Supporting Information for

Combining Measurements from Mobile Monitoring and a Reference Site to Develop Models of Ambient Ultrafine Particle Number Concentration at Residences

Matthew C. Simon[a,b,*], Allison P. Patton[c], Elena N. Naumova[b,d], Jonathan I. Levy[a], Prashant Kumar[e], Doug Brugge[b,f,g], John L. Durant[b]

[a] Department of Environmental Health, Boston University School of Public Health, 715 Albany Street, Boston, MA 02118, USA

[b] Department of Civil and Environmental Engineering, Tufts University, 200 College Avenue, Medford, MA 02155, USA

[c] Health Effects Institute, 75 Federal Street, Suite 1400, Boston, MA 02110, USA

[d] Friedman School of Nutrition Science and Policy, Tufts University, 150 Harrison Avenue, Boston, MA 02111, USA

[e] Global Centre for Clean Air Research (GCARE), Department of Civil and Environmental Engineering, University of Surrey, Guildford GU2 7XH, United Kingdom

[f] Department of Public Health and Community Medicine, Tufts University, 136 Harrison Avenue, Boston, MA 02111, USA

[g] Jonathan M. Tisch College of Civil Life, Tufts University, 10 Upper Campus Road, Medford, MA 02155, USA

Contents of the Supporting Information (21 pages):

*Corresponding author: Dr. Matthew Simon; e-mail: simonmattc@gmail.com

1. Figures

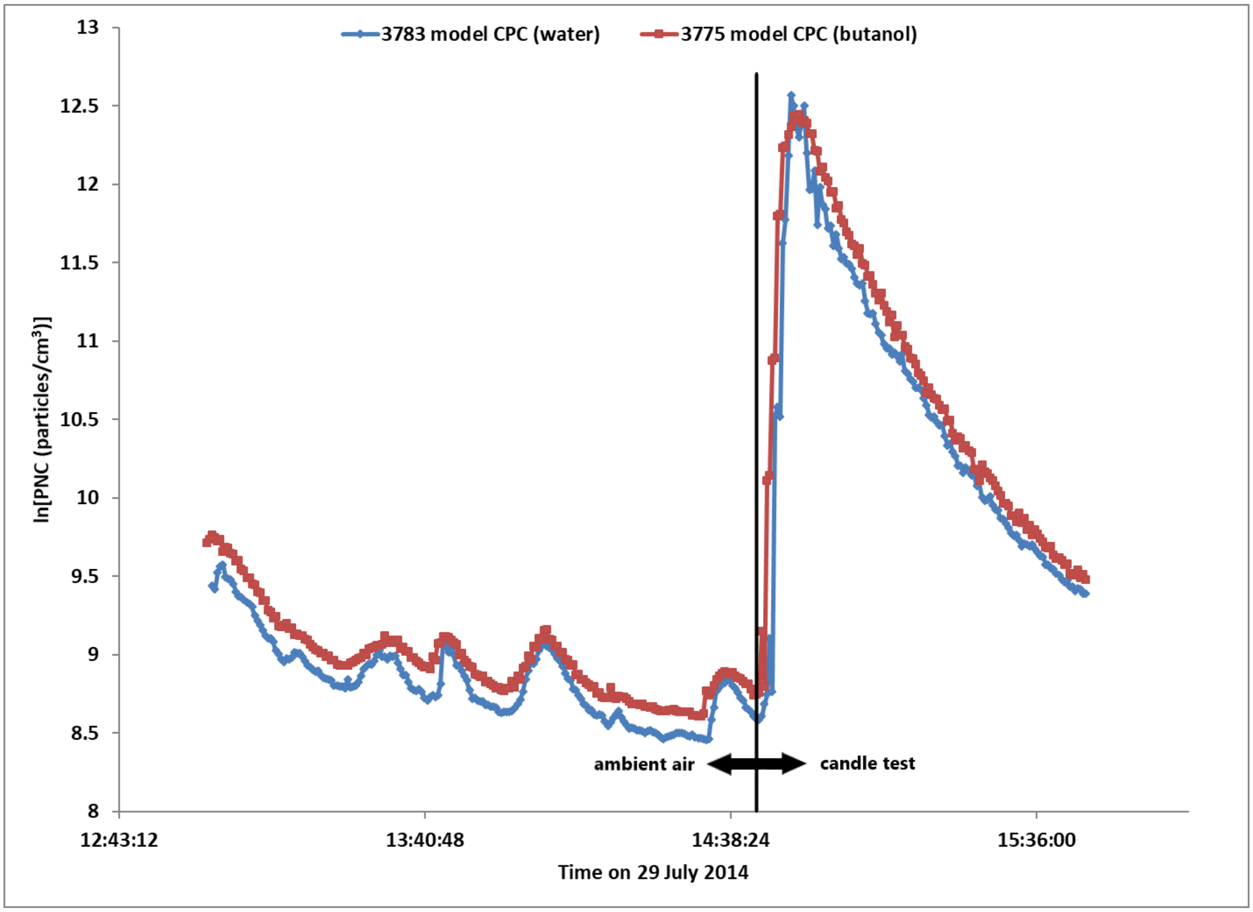


Figure S1: Side-by-side comparison of the water-based 3783 model CPC and butanol-based 3775 model CPC. Tests were completed in both ambient air and with a concentration spike using a candle.

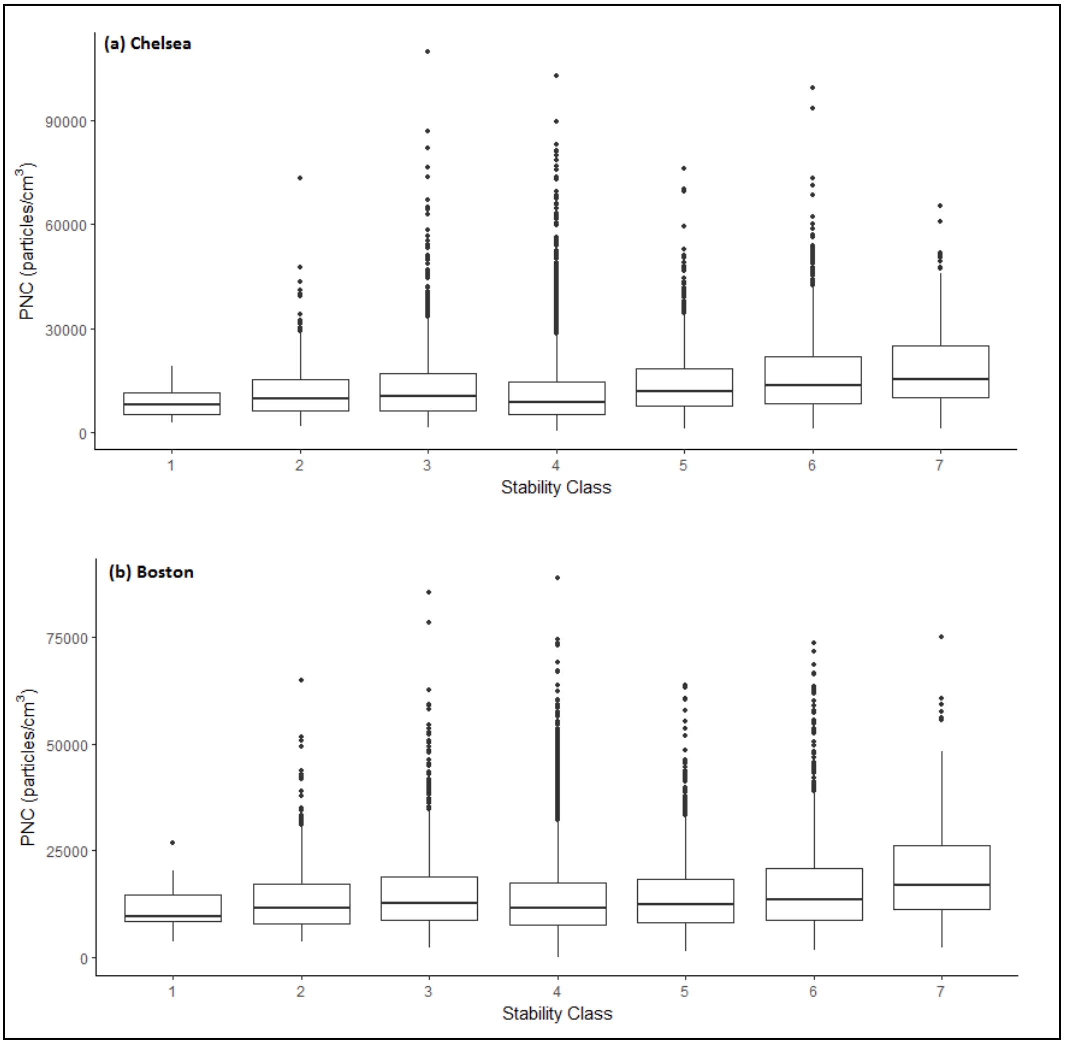


Figure S2: Central-site hourly fifth-percentile PNC by Pasquill stability class for (a) Chelsea and (b) Boston. Stability classes range from 1 (extremely unstable) to 7 (extremely stable).[1]

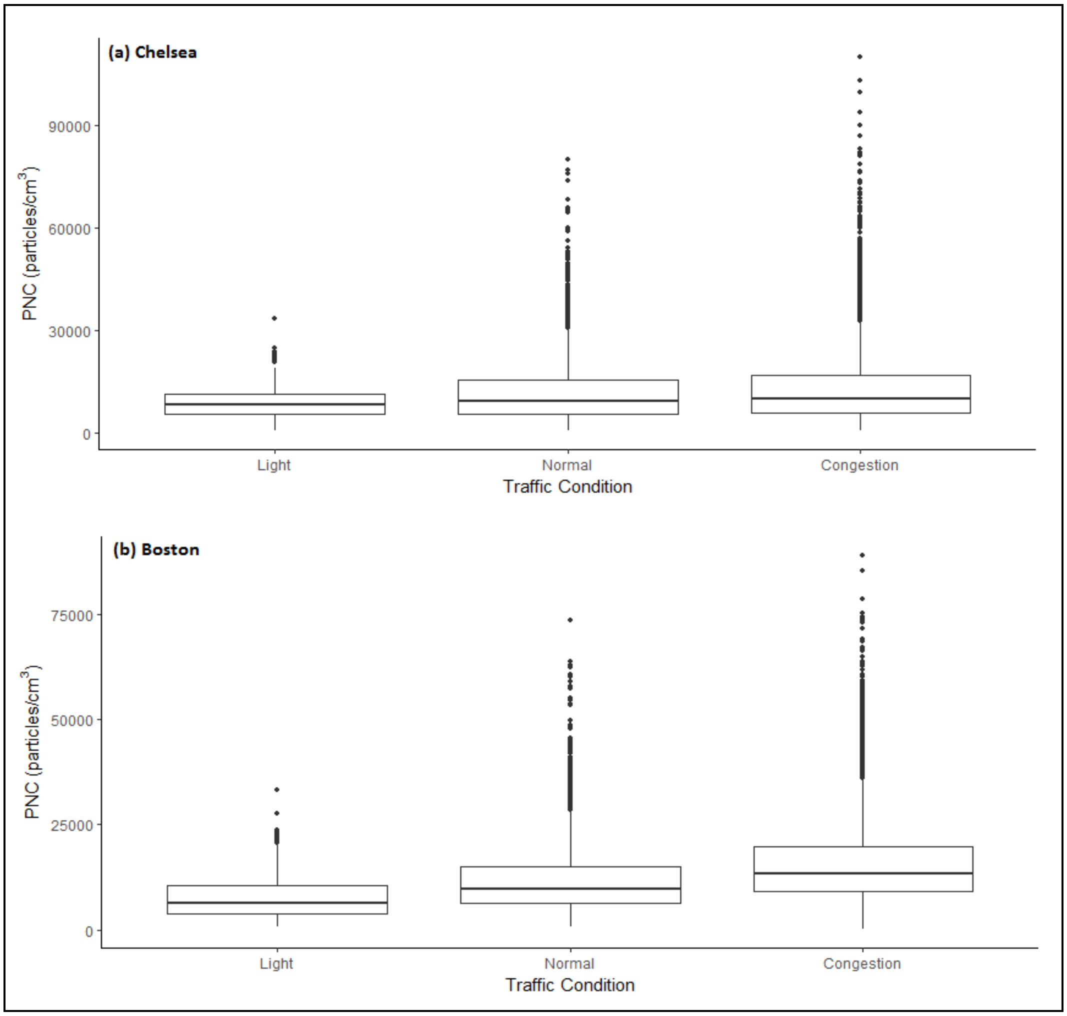
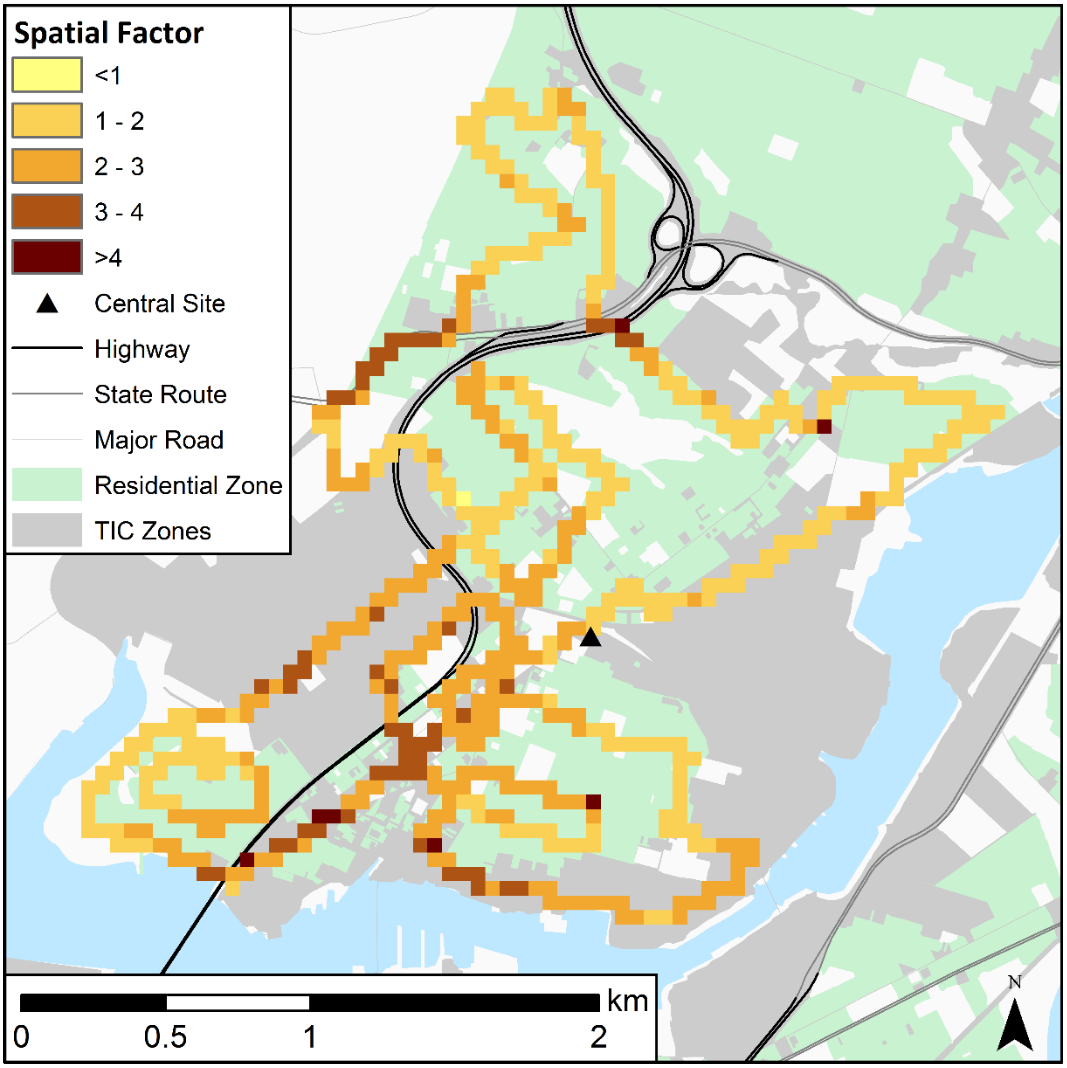


Figure S3: Central-site hourly fifth-percentile PNC by I-93 traffic condition for (a) Chelsea and (b) Boston. Hourly traffic was a log-transformed ratio of traffic volume (vehicles/h) to traffic speed (km/h). Light = ratio <-1, Normal = -1< ratio <1, Congestion = ratio >1.



Figures S4. Calculated *spatial factor* values from measured data in the Chelsea study area averaged across the entire mobile-monitoring campaign. Grid squares are 50 x 50 m. TIC zones are aggregated transportation, industrial, and commercial land-use zones.[2]

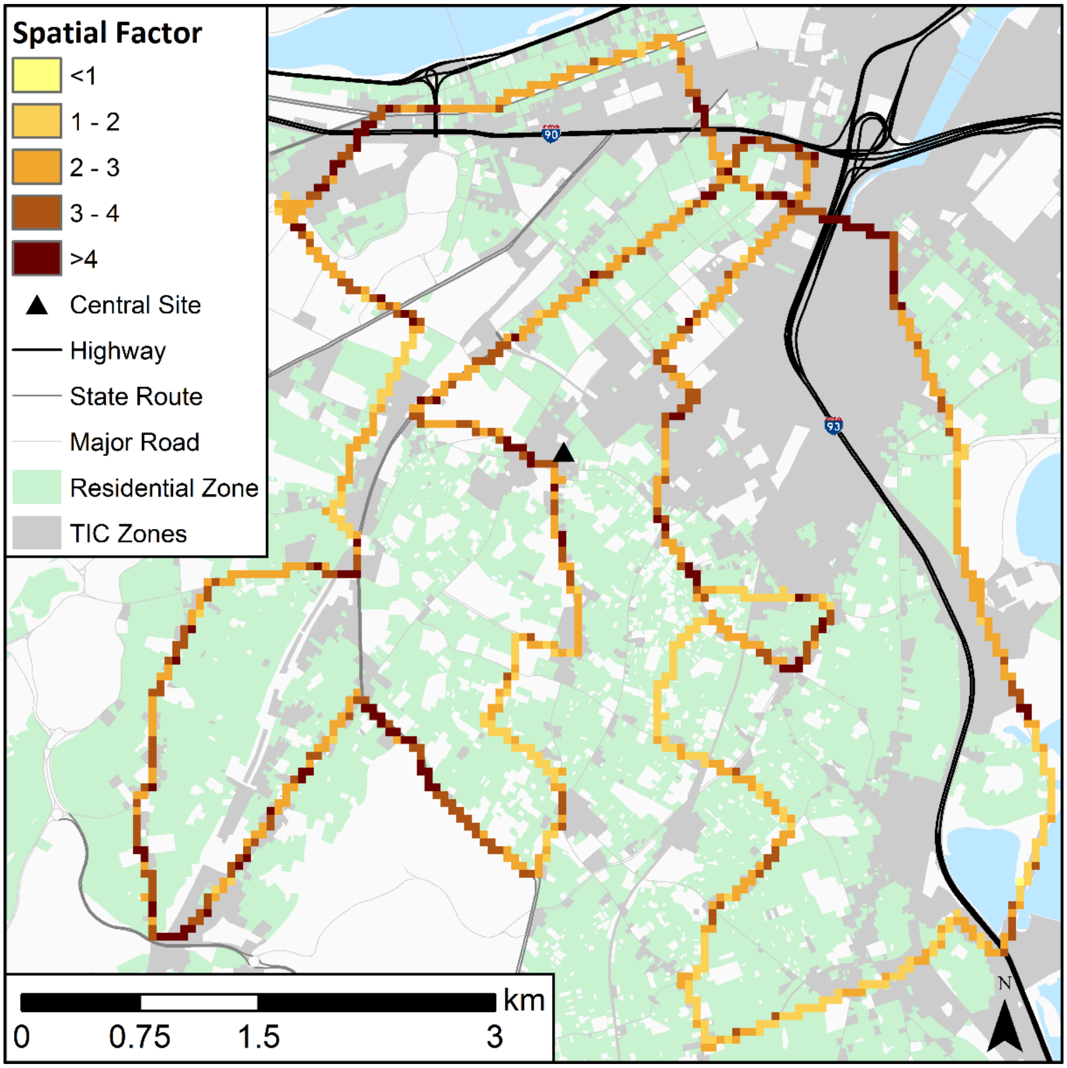


Figure S5. Calculated *spatial factor* values from measured data in the Boston study area averaged across the entire mobile-monitoring campaign. Grid squares are 50 x 50 m. TIC zones are aggregated transportation, industrial, and commercial land-use zones.[2]

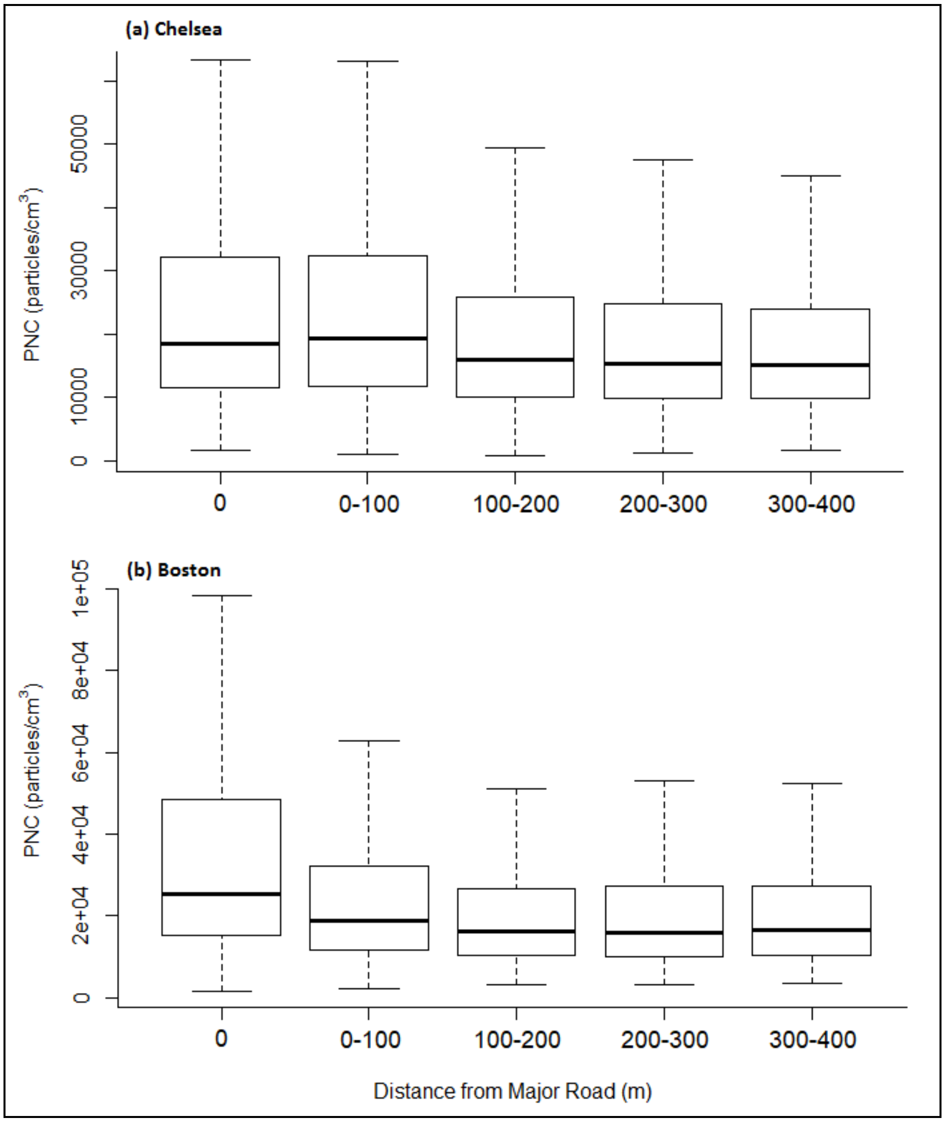


Figure S6: Particle number concentration versus distance from major roadways as measured by the mobile laboratory in (a) Chelsea and (b) Boston.

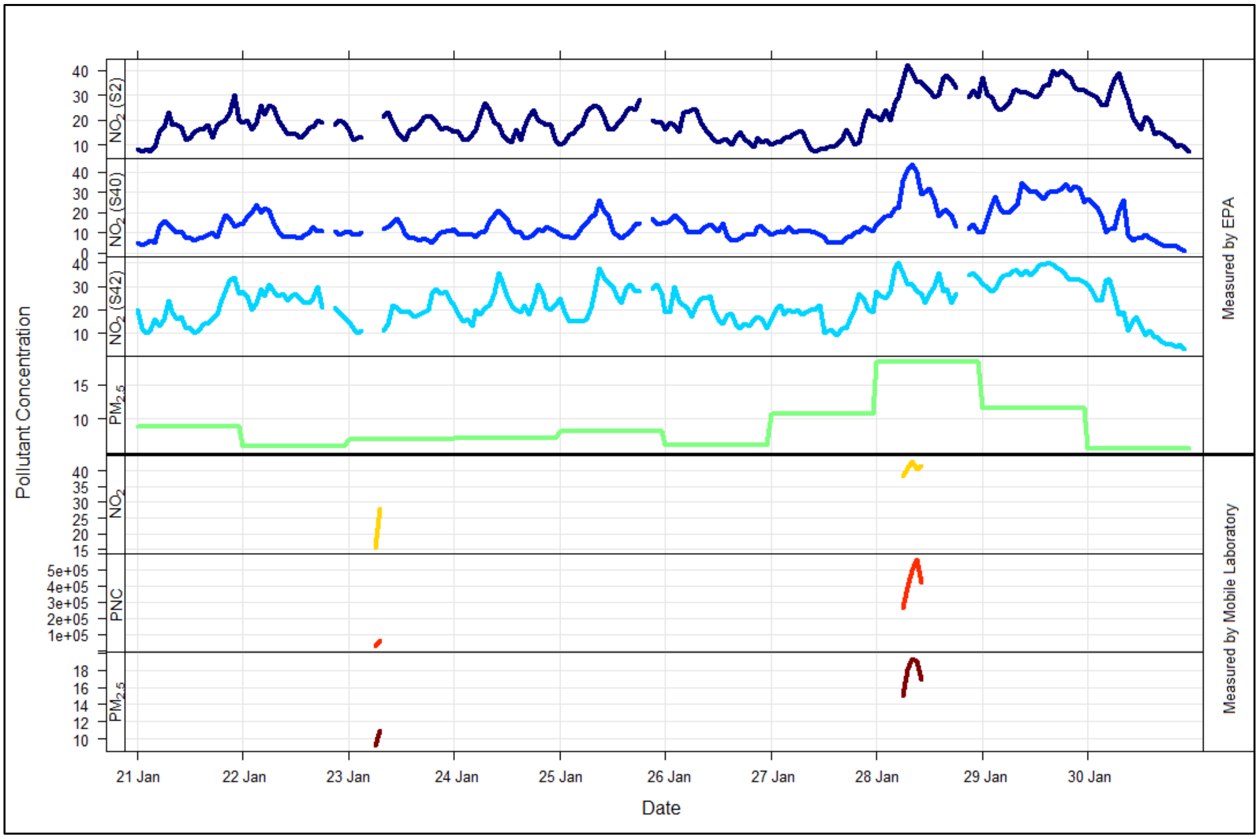


Figure S7. Hourly time series comparison of measured pollutants by the U.S. EPA in Boston to measured pollutants made by the mobile laboratory.[3] Units of NO_2 are in parts per billion, units of $PM_{2.5}$ are in micrograms per cubic meter, and units of PNC are in particles per cubic centimeter. $PM_{2.5}$ measurements by the EPA are 24-h measurements. Three different sites in Boston were used to compare hourly NO_2 concentrations: S2 = Kenmore Square, S40 = 531A East First Street, and S42 = Dudley Square.

Discussion on removing 28 January 2013 data:

28 January 2013 had unusually high PNC measured in the Boston study area with the mobile laboratory. To ensure measurements made by the mobile laboratory were accurate, we reviewed other particle and gas measurements made by the mobile laboratory that day – all pollutant concentration measurements were high. Additionally, we compared our data to measurements made by the EPA (Figure S7). To test the impact of this day on the models, we conducted a sensitivity analysis around the removal of data from 28 January 2013 in both mobile and hybrid models. As reported in the main text, the result of removing these data on Mobile-B was a reduced adjusted-R^2 of 0.31 (vs. 0.43 when this data was retained in the model). The β coefficients remained largely the same except for wind direction effects: when data from 28 January 2013 were removed, β coefficients for the south-southwest and west-southwest wind directions, the dominant wind directions that day, decreased two fold. We observed the boundary layer was <80 m at 05:30 increasing to only 470 m by 10:00 (mobile monitoring was from 05:30-10:00). Additionally, 28 January was the eighth consecutive day where Boston’s highest temperature of the day was below freezing. The sub-freezing temperatures combined with a very low boundary layer height likely favored the high measured particle concentrations. We also measured unusually high levels of NO_2 and $PM_{2.5}$ with the mobile laboratory on 28 January, which is consistent with measurements by the U.S. EPA (Figure S7).

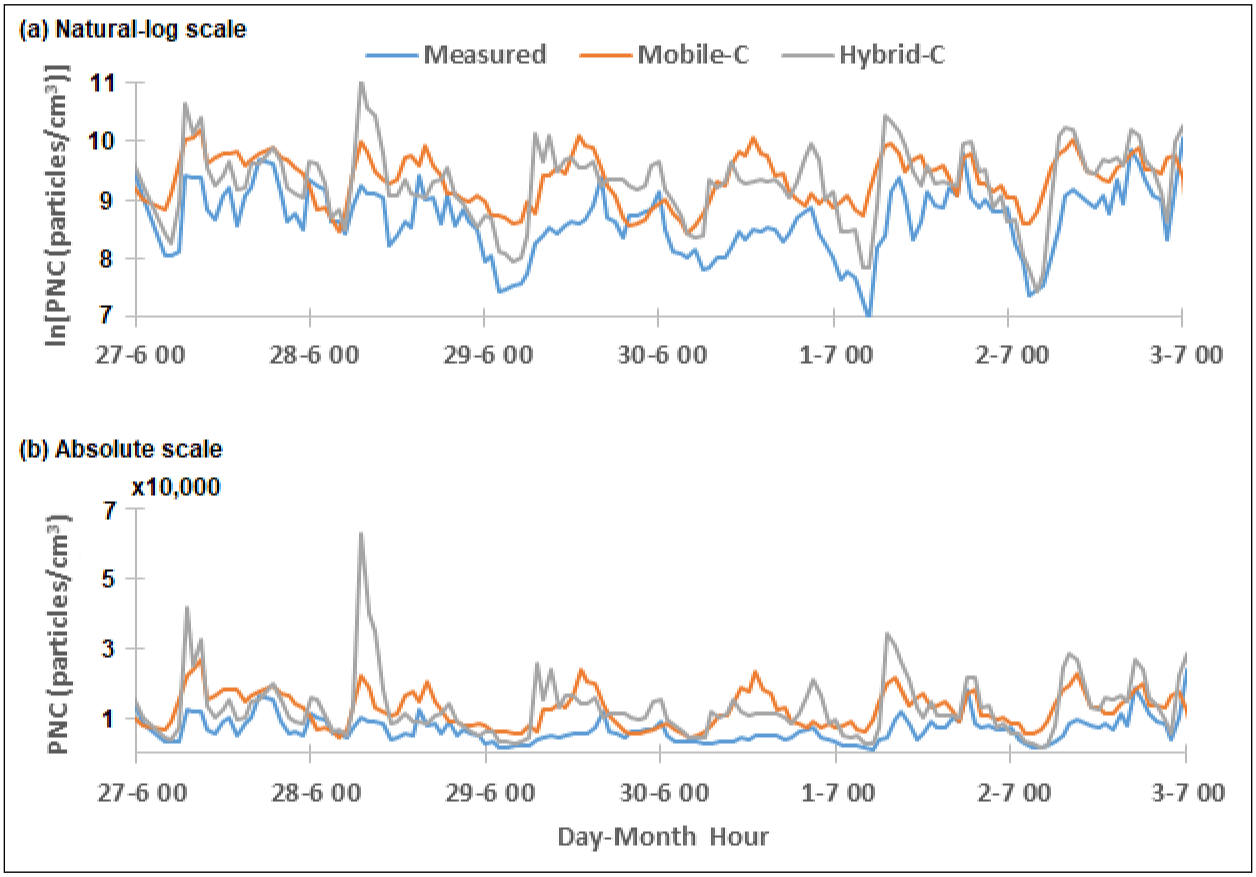


Figure S8. Comparisons of Mobile-B and Hybrid-B predictions versus measured (a) ln(PNC) and (b) PNC at Home 8 in Boston from 27 Jun-3 Jul 2013.

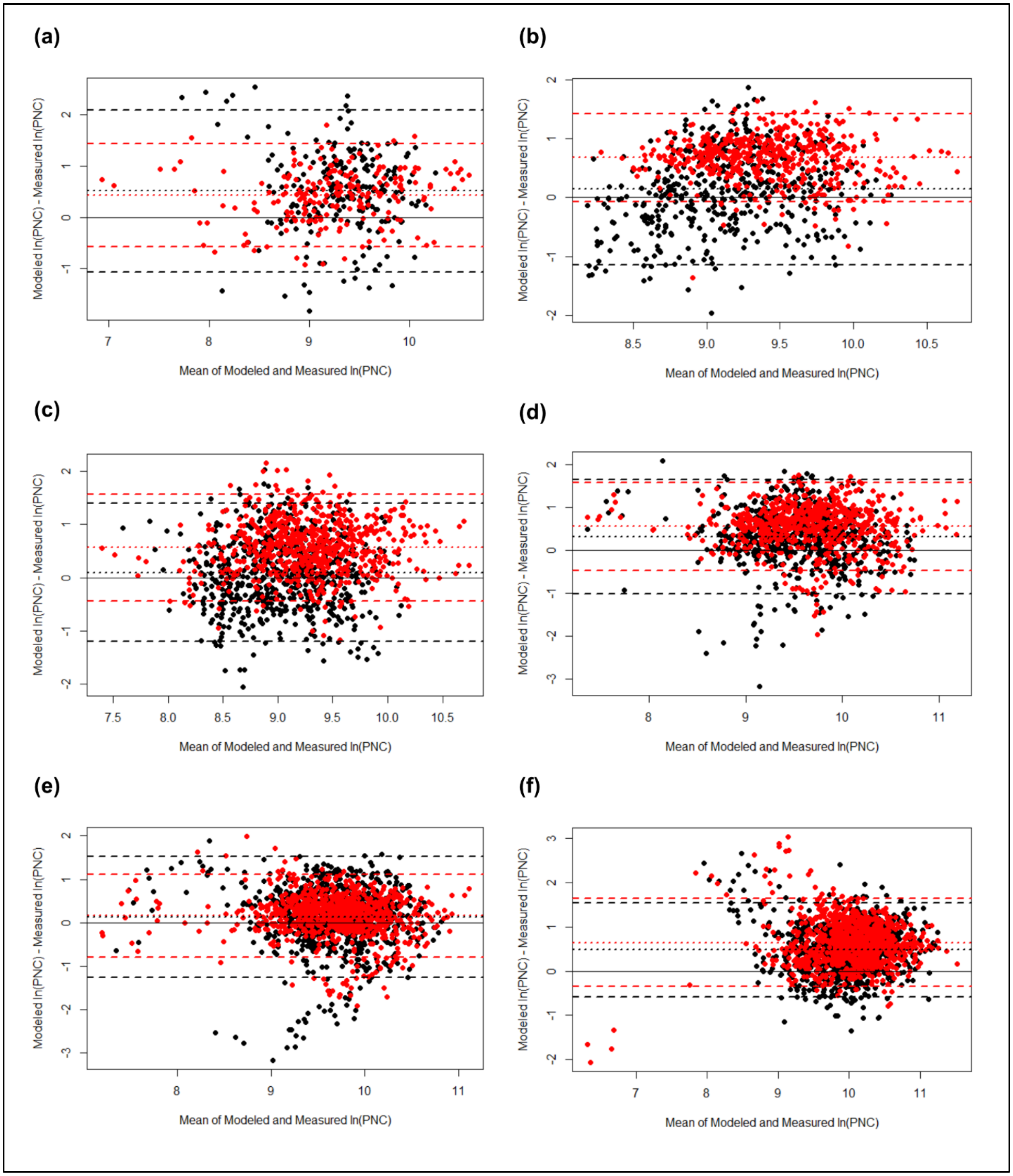


Figure S9: Bland-Altman plots comparing modeled to measured ln(PNC) at Homes 1-6 (a-f, respectively) in the Boston study area. Black dots = Mobile-B; red dots = Hybrid-B.

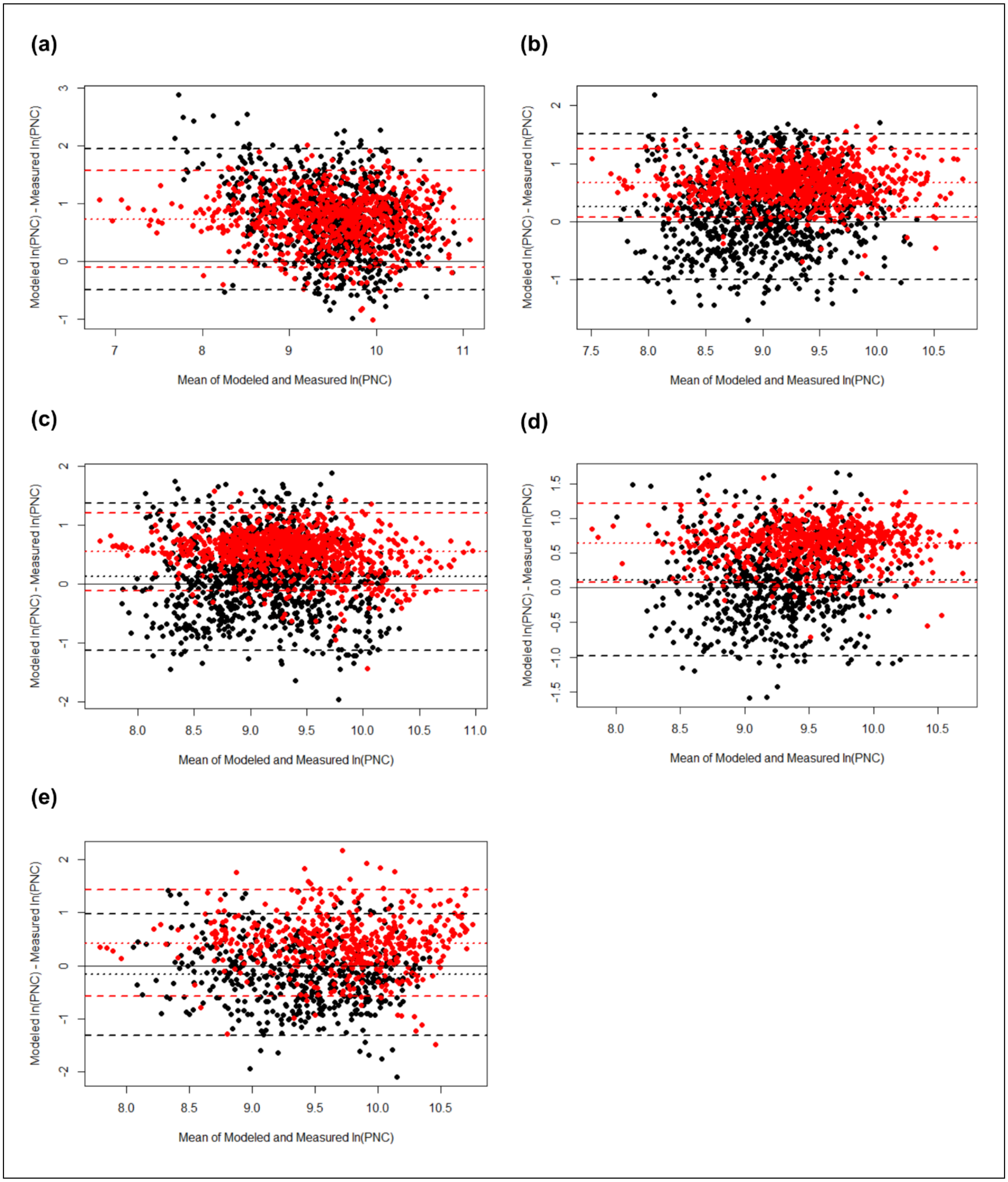


Figure S10: Bland-Altman plots comparing modeled to measured ln(PNC) at Homes 7-11 (a-e, respectively) in the Boston study area. Black dots = Mobile-B; red dots = Hybrid-B.

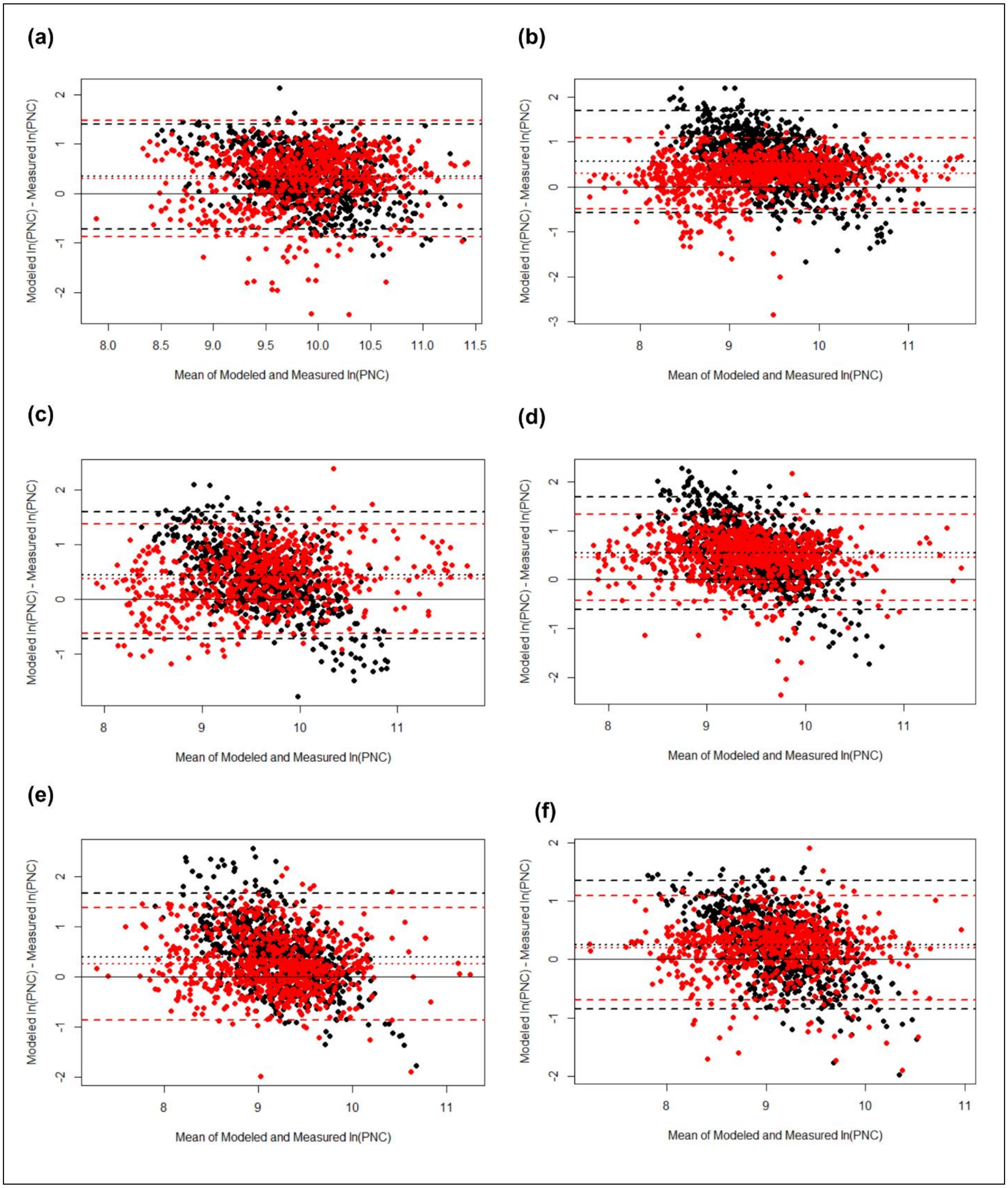


Figure S11: Bland-Altman plots comparing modeled to measured ln(PNC) at Homes 12-17 (a-f, respectively) in the Chelsea study area. Black dots = Mobile-C; red dots = Hybrid-C.

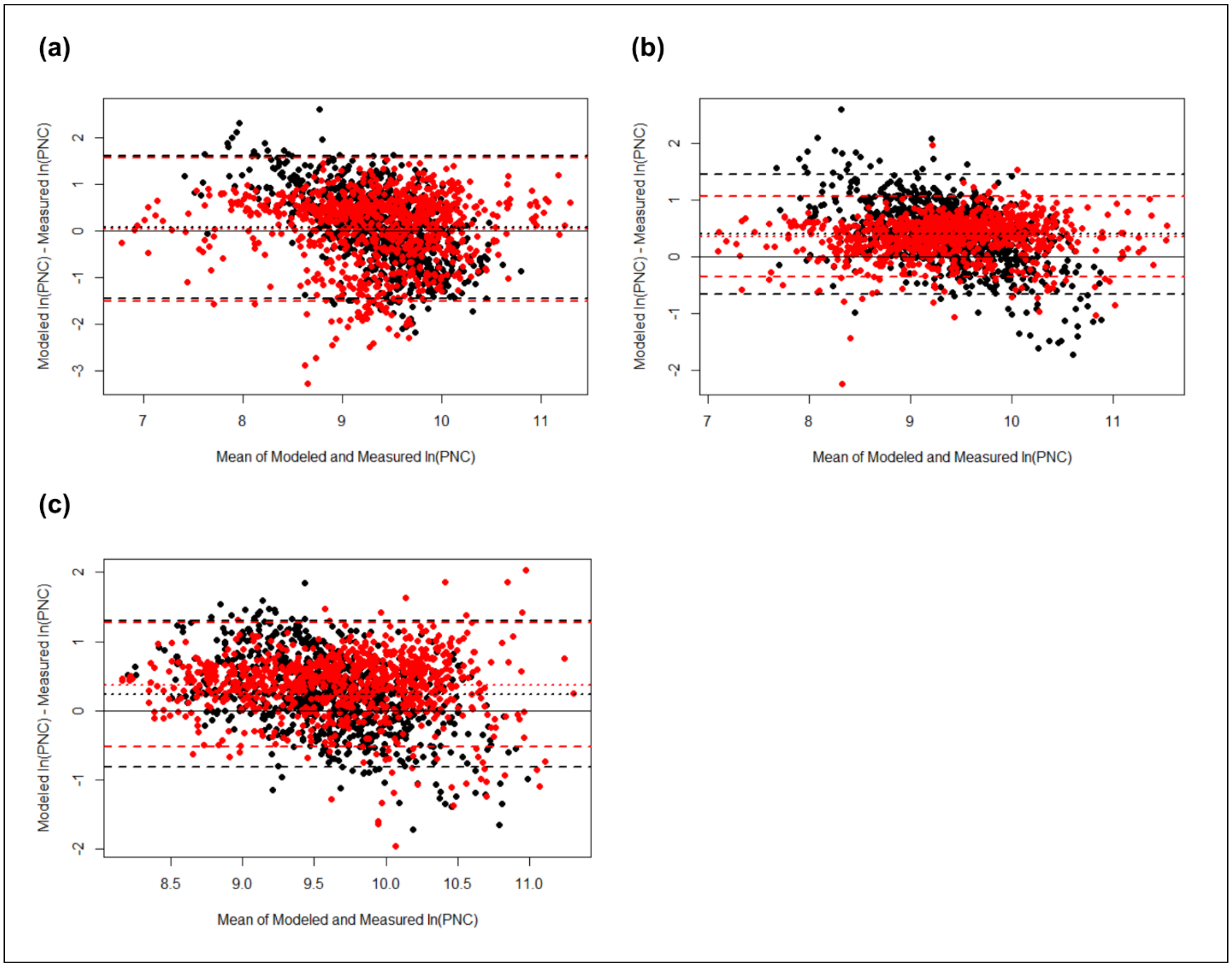


Figure S12: Bland-Altman plots comparing modeled to measured ln(PNC) at Homes 18-20 (a-c, respectively) in the Chelsea study area. Black dots = Mobile-C; red dots = Hybrid-C.

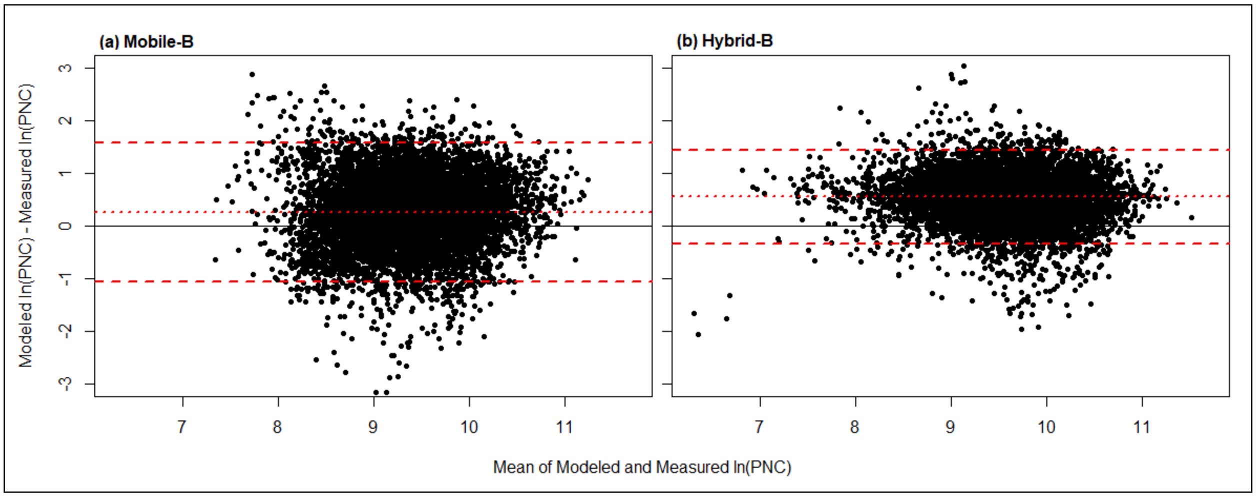


Figure S13. Bland-Altman plot comparing modeled to measured hourly ln(PNC) at all residential sites in the Boston study area using (a) Mobile-B and (b) Hybrid-B models. The y-axis represents the difference between hourly modeled and measured ln(PNC), and the x-axis represents the mean of hourly modeled and measured ln(PNC). Mean differences (dotted red lines) indicate systematic positive differences of modeled ln(PNC) relative to measured ln(PNC). The solid black (reference) line represents a mean difference of zero. Outer dashed red lines represent ± two standard deviations from the mean difference.

2. Tables

Table S1: List of potential independent variables used in model building.

Temporal Variables: temperature, relative humidity, atmospheric pressure, wind speed, wind direction (degrees), wind direction (octant), ENE winds (y/n), solar radiation, solar radiation (lagged one hour), precipitation amount, precipitation rate, boundary layer mixing height, Monin-Obkhov length, heat flux, hour of day, morning rush (06:00-09:00), evening rush (15:00-18:00), hour of day category (morning rush, evening rush, midday, overnight), visibility, cloud cover, cloud height, cloud cover/cloud height, weekday (y/n), weekend (y/n), no school (y/n), I-93 volume/speed (station 1), I-93 volume/speed (station 2), I-90 volume/speed, atmospheric stability, stability class (extremely stable, stable, neutral, unstable, extremely unstable), stability class (stable, neutral, extremely unstable), cloud height (category).
Spatial Variables: on bus route, ≤25 m bus stop, ≤50 m bus stop, ≤100 m bus stop, ≤150 m bus stop, ≤200 m bus stop, distance from commercial land use, distance from industrial land use, distance from residential land use, distance from transportation land use, distance from class 1-4 road, ≤100 m class 1-4 road, ≤200 m class 1-4 road, distance from interstates, distance from I-93, distance from I-90, distance from Tobin Bridge (US-1), distance from elevated section of US-1, distance from US-1 (sans elevated sections), distance from US-1 (grouped as single line source), ≤50 m major intersection, ≤100 m major intersection, ≤150 m major intersection, ≤200 m major intersection, ≤25 m bus-route intersection, ≤50 m bus-route intersection, ≤100 m bus-route intersection, ≤150 m bus-route intersection, ≤200 m bus-route intersection, vehicle distance traveled (by zone), vehicle hours traveled (by zone), on major road, distance from train line, ≤50 m train line, ≤100 m train line, ≤100 m train station, ≤200 m train station, distance from open space (≥0 m^2), distance from open space (≥5000 m^2), distance from open space (≥10,000 m^2), distance from open space (≥15,000 m^2), distance from open space (≥20,000 m^2), distance from open space (≥50,000 m^2),
Spatial-Temporal Variables (15): downwind of Logan Airport, downwind of downtown Boston, downwind Logan x wind speed, downwind Boston x wind speed, downwind of commercial zone, downwind of industrial zone, downwind of residential zone, downwind of transportation zone, downwind of I-93[b], downwind of I-90[b], downwind of US-1[c], downwind of Tobin Bridge (US-1)[c], downwind of elevated section of US-1[c], downwind nearest class 1-4 road, downwind nearest class 1-4 road x distance to class 1-4 road

[b] Boston-specific variables. [c] Chelsea-specific variables.

Discussion on solar radiation terms:

Both the concurrent hourly solar radiation and a one-hour lag of hourly solar radiation were tested in the models; the non-lagged term was used in Mobile-C since it accounted for more of the variability in PNC than the lagged term, while the opposite was true for Mobile-B. The β coefficient for the lagged solar radiation term in Mobile-B was 4-fold higher than the non-lagged solar radiation term in Mobile-C, possibly reflecting increased particle bloom activity in Boston compared to Chelsea. The coefficients did not change substantially in either model when the other solar radiation term was forced into the model. This is likely because the solar radiation was not substantially different from one hour to another.

Table S2. Summary of Boston and Chelsea *spatial factor* models incorporating temporal and spatio-temporal covariates into the model-building process ($Hybrid_{temporal}$). Models predict ln(*SF*), where *SF* is the unitless *spatial factor*. Variables are significant at p<0.05 unless specified otherwise.

Variable	**Chelsea (adj. R^2 = 0.21)**		**Boston (adj. R^2 = 0.17)**	
	Beta Coefficient	Standard Error	Beta Coefficient	Standard Error
(Intercept)	0.280	0.010	0.557	0.016
Wind Speed (per m/s)	n/a	n/a	-0.0684	0.0009
Temperature (per °C)	n/a	n/a	-0.0141	0.0002
Monin-Obukhov Length (per km)	-0.0155	0.0004	-0.073	0.002
Wind Direction (rel. to no wind)				
NNE (0°-45°)	0.535	0.010	0.479	0.016
ENE (45°-90°)	0.843	0.010	0.424	0.015
ESE (90°-135°)	0.544	0.010	0.426	0.016
SSE (135°-180°)	0.405	0.010	0.608	0.018
SSW (180°-225°)	0.393	0.010	0.701	0.016
WSW (225°-270°)	0.200	0.010	0.608	0.016
WNW (270°-315°)	0.324	0.009	-0.024ǂ	0.016
NNW (315°-360°)	0.405	0.009	0.383	0.015
Type of Hour/Day (rel. to weekdays 18:00-06:00)				
Weekday Morning Rush (06:00-09:00)	0.011	0.003	0.241	0.007
Weekday Midday (09:00-15:00)	-0.036	0.003	0.058	0.007
Weekday Evening Rush (15:00-18:00)	0.098	0.003	0.039	0.008
Saturday or Sunday (all hours)	-0.146	0.003	0.434	0.008
Distance from Open Spaces >5000 m^2 (per km)	0.508	0.008	n/a	n/a
Distance from Bus Routes (per km)	n/a	n/a	-0.918	0.015
Distance from Road Classes 1-4 (per km)	-0.676	0.006	-1.048	0.026
Near Interstate (<400 m; binary covariate)	n/a	n/a	0.223	0.004
Distance from Elevated US-1 (per km)	-0.178	0.001	n/a	n/a

ǂ Not significant at p<0.05.

Removing the 28 Jan 2013 mobile monitoring day from the data set resulted in a reduced adjusted-R^2 of 0.14. All variables remained significant at p<0.05 except for categorical WNW wind direction.

Table S3. Pearson correlation coefficients, Bland-Altman statistics, and RMSE of predicted ln(PNC) based on alternative hybrid-model formulations (i.e., allowing temporal covariates into the model ($Hybrid_{temporal}$)) and measured ln(PNC) at homes.

Boston						**Chelsea**					
Home	**Pearson's r**	**Bland-Altman Statistics** (Mean Difference from Zero)			**RMSE**	**Home**	**Pearson's r**	**Bland-Altman Statistics** (Mean Difference from Zero)			**RMSE**
		Difference	*95% CI*	*Std. Dev.*				*Difference*	*95% CI*	*Std. Dev.*	
1	0.61	0.55	(0.47-0.64)	1.22	0.83	12	0.65	0.31	(0.27-0.35)	1.07	0.63
2	0.57	0.50	(0.46-0.54)	0.96	0.70	13	0.85	0.32	(0.30-0.35)	0.79	0.52
3	0.49	0.63	(0.59-0.68)	1.14	0.86	14	0.71	0.43	(0.40-0.46)	1.03	0.68
4	0.55	0.56	(0.51-0.61)	1.29	0.86	15	0.71	0.47	(0.44-0.50)	0.89	0.66
5	0.61	0.26	(0.21-0.30)	1.20	0.66	16	0.58	0.26	(0.22-0.30)	1.13	0.63
6	0.60	0.72	(0.68-0.76)	1.24	0.96	17	0.72	0.26	(0.22-0.29)	0.90	0.53
7	0.79	0.68	(0.64-0.71)	0.97	0.84	18	0.61	0.04	(-0.01-0.08)	1.36	0.70
8	0.77	0.56	(0.54-0.58)	0.75	0.68	19	0.87	0.34	(0.31-0.36)	0.75	0.51
9	0.69	0.49	(0.46-0.52)	0.92	0.68	20	0.79	0.40	(0.37-0.43)	0.81	0.58
10	0.65	0.48	(0.45-0.52)	0.92	0.67						
11	0.65	0.38	(0.33-0.42)	1.07	0.66						
All Homes	**0.69**	**0.53**	**(0.52-0.55)**	**1.09**	**0.77**	**All Homes**	**0.74**	**0.32**	**(0.31-0.33)**	**1.01**	**0.61**

Table S4. Pearson correlation coefficients of mobile- and hybrid-model-predicted ln(PNC) and measured ln(PNC) at homes.

Home	Pearson Correlation - Boston		Home	Pearson Correlation - Chelsea	
	Mobile-B	Hybrid-B		Mobile-C	Hybrid-C
1	0.28	0.72	12	0.52	0.58
2	0.25	0.66	13	0.58	0.86
3	0.28	0.58	14	0.49	0.74
4	0.39	0.64	15	0.37	0.73
5	0.35	0.67	16	0.35	0.61
6	0.56	0.63	17	0.49	0.75
7	0.50	0.82	18	0.31	0.52
8	0.37	0.84	19	0.66	0.89
9	0.41	0.83	20	0.52	0.74
10	0.35	0.84			
11	0.55	0.64			
All Homes	**0.47**	**0.74**	**All Homes**	**0.51**	**0.73**

Table S5. Spearman rank correlation coefficients of mobile- and hybrid-model-predicted PNC and measured PNC at homes.

Home	Spearman Correlation - Boston		Home	Spearman Correlation - Chelsea	
	Mobile-B	Hybrid-B		Mobile-C	Hybrid-C
1	0.25	0.68	12	0.47	0.59
2	0.25	0.68	13	0.56	0.87
3	0.30	0.54	14	0.52	0.70
4	0.42	0.62	15	0.39	0.71
5	0.37	0.61	16	0.36	0.58
6	0.56	0.60	17	0.48	0.76
7	0.47	0.79	18	0.25	0.42
8	0.37	0.84	19	0.67	0.89
9	0.43	0.82	20	0.50	0.77
10	0.36	0.82			
11	0.55	0.62			
All Homes	**0.48**	**0.73**	**All Homes**	**0.50**	**0.73**

Table S6. Bland-Altman statistics on the mean difference between mobile- and hybrid-model-predicted ln(PNC) and measured ln(PNC) at homes.

Home		**Mobile Model**			**Hybrid Model**		
		Mean Difference from Zero		Standard Deviation	Mean Difference from Zero		Standard Deviation
		Difference	*95% CI*		*Difference*	*95% CI*	
Boston	1	0.33	(0.29-0.38)	1.45	0.43	(0.36-0.49)	1.00
	2	0.11	(0.05-0.17)	1.30	0.67	(0.64-0.71)	0.74
	3	0.09	(0.05-0.12)	1.23	0.57	(0.53-0.60)	0.99
	4	0.34	(0.30-0.39)	1.29	0.56	(0.52-0.60)	1.03
	5	0.16	(0.11-0.20)	1.34	0.18	(0.15-0.22)	0.95
	6	0.48	(0.44-0.51)	1.06	0.64	(0.61-0.67)	0.99
	7	0.67	(0.63-0.71)	1.26	0.73	(0.70-0.76)	0.83
	8	0.25	(0.21-0.29)	1.26	0.66	(0.64-0.68)	0.59
	9	0.12	(0.08-0.16)	1.25	0.55	(0.53-0.57)	0.66
	10	0.09	(0.05-0.13)	1.12	0.65	(0.63-0.67)	0.56
	11	-0.16	(-0.20- -0.12)	1.09	0.40	(0.36-0.44)	1.01
	All Homes	**0.24**	**(0.22-0.25)**	**1.31**	**0.56**	**(0.55-0.57)**	**0.90**
Chelsea	12	0.33	(0.30-0.37)	1.07	0.30	(0.25-0.34)	1.16
	13	0.56	(0.52-0.60)	1.13	0.29	(0.27-0.32)	0.79
	14	0.44	(0.40-0.48)	1.16	0.37	(0.34-0.40)	0.98
	15	0.53	(0.50-0.57)	1.15	0.46	(0.43-0.49)	0.88
	16	0.47	(0.42-0.51)	1.28	0.26	(0.22-0.30)	1.12
	17	0.24	(0.20-0.27)	1.14	0.20	(0.17-0.24)	0.89
	18	0.10	(0.05-0.15)	1.53	0.05	(0.00-0.10)	1.53
	19	0.37	(0.34-0.41)	1.09	0.36	(0.34-0.38)	0.71
	20	0.28	(0.25-0.32)	1.06	0.38	(0.35-0.41)	0.90
	All Homes	**0.37**	**(0.36-0.38)**	**1.22**	**0.30**	**(0.29-0.31)**	**1.04**

Table S7. RMSE between mobile- and hybrid-model-predicted ln(PNC) and measured ln(PNC) at homes.

Home	**RMSE – Boston**		**Home**	**RMSE – Chelsea**	
	Mobile-B	Hybrid-B		Mobile-C	Hybrid-C
1	0.80	0.66	12	0.64	0.66
2	0.67	0.77	13	0.81	0.50
3	0.63	0.76	14	0.74	0.62
4	0.74	0.77	15	0.79	0.64
5	0.70	0.52	16	0.80	0.63
6	0.72	0.82	17	0.63	0.50
7	0.93	0.85	18	0.79	0.78
8	0.69	0.73	19	0.67	0.51
9	0.65	0.64	20	0.61	0.59
10	0.58	0.71			
11	0.58	0.65			
All Homes	**0.71**	**0.72**	**All Homes**	**0.72**	**0.61**

3. References

(1) Turner, D. B. Relationships Between 24-Hour Mean Air Quality Measurements and Meteorological Factors in Nashville, Tennessee. *J. Air Pollut. Control Assoc.* **1961**, *11* (10), 483–489.

(2) MassGIS. MassGIS Data - Land Use http://www.mass.gov/anf/research-and-tech/it-serv-and-support/application-serv/office-of-geographic-information-massgis/datalayers/layerlist.html (accessed Mar 18, 2016).

(3) Carslaw, D. C.; Ropkins, K. Openair — An R Package for Air Quality Data Analysis. *Environ. Model. Softw.* **2012**, *27–28*, 52–61.