

Real-Time Prediction of Size-Resolved Ultrafine Particulate Matter on Freeways

Srijan Aggarwal,[†] Ricky Jain,[‡] and Julian D. Marshall^{*,†}

[†]Department of Civil Engineering, University of Minnesota, Minneapolis, Minnesota 55455, United States

[‡]Department of Mechanical Engineering, University of Minnesota, Minneapolis, Minnesota 55455, United States

S Supporting Information

ABSTRACT: Ultrafine particulate matter (UFP; diameter $<0.1 \mu\text{m}$) concentrations are relatively high on the freeway, and time spent on freeways can contribute a significant fraction of total daily UFP exposure. We model real-time size-resolved UFP concentrations in summer time on-freeway air. Particle concentrations (32 bins, 5.5 to 600 nm) were measured on Minnesota freeways during summer 2006 and 2007 (Johnson, J. P.; Kittelson, D. B.; Watts, W. F. *Environ. Sci. Technol.* **2009**, *43*, 5358–5364). Here, we develop and apply two-way stratified multilinear regressions, using an approach analogous to mobile-monitoring land-use regression but using real-time meteorological and traffic data. Our models offer the strongest predictions in the 10–100 nm size range (adj- R^2 : 0.79–0.89, average adj- R^2 : 0.85) and acceptable but weaker predictions in the 130–200 nm range (adj- R^2 : 0.41–0.62, average adj- R^2 : 0.52). The aggregate model for total particle counts performs well (adj- R^2 = 0.77). Bootstrap resampling ($n = 1000$) indicates that the proposed models are robust to minor perturbations in input data. The proposed models are based on readily available real-time information (traffic and meteorological parameters) and can thus be exploited to offer spatiotemporally resolved prediction of UFPs on freeways within similar geographic and meteorological environments. The approach developed here provides an important step toward modeling population exposure to UFP.



INTRODUCTION

The United States Environmental Protection Agency (USEPA) began regulating ambient PM_{10} (particulate matter smaller than $10 \mu\text{m}$) in 1988 and $\text{PM}_{2.5}$ (particulate matter smaller than $2.5 \mu\text{m}$) in 1997. While there are no US regulations for $\text{PM}_{0.1}$ (“ultrafine particles”, UFP; diameter less than $0.1 \mu\text{m}$), recent research raises the concern that these particles may be especially toxic.^{1,2} UFP can penetrate deeply into the lung and can cross the lung lining,^{3,4} which is ~ 0.1 – $20 \mu\text{m}$ thick.⁵ The European Union has proposed to regulate tailpipe number concentrations as part of Euro 5 and Euro 6 standards for light passenger and commercial vehicles. In typical ambient and on-roadway conditions, UFPs have high number concentrations but low mass concentrations, relative to other particles.⁶

Vehicles and other combustion sources are important contributors to urban UFP. UFP concentrations (particle number per volume of air) can be an order of magnitude higher on freeways than in background urban air.^{7–9} Variations in vehicle speed and density, type and age of vehicles, roadway topography, meteorology, and particle dynamics create spatially and temporally heterogeneous distributions of UFPs. Real-time estimation of UFP concentration on freeways is important for understanding UFP exposures and for identifying UFP hotspots.

Several studies have reported UFP concentrations on roadways. These studies evaluated aspects of UFPs such as

dispersion on freeways,^{10–12} correlation to vehicle types^{13–15} and to other pollutants,⁸ regional variation,^{12,16} and aerosol characteristics.¹⁷ Table 1 summarizes the main results of several extant studies; further details are available in an expanded version available online (Table S1, Supporting Information). None of the extant studies using size-resolved concentrations investigated how to predict on-roadway concentrations at times or locations other than when and where measurements occurred; such models are critical for estimating population exposure to UFP. Our investigation aims to fill this gap by providing real-time traffic- and weather-based prediction of UFPs on freeways.

Parameters such as traffic volume and speed are routinely measured on freeways in many urban areas; the data tend to be publicly available, often in near-real time, and thus can potentially be used for estimating concentrations of UFPs or other traffic-dominated pollutants at large spatial and temporal scales. The modeling approach developed and applied here is broadly similar to land-use regression^{18–20} but builds on prior research by employing real-time measurements of pollution, traffic, and meteorological parameters. Our approach could be

Received: September 19, 2011

Revised: December 11, 2011

Accepted: December 20, 2011

Published: December 20, 2011

Table 1. Summary^a of Studies of Real-Time on-Road Measurement of Ultrafine Particles (UFPs)

reference	sampling year	location	pollutants measured	sampling duration
Bukowiecki 2002a ¹⁷	1999–2000	Minneapolis, MN; Columbus, IN	UFP	N/A ^b
Bukowiecki 2002b ¹⁶	2001	Zürich, Switzerland	UFP	4 h
Canagratna 2004 ¹⁵	2000–2001	New York, NY	UFP, PM, NO, NO ₂ , CO, N ₂ O, CH ₄ , SO ₂ , and HCHO	~24 h
Weijers et al. 2004 ¹²	1999–2000	Amsterdam, Netherlands	UFP, PM	3 days
Kittelson et al. 2004 ¹¹	2000	Minneapolis, MN	UFP, CO, CO ₂ , NO _x	~20 h
Westerdahl et al. 2005 ⁸	2003	Los Angeles, CA	BC, NO, PM-PAH, UFP, NO ₂ , CO, CO ₂ , PM _{2.5}	12 h
Pirjola et al. 2006 ¹⁰	2003–04	Helsinki, Finland	CO, NO, UFP, NO _x , PM	384 h
Fruin et al. 2008 ¹⁴	2003	Los Angeles, CA	BC, NO, PM-PAH, UFP, NO ₂ , CO, CO ₂ , PM _{2.5}	15 h
Johnson et al. 2009 ²¹	2006–2007	Minneapolis, MN	UFP, CO, CO ₂ , NO _x	85 h
Int Panis et al. 2010 ¹³	2009	Belgium	UFP, PM _{2.5} , PM ₁₀	~30 h

^aA more detailed version of this table is in Supporting Information (Table S3). ^bInformation not available.

potentially useful for identifying high-risk times and locations, in prioritizing among options to mitigate air pollution exposure and in modeling exposures for epidemiology.

METHODS

Details of the equipment and data in this study have been discussed previously.²¹ Briefly, the University of Minnesota mobile emissions laboratory (MEL)^{11,17} was used to collect aerosol data while operating on a pre-established, interstate highway loop around Minneapolis, MN (Figure S1, Supporting Information). Sample air was drawn into the MEL through a probe located above the cab at a height of 3.5 m above the roadway at a flow rate of approximately 1000 L/min. Air was distributed to the instruments through a sampling manifold. Size-resolved real-time measurements were made using an Engine Exhaust Particle Sizer (EEPS) (Model 3090; TSI, Inc., Shoreview, MN), which counts 6–500 nm particles. The instrument has 32 channels (size bins). We consider here the 40 h of on-freeway data collected during June–July 2006 and June–July 2007 (total: 19 days of sampling). Because measurements occurred in summer, models generated here are most applicable for summer conditions (see below).

We obtained publicly available weather²² and on-freeway traffic²³ conditions matched to the time and location of the MEL. Time resolution of the data was 1 h averages for weather and 30 s averages for on-freeway traffic conditions. We employed weather conditions at Minneapolis Airport, which is within the travel loop. Traffic data provide real-time averaged speed (km h⁻¹) and vehicular volume (vehicles h⁻¹) for all lanes in both directions. By design, the measurement vehicle drove at a more constant speed than surrounding traffic (range [km h⁻¹]: 85–96 for MEL, 24–105 for surrounding traffic).

Overall, ~145 000 second-by-second data-points were collected for each of the 32 EEPS channels. We computed 1 min averages of the concentration data to match with the traffic data (measured every half minute) and to increase the signal-to-noise ratio. In generating 1 min averages, we removed 3.6% of the data because of missing or incomplete concentration values (requirement: more than 54 s of data per minute). We removed an additional 4.6% owing to missing traffic data, leaving the 2222, 1-min concentration estimates used here.

A two-way variable-width stratification scheme, based on traffic speed and volume, was applied to the minute-averaged data. Strata were selected so as to achieve an approximately even distribution of data among the strata. To ensure that our results are not contingent on the stratification scheme, we conducted several sensitivity analyses to systematically vary the strata; in total, twenty stratification schemes were evaluated.

For each alternative stratification scheme, we generated regression models as described below. (The resulting speed and volume coefficients in the models exhibited low coefficients of variability among stratification schemes, and model predictions were highly correlated with each other. Thus, we did not find evidence that model results are highly sensitive to stratification scheme.) The most efficient stratification scheme was selected on the basis of three criteria: model performance (measured by adj-R²), statistical significance of the model coefficients, and the percentage of the minute-averaged data set employed in the model.

The stratification scheme selected had 42 strata (six categories for traffic volumes and seven categories for traffic speeds) and used 96% of the valid data obtained (Table S2, Supporting Information). For each strata, we calculate mean traffic parameters (speed, volume); mean weather parameters (temperature, wind speed, relative humidity [RH]); and several percentiles (10th, 25th, 50th, 75th, and 90th percentiles), mean, and standard deviation of log-transformed particle counts. (The 10th and 90th percentiles, respectively, can be taken as representing exposure-relevant estimates of “on-freeway background” and “on-freeway in-plume” conditions.) Median of log-transformed particles counts were determined for each of the 32 EEPS channels. Example measurements are shown in Table S3, Supporting Information, for total PNC.

Linear regressions were performed in Matlab (ver. R2009b, Mathworks, Inc., Natick, MA) to predict EEPS-derived concentration measurements (dependent variable) based on traffic and weather parameters (independent variables). Regressions were performed with and without the weather parameters to investigate the effect of weather on the model. Finally, stepwise regressions were performed to optimally select predictors from the traffic and weather parameters. Stepwise regressions allow the insertion and removal of parameters to improve the fit of the model, keeping only those parameters that contribute to a significant increase in model performance.

EEPS provides size-resolved measurements (5.5–600 nm) in 32 bins, equi-spaced on a log scale. For size resolved measurements overall, 96 regressions were carried out, to estimate central tendencies (median of log-transformed PNC) for each of the 32 bins, employing three schemes: (1) including traffic and weather parameters, (2) including only traffic parameters, and (3) selecting parameters by stepwise regression. Results for the first and third schemes were similar; thus, results from only the first (“with weather parameters”) and second (“excluding weather parameters”) are presented here.

We employed bootstrap resampling ($n = 1000$) to examine uncertainty and the models' statistical robustness. Bootstrap involves generating hypothetical data sets for input to the regression, based on "random sampling with replacement" from the original data set; this approach helps quantify the uncertainty of each estimator. As mentioned above, 20 sensitivity analyses were conducted to explore the impact of stratification scheme. Additional analyses were performed to test (1) the effect of wind direction, by stratifying by relative location (MEL is on the upwind/downwind side of the highway), and (2) changes over time, by stratifying by measurement year (2006, 2007).

RESULTS AND DISCUSSION

Modeling was performed on log-transformed metrics because particle number concentrations (PNCs) are log-normal (Figure 1). Weather data are summarized in Figure 2 and in Supporting

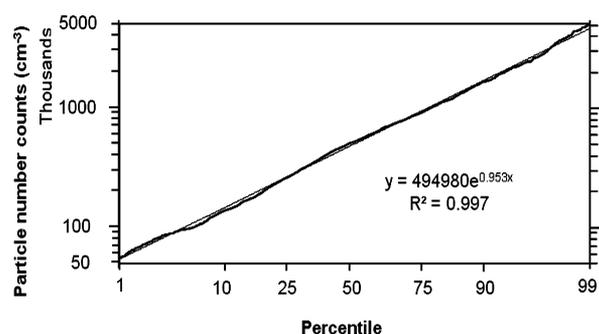


Figure 1. Cumulative distribution of total particle number counts (PNCs) for the 40 h of measurement. The plot is nearly a straight line, indicating that the data are nearly log-normal (geometric mean: $495\,000\text{ cm}^{-3}$, geometric standard deviation: 2.60).

Information (Table S4). Linear regressions were performed for the mean, standard deviation, and several percentiles (10th, 25th, 50th, 75th, 90th) of log PNC (Figure 3). Performance trends among these seven models are similar with- versus without-weather parameters included in the model. Specifically, models with weather parameters perform slightly better than models without weather parameters (less than 10% mean difference in performance) for the median, 75th percentile, 90th percentile and standard deviation. Model performance is similar between with- and without-weather models ($\text{adj-}R^2$: 0.51–0.82) for 10th percentile, 25th percentile and the mean. Correlations for observed and predicted PNC using the developed percentile models are presented in Supporting Information (Figure S2). R^2 values for the correlations range from 0.65 to 0.83.

Coefficients for speed and volume, for size-resolved models, are in Figure 4. The difference in performance between the "including weather" and the "excluding weather" models is greatest for the size range 200–400 nm. Particles in this size range ("accumulation mode") are influenced by urban and regional background concentrations, which are often correlated with weather parameters; thus, it seems reasonable that including weather improves model performance for this size range. For example, a recent study by Riddle et al.²⁴ provides evidence that ultrafine particles near a freeway can be accounted for primarily by on-road diesel fuel contribution, while particles with size >180 nm may reflect contributions from background secondary organic aerosols. Overall, model

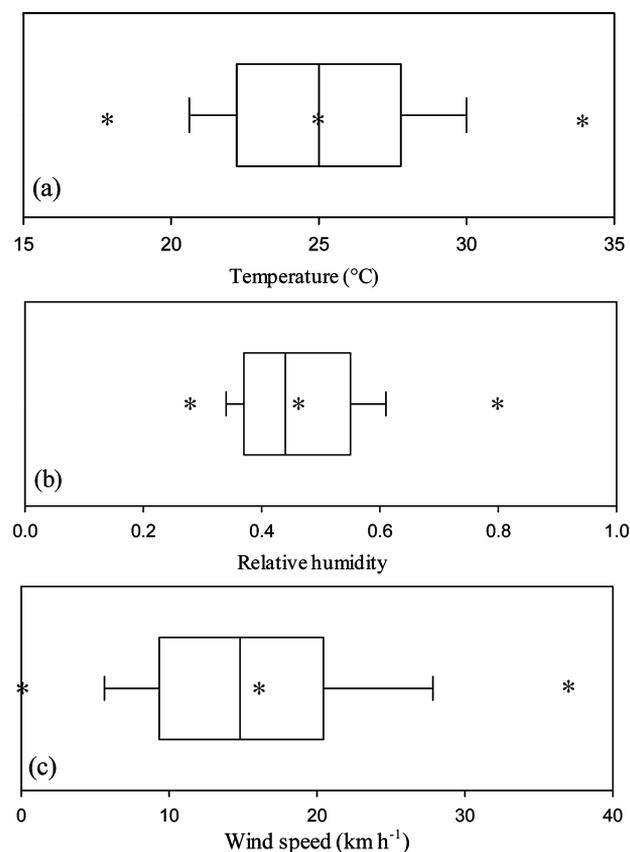


Figure 2. Summary of input data sets for (a) temperature, (b) relative humidity, and (c) wind speed. Shown are 10th, 25th, 50th, 75th, and 90th percentiles (box plot) and minimum, mean, and maximum values (asterisks), reflecting conditions during on-highway measurements.

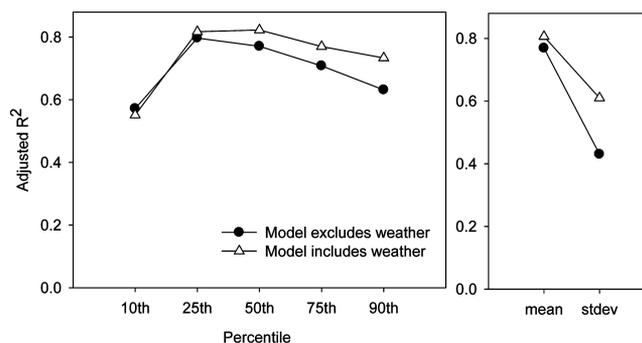


Figure 3. Model performance ($\text{adj-}R^2$) for prediction of percentiles, mean, and standard deviations of log transformed particle number concentration measurements.

performance is best for 10–100 nm particles ($\text{adj-}R^2$: 0.79–0.89, average $\text{adj-}R^2$: 0.85).

Additional analysis (not shown) revealed that model performance ($\text{adj-}R^2$) was generally better for high-concentration channels than for low-concentration channels, a finding that is consistent with Figures 4 and 5. Among the size bins, the speed coefficient was negative and the volume coefficient was generally positive (Figure 4), indicating that higher traffic speeds yield lower PNC while higher traffic volumes generally yield higher PNC. The latter finding is intuitive: increasing vehicle volumes generally increases total emissions. The former finding is likely a result of the increased vehicle spacing (and therefore dilution) and increased turbulent mixing found at

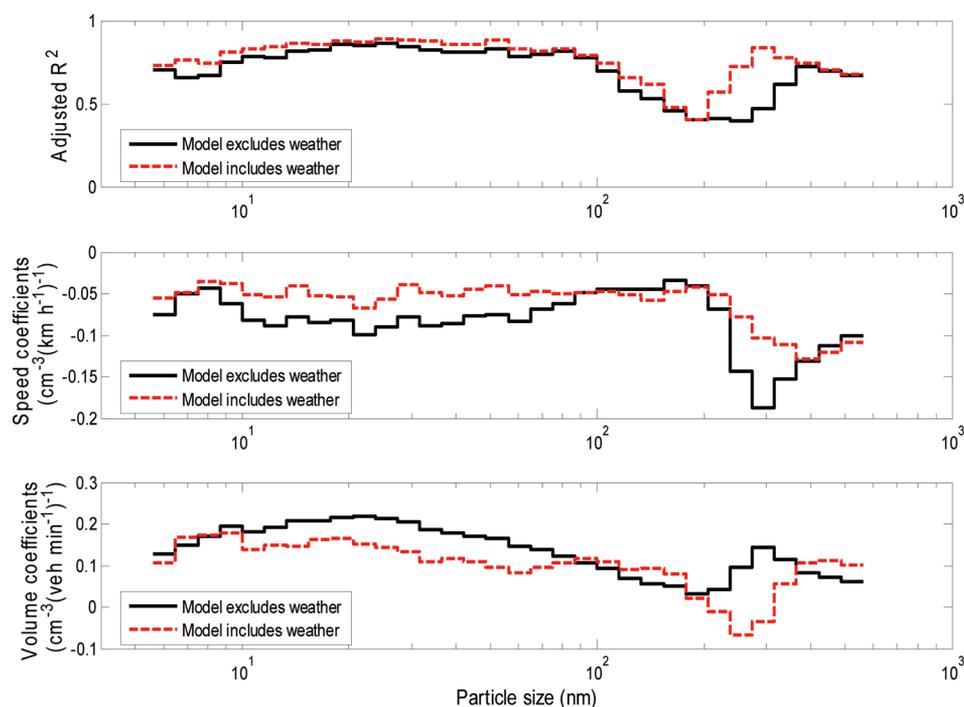


Figure 4. Modeling parameters for size-resolved prediction of median particle count measurements: (a) adj- R^2 , (b) speed coefficients, and (c) traffic volume coefficients for the 32 models.

higher speeds. (At low speeds [$\sim 50 \text{ km h}^{-1}$], the negative correlation between speed and PNC no longer holds; see Figure S3, Supporting Information, and the first two rows of Table S3, Supporting Information.)

A prior study by Kittelson et al.¹¹ reported a positive correlation between speed and PNC, which is in contrast with the negative correlation reported here. Several factors may explain the different findings: temperature and season (warmer weather/summer [current study] versus colder weather [November 2000 for the prior study by Kittelson et al.¹¹]); sampling height above roadway (3.5 m [current study] versus 0.7 m [prior study]); and speed/volume data (speed and volume for surrounding traffic [current] versus test-vehicle speed, no volume data [prior]). Particle dynamics (e.g., rates of physical and chemical processes) may vary with temperature; engine performance, and the composition of fuels and vehicles, may differ between summer and winter; and, changing the sampling height could impact the relative importance of cars versus diesel trucks. As mentioned above, the relationship between speed and PNC may differ at slower speeds (see Figure S3, Supporting Information).

To examine interactions between traffic speed and volume, observed and predicted median values of PNC for four traffic conditions are presented in Figure 5 (upper). As expected, number concentrations are generally higher for smaller particles than for larger particles. For all particle sizes, “denser, slower” traffic exhibits the highest particle counts, while “lighter, faster” traffic exhibits the lowest particle counts. The other two cases (“denser, faster” and “lighter, slower”) exhibit differing behavior: for ultrafine particles, the two cases are similar to each other; for larger particles, concentrations are lower for “denser, faster” than for “lighter, slower”. For ultrafine particle concentrations, traffic speed and volume appear to be comparable in importance; for other particles (greater than 100 nm), traffic speed appears to be more important than

traffic density. A box-plot comparison of observed and model-predicted PNC distribution is presented in Figure 5 (lower). For the four traffic conditions shown, modeled UFP concentrations provide reasonable but not perfect estimates of within- and between-group variability. The percentiles in the boxplots represent model predictions from the models we created for the P10/P25/P50/P75/P90/mean. For example, consider one specific condition (e.g., denser/slower). We calculated the mean speed (S^*) and volume (V^*) among data in the “denser/slower” strata, as representative of that strata (Table S3, Supporting Information). Then, the mean S^* and V^* were used as inputs to the P10/P25/P50/P75/P90/mean models; the results are the model-derived prediction box plots in Figure 5 (lower). Figure S2, Supporting Information, provides additional model-measurement comparisons as scatter-plots; Figure S1, Supporting Information, provides maps of measured and modeled PNC; Figure S4, Supporting Information, provides maps of traffic speeds and volume along the test route.

Results in Table 2 include either or both of two meteorology variables (temperature and RH) in all models except 10th percentile. RH is generally more important as a predictor than temperature or wind speed, but results vary by particle size (see Figures 6 and S1 and Table S5, Supporting Information). In Figure 6, several coefficients border on statistical significance (just over or just under). Of the seven models in Table 2, traffic volume is a predictor in all of the models while traffic speed, RH, and temperature are a predictor in five of the seven models.

Tables S6 and S7, Supporting Information, provide results of the stratified models (stratified by year and by MEL location relative to wind [e.g., MEL location is upwind of nearby traffic, MEL location is downwind of nearby traffic]). Model performance (R^2) is generally better for models presented here than for the “MEL location relative to wind” stratified

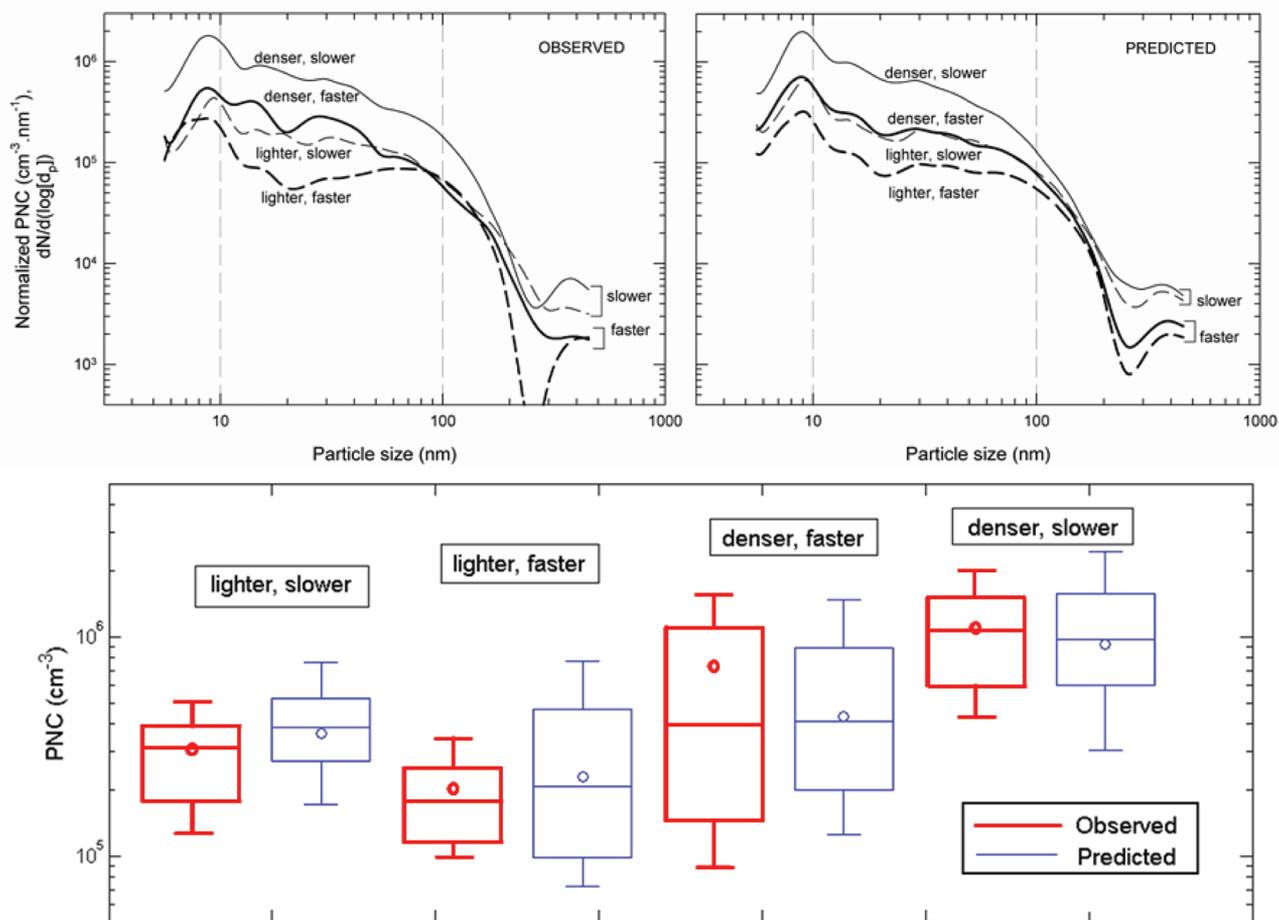


Figure 5. Ultrafine particle concentrations for four traffic conditions: denser, slower (traffic speed: 64–80 km h⁻¹, traffic volume: >80 veh min⁻¹); denser, faster (>105 km h⁻¹, 60–70 veh min⁻¹); lighter, slower (24–64 km h⁻¹, <40 veh min⁻¹); lighter, faster (>105 km h⁻¹, <40 veh min⁻¹). Size distributions (upper plots) present observations (left plot) and model predictions based on median values (right plot). Total particle number concentration is ~6× greater for the upper line (denser, slower; PNC: 980 000 cm⁻³) than for the lower line (lighter, faster; PNC: 171 000 cm⁻³). Lower plot compares observed and model-predicted distributions for total particle counts. Box plots show 10th, 25th, 50th, 75th, and 90th percentiles. Mean values are represented by open circles.

Table 2. Multilinear Regression Models for Prediction of Particle Number Concentration

dependent variable ^a	regression model ^b					RMSE ^c	R ²	adj-R ²	F-statistic
	constant term	SP coefficient	VOL coefficient	TEMP coefficient	RH coefficient				
10th percentile	5.25	-0.006	0.008			0.128	0.60	0.57	19.0
25th percentile	6.75	-0.007	0.012	-0.057		0.090	0.84	0.82	43.4
median	7.86	-0.002	0.007	-0.068	-1.55	0.077	0.85	0.83	33.5
mean	7.86		0.007	-0.078	-1.44	0.076	0.82	0.80	37.5
75th percentile	7.31		0.007	-0.047	-1.30	0.080	0.80	0.77	31.8
90th percentile	8.65	0.003	0.006	-0.086	-2.04	0.089	0.78	0.74	20.7
standard deviation	0.57	0.004	-0.003		-0.77	0.037	0.66	0.62	15.6

^aDependent variable: particle number concentrations (PNC; units, cm⁻³). Independent variables: strata average of loop speed (SP; units, km h⁻¹); total traffic volume (VOL; number of vehicles per minute); urban temperature (TEMP; °C); relative humidity (RH). All coefficients have *p* < 0.10.

^bWithin each strata of traffic speed and volume, 25th percentile, median, mean, and standard deviation were computed for log transformed data set for PNC. ^cRMSE = root-mean-square error.

models. Models stratified by year indicate that on-road concentrations measured in 2007 were lower than those measured in 2006, because of the reduction in emission factors (which Johnson et al.²¹ pointed out) and also because of differences in traffic conditions during field campaigns (less congestion, i.e., faster speeds and lower volumes, in 2007 than in 2006). On-road concentrations were slightly less dependent on traffic conditions in 2007 than in 2006; this finding is

manifested via (in 2007) a larger constant-term (intercept) in the model and smaller coefficients (beta values) for traffic speed and traffic volume (Table S6, Supporting Information).

Bootstrap results (Table S7, Supporting Information) indicate that standard errors of the coefficients from the original data set match well with the standard deviation of the bootstrap results. Low CV values for traffic speed and volume coefficients (9% to 29%) in bootstrap models indicate the

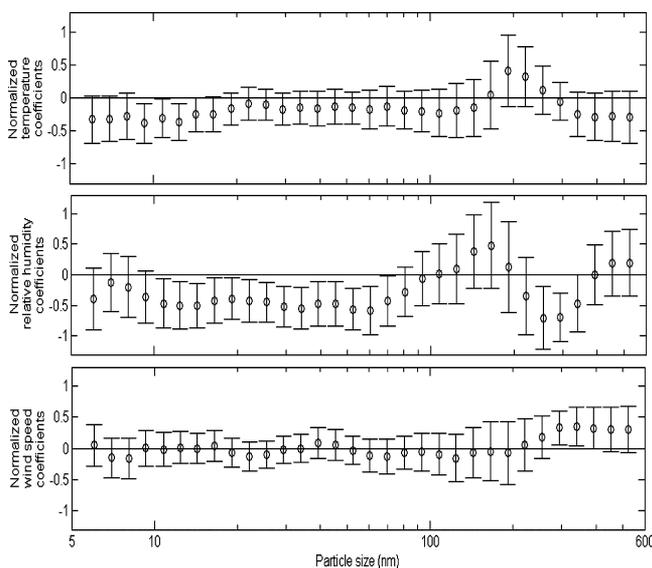


Figure 6. Normalized coefficients of meteorological parameters for size resolved prediction of median particle count measurements, from the weather-inclusive models: (a) normalized temperature coefficients, (b) normalized relative humidity coefficients, and (c) normalized wind speed coefficients for the 32 models. The error bars represent 95% confidence intervals. The coefficients and error bars are normalized to the ratio of the standard deviation of the parameter to the standard deviation of the particle counts in the size bin. For example, a y-axis value of +0.5 in the upper plot (temperature) would mean that a 1 standard-deviation increase in temperature would yield a 0.5 standard-deviation increase in concentration for that size bin.

stability of these two parameters as PNC predictors; the comparatively high CV for temperature (111%) in the 25th percentile model suggests that the model coefficients for temperature are less robust. Relatively low CV values for adj- R^2 (10–14%) for the models (exception: the standard deviation model) suggests general robustness of these models. As mentioned above, we employed 20 alternative stratification schemes; results were similar among all 20 schemes, suggesting that results presented here are robust to the stratification scheme employed.

This study has several additional uncertainties, especially for applying results to exposure analysis. Concentrations were measured at 3.5 m, which may or may not be representative of concentrations in vehicles²⁵ or at car height (typically, 1–2 m). Concentrations were measured on-freeway in one urban area (Minneapolis) during one season (summer) during daytime only (9 am to 3 pm); results may not apply to other locations or times. To interpolate rather than extrapolate from measurements, the model generally should only be applied if the values for independent variables are inside the range of values in the input data. (For example, during the 40 h of on-road measurements, temperature and wind speed were 18–34 °C and 0–37 km h⁻¹, respectively. During summer 2008, meteorological conditions in Minneapolis were in this range during 75% of the time.) Additional sampling (e.g., different times-of-day, seasons, cities, road types) would usefully extend the applicability of the models. Particle emissions may depend on characteristics such as altitude, topography, driver characteristics, and vehicle type. For example, Paatero et al.²⁶ modeled particle number concentrations (PNC) in five European cities based on other air-pollution parameters and suggested different models for each city. Model inputs employed here (e.g., traffic

volume) are based on available real-time data but are not, for example, subclassified by vehicle type (e.g., heavy duty versus light duty vehicles).^{14,21,27,28}

On the basis of traffic camera records, heavy duty vehicles comprised 7% of on-highway traffic during data collection.^{21,29} We do not expect major shifts in the vehicle mix during the sampling hours (9 am–3 pm), but the real-time truck counts needed to test this hypothesis are not available for the times and locations where concentration measurements occurred. If real-time data on vehicle type were available, we would be able to include it in our approach; this issue represents an opportunity for future model improvement. Johnson²⁹ reported that particle-number emissions per kg fuel (per minute) are 7× (44×) greater for heavy- than for light-duty vehicles.

As mentioned above, few studies have investigated size-resolved models of UFPs based on traffic parameters, but prior studies have investigated (nonsize-resolved) UFPs on and near roadways. For example, Fruin et al.¹⁴ conducted real-time measurements of UFP on California freeways and suggested that truck density ($R^2 = 0.58$) and hour of day ($R^2 = 0.26$) are good predictors of UFP concentrations. Nanzetta and Holmen²⁷ showed that road-side measurements of UFPs in California were correlated with traffic counts and meteorological parameters but the observed correlations were not strong (correlation coefficients in the range of 0.21–0.46). Similar results, that traffic density and meteorology are important predictors of UFP, were reported by Voigtländer et al.³⁰ for a street canyon in Germany ($R^2 = 0.68$). Measurements in the latter two studies were conducted less than 10 m from the roadway. Paatero et al.²⁶ used more than 50 explanatory variables (including traffic related air pollutants and meteorological parameters) to predict PNCs in five European cities ($R^2 = 0.58–0.77$); our results suggest that more parsimonious models can also yield reasonable results. In a recent study, Kaur and Nieuwenhuijsen³¹ used traffic counts, meteorological parameters, and mode of transport to model personal UFP exposure in UK ($R^2 = 0.62$). One study that did consider size-resolved UFP exposures on-roadway emphasized that concentrations may differ between on-roadway and in-vehicle air.²⁵

Our findings are in agreement with some but not all previous studies. Given the different goals of our study versus previous studies, direct comparison of model coefficients or model results is not feasible in most cases. Only one of the studies³⁰ investigated size resolved correlations of PNCs with traffic counts (passenger cars). Trends of size-resolved prediction are similar between their and our study: model performance is best for 10–100 nm particles (ultrafines) and comparatively poorer for ~200–500 nm particles. Knibbs and de Dear³² measured UFP concentrations in four modes (train, bus, ferry, automobile) in Sydney, Australia, and found wind speed to be a weak determinant of UFP concentrations. On the other hand, some prior studies^{30,31} found that meteorological parameters (especially wind speed) are important for predicting PNCs; however, those studies investigated surface streets rather than on-freeway air. Our result that RH and PNC are inversely related (Figure 6) is consistent with some²⁷ but not all³¹ extant research.

The approach developed here, involving two-way stratified multilinear regression, could usefully be applied to other locations, times, or pollutants. Unique features of our approach include providing a simple model with good prediction of central tendency (adj- $R^2 = 0.80$ and 0.83 for mean and median, respectively), based on readily available and already measured

parameters. The proposed models are useful for determining time-, location-, and size-resolved concentrations of fine particles on freeways, including for identifying concentration hotspots. This information could be applied to identifying susceptible subpopulations; for estimating exposures (see ref 25), including as part of large-scale epidemiological studies; for testing potential control strategies; and eventually for estimating public health risks from exposure to ultrafine particles on highways.

■ ASSOCIATED CONTENT

📄 Supporting Information

Additional details about the stratification scheme and bootstrap results are provided in the following tables and figures: Table S1, an expanded version of Table 1; Table S2, the stratification scheme used in this work; Table S3, speeds, traffic volumes, and particle count information for the strata included in the study; Table S4, summary statistics for the input variables (weather; traffic) used in this study; Table S5, an expanded list of stepwise regression models subdivided by wind-direction; Table S6, an expanded list of stepwise regression models subdivided by measurement year; Table S7, the results from bootstrap analyses; Figure S1, four color-coded maps for measured and predicted (based on the mean and median predicted models) PNC along the test route for a single day and for ten days, respectively (to contrast short-term versus long-term prediction results); Figure S2, plots for correlations between predicted and observed PNC values for 10th percentile, 25th percentile, median, 75th percentile, 90th percentile, mean, and standard deviation; Figure S3, correlations for observed PNC values with average loop speed and traffic loop volume; Figure S4, color-coded maps for average loop speeds and total loop volumes along the test route, averaged for the same ten days as in Figure S1; Figure S5, the non-normalized coefficients of weather parameters for size resolved predictive models. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +1-612-625-2397; fax: +1-612-626-7750; e-mail: julian@umn.edu.

■ ACKNOWLEDGMENTS

We thank David Kittelson, Win Watts, and Jason Johnson for providing the on-roadway data and for helpful comments on an earlier draft of this article. We also thank the anonymous reviewers for helpful suggestions. This research has been supported by a grant from the National Center for Environmental Research (NCER) STAR Program, EPA.

■ REFERENCES

- (1) Oberdörster, G. Pulmonary effects of inhaled ultrafine particles. *Int. Arch. Occup. Environ. Health* **2000**, *74*, 1–8.
- (2) Ibaldo-Mulli, A.; Wichmann, H. E.; Kreyling, W.; Peters, A. Epidemiological evidence on health effects of ultrafine particles. *J. Aerosol Med.* **2002**, *15*, 189–201.
- (3) Geiser, M.; Rothen-Rutishauser, B.; Kapp, N.; Schürch, S.; Kreyling, W.; Schulz, H.; Semmler, M.; Im Hof, V.; Heyder, J.; Gehr, P. Ultrafine particles cross cellular membranes by nonphagocytic mechanisms in lungs and in cultured cells. *Environ. Health Perspect.* **2005**, *113*, 1555–60.
- (4) Choi, H. S.; Ashitate, Y.; Lee, J. H.; Kim, S. H.; Matsui, A.; Insin, N.; Bawendi, M. G.; Semmler-Behnke, M.; Frangioni, J. V.; Tsuda, A.

Rapid translocation of nanoparticles from the lung airspaces to the body. *Nat. Biotechnol.* **2010**, *28*, 1300–1303.

- (5) Crapo, J. D.; Barry, B. E.; Gehr, P.; Bachofen, M.; Weibel, E. R. Cell number and cell characteristics of the normal human lung. *Am. Rev. Respir. Dis.* **1982**, *126*, 332–337.
- (6) Tuch, T.; Brand, P.; Wichmann, H. E.; Heyder, J. Variation of particle number and mass concentration in various size ranges of ambient aerosols in Eastern Germany. *Atmos. Environ.* **1997**, *31*, 4193–4197.
- (7) Hu, S.; Fruin, S.; Kozawa, K.; Mara, S.; Paulson, S. E.; Winer, A. M. A wide area of air pollutant impact downwind of a freeway during pre-sunrise hours. *Atmos. Environ.* **2009**, *43*, 2541–2549.
- (8) Westerdahl, D.; Fruin, S.; Sax, T.; Fine, P. M.; Sioutas, C. Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los Angeles. *Atmos. Environ.* **2005**, *39*, 3597–3610.
- (9) Zhu, Y.; Hinds, W. C.; Kim, S.; Shen, S.; Sioutas, C. Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmos. Environ.* **2002**, *36*, 4323–4335.
- (10) Pirjola, L.; Paasonen, P.; Pfeiffer, D.; Hussein, T.; Hämeri, K.; Koskentalo, T.; Virtanen, A.; Rönkkö, T.; Keskinen, J.; Pakkanen, T. A. Dispersion of particles and trace gases nearby a city highway: mobile laboratory measurements in Finland. *Atmos. Environ.* **2006**, *40*, 867–879.
- (11) Kittelson, D. B.; Watts, W. F.; Johnson, J. P. Nanoparticle emissions on Minnesota highways. *Atmos. Environ.* **2004**, *38*, 9–19.
- (12) Weijers, E. P.; Khlystov, A. Y.; Kos, G. P. A.; Erismann, J. W. Variability of particulate matter concentrations along roads and motorways determined by a moving measurement unit. *Atmos. Environ.* **2004**, *38*, 2993–3002.
- (13) Int Panis, L.; de Geus, B.; Vandenbulcke, G.; Willems, H.; Degraeuwe, B.; Bleux, N.; Mishra, V.; Thomas, I.; Meeusen, R. Exposure to particulate matter in traffic: a comparison of cyclists and car passengers. *Atmos. Environ.* **2010**, *44*, 2263–2270.
- (14) Fruin, S.; Westerdahl, D.; Sax, T.; Sioutas, C.; Fine, P. M. Measurements and predictors of on-road ultrafine particle concentrations and associated pollutants in Los Angeles. *Atmos. Environ.* **2008**, *42*, 207–219.
- (15) Canagaratna, M. R.; Jayne, J. T.; Ghertner, D. A.; Herndon, S.; Shi, Q.; Jimenez, J. L.; Silva, P. J.; Williams, P.; Lanni, T.; Drewnick, F. Chase studies of particulate emissions from in-use New York City vehicles. *Aerosol Sci. Technol.* **2004**, *38*, 555–573.
- (16) Bukowiecki, N.; Dommen, J.; Prevot, A. S. H.; Richter, R.; Weingartner, E.; Baltensperger, U. A mobile pollutant measurement laboratory—measuring gas phase and aerosol ambient concentrations with high spatial and temporal resolution. *Atmos. Environ.* **2002**, *36*, 5569–5579.
- (17) Bukowiecki, N.; Kittelson, D. B.; Watts, W. F.; Burtscher, H.; Weingartner, E.; Baltensperger, U. Real-time characterization of ultrafine and accumulation mode particles in ambient combustion aerosols. *J. Aerosol Sci.* **2002**, *33*, 1139–1154.
- (18) Larson, T.; Henderson, S. B.; Brauer, M. Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. *Environ. Sci. Technol.* **2009**, *43*, 4672–4678.
- (19) Henderson, S. B.; Beckerman, B.; Jerrett, M.; Brauer, M. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environ. Sci. Technol.* **2007**, *41*, 2422–2428.
- (20) Ross, Z.; English, P. B.; Scalf, R.; Gunier, R.; Smorodinsky, S.; Wall, S.; Jerrett, M. Nitrogen dioxide prediction in Southern California using land use regression modeling: potential for environmental health analyses. *J. Exposure Sci. Environ. Epidemiol.* **2005**, *16*, 106–114.
- (21) Johnson, J. P.; Kittelson, D. B.; Watts, W. F. The effect of federal fuel sulfur regulations on in-use fleets: on-road heavy-duty source apportionment. *Environ. Sci. Technol.* **2009**, *43*, 5358–5364.
- (22) Online Weather Information, Weather Underground; Available: <http://www.wunderground.com>.
- (23) Minnesota Department of Transportation Traffic Data; Available: <http://data.dot.state.mn.us>.

(24) Riddle, S. G.; Robert, M. A.; Jakober, C. A.; Hannigan, M. P.; Kleeman, M. J. Size-resolved source apportionment of airborne particle mass in a roadside environment. *Environ. Sci. Technol.* **2008**, *42*, 6580–6586.

(25) Zhu, Y.; Eiguren-Fernandez, A.; Hinds, W. C.; Miguel, A. H. In-cabin commuter exposure to ultrafine particles on Los Angeles freeways. *Environ. Sci. Technol.* **2007**, *41*, 2138–2145.

(26) Paatero, P.; Aalto, P.; Picciotto, S.; Bellander, T.; Castaño, G.; Cattani, G.; Cyrys, J.; Kulmala, M.; Lanki, T.; Nyberg, F. Estimating time series of aerosol particle number concentrations in the five HEAPSS cities on the basis of measured air pollution and meteorological variables. *Atmos. Environ.* **2005**, *39*, 2261–2273.

(27) Nanzetta, M. K.; Holmen, B. A. Roadside particle number distributions and relationships between number concentrations, meteorology, and traffic along a northern California freeway. *J. Air Waste Manage. Assoc.* **2004**, *54*, 540–554.

(28) Hussein, T.; Karppinen, A.; Kukkonen, J.; Härkönen, J.; Aalto, P. P.; Hämeri, K.; Kerminen, V. M.; Kulmala, M. Meteorological dependence of size-fractionated number concentrations of urban aerosol particles. *Atmos. Environ.* **2006**, *40*, 1427–1440.

(29) Johnson, J. P. The impact of ultra low sulfur diesel fuel on real-world fleet particle emissions. M.S. Thesis, University of Minnesota, Minneapolis, MN, 2009.

(30) Voigtländer, J.; Tuch, T.; Birmili, W.; Wiedensohler, A. Correlation between traffic density and particle size distribution in a street canyon and the dependence on wind direction. *Atmos. Chem. Phys.* **2006**, *6*, 4275–4286.

(31) Kaur, S.; Nieuwenhuijsen, M. J. Determinants of personal exposure to PM_{2.5}, ultrafine particle counts, and CO in a transport microenvironment. *Environ. Sci. Technol.* **2009**, *43*, 4737–4743.

(32) Knibbs, L. D.; de Dear, R. J. Exposure to ultrafine particles and PM_{2.5} in four Sydney transport modes. *Atmos. Environ.* **2010**, *44*, 3224–3227.

Supplementary Information

Online supplement for

Real time prediction of size resolved ultrafine PM on freeways

Srijan Aggarwal, Ricky Jain, Julian D. Marshall

This supplement provides a more detailed description of the stratification method, then the following tables and figures. Table S1 is an expanded version of Table 1. Table S2 shows the stratification scheme used in this work. Table S3 shows speeds, traffic volumes and particle count information for the strata included in the study. Table S4 provides summary statistics for the input variables (weather; traffic) used in this study. An expanded list of stepwise regression models subdivided by wind-direction and measurement year are presented in Table S5 and S6, respectively. Table S7 presents the results from bootstrap analyses.

Figure S1 shows four color-coded maps for measured and predicted (based on the mean and median predicted models) PNC along the test route for a single day and for ten days respectively (to contrast short-term versus long term prediction results). Figure S2 shows plots for correlations between predicted and observed PNC values for 10th percentile, 25th percentile, median, 75th percentile, 90th percentile, mean and standard deviation. Figure S3 shows correlations for observed PNC values with average loop speed and traffic loop volume. Figure S4 shows color coded maps for average loop speeds and total loop volumes along the test route, averaged for the same ten days as in Figure S1. Figure S5 shows the non-normalized coefficients of weather parameters for size resolved predictive models.

Stratification Method

Strata were selected so as to achieve even distribution of data among the strata. To ensure that our results are not contingent on the stratification scheme, we experimented with systematically varying the strata as a sensitivity analysis; twenty different stratification schemes were evaluated. For each alternative stratification scheme, we generated regression models. Speed and volume coefficients in the models exhibited low coefficients of variability, and model predictions were highly correlated with each other. Thus, we did not find evidence that model results are highly sensitive to stratification scheme. The most efficient stratification scheme was selected based on three criteria: model performance (measured by adjusted R^2), statistical significance of the model coefficients, and the percentage of minute-averaged data set used. The optimal stratification scheme selected had forty-two strata (six categories for traffic volumes and seven categories for traffic speeds) and used 96% of the valid data obtained. For each strata, we calculate mean values for traffic parameters (speed, volume); weather parameters (temperature, wind speed and relative humidity [RH]); and, for the dependent variable (log-transformed particle counts), several percentiles (10th, 25th, 50th, 75th and 90th percentiles), mean and standard deviation. Also, median of log-transformed particles counts were determined for each of the 32 EEPS channels. Example measurements are shown in Table S3 for total PNC.

Table S1: Summary of studies on real time on-road measurement of ultrafine particles (UFPs).

Reference	Sampling year	Location	Methods	Pollutants	Instruments	Number of hours	Key results
Bukowiecki 2002a [17]	1999-2000	Minneapolis, Minnesota. Columbus, Indiana	Average surface diameter and PAS/DC versus Avg surface diameter scatter plots,	UFP	photoelectric aerosol sensor(PAS), diffusion charging sensor (DC), CPC, SMPS.		Theoretical conclusions related to aerosols.
Bukowiecki 2002b [13]	2001	Zürich (Switzerland)	Design and construction of mobile laboratory, diurnal variation of aerosols.	UFP	CPC*	4 hours	1. This case study confirms that there is a large diurnal and regional variation of ultrafine particles for both urban and rural areas. 2. Neither the UFP nor the total particle number concentration is an exclusive indicator of primary traffic emissions.
Canagratna 2004 [15]	2000-2001	New York city, New York.	Vehicle chase studies, chemical composition, size distribution	UFP, PM, NO, NO ₂ , CO, N ₂ O, CH ₄ , SO ₂ , and HCHO.	Aerodyne, aerosol mass spectrometer (AMS), CPC, Tunable Diode Lasers (TDL), LiCOR	~24 hours	1. The nonrefractory diesel exhaust PM appears to be dominated by lubricating oil and the typical measured mass distribution of organic as well as sulfate species. 2. Order of PM emissions in diesel operated engines : 6V-92 engines> series-50 engines > CNG engines
Weijers et al. 2004 [12]	1999-2000	Amsterdam, Netherlands	Dispersion study, particle size distributions, regional variability, within city variability	UFP, PM	CPC and an optical aerosol spectrometer	3 days	1. Aerosol concentrations decrease exponentially with increasing distance from the road. 2. Number concentrations are more sensitive than mass concentrations due to dominance of number of UFPs. 3. Number concentration in city change on scale of a hundred meters, correlating with the local traffic intensity and driving conditions

Kittleson et al. 2004 [11]	2000	Minneapolis, Minnesota	Size distributions, fuel-specific emissions, dispersion study.	UFP, CO, CO ₂ , NO _x .	CPC, SMPS, IR CO analyzer, IR CO ₂ analyzer, chemiluminescence NO _x analyzer.	~ 20 hours	<ol style="list-style-type: none"> 1. High UFPs correlated with high speed traffic. 2. Most of the particles added by the on-road fleet were below 50nm in diameter. 3. Number concentrations measured in residential areas, 10–20m from the highway, considerably lower than on-road concentrations, and much lower concentrations for areas 500-700 m from the highway.
Westerdahl et al. 2005 [8]	2003	Los Angeles, CA	Pollutant concentration differences by location, pollutant concentration correlations, UFP size distributions, time series plots	BC, NO, PM-PAH,UFP, NO ₂ , CO, CO ₂ , PM _{2.5}	CPC, SMPS, Aethalometer, DMA, PAH Analyzer, NO _x analyzer, Q-trak plus monitor, TSI DusTrak	12 hours	<ol style="list-style-type: none"> 1. Good correlation between UFP concentrations and BC, NO and PM-PAHs. 2. Freeway concentrations an order of magnitude higher than on residential streets for UFP, NO, BC and CO. 3. Average concentrations of UFP and related pollutants varied strongly by location, road type, and truck traffic volumes, suggesting a relationship between these concentrations and truck traffic density.
Pirjola et al. 2006 [10]	2003-04	Helsinki, Finland	Seasonal (summer vs winter analysis).Pollutant dispersion measurements upto 140 m distance from the road side. Particle size distribution analysis in the range of 3 nm -10 um.	CO, NO, UFP, NO _x ,PM	CPC, SMPS, CO and NO _x monitors	384 hours	<ol style="list-style-type: none"> 1. Average concentrations 2-3 times higher in winter than in summer. 2. Concentrations fell to 20-40% as far as only 65 m away from the road, still more than double the background urban concentrations. 3. 85% or more particles were smaller than 50 nm. Observed distribution was multi-modal.
Fruin et al. 2008 [17]	2003	Los Angeles, CA	DVD analysis, multiple regression, ANOVA	BC, NO, PM-PAH,UFP, NO ₂ , CO, CO ₂ , PM _{2.5}	CPC, SMPS, Aethalometer, DMA, PAH Analyzer, NO _x analyzer, Q-trak plus monitor, TSI dustrak	15 hours	<ol style="list-style-type: none"> 1. Arterial concentrations one-third to freeway concentrations. 2. Freeways responsible for 33-45% UFP exposure in LA 3. Diesel powered vehicles primary sources of UFPs, NO, BC, PM-PAH

Johnson et al. 2009 [21]	2006-2007	Minneapolis, MN	Fuel specific apportionment of particle number concentration on highways, size distribution statistics	UFP, CO, CO ₂ , NO _x	CPC, SMPS, EEPS, TSI DustTrak, NDIR QTrak	85 hours	Decrease in fuel sulfur content led to reduced particle counts from 2002 to 2007.
Int Panis et al. 2010 [13]	2009	Belgium	Comparison of fine particle exposure to car riders and cyclists in three Belgian cities	UFP, PM _{2.5} , PM ₁₀	TSI P-Trak, TSI DustTrak	~30 hours	Mean bicycle/car ratio for PNC and PM are close to 1 and rarely significant.

*Other instrumentation to measure gaseous pollutants etc. were also used but the study focuses on only the data collected from the instrument(s) listed here

Table S2: Two-way stratification scheme, showing strata serial numbers (see Table S2 for data per serial number), based on average loop speed and total loop volume

Total loop volume (veh min ⁻¹)	Average loop speed (km h ⁻¹)						
	24 to 64	64 to 80	80 to 87	87 to 93	93 to 100	100 to 105	> 105
< 40	1	7	13	19	25	31	37
40-50	2	8	14	20	26	32	38
50-60	3	9	15	21	27	33	39
60-70	4	10	16	22	28	34	40
70-80	5	11	17	23	29	35	41
>80	6	12	18	24	30	36	42

Table S3. Speed, traffic volumes and particle count information for the strata included in the study

Strata serial number ^a	Number of data points	Average loop speed ^b (km h ⁻¹)	Total loop volume (veh min ⁻¹)	PNC ^c (Median Log)
1	20	53	35	5.50
2	25	53	45	5.60
9	31	76	55	5.65
10	52	76	65	5.75
11	37	76	75	5.77
12	22	76	88	6.02
14	44	84	46	5.61
15	84	84	56	5.78
16	142	84	65	5.84
17	81	84	74	5.88
18	21	84	86	5.86
20	70	90	46	5.58
21	213	90	55	5.78
22	258	90	65	5.83
23	110	90	74	5.81
24	29	90	86	5.70
25	21	96	36	5.31
26	84	96	46	5.49
27	181	96	56	5.67
28	175	96	65	5.73
29	53	96	74	5.79
32	42	102	46	5.47
33	77	102	55	5.46
34	52	102	65	5.72
37	20	110	33	5.27
38	54	110	45	5.35
39	81	110	55	5.46
40	43	110	63	5.60

^a See Table S1. For example, Strata #1 (first row) refers to times when the loop volume is <40 veh min⁻¹ and traffic speed is 24-64 km h⁻¹. Only the cells with greater than 20 data points were included in the analysis.

^b Ensemble averages for the cells with same speed range were calculated, because individual cell averages for the same speed range were similar.

^c Values are the base-10 logarithm of the median. For example, in the first row log median is 5.50, indicating a particle concentration of 316,000 particles cm⁻³.

Table S4: Summary statistics for the independent variables

Statistical parameter	Temperature (deg C)	RH	Wind speed (km h ⁻¹)	Traffic speed (km h ⁻¹)	Traffic volume (veh min ⁻¹)
Mean	25	0.46	16.0	91.3	60.1
Median	25	0.44	14.8	91.6	60.1
Std Dev	3.6	0.11	8.3	11.3	11.7
SE	0.08	0	0.18	0.24	0.25
CV (%)	14%	23%	52%	12%	19%
Minimum	17.8	0.28	0	36.7	23.4
10 th percentile	20.6	0.34	5.6	80.3	44.9
25 th percentile	22.2	0.37	9.3	86.3	52.3
75 th percentile	27.8	0.55	20.4	97.4	67.4
90 th percentile	30	0.61	27.9	104	74
Maximum	33.9	0.8	37.0	130	101

Table S5: Multi-linear regression models for prediction of particle number concentration (PNC) based on wind speed classification.

Dependent variable ^a	Condition ^e	Un-aided stepwise regression ^b			Aided stepwise regression ^c		
		Parameter ^d	R ²	adj-R ²	Parameter	R ²	adj-R ²
Median	upwind	RH	0.66	0.64	RH, WS	0.70	0.68
	downwind	Vol, Temp, RH , WS	0.73	0.69	Same as un-aided model		
	niether	RH	0.72	0.71	Same as un-aided model		
	total	RH, WS	0.78	0.76	Speed, Vol, Temp, RH	0.85	0.83
					Vol, Temp, RH	0.83	0.81
					Vol, RH , WS	0.81	0.78
Mean	upwind	Vol, RH, WS	0.76	0.73	Speed, Vol , RH, WS	0.79	0.76
	downwind	Vol	0.46	0.44	Vol , RH, WS	0.76	0.73
	niether	Vol, RH	0.73	0.70	Vol , Temp, WS	0.66	0.62
	total	Speed, Vol	0.79	0.77	Same as un-aided model		
					Speed, Vol , Temp	0.81	0.79
10th	upwind	Speed, Vol, WS	0.73	0.69	Vol , Temp, RH	0.82	0.80
				Same as un-aided model			

Percentile	downwind	Vol, Temp	0.41	0.36	Same as un-aided model		
	niether	RH	0.36	0.34	Speed, Vol	0.53	0.49
	total	Speed, Vol	0.60	0.57	Same as un-aided model		
25th Percentile	upwind	Speed, Vol, WS	0.75	0.72	Vol, RH, WS	0.69	0.65
	downwind	Vol, Temp	0.60	0.57	Vol, RH	0.64	0.61
	niether	RH, WS	0.76	0.74	Same as un-aided model		
75th Percentile	total	Speed, Vol, Temp	0.84	0.82	Speed, Vol	0.78	0.76
	upwind	RH	0.55	0.54	Same as un-aided model		
	downwind	RH	0.46	0.44	RH, WS	0.61	0.58
90th Percentile	niether	RH	0.63	0.62	Temp, RH, WS	0.64	0.59
	total	Vol, RH	0.77	0.75	Speed, RH, WS	0.72	0.68
	upwind	RH, WS	0.65	0.62	Vol, Temp, RH	0.80	0.77
Standard Deviation	downwind	RH	0.31	0.29	Vol, RH, WS	0.69	0.65
	niether	Vol	0.46	0.44	Temp, RH, WS	0.54	0.48
	total	Vol	0.66	0.64	Speed, RH	0.38	0.34
Standard Deviation	upwind	None			RH	0.40	0.38
	downwind	WS	0.17	0.14	Vol, Temp, RH	0.74	0.71
	niether	Speed	0.41	0.39	Speed, Vol, Temp, RH	0.78	0.74
Standard Deviation	total	Speed	0.46	0.44	RH	0.54	0.53
	upwind	None			Speed, Temp, RH	0.43	0.36
	downwind	WS	0.17	0.14	Vol, Temp, WS	0.38	0.30
Standard Deviation	niether	Speed	0.41	0.39	Vol, Temp	0.28	0.22
	total	Speed	0.46	0.44	Vol, WS	0.28	0.22
	upwind	None			Speed, Vol, RH, WS	0.64	0.58
Standard Deviation	downwind	WS	0.17	0.14	Speed, Vol, RH	0.66	0.62
	niether	Speed	0.41	0.39			
	total	Speed	0.46	0.44			

^aDependent variable: Particle number concentrations (PNC; units: cm^{-3}); ^bUn-aided stepwise regression refers to forward stepwise regression starting with a null model. ^cAided stepwise regression refers to a combination of forward and backward (selective removal starting with a full model) stepwise approach. ^dIndependent variables: strata average of loop speed (Speed; units: km h^{-1}); total traffic volume (Vol; number of vehicles per minute); urban temperature (TEMP; $^{\circ}\text{C}$); relative humidity (RH); and, wind speed (WS; km h^{-1}). All coefficients have $p < 0.10$. Within each strata of traffic speed and volume, 25th percentile, median, mean and standard deviation were computed for log transformed data set for PNC. ^eWind direction condition.

Table S6: Multi-linear regression models for prediction of particle number concentration (PNC) based on different measurement years.

	median log(PNC) model ^{a,c}		mean log(PNC) model ^{a,c}	
	Year 2006	Year 2007	Year 2006	Year 2007
constant term	5.40	5.42	5.35	5.46
SP coefficient ^b	-0.009	-0.002	-0.006	-0.003
VOL coefficient ^b	0.014	0.006	0.012	0.005
Adj-R ²	0.73	0.50	0.71	0.50

^aDependent variable: Particle number concentrations (PNC; units: cm⁻³); ^bIndependent variables: strata average of loop speed (SP; units: km h⁻¹); total traffic volume (VOL; number of vehicles per minute); ^cWithin each strata of traffic speed and volume, median and mean were computed for log transformed data set for PNC.

Table S7: Bootstrap re-sampling results for the particle count dataset used in this study

	Original dataset		Bootstrap datasets ^a				
	Values	SE ^b	min	Mean	max	SD ^c	CV ^d
25th percentile							
log PNC							
Constant	2.578	1.324	0.122	4.278	9.195	1.303	30%
Speed coeff	-0.011	0.002	-0.017	-0.011	-0.002	0.002	20%
Volume coeff	0.010	0.001	0.005	0.010	0.013	0.001	14%
Temp coeff	0.037	0.017	-0.054	0.015	0.069	0.017	111%
Adj R ²	0.823		0.285	0.699	0.885	0.096	14%
50th percentile							
log PNC							
Constant	5.441	0.128	5.008	5.425	5.809	0.116	2%
Speed coeff	-0.007	0.002	-0.011	-0.006	0.000	0.002	29%
Volume coeff	0.010	0.001	0.006	0.010	0.014	0.001	12%
Adj-R ²	0.771		0.424	0.685	0.878	0.075	11%
Mean							
log PNC							
Constant	5.379	0.118	5.099	5.382	5.678	0.094	2%
Speed coeff	-0.005	0.002	-0.009	-0.005	-0.001	0.001	28%
Volume coeff	0.009	0.001	0.006	0.009	0.012	0.001	9%
Adj R ²	0.769		0.454	0.698	0.858	0.067	10%
SD log PNC							
Constant	0.132	0.053	-0.075	0.115	0.351	0.076	66%
Speed coeff	0.004	0.001	0.001	0.005	0.008	0.001	29%
Adj R ²	0.442		-0.033	0.316	0.719	0.159	50%

^a Bootstrap sample size: n=1000. ^b SE is standard error. ^c SD is standard deviation of mean. ^d CV is the coefficient of variation.

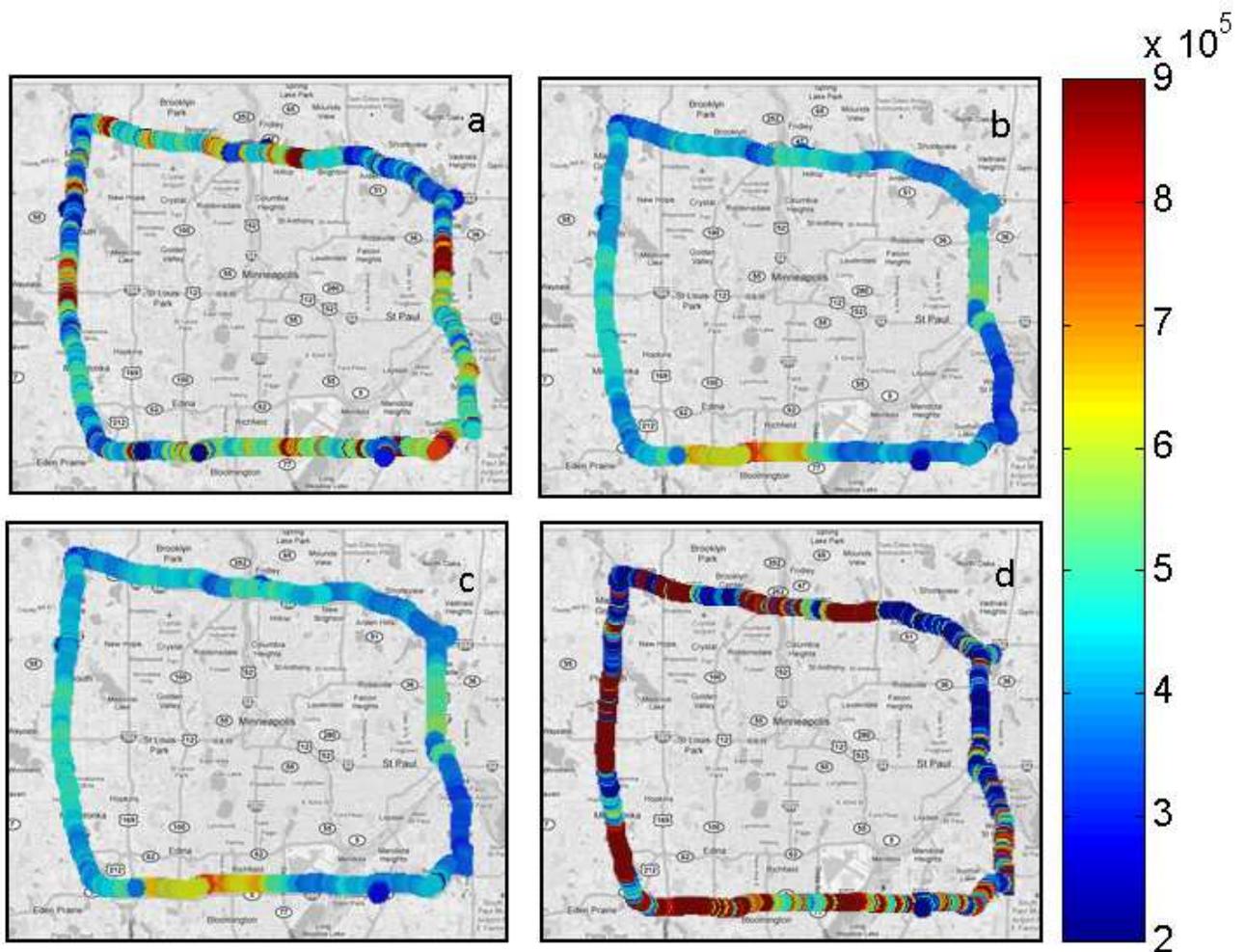


Figure S1: Particle number concentrations along the test route (a) measured concentrations averaged over ten summer days in June 2007 (b) median predicted concentrations averaged over ten summer days in June 2007 (c) mean predicted concentrations averaged over ten summer days in June 2007 (d) median predicted concentrations along the test route on a representative day (28th June, 2007; 9:00 AM to 1:30 PM). MN-DoT measured vehicle speeds and volumes were used to make predictions using the proposed models.

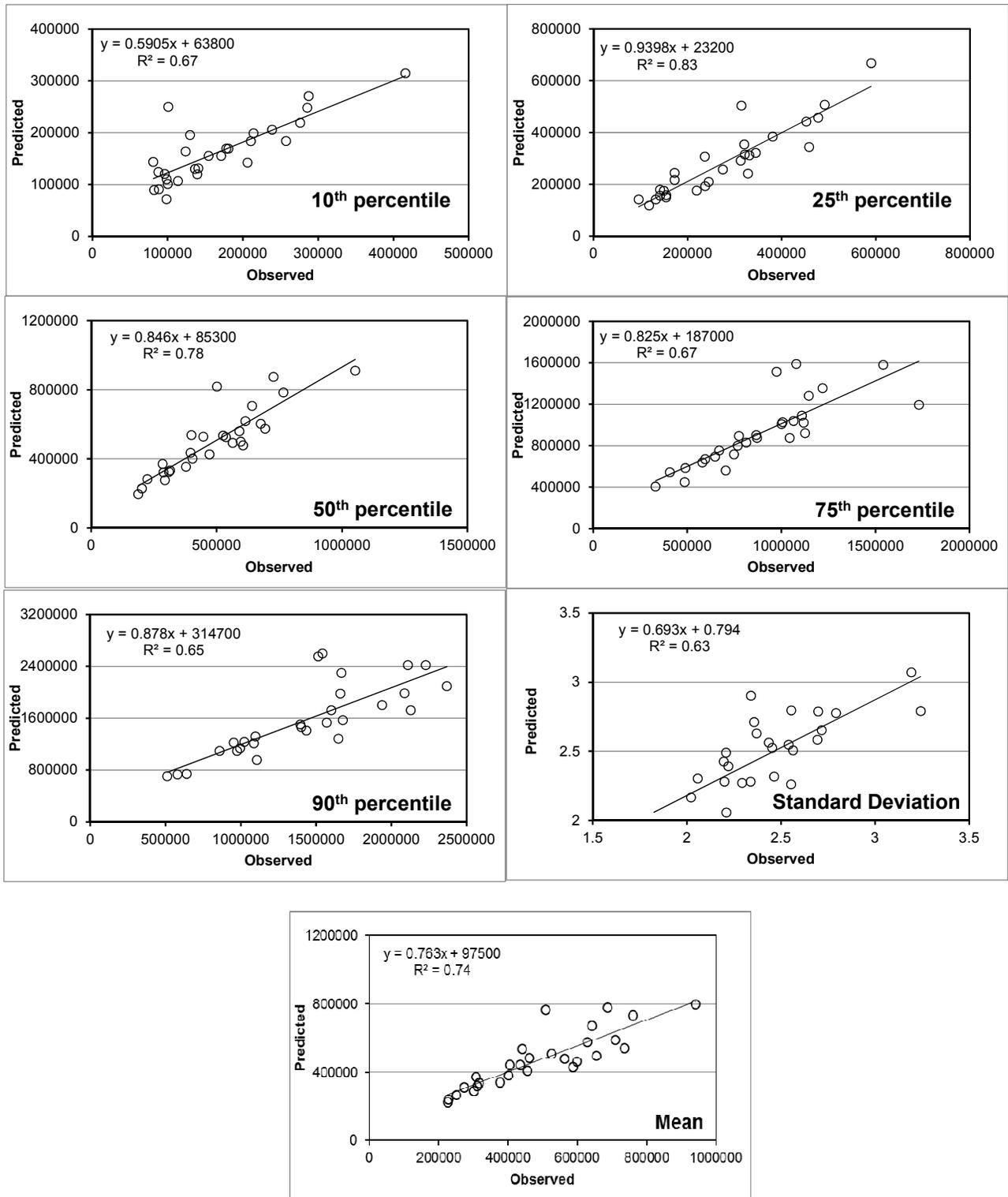


Figure S2: Regression plots for predicted versus observed PNC values for each of 28 strata (as shown in Table S3), for 10th, 25th, 50th, 75th, 90th percentiles; mean and standard deviation.

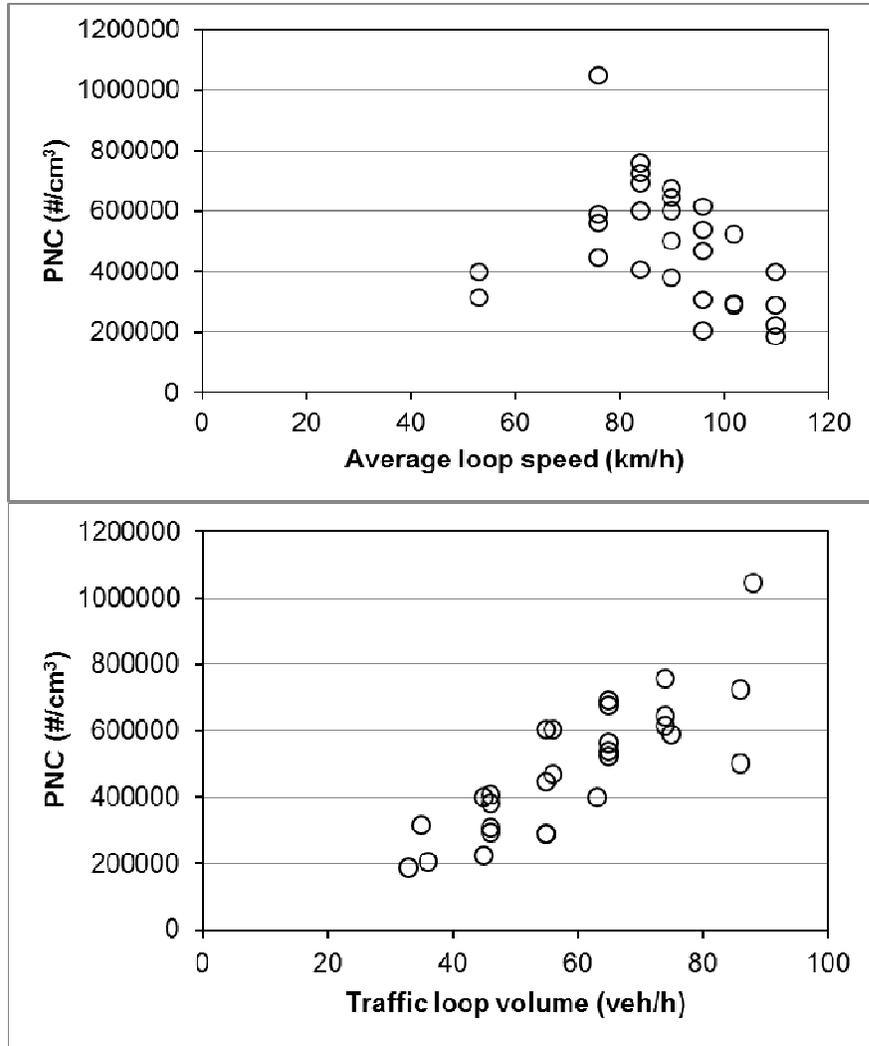


Figure S3: Correlations for observed PNC values with average loop speed and traffic loop volume for each of the 28 strata (as shown in Table S3).

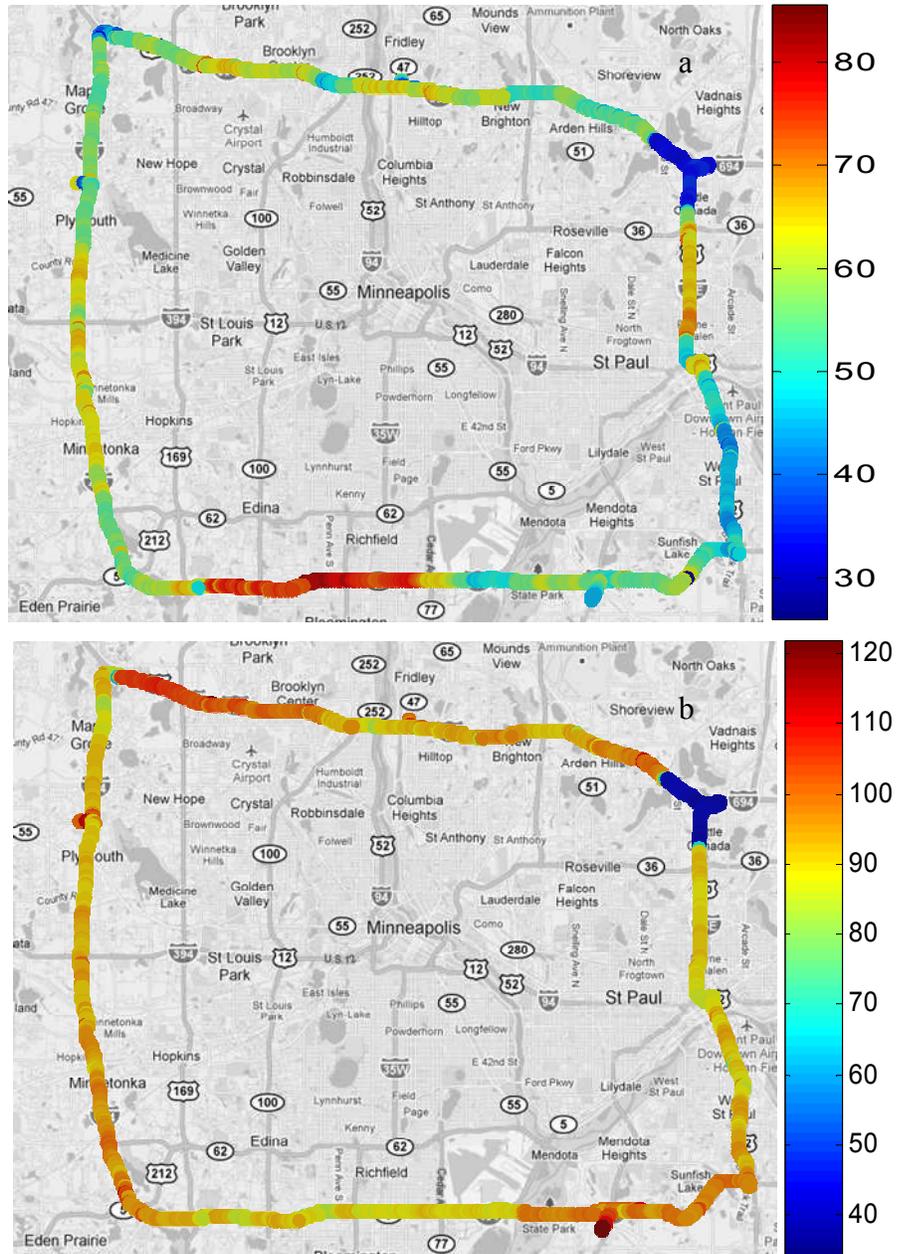


Figure S4: Measured values of (a) total traffic volumes, and (b) mean loop speeds, averaged over ten summer days in June 2007. These speeds and volumes were used to make prediction plots in Figure S1.

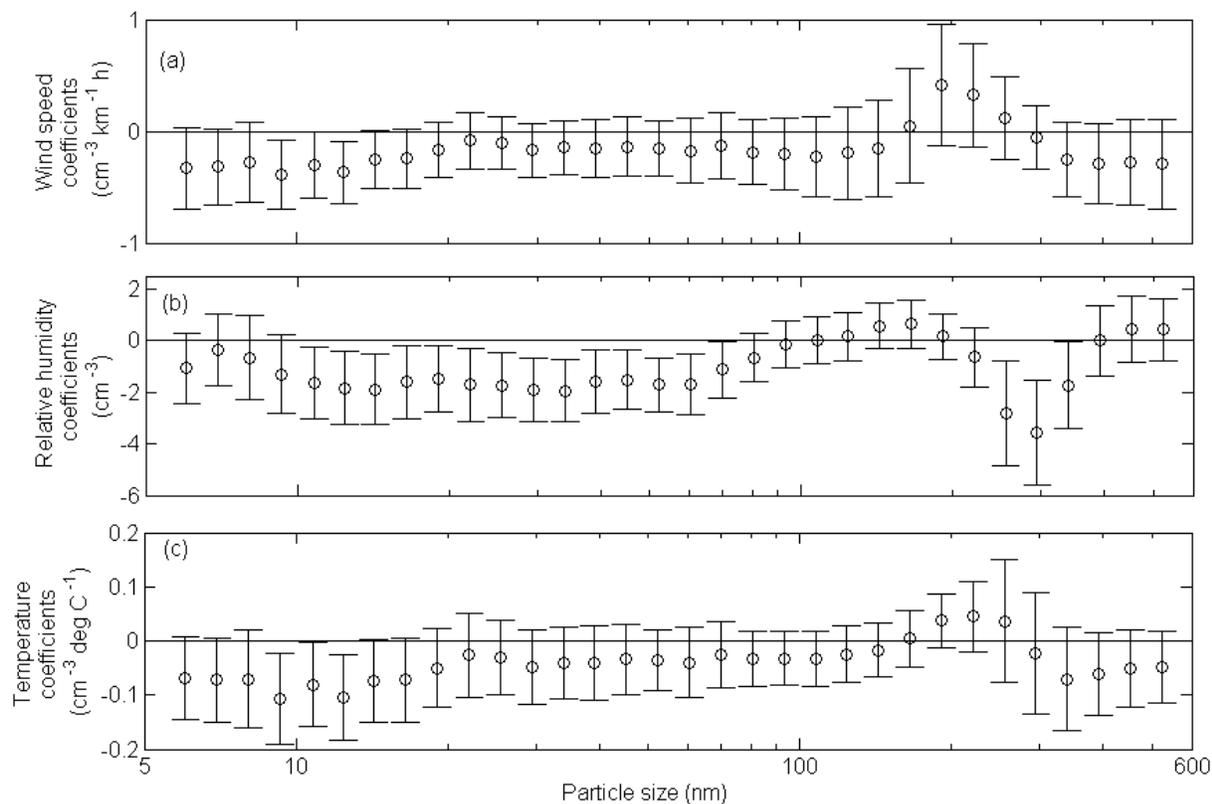


Figure S5: Coefficients (non-normalized) of meteorological parameters for size resolved prediction of median particle count measurements, from the ‘inclusive weather’ models. (a) temperature coefficients, (b) relative humidity coefficients, and (c) wind speed coefficients for the thirty-two models. The error bars represent 95% confidence intervals.