Supporting Information

Variation in gravimetric correction factors for nephelometer-derived estimates of personal exposure to $PM_{2.5}$

Jessica Tryner[†], Nicholas Good[‡], Ander Wilson,[§] Maggie L. Clark[‡], Jennifer L. Peel[‡], and John Volckens[†‡]

[†] *Department of Mechanical Engineering, Colorado State University, Fort Collins, Colorado, USA*

[‡] *Department of Environmental & Radiological Health Sciences, Colorado State University, Fort Collins, Colorado, USA*

[§] *Department of Statistics, Colorado State University, Fort Collins, Colorado, USA*

S1 Methods

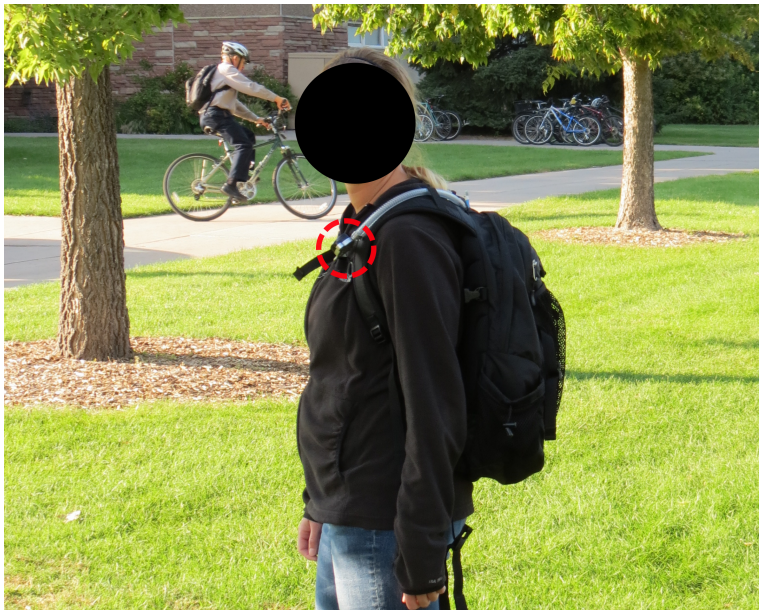

Figure S1. A photograph illustrating how the backpack was worn by participants. The $PM_{2.5}$ inlet is circled with a dashed red line.

S1.1 Quality assurance

As a result of applying the RH correction given by Equation 1, the 33-hour average nephelometer-derived $PM_{2.5}$ concentration decreased by less than 15% for 317 of the 333 samples for which nephelometer data were available (**Figure S2**). The concentration decreased by less than 5% for 151/333 samples.

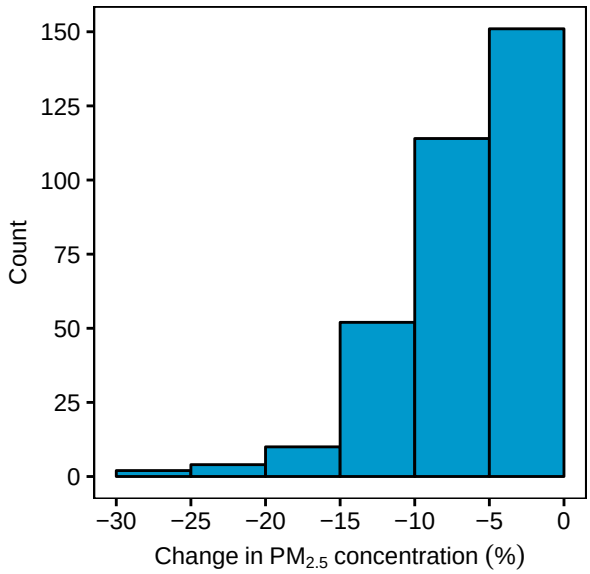

Figure S2. The percent change in the 33-hour average $PM_{2.5}$ concentration that resulted from correcting the 10-s average nephelometer readings for relative humidity using Equation 1.

S1.2 Intraclass correlation coefficient

The intraclass correlation coefficient (ICC) was used to assess the fraction of the variance in the nephelometer/filter ratio that could be explained by (1) differences between versus within participants and (2) differences between versus within sample dates. The ICC was determined using a one-way random effects model (McGraw, Wong, 1996):

$$\begin{aligned} x_{ij} &= \mu + r_i + w_{ij} \\ r_i &\overset{iid}{\sim} N(0, \sigma_r^2) \\ w_{ij} &\overset{iid}{\sim} N(0, \sigma_w^2) \end{aligned} \tag{S1}$$

where x_{ij} was observation j for group i, μ was the population mean, r_i was the random effect associated with group i, and w_{ij} represented the random residual effects. Data were grouped by participant or by sample day. Both r_i and w_{ij} were assumed to be independent and normally distributed with means of zero and constant variances of σ_r^2 and σ_w^2, respectively. The nephelometer/filter ratio was log-transformed to satisfy model assumptions.

Using this model, the ICC was defined as shown in Equation S2 (McGraw, Wong, 1996):

$$ICC = \frac{\sigma_r^2}{\sigma_r^2 + \sigma_w^2} \tag{S2}$$

S1.3 Sensitivity analyses

The sensitivity analyses were used to examine how the results were affected by: (a) not replacing 10-second average nephelometer readings below the LOD with $\text{LOD}/\sqrt{2}$ and (b) filtering the data based on different criteria. For the data used in the sensitivity analyses, raw 10-second average concentrations recorded by the nephelometer were not adjusted based on the LOD of the instrument. Instead, all 10-second average as-recorded values (regardless of concentration) were RH-corrected and then averaged over the 33-hour sample

period. The full dataset (377 samples) was then filtered in four steps based on progressively more stringent criteria.

In the first filtering step, samples were only retained if: (1) 10-second average nephelometer measurements were available for at least 85% of the sample period (309/377), (2) the mass accumulated on the filter was above the LOD (292/377 samples), and (3) the filter-derived $PM_{2.5}$ concentration was less than 145 $\mu g \cdot m^{-3}$ (375/377). These were the same filtering criteria used for the main analysis. A total of 249 samples were retained after these three criteria were applied.

In the second filtering step, samples were only retained if fewer than 25% of the 10-s average nephelometer readings were equal to zero. Nephelometer readings of zero bias the nephelometer/filter ratio because they contribute nothing to the total $PM_{2.5}$ mass measured by the nephelometer, while the integrated filter simultaneously measures some (probably) small, but (almost certainly) non-zero, concentration. There were 159 samples retained after this additional criterion was applied.

In the third filtering step, samples were only retained if fewer than 50% of the 10-s average nephelometer readings were below the LOD. These samples were excluded due to the uncertainty associated with readings below the LOD. There were 81 samples retained after this additional criterion was applied.

In the fourth filtering step, samples were only retained if the 33-hour average nephelometer-derived concentration was greater than 4 $\mu g \cdot m^{-3}$ (the LOD for the filter-derived concentrations). There were 68 samples retained after this additional criterion was applied.

S2 Results

All participants were mobile adults who worked outside their homes. Data on the distribution of participant age, gender, and time spent in five different microenvironment categories are shown in **Table S1**. On average, participants spent 58% of their time at home, 20% of their time at work, 6% of their time in transit, 2% of their time at an eatery, 8% of their time in another microenvironment, and nephelometer data were not available 6% of the time.

Table S1. Age, gender, and fraction of time spent in each microenvironment category for 249 samples collected by 44 participants.

Gender			**Age**			**Time spent in each microenvironment** (%, by sample)				
	By particpant	By sample	Range (years)	By participant	By sample	Quartile	Home	Work	Transit	Eatery	Other
Male	19	89	20–29	12	66	Min	9	0	1	0	0
Female	25	160	30–39	12	73	25^{th}	54	18	4	0	1
			40–49	5	23	50^{th}	59	22	5	0	5
			50–59	8	51	75^{th}	64	25	7	2	10
			60–69	1	4	Max	97	43	14	22	68
			NA	6	32	Mean	58	20	6	2	8
Total	44	249	Total	44	249	n	249	249	249	249	249

For the adjusted data set (i.e, the main analysis), filter-derived personal $PM_{2.5}$ concentrations ranged from 5 to 83 $\mu g \cdot m^{-3}$, with a median of 8 $\mu g \cdot m^{-3}$ (**Figure S3**). This median concentration was similar to the annual average ambient $PM_{2.5}$ concentration in Fort Collins (7 $\mu g \cdot m^{-3}$) (CDPHE, 2013, 2014); however, the filter-derived 33-hour personal $PM_{2.5}$ concentration was only weakly correlated with the average ambient $PM_{2.5}$ concentration measured by the TEOM over the 33-hour sample duration (Spearman's $\rho = 0.34$; **Figure S4**).

The absolute difference between the nephelometer- and filter-derived 33-hour average concentrations was $\leq$ 5 $\mu g \cdot m^{-3}$ for 73% of samples (**Figure S5b**). A difference of 5 $\mu g \cdot m^{-3}$ is small from an absolute standpoint but represents a percent difference of 63% for the median concentration of 8 $\mu g \cdot m^{-3}$ (**Figure S5c**). The median percent difference between the nephelometer- and filter-derived concentrations was 49%.

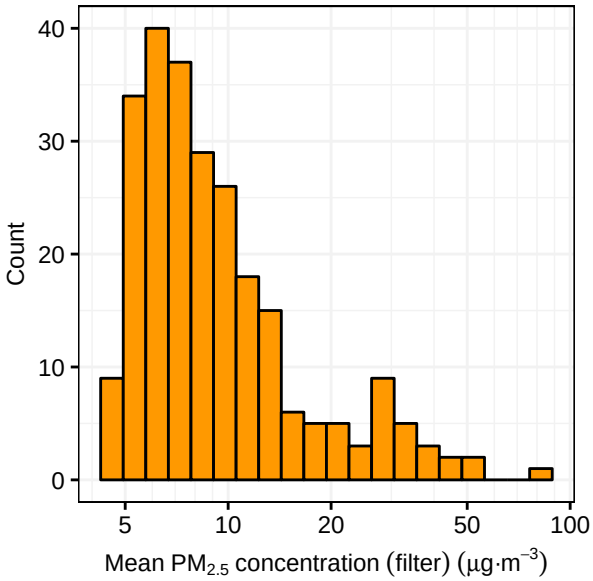

Figure S3. Histogram of filter-derived 33-hour average personal $PM_{2.5}$ concentrations for 249 samples for which nephelometer data were available for at least 85% of the sample period, the change in filter mass was above the LOD, and the filter-derived concentration was less than 145 $\mu g \cdot m^{-3}$.

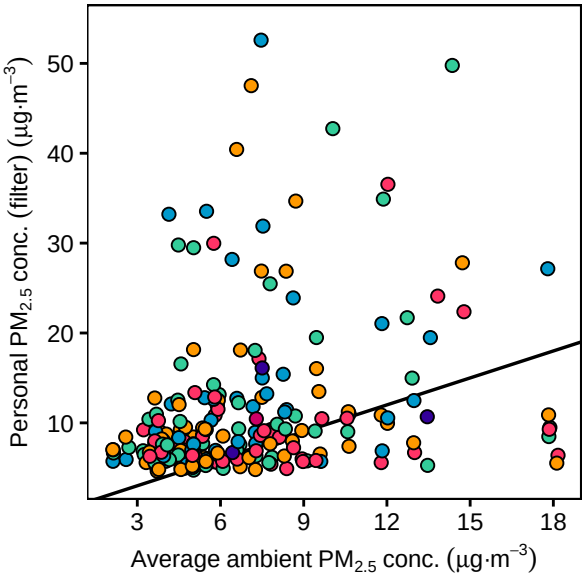

Figure S4. Filter-derived 33-hour average personal $PM_{2.5}$ concentrations compared to average ambient $PM_{2.5}$ concentrations in Fort Collins during each 33-hour sample period. Data are shown for 249 samples for which nephelometer data were available for at least 85% of the sample period, the change in filter mass was above the LOD, and the filter-derived concentration was < 145 $\mu g \cdot m^{-3}$. One data point (5,83) is not shown. The diagonal line is y = x.

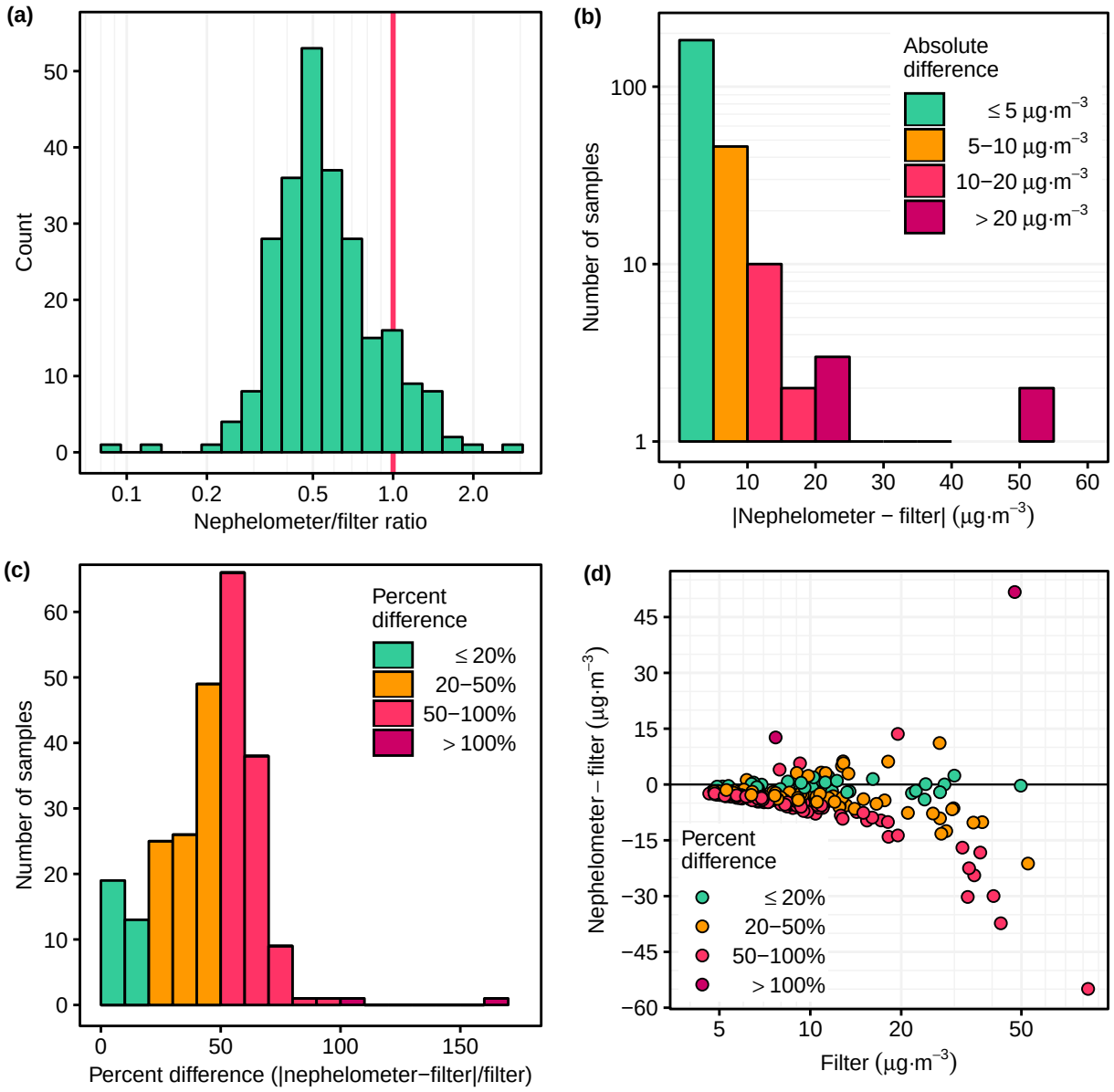

Figure S5. Comparisons of the nephelometer- and filter-derived 33-hour average personal $PM_{2.5}$ mass concentrations (n = 249): **(a)** a histogram of nephelometer/filter ratios (also Figure 4 in the main text), **(b)** a histogram of the absolute difference between the two concentrations, **(c)** a histogram of the percent difference between the two concentrations, and **(d)** the difference between the two concentrations compared to the filter-derived concentration.

S2.1 Comparisons of collocated filter samples

The extent to which imprecision in the filter measurements contributes to the variability in the nephelometer/filter ratio shown in **Figure 4** is unknown. For comparison, ratios of $PM_{2.5}$ concentrations measured in previous studies using collocated filter measurements are shown in **Figure S6**.

The ratios of 43 paired 24-hour personal $PM_{2.5}$ concentrations measured during the Honduras Cookstove Intervention Study using an Ultrasonic Personal Aerosol Sampler (UPAS) and a conventional personal filter sampling train had a median value of 0.88 and varied by a factor of 1.5 between the 10th percentile (0.72) and the 90th percentile (1.1) (Pillarisetti et al., 2019). Note, however, that the median 24-hour personal $PM_{2.5}$ concentration measured during the Cookstove Intervention Study (52 $\mu g \cdot m^{-3}$; range = 23–131 $\mu g \cdot m^{-3}$) was much higher than the median 33-hour personal $PM_{2.5}$ concentration measured during the Commuter Study (8 $\mu g \cdot m^{-3}$; range = 5–83 $\mu g \cdot m^{-3}$). As a result, even though the median absolute difference between the $PM_{2.5}$ concentrations measured using the UPAS and the conventional filter sample (5.5 $\mu g \cdot m^{-3}$) was higher than the median absolute difference between the nephelometer- and filter-derived $PM_{2.5}$ concentrations measured during the Commuter Study (3.6 $\mu g \cdot m^{-3}$), the median percent difference between the UPAS and the conventional filter sample was only 12% (compared to 49% for the nephelometer- and filter-derived $PM_{2.5}$ concentrations).

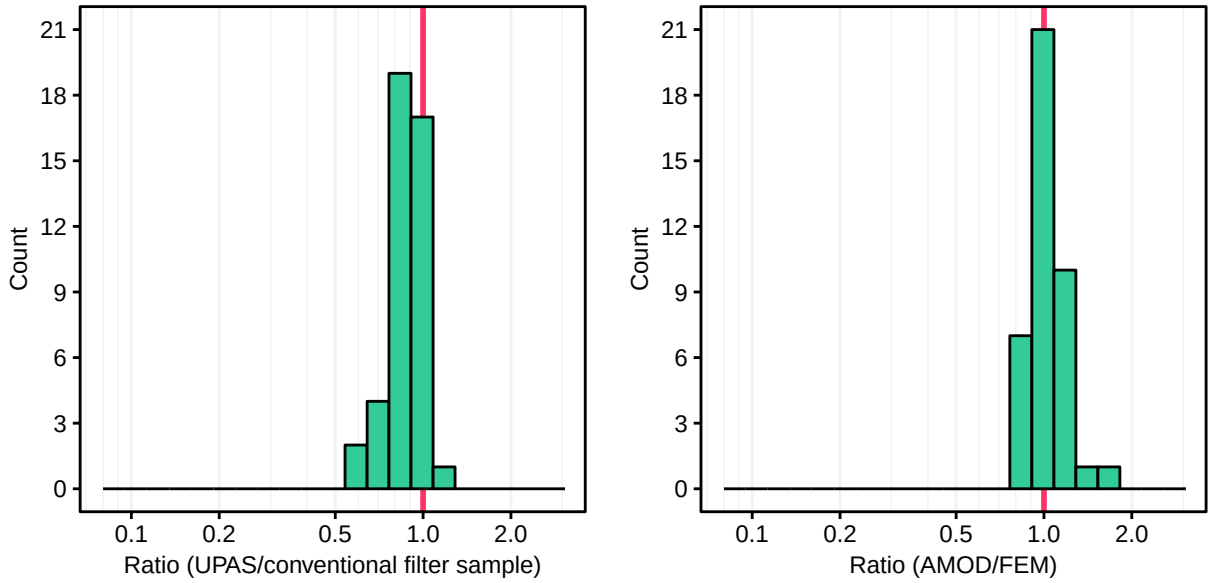

Figure S6. Left: A histogram of the ratio of 24-hour personal $PM_{2.5}$ concentrations measured using the Ultrasonic Personal Aerosol Sampler and a conventional filter sample during the Honduras Cookstove Intervention Study (n = 43). Right: A histogram of the ratio of 48-hour ambient $PM_{2.5}$ concentrations measured during collocations of the Aerosol Mass and Optical Depth (AMOD) sampler and an FEM filter sampler in Fort Collins (n = 40).

The ratios of 40 paired 48-hour ambient $PM_{2.5}$ concentrations measured during collocations of the Aerosol Mass and Optical Depth (AMOD) sampler and a conventional 16.7 $L \cdot min^{-1}$ filter sampler in Fort Collins, Colorado had a median value of 1.01 and varied by a factor of 1.3 between the 10th percentile (0.90) and the 90th percentile (1.2) (Wendt, 2018). The median 48-hour ambient $PM_{2.5}$ concentration measured during these collocations (8 $\mu g \cdot m^{-3}$; range = 4–13 $\mu g \cdot m^{-3}$) was similar to the median 33-hour personal $PM_{2.5}$ concentration measured during the Commuter Study (8 $\mu g \cdot m^{-3}$; range = 5–83 $\mu g \cdot m^{-3}$); however, Wendt (2018) collected stationary measurements of ambient $PM_{2.5}$ concentrations. During personal sampling, participant motion and short-term variations in aerosol concentrations and properties may lead to reduced measurement precision.

S2.2 Intraclass correlation coefficient

The nephelometer/filter ratios for samples collected by different participants on the same day are shown in **Figure S7**. On 2013/06/26, the nephelometer/filter ratios for four participants varied by a factor of 1.1; however, on 2013/02/11, the nephelometer/filter ratios for four participants varied by a factor of 5.4. The ICC for the data grouped by sample date (0.14, 95% CI = 0–0.29) indicated that differences between sample dates accounted only 14% of the variance in the log-transformed ratio.

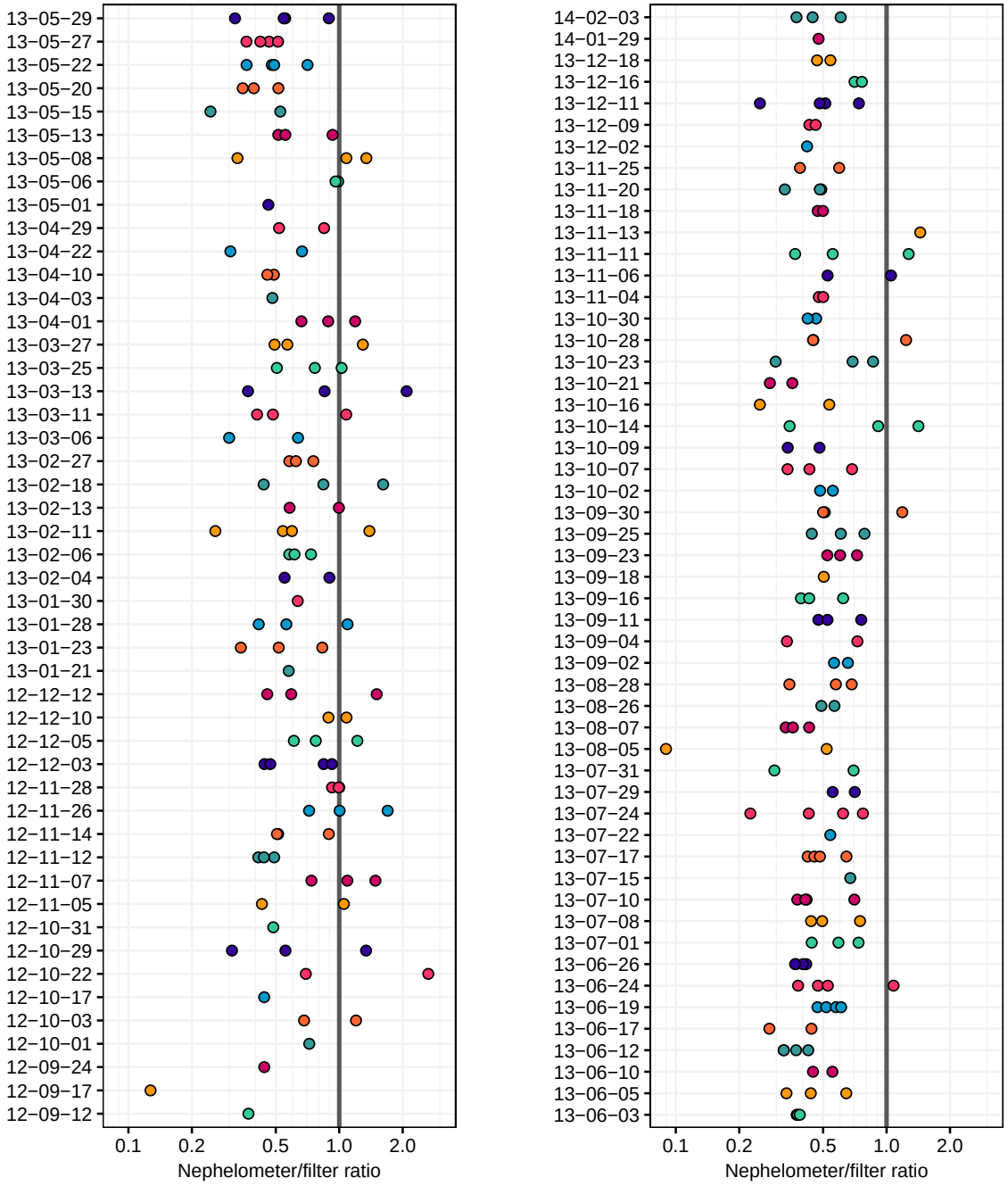

Figure S7. Nephelometer/filter ratio vs. sample collection date. The solid vertical line represents a ratio of 1.

S2.3 Sensitivity analyses

As filtering progressed from the 1st to 4th steps in the sensitivity analysis, more samples with low average $PM_{2.5}$ concentrations were removed from the data set, and the median filter-derived personal $PM_{2.5}$ concentration increased from 8 to 12 $\mu g \cdot m^{-3}$ (**Figure S8**). As a greater number of samples with low concentrations were removed, many samples for which the nephelometer underestimated the filter-derived $PM_{2.5}$ concentration were removed, and the median nephelometer/filter ratio increased from 0.44 to 0.74. In addition, the fraction of nephelometer/filter ratios equal to 1.0 ± 0.2 increased from 11% to 29% (**Figure S10a**).

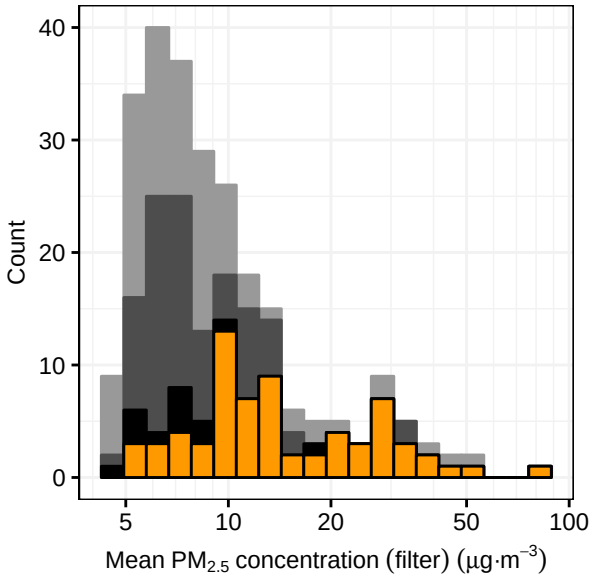

Figure S8. Histogram of filter-derived personal $PM_{2.5}$ concentrations retained after each filtering step. Samples excluded by the second, third, and fourth filtering steps are shown in light gray, dark gray, and black, respectively.

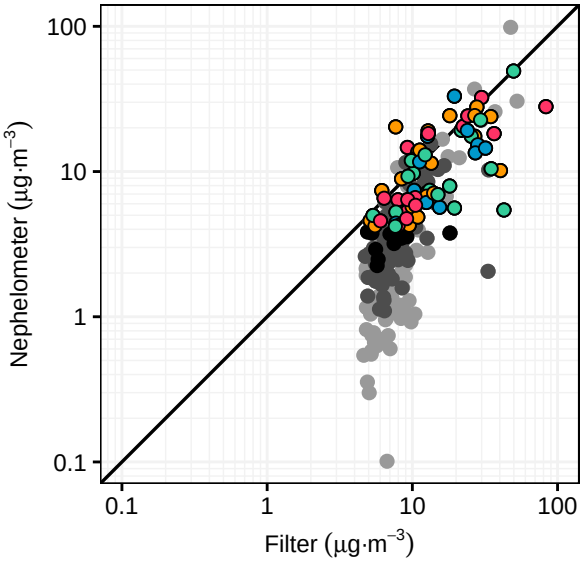

Figure S9. A comparison of the average filter- and nephelometer-derived 33-hour average personal $PM_{2.5}$ concentrations retained after each filtering step. Samples excluded by the second, third, and fourth filtering steps are shown in light gray, dark gray, and black, respectively. The solid line is y = x.

For all five data sets shown in **Table 1** (main manuscript), the absolute difference between the nephelometer- and filter-derived $PM_{2.5}$ concentrations was less than 5 $\mu g \cdot m^{-3}$ for 57% to 73% of samples (**Figure S5b** and **Figure S10b**), the median absolute difference ranged from 3.6 to 4.5 $\mu g \cdot m^{-3}$, the percent difference between the nephelometer- and filter-derived $PM_{2.5}$ concentrations was less than 20% for 11% to 29% of samples (**Figure S5c** and **Figure S10c**), the median percent difference ranged from 36% to 57%, and the fraction of nephelometer/filter ratios less than one ranged from 71% to 89% (**Figure S5a** and **Figure S10a**).

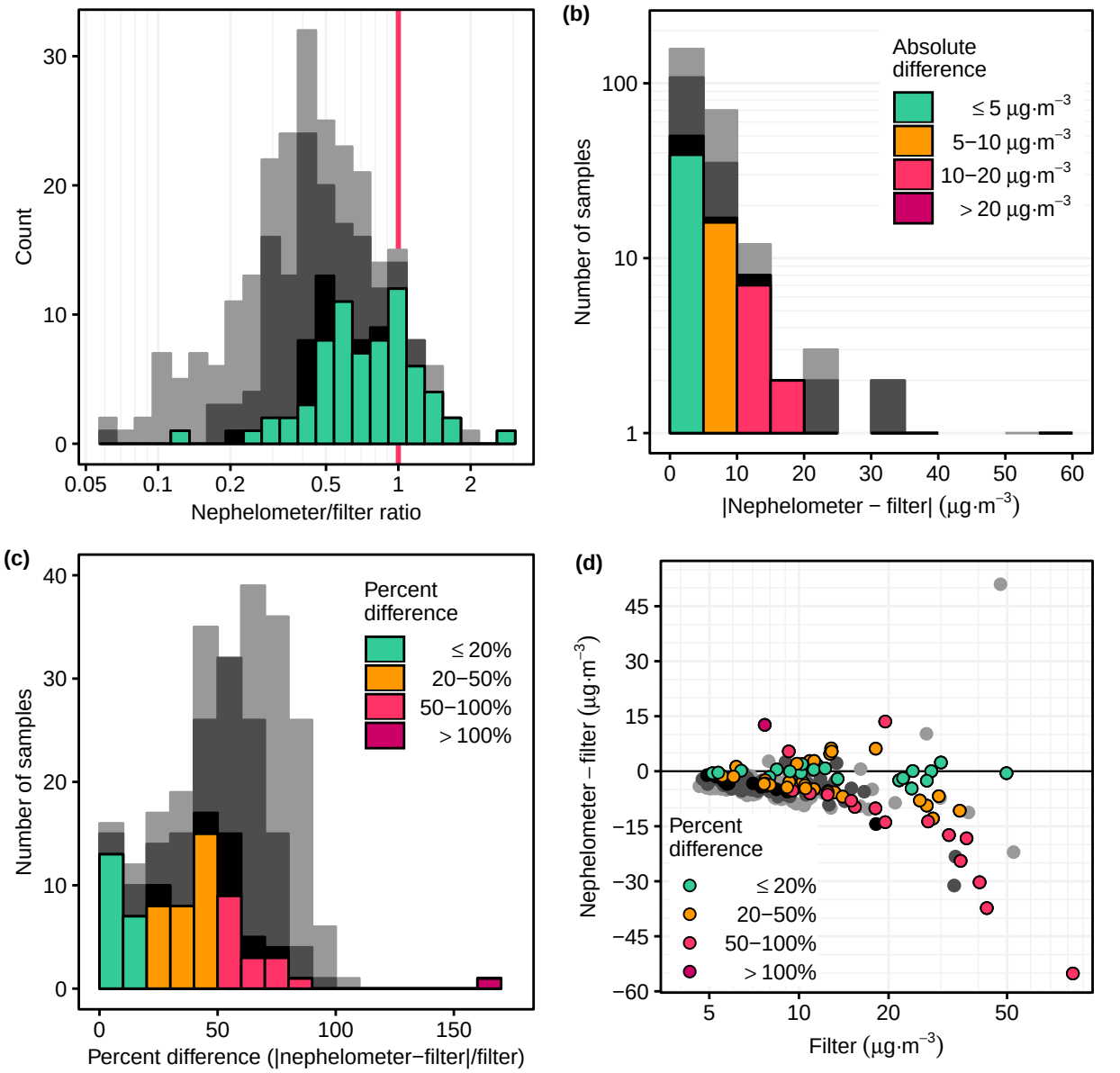

Figure S10. Comparisons of the nephelometer- and filter-derived 33-hour average personal $PM_{2.5}$ mass concentrations: **(a)** a histogram of nephelometer/filter ratios, **(b)** a histogram of the absolute difference between the two concentrations, **(c)** a histogram of the percent difference between the two concentrations, and **(d)** the difference between the two concentrations compared to the filter-derived concentration. All samples retained after the first filtering step (n = 249) are shown. Samples excluded by the second, third, and fourth filtering steps are shown in light gray, dark gray, and black, respectively. One nephelometer/filter ratio of 0.015 is not shown; that sample was excluded by the second filtering step.

For all five data sets listed in **Table 1**, the ICC calculated from a linear mixed model using repeated samples for each participant ranged from 0.23 to 0.30 (**Figure S11**). The ICC calculated from a linear mixed model using repeated samples for each date ranged from 0.00 to 0.20 (**Figure S12**). For every data set, the 95% confidence interval for the ICC calculated from a linear mixed model using repeated samples for each date included zero.

Results were not affected by restricting the four data sets listed in **Table 1** to samples for which 10-s average nephelometer measurements were available for at least 90% and 95% of the sample period.

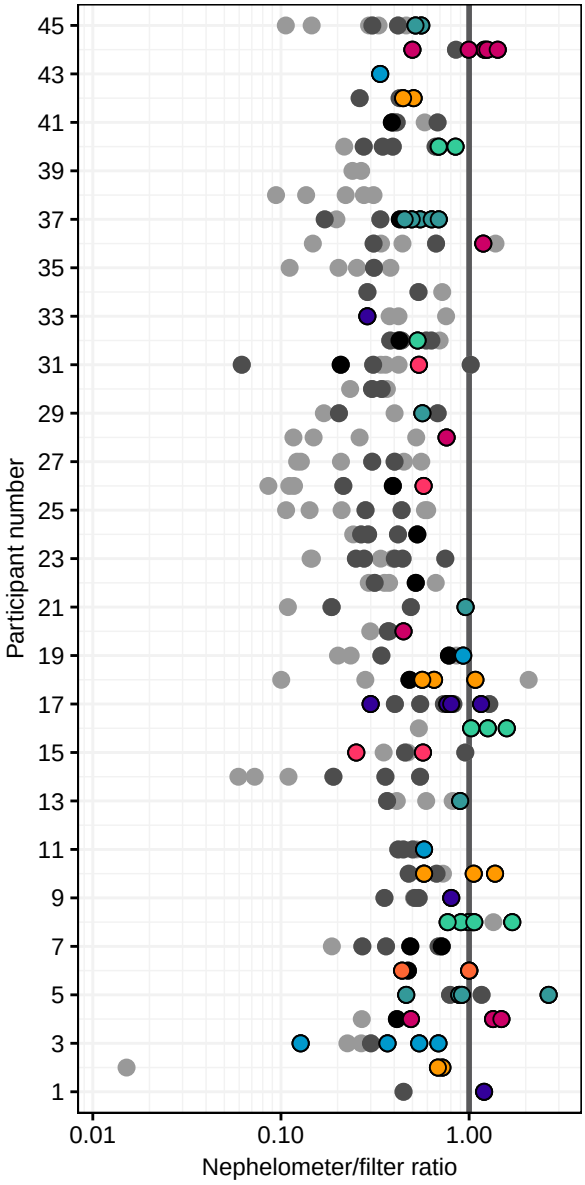

Figure S11. Nephelometer/filter ratio vs. participant number (n=44) for samples retained after the first filtering step (n = 249). Samples excluded by the second, third, and fourth filtering steps are shown in light gray, dark gray, and black, respectively. The solid vertical line represents a ratio of 1.

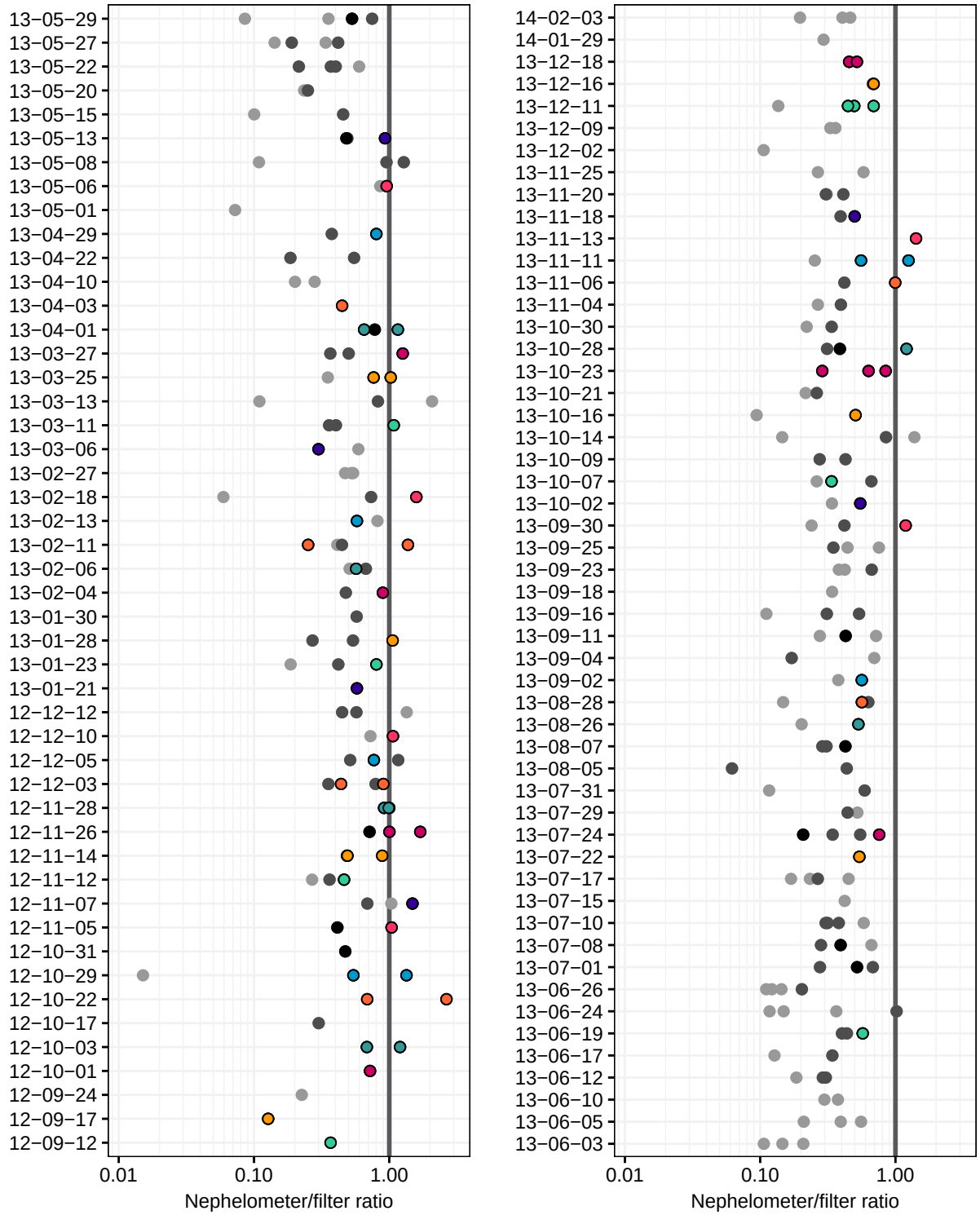

Figure S12. Nephelometer/filter ratio vs. sample date for samples retained after the first filtering step (n = 249). Samples excluded by the second, third, and fourth filtering steps are shown in light gray, dark gray, and black, respectively. The solid vertical line represents a ratio of 1.

S2.4 Linear mixed model development

Fifteen metrics were tested as predictors in linear mixed models (which included a single metric of interest as the fixed effect, a random participant-specific intercept, and the logarithm of the nephelometer/filter ratio as the response variable). In **Figure S13**, the fixed-effect coefficients have been transformed to represent the percent change in the nephelometer/filter ratio per one standard deviation change in the predictor. There were two predictors with fixed-effect coefficient 95% confidence intervals (CIs) that did not include zero for any of the five sets of training data: participant age and the fraction of exposure received in transit. There were two predictors with fixed-effect coefficient 95% CIs that did not include zero for 4/5 sets of training data: fraction of exposure received at work and fraction of exposure received in an eatery. Predictors with fixed-effect coefficient 95% CIs that did not include zero for one or two sets of training data were: mean ambient temperature, fraction of exposure received at home, fraction of time spent at home, fraction of time spent in another microenvironment, and fraction of $PM_{2.5}$ mass that was black carbon.

The correlation between the 15 metrics tested as predictors (Pearson's r) is illustrated in **Figure S14**. As expected, the fraction of time spent in a given microenvironment was often negatively correlated with the fraction of time spent in other microenvironments. The fraction of time spent in a given microenvironment was always positively correlated with the fraction of exposure received in that environment. This correlation was observed because the amount of exposure received in a given microenvironment depended on both the time spent in that microenvironment and the $PM_{2.5}$ concentration to which the participant was exposed in that microenvironment. Additionally, the fraction of exposure received at home was negatively correlated with the fraction of exposure received in the other four microenvironment categories. Finally, participant age was moderately positively correlated with the fraction of exposure received at work.

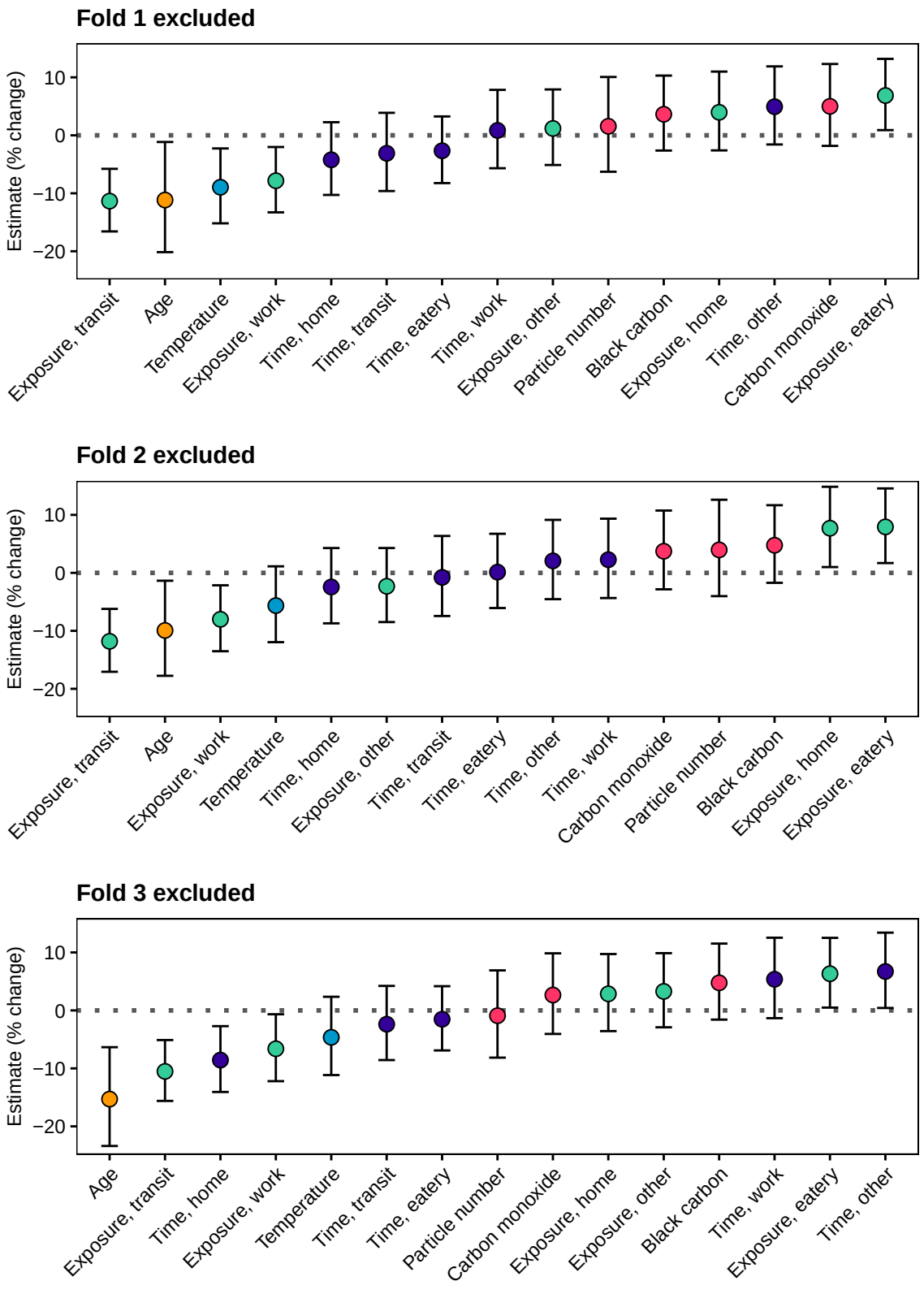

Figure continued on next page...

Figure continued from previous page

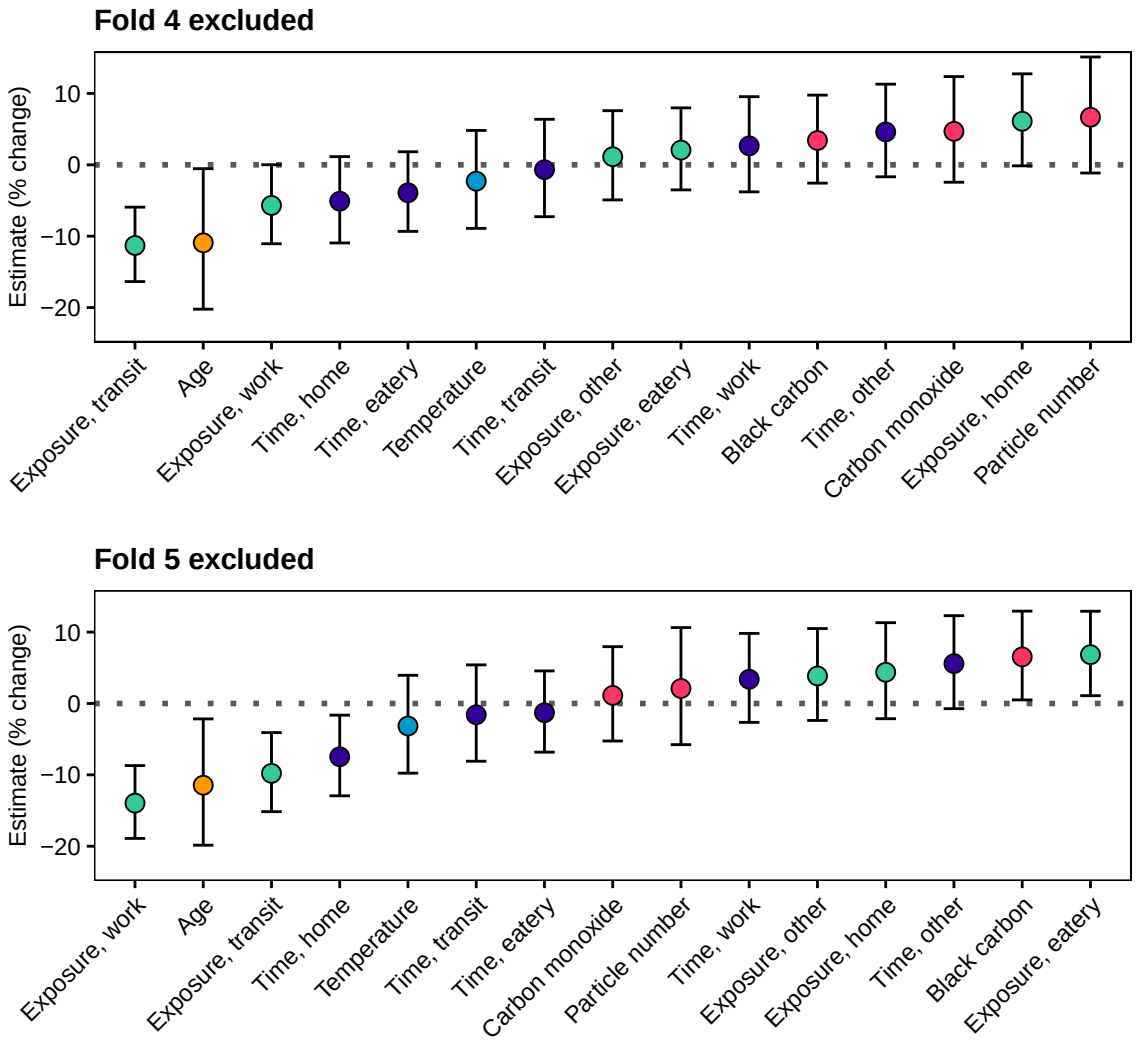

Figure S13. Estimates and 95% confidence intervals (CIs) for the fixed-effect coefficients from the 15 linear mixed models fit to the training data during each step in the K-fold cross-validation process ($K = 5$). All predictor variables were standardized to have a mean of zero and unit variance, and each estimate has been transformed to represent the percent change in the nephelometer/filter ratio per one standard deviation increase in the predictor. Only one fixed-effect term was included in each model. Colors represent different categories of predictors (gold = participant characteristics, teal = fraction of exposure received in a microenvironment category, purple = time spent in a microenvironment category, blue = weather-related, pink = exposure to other pollutants).

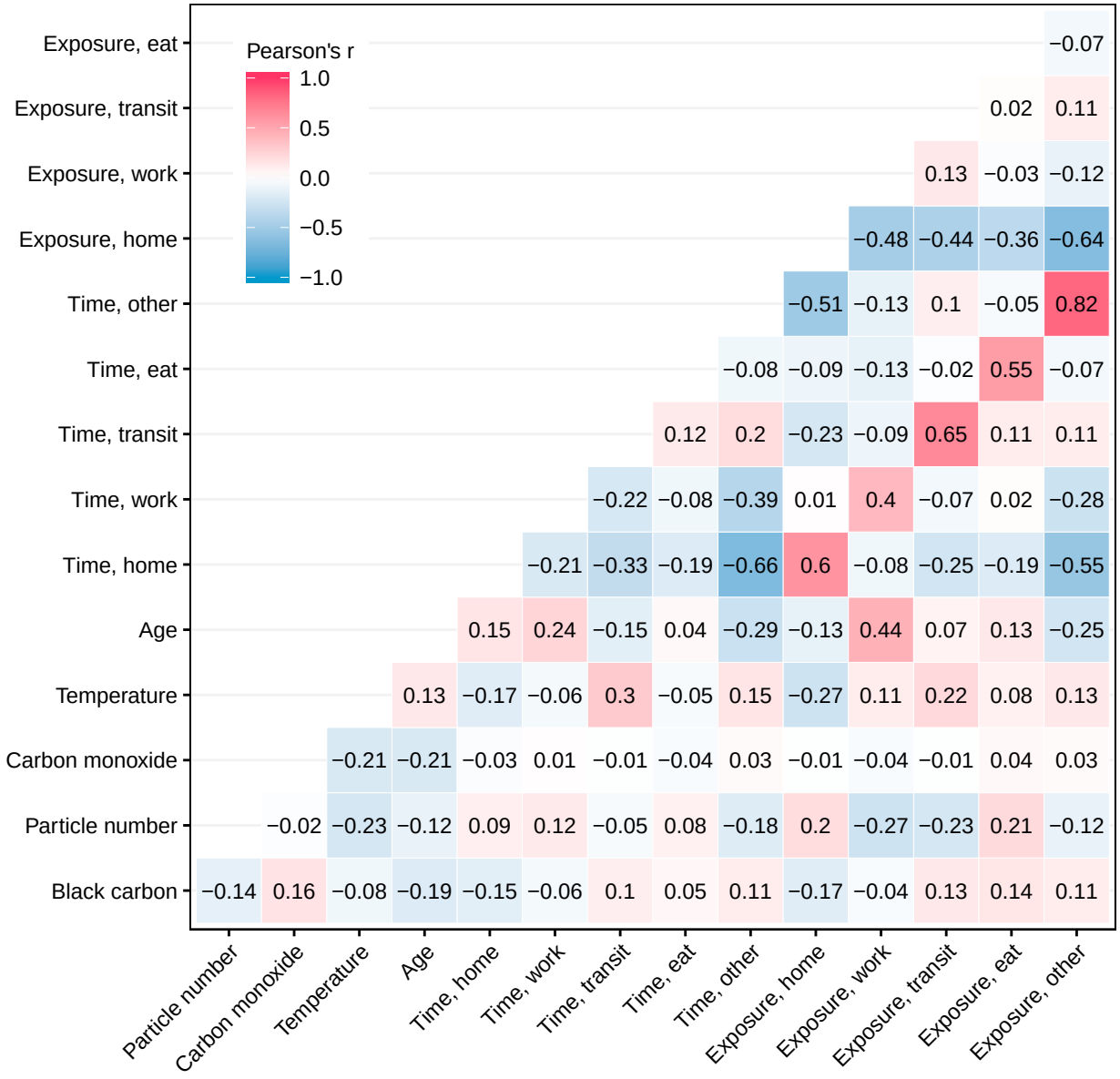

Figure S14. Correlation between the 15 continuous metrics tested as predictors in linear mixed models. The number in each box is the Pearson correlation coefficient.

References

CDPHE, 2013. Colorado Air Quality Data Report 2012. Colorado Department of Public Health and Environment, Air Pollution Control Division, Denver, CO.

CDPHE, 2014. Colorado Air Quality Data Report 2013. Colorado Department of Public Health and Environment, Air Pollution Control Division, Denver, CO.

McGraw, K. O., Wong, S. P., 1996. Forming inferences about some intraclass correlation coefficients. Psychological Methods 1(1), 30–46. DOI: 10.1037/1082-989X.1.1.30.

Pillarisetti, A., Carter, E., Rajkumar, S., Young, B. N., Benka-Coker, M. L., Peel, J. L., Johnson, M., Clark, M. L., 2019. Measuring personal exposure to fine particulate matter (PM2.5) among rural Honduran women: A field evaluation of the Ultrasonic Personal Aerosol Sampler (UPAS). Environment International 123 50–53. DOI: 10.1016/j.envint.2018.11.014.

Wendt, E., 2018. A low-cost monitor for simultaneous measurement of fine particulate matter and aerosol optical depth. Master's thesis. Colorado State University, Fort Collins, CO. URL: https://hdl.handle.net/10217/189408.