdoor exposure factor for individual homes and persons. *Environ Sci Technol* **2005,** *39*,1707–1714. doi:10.1021/es049547u https://pubs.acs.org/doi/10.1021/es049547u

19. Williams, R.; Suggs, J.; Rea, A.; Sheldon, L.; Rodes, C.; Thornburg, J. The Research Triangle Park particulate matter panel study: Modeling ambient source contribution to personal and residential PM mass concentrations. *Atmos Environ* **2003,** *37*, 5365–5378. doi:10.1016/ *Atmos Environ*. 2003.09.010.https://www.sciencedirect.com/science/article/abs/pii/S1352231003007507

20. Kearney, J.; Wallace, L.; MacNeill, M.; Xu, X.; Van Ryswyk, K.; You, H.; Kulka, R.; Wheeler, A.J. Residential indoor and outdoor ultrafine particles in Windsor, ON. *Atmos Environ* **2011,** *45*, 7583-7593. https://www.sciencedirect.com/science/article/abs/pii/S1352231010009428

21. MacNeill, M.; Wallace, L.; Kearney, J.;, Allen R.W.; Van Ryswyk, K.; Judek S.; Xu, X.; Wheeler A.J. Factors Influencing Variability in the Infiltration of $PM_{2.5}$ Mass and its Components. *Atmospheric Environ,* **2012,** *61*, 518-532. https://www.sciencedirect.com/science/article/abs/pii/S1352231012006772

22. MacNeill, M.; Kearney, J.; Wallace, L.; Gibson, M.; Heroux, M.E.; Kuchta, J.; J. R. Guernsey, J.R.; Wheeler, A.J. Quantifying the contribution of ambient and indoor-generated fine particles to indoor air in residential environments. *Indoor Air* **2014,** *24*, 362-375. doi:10.1111/ina.12084. https://pubmed.ncbi.nlm.nih.gov/24313879/

23. Kearney, J.; Wallace, L.; MacNeill, M.; Heroux, M.-E.; Kindzierski, W.; Wheeler, A. Residential infiltration of fine and ultrafine particles in Edmonton. *Atmos Environ* **2014,** *94*, 793-805. https://www.sciencedirect.com/science/article/pii/S1352231014003677

24. Switzer, P.; Ott, W. Derivation of an indoor air averaging time model from the mass balance equation for the case of independent source inputs and fixed air exchange rates. *J Expos Anal Environ Epidemiol* **1992,** *2*, 113–135. https://cfpub.epa.gov/si/si_public_record_Report.cfm?Lab=ORD&dirEntryID=50803

25. Allen, R.; Wallace, L.A.; Larson, T.; Sheppard, L.; Liu, L-J S. Evaluation of the recursive model approach for estimating particulate matter infiltration efficiencies using continuous light scattering data. *J Expos Sci Environ Epidemiol* **2007,** *17*, 468-477. https://www.nature.com/articles/7500539

26. Thatcher, T.L.; Layton, D.W. Deposition, resuspension, and penetration of particles within a residence. *Atmos Environ* **1995,** *29*, 1487–1497. doi:10.1016/1352-2310(95)00016-R. https://www.sciencedirect.com/science/article/abs/pii/135223109500016R

27. Ferro, A.R., Kopperud, R.J.; Hildemann, L.M. Source strengths for indoor human activities that resuspend particulate matter. *Environ Sci Technol*, **2004,** *38*, 1759–1764. doi:10.1021/es0263893. https://pubs.acs.org/doi/10.1021/es0263893

28. Hänninen, O.O., Lebret, E.; Ilacqua, V.; Katsouyanni, K.; Kunzli, N.; Sram, R.J.; Jantunen, M. Infiltration of ambient $PM_{2.5}$ and levels of indoor generated non-ETS $PM_{2.5}$ in residences of four European cities. *Atmos Environ* **2004,** *38*, 6411–6423. doi:10.1016/j.atmosenv.2004.07.015 https://www.sciencedirect.com/science/article/abs/pii/S1352231004007010

29. Kopperud, R.J., Ferro, A.R.; Hildemann, L.M. Outdoor versus indoor contributions to indoor particulate matter (PM) determined by mass balance methods. *J Air Waste Manage Assoc* **2004** 54:1188–1196. doi:10.1080/ 10473289.2004.10470983. https://pubmed.ncbi.nlm.nih.gov/15468671/

30. Bi, J.; Wallace, L.; Sarnat, J.A.; Liu, Y. Characterizing outdoor infiltration and indoor contribution of $PM_{2.5}$ with citizen-based low-cost monitoring data. *Environ Pollut* **2021,** *276:*116793. https://pubmed.ncbi.nlm.nih.gov/33631689/

31. Ott, W., L. Wallace, and D. Mage. Predicting particulate (PM10) personal exposure distributions using a random component superposition statistical model. *J Air Waste Manage Assoc* **2000,** 50:1390–1406. doi:10.1080/ 10473289.2000.10464169. https://www.tandfonline.com/doi/pdf/10.1080/10473289.2000.10464169

32. Repace, J.L.; Lowrey, A.H. Indoor air pollution, tobacco smoke, and public health *Science* **1980** 208, 464. https://pubmed.ncbi.nlm.nih.gov/7367873/

33. Ott, W.R.; Wallace, L.A.; Cheng, K-C,; Hildemann, L.M. Measuring $PM_{2.5}$ concentrations from secondhand tobacco vs. marijuana smoke in 9 rooms of a detached 2-story house, *Sci Tot Environ* **2022**, https://pubmed.ncbi.nlm.nih.gov/36037897/

34. Ott, W.R.; Zhao, T.; Cheng, K-C.; Wallace, L.A.; Hildemann, L.M. Measuring indoor fine particle concentrations, emission rates, and decay rates from cannabis use in a residence. *Atmos Environ X* **2021,** *10* 100106. https://doi.org/10.1016/j.aeaoa.2021.100106

35. Wallace, L.; Ott, W.; Zhao, T., Cheng, K-C.; Hildemann, L. Secondhand exposure from vaping marijuana: Concentrations, emissions, and exposures determined using both research-grade and low-cost monitors, *Atmos Environ X*, **2020,** *10***,** 100093. https://doi.org/10.1016/j.aeaoa.2020.100093 (Accessed 11/19/20).

36. Zhao, T.; Cheng, K-C; Ott, W.R.; Wallace L.A.; Hildemann, L.M. Characteristics of secondhand cannabis smoke from common smoking methods: calibration factor, emission rate, and particle removal rate. *Atmos Environ* **2020**, *242* 117731. https://doi.org/10.1016/j.atmosenv.2020.117731 (Accessed 11/19/20)

37. Wallace, L.; Jeong, S-G.; Rim D. Dynamic behavior of indoor ultrafine particles (2.3-64 nm) due to burning candles in a residence. *Indoor Air*. **2019** *29*, 1018–1027. https://doi.org/10.1111/ina.12592

38. Wallace, L.A.; Emmerich, S.J.; Howard-Reed, C. Source strengths of ultrafine and fine particles due to cooking with a gas stove. *Environ Sci Tech* **2004a,** *38*, 2304-2311. https://pubs.acs.org/doi/10.1021/es0306260.

39. Wallace, L.A.; Emmerich, S.J.; Howard-Reed, C. Effect of central fans and in-duct filters on deposition rates of ultrafine and fine particles in an occupied townhouse. *Atmos Environ* **2004b,** *38*, 405-413. https://www.researchgate.net/publication/222528671_Effect_of_central_fans_and_in-duct_filters_on_deposition_rates_of_ultrafine_and_fine_particles_in_an_occupied_townhouse

40. Howard-Reed, C.; Wallace, L.A.; Emmerich, S.J. Effect of ventilation systems and air filters on decay rates of particles produced by indoor sources in an occupied townhouse. *Atmos Environ* **2003,** *37*, 5295-5306. https://www.sciencedirect.com/science/article/abs/pii/S1352231003007532

41. Wallace, L.A.; Howard-Reed, C.H.; Emmerich, S.J. Continuous measurements of air change rates in an occupied house for one year: the effect of temperature, wind, fans, and windows. *Expos Anal Environ Epidemiol* **2002,** *12*, 296-306. https://www.nature.com/articles/7500229

42. Howard-Reed, C.H., Wallace, L.A., and Ott, W.R. The effect of opening windows on air change rates in two homes. *J Air Waste Manage Assoc* **2002** *52*, 147-159. https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=65288&Lab=NERL

43. AQ-SPEC 2016. http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/purpleair—field-evaluation.pdf. Accessed 19 December 2020

44. Barkjohn, K.; Gantt, B.; Clements, A.L. Development and application of a United States wide correction for $PM_{2.5}$ data collected with the PurpleAir sensor. *Atmos Meas. Techniques* **2020,** *4*, 10.5194. https://doi.org/10.5194/amt-2020-413 Accessed July 6, 2022

45. Liang, Y.; Sengupta, D.; Campmier, M.J.; Lunderberg, D.M.; Apte, J.S.; Goldstein, A. Wildfire smoke impacts on indoor air quality assessed using crowdsourced data in California. *Proc National Academy Sciences* **2021**,*118,* e2106478118. https://www.pnas.org/doi/full/10.1073/pnas.2106478118. Accessed July 6, 2022.

46. Wallace, L.; Bi, J.; Ott, W.R.,; Sarnat, J.A.; Liu, Y. Calibration of low-cost PurpleAir outdoor monitors using an improved method of calculating $PM_{2.5}$. *Atmos Environment*, **2021,** *256* 118432. https//doi.org/10.1016/j.atmosenviron.2021.118432. https://www.sciencedirect.com/science/article/abs/pii/S135223102100251X

47. Wallace, L.; Zhao, T.; Klepeis, N.E.. Calibration of PurpleAir PA-I and PA-II monitors using daily mean $PM_{2.5}$ concentrations measured in California, Washington, and Oregon from 2017 to 2021. *Sensors* 2022, *22*, 4741. https://doi.org/10.3390/ s22134741 https://pubmed.ncbi.nlm.nih.gov/35808235/

48. Wallace, L. Intercomparison of PurpleAir sensor performance over three years indoors and outdoors at a home: bias, precision, and limit of detection using an improved algorithm for calculating $PM_{2.5}$. *Sensors* **2022**, *22*, 2755. https://doi.org/10.3390/s22072755

49. Wallace, L.A.; Wheeler, A.; Kearney, J.; Van Ryswyk, K.; You, H.; Kulka, R.; Rasmussen, P.; Brook, J.; Xu, X. Validation of continuous particle monitors for personal, indoor, and outdoor exposures. *J Expos Sci & Environ Epidemiol* **2010,** *21*, 49-64. https://www.nature.com/articles/jes201015

50. Delp, W.; Singer B. Wildfire smoke adjustment factors for low-cost and professional $PM_{2.5}$ monitors with optical sensors. *Sensors* **2020**, *20*, 3683; doi:10.3390/s20133683. https://www.mdpi.com/1424-8220/20/13/3683

51. Wallace, L.A.; Zhao, T.; Klepeis, N.R. Indoor contribution to $PM_{2.5}$ exposure using all PurpleAir sites in Washington, Oregon, and California. *Indoor Air* **2022,** *32*,13105. https://onlinelibrary.wiley.com/doi/abs/10.1111/ina.13105.