



Published in final edited form as:

Environ Sci Technol. 2006 July 1; 40(13): 4226–4232.

Modeling Particle Exposure in US Trucking Terminals

ME Davis^{1,2,*}, **TJ Smith**¹, **F Laden**^{1,3}, **JE Hart**^{1,3}, **LM Ryan**⁴, and **E Garshick**^{3,5}

1 Department of Environmental Health, Harvard School of Public Health, 401 Park Drive, Boston, Massachusetts 02215

2 Environmental Earth and Ocean Sciences Department, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, Massachusetts 02125

3 Channing Laboratory, Brigham and Women's Hospital and Harvard Medical School, 181 Longwood Avenue, Boston, Massachusetts 02115

4 Department of Biostatistics, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115

5 Pulmonary and Critical Care Medicine Section, VA Boston Healthcare System, 1400 VFW Parkway, West Roxbury, Massachusetts 02132

Abstract

Multi-tiered sampling approaches are common in environmental and occupational exposure assessment, where exposures for a given individual are often modeled based on simultaneous measurements taken at multiple indoor and outdoor sites. The monitoring data from such studies is hierarchical by design, imposing a complex covariance structure that must be accounted for in order to obtain unbiased estimates of exposure. Statistical methods such as structural equation modeling (SEM) represent a useful alternative to simple linear regression in these cases, providing simultaneous and unbiased predictions of each level of exposure based on a set of covariates specific to the exposure setting. We test the SEM approach using data from a large exposure assessment of diesel and combustion particles in the US trucking industry. The exposure assessment includes data from 36 different trucking terminals across the United States sampled between 2001 and 2005, measuring PM_{2.5} and its elemental carbon (EC), organic carbon (OC) components, by personal monitoring, and sampling at two indoor work locations and an outdoor “background” location. Using the SEM method, we predict: 1) personal exposures as a function of work related exposure and smoking status; 2) work related exposure as a function of terminal characteristics, indoor ventilation, job location, and background exposure conditions; and 3) background exposure conditions as a function of weather, nearby source pollution, and other regional differences across terminal sites. The primary advantage of SEMs in this setting is the ability to simultaneously predict exposures at each of the sampling locations, while accounting for the complex covariance structure among the measurements and descriptive variables. The statistically significant results and high R² values observed from the trucking industry application supports the broader use of this approach in exposure assessment modeling.

*Corresponding author phone: (617) 384-8809; Fax: (617) 384-8728; Email: medavis@hsph.harvard.edu.

Supporting Information

Map of US study locations, terminal layout, pathway graph for statistical analysis, summary statistics for model covariates, results of principle component analyses for terminal characteristics, cross-validation results for background measurements compared to EPA AIRS data, SEM results for organic carbon and PM_{2.5}, SEM results for PCA factors, OLS regression results.

Introduction

Diesel exhaust contains respirable particles with mutagenic and carcinogenic compounds (1) and has been linked to increased risk of lung cancer mortality in over 35 epidemiologic studies (2,3). Of particular interest are the rates in occupational groups exposed to diesel exhaust, such as the trucking and railroad industries. Although epidemiologic studies illuminating the increased cancer risks are numerous, none of these studies have included detailed exposure assessments, and for this reason the dose-response relationship remains uncertain (4). While multiple national and international health organizations have declared diesel exhaust to be a “probable” human carcinogen (5,6,7,8), the lack of quantitative exposure data has been consistently cited as the fundamental problem in determining causality from the existing occupational health studies of diesel exposure (2,9).

In an attempt to address this issue, the Trucking Industry Particle Study was carefully planned as a joint effort in exposure assessment and epidemiology (10). Our hypothesis was that trucking industry workers are highly exposed to diesel exhaust because of their close proximity to operating diesel trucks. Elemental carbon (EC), which is a major component of diesel engine emissions, was chosen as a marker for diesel exposure. A national exposure assessment collecting PM_{2.5}, EC, and OC data has been conducted to complement epidemiologic data on lung cancer mortality for workers in this industry. The epidemiologic study is being conducted with cooperation from the International Brotherhood of Teamsters and four large unionized trucking companies. This retrospective cohort study contains approximately 55,000 people working in 1985, whose lung cancer mortality experience has been determined through the year 2000. The exposure assessment is intended to provide data for the epidemiology study, monitoring exposures at the four trucking companies to define current exposure levels in the industry. These estimates, coupled with information on policy and regulatory changes across time, will enable us to predict historical levels and assign cumulative exposure to members in the epidemiology cohort, and to ultimately estimate lung cancer risk for this occupational group.

Although we are using EC as a surrogate, or marker of diesel emissions, other combustion sources also emit EC and OC, and the amounts of EC and OC emitted by diesels varies with driving conditions (11). Lightly loaded and idling diesel emissions have low EC and high OC and are indistinguishable from general gasoline engine emissions. Previous measurements made by our study team indicated that diesel exhaust is the dominant source of EC in the terminal work locations (representing all engine operating conditions). Furthermore, EC is not a significant component of cigarette smoke (0.49%), and is a small component of gasoline car engine emissions (2–5 mg/km) and propane emissions, when compared to heavy-duty truck engine emissions (early 1990’s technology, 164 mg/km) (12). Therefore, the EC in terminal locations is primarily an indicator of diesel emissions with some influence from car and forklift exhaust, and other combustion particles, such as from home oil heating and outdoor burning in some locations and seasons.

Methods and Data Collection

Our sampling plan was designed to measure particle exposures of fine particle mass (PM_{2.5}) and two of its components, elemental carbon (EC) and organic compounds (OC) at a randomly selected and regionally representative set of large US truck freight terminals (>100 employees) from four participating companies. Samples were taken during a 5-day period, and a new terminal was visited approximately every month during 2001–2005 for a total of 36 site visits. The terminal locations are widely scattered across the United States (see Figure A-1, Supporting Information), and due to relatively strict union work rules, the occupational groups within our cohort are well defined and homogeneous. The freight dock operates 24 hr/day,

7 days a week, and the major onsite jobs at trucking terminals are dockworkers and mechanics, along with office workers and hostlers (onsite drivers). The truck transport operations are separated into drivers making local (P&D) and long distance (LH) trips. This paper exclusively models onsite worker exposures in the shop (mechanics) and dock (dockworkers), which represent the major onsite worker groups exposed to combustion particles in this occupational cohort.

A visual representation of the trucking terminal work locations is provided in Figure A-2, Supporting Information. At the center of the trucking terminal is the loading dock, which is an elongated warehouse building where freight is moved from one trailer to another. The trailers are backed up to a raised platform on the dock with a series of doors along each side. A large terminal may have hundreds of doors and have a dock that is hundreds of meters long. Freight is moved between trailers by dockworkers driving small liquid propane powered forklifts within the semi-enclosed dock area. Truck tractors and trailers that have been damaged or need to be serviced are taken to an onsite shop to be repaired by a mechanic, while those terminals without shops contract out their repair and maintenance needs. The fenced in area of the terminal property outside the dock and shop is known as the “yard”.

A worker’s exposure was hypothesized to be defined by three components, which each have a set of factors associated with them: personal factors (work area, cigarette smoking), work area factors (terminal size, numbers of mechanics and pickup and delivery drivers, job, ventilation, yard background), and yard background factors (relative humidity, temperature, wind speed, distance to highways, land use, region of US). Personal exposures were monitored through the use of particle samplers placed in a special vest, which ran for the entire duration of the work shift (approximately 8–12 hours); very few workers were sampled more than once. Work area concentrations were monitored with stationary box samplers in the indoor work locations, and were collected at consecutive twelve-hour intervals that generally overlapped with the personal samples. Background measurements were made in an upwind location in the yard for each terminal sampled, and were also collected in consecutive twelve-hour intervals. For modeling purposes, personal measurements were matched to the work area and yard measurements that most closely overlap the time period of a given sample.

Measurement Methods

Airborne particles in outdoor and occupational settings are a complex mixture and the components vary as a function of the sources. Three aspects of fine particle matter were assessed that might together serve as indicators of exposure to diesel exhaust and other combustion products: mass of particles less than 2.5 microns in diameter (PM_{2.5}), and elemental carbon (EC) and organic compounds (OC) in particles less than 1 micron in diameter (PM₁). EC was chosen as our primary marker of diesel exposure in this study, but its relationship with mass of diesel emissions varies with operating conditions.

We also noted an artifact associated with OC data obtained from area and personal samples collected where there was cigarette smoking. It is known that quartz fiber filters will adsorb polar hydrocarbon vapors from the air stream and cigarette smoke has a high concentration of polar vapors (13,14). In those settings, the OC content frequently exceeded the PM_{2.5}, and we have not used this data in OC exposure models.

The particle collectors, their pumps, and a real time monitor for temperature and humidity (HOBO, Onset Computer Corp, Bourne, MA) were all mounted in a box housing (or jacket worn by employee volunteers) connected to an external battery. PM_{2.5} was collected on a 37 mm Teflon filter, 0.2 μm diameter pore size, after passing through a precision machined, cyclone separator (GK2.05 SH (KTL), BGI, Inc., Waltham, MA) to remove particles greater than 2.5 μm aerodynamic diameter. Mass collected on the filter was determined by gravimetric

analysis using an analytical balance (Mettler Micro-Gravimetric No. M5; Mettler Instruments Corporation, Hightstown, NJ). The filters were weighted after humidity equilibrium (>48 h) in a chamber. At the end of sampling, the filter was taken back to the laboratory and reweighed (after humidity equilibrium, >48 h) to determine weight gain. EC and OC were determined by the NIOSH 5040 method (15). PM₁ was collected on a 22 mm Quartz tissue filter, preceded by a precision machined cyclone separator (SCC1.062 Triplex, BGI, Inc., Waltham, MA) to remove particles greater than 1.0 μm aerodynamic diameter.

Finally, a small recording anemometer and wind vane were set up next to the yard sampler, and we attempted to place the sampler upwind of the terminal within about 45° of the prevailing wind. As expected, the wind direction was rarely stable, and in several cases, it was necessary to move the sampler to another location when a major weather system moved through the area. However, it was not feasible to move the upwind sampler during frequent minor wind shifts, $\pm 45^\circ$.

Data Description

The primary response variable for this analysis is personal exposure to EC as measured by the sampling jackets worn by employees on the dock and in the shop during their work shifts. OC and PM_{2.5} were also modeled. There are a total of 689 personal observations, and the group arithmetic and geometric means and standard deviations are listed in Table 1. For the statistical model, the total number of observations were reduced somewhat by missing data on smoking status, terminal characteristics, and matching session data for work area, yard, and weather measurements. However, this subset of the data (n=547) did not vary significantly from the complete dataset, and was therefore determined to be a representative sample. We cross-validated our background exposure measurements with nearby monitoring data from the EPA Air Quality System (see Discussion A-1, Table A-1, Supporting Information).

Detailed data on terminal operations during our visit was collected in order to accurately account for the impact of terminal activities on measured work area particle concentrations. This information included 16 different variables, such as the number of employees by job, equipment data, and building dimensions (see Table A-2, Supporting Information, for all covariate summary statistics). There were too many highly correlated variables to use them directly, so a principal component analysis (PCA) was performed to generate potential covariates for the statistical model (see Discussion A-2, Table A-3, Supporting Information).

Work location (shop or dock) and indoor ventilation rates were also identified as important predictors of particle exposure. Particles in these locations came from sources within and outside the work areas. For both the shop and dock, emissions from activities in the yard and local air pollution may enter the work areas through open doors. Both areas also have internal emission sources: in the dock area groups of 3–10 forklifts (each with limited emissions) move freight in and out of trailers; and in the shop, tractors are driven in and out (brief intense emissions), but rarely operated indoors during repairs. All of the shops had tight fitting doors, which were closed during cold weather. The freight docks usually did not have closing doors, but 60–95% of the doors will be blocked by trailers. Particle concentrations in shops have been shown to be significantly higher than those in the dock due to more emissions in a smaller space (16,17,18). Also, cold temperatures can impact indoor dock and shop particle exposures thru door closures (or the blockage of openings with trailers), resulting in indoor particle concentrations several-fold higher, with a particularly strong effect in the shop (17,18). To control for these effects, an indicator variable designating work location, and an interaction variable of work location and outdoor temperature, are included in the model to identify the strength of the winter closure effect and to test for differences across the two work locations.

In order to predict background particle levels, data on weather, as well as regional and location-specific characteristics were collected to describe the area immediately surrounding each of the study locations. Although the true effects of weather variables are difficult to identify in a regression modeling context and can vary widely by season and location (19), site-specific weather data are included in the model to control for their effects. Weather data on wind speed, relative humidity, precipitation, temperature, and barometric pressure were observed from the closest monitoring station matched to the specific sampling time periods using an online source (20). The weather variables from this online source correlated very highly with a limited set of weather observations made by our onsite weather station.

A series of background characteristics were constructed around the point of the geocoded terminal address using GIS software (ArcGIS 9.0, ESRI, Redlands, CA, USA), including industrial land use categorizations and distance to a major road. In the first, raster images from the 1992 National Land Cover Data (NLCD) of the U.S. Geological Survey were used to generate a variable for the percent of industrial, commercial, and transportation land uses in the immediate vicinity of the terminal (1 kilometer buffer). Another variable was generated to indicate the distance to a major road (any type), but it had little variability across terminals because all are located at major road intersections. However, a more interesting and variable feature was the terminal's proximity to heavily trafficked interstate highways. Studies suggest a rapid exponential decay of particles starting within 100 meters of the traffic sources, leading to concentrations statistically indistinguishable from background levels at a distance of 300–500 meters depending on wind patterns (21,22,23,24,25). Although there are currently an insufficient number of observations in our dataset to detect this exponential decay function, eleven of the 36 terminals are located within 500 meters of an interstate highway and several are in close proximity to multiple interstate highways.

Finally, a categorical variable for Census Bureau Regions is included in the model of to control for regional differences in particle concentrations left unexplained by the earlier covariates. The United States is separated into four regions and nine divisions by the Census Bureau (See Figure A-3, Supporting Information), which also corresponds to the regional categorizations used for the ongoing diesel epidemiology study of lung cancer mortality. Census regions are used (as opposed to divisions) due to lack of variability within each of the nine divisions – relatively few site visits were observed within each division to provide stable estimates.

Statistical Model

All of the statistical analyses were performed using STATA Version 8.2, College Station, TX. The concentration data are approximately lognormal, and have been log-transformed to meet normality assumptions. A structural equation model (SEM) approach was developed to analyze the data, a method that is becoming increasingly popular among environmental epidemiologists as a way of handling high dimensional data (26). SEM methods are useful in understanding causal pathways and identifying the indirect effects of intermediate variables on a primary dependent variable (such as occupational exposure to particles). In our setting, SEMs provide a way to analyze the data that reflects the natural hierarchy present in our sampling scheme, namely background, work area, and personal exposures. This relationship is exemplified in the pathway diagram presented in Figure A-4, Supporting Information.

In particular, the nature of the sampling plan imposes a complex covariance structure on the collected data since the concurrent measurements taken by personal sampling jackets, stationary work area samplers, and external measurements of background conditions are not independent. Different sources contribute simultaneously to the measurements observed at different locations (personal, area, yard) within the terminals during the same time periods. Of particular statistical concern is the correlation among the error terms, as well as the correlation

between the response variables and the error terms. Both of these conditions violate necessary assumptions for ordinary least squares.

Therefore, instead of trying to fit one large model encompassing all covariates simultaneously, we fit three related models. Using the SEM method, we simultaneously predict personal exposures as a function of work related exposure and smoking status; work related exposure as a function of terminal characteristics, indoor ventilation, job location, and background exposure conditions; and background exposure conditions as a function of weather, nearby source pollution, and other regional differences across terminal sites. This multi-layered structure allows us to use the statistical technique known as three stage least squares (3SLS), a common SEM approach in econometrics (27,28). The advantage of this particular method is that it provides coefficient estimates for all of the covariates in the model, along with equation specific R^2 values to interpret each level of exposure data.

Results

The current SEM model (shown below) estimates exposure to EC for person i in job location j at terminal k , as measured by the concentrations collected from sampling jackets worn by the employees during their work shifts. There were very few repeated measurements across individuals for the personal samples, and due to the timing of the measurements (by shift), the correlation in these samples across time was negligible. The data are formatted to match the personal samples by time period (session) with indoor work exposures and outdoor background conditions, as well as job location, terminal characteristics, smoking status, and other covariates for each subject.

$$\log(\text{PersonalEC}_{ijk}) = \beta_{10} + \beta_{11} \log(\text{WorkAreaEC}_{jk}) + \beta_{12}(\text{Smoking}_{ijk}) + \epsilon_{ijk} \quad \text{Equation 1}$$

$$\log(\text{WorkAreaEC}_{jk}) = \beta_{20} + \beta_{21}(\text{Terminal Size}_k) + \beta_{22}(\text{P \& D}_k) + \beta_{23}(\text{Shop}_k) + \beta_{24}(\text{Ventilation}_{jk}) + \beta_{25} \log(\text{YardEC}_k) + \beta_{26}(\text{Job}_{jk}) + \gamma_{ijk} \quad \text{Equation 2}$$

$$\log(\text{YardOC}) = \beta_{30} + \beta_{31}(\text{Relative Humidity}) + \beta_{32}(\text{Temperature}) + \beta_{33}(\text{Windspeed}) + \beta_{34}(\text{Interstate}) + \beta_{35}(\text{Industrial}) + \beta_{3(6-9)}(4 \text{ Regional Dummies}) + \eta_{ijk} \quad \text{Equation 3}$$

where ϵ_{ijk} , γ_{ijk} , and η_{ijk} are i.i.d. and normally distributed correlated error terms.

Personal particle concentrations are predicted by matching work area exposures and smoking status. The work area exposures are predicted by terminal-specific characteristics [terminal size (yard dimensions), number of local (P&D) drivers and number of mechanics], ventilation (interaction term of job location and outdoor temperature), matching background exposures observed in the yard, and the job identification [dockworker (0) or mechanic(1)]. Finally, background exposures observed in the yard are predicted by weather (relative humidity, temperature, and wind speed), proximity to a major road (0=less than 500m to interstate, 1 otherwise), industrial land use characteristics (% of industrial, commercial, and transportation land uses within a 1km radius), and regional location within the United States.

The logarithmic transformation of the concentration variables and the subsequent multiplicative regression model requires careful interpretation of the coefficients. For this reason, each of the results is discussed separately in terms of the estimated effect of a one standard deviation increase on personal exposure to particles for dockworkers and mechanics. As the primary marker of diesel exposure in this study, the results for EC (Table 2) are presented in detail, followed by additional comments on the model results for OC (Table A-4, Supporting Information) and PM2.5 (Table A-5, Supporting Information).

Personal Exposure

$(\log(\text{PersonalEC}_{ijk}) = \beta_{10} + \beta_{11}\log(\text{WorkAreaEC}_{jk}) + \beta_{12}(\text{Smoking}_{ijk}) + \varepsilon_{ijk})$ As shown in Table 2, personal exposure to EC is significantly predicted by work area concentrations and smoking status ($R^2=0.64$). Based on the coefficient and summary statistics for work area EC, a one standard deviation increase above average work area levels will lead to an increase in estimated personal exposures of 32.7% in the shop and 79% in the dock. The predicted value of EC exposures for smokers to non-smokers increases more modestly by 19% when work area exposures are held constant.

Area Exposure

$(\log(\text{WorkAreaEC}_{jk}) = \beta_{20} + \beta_{21}(\text{Terminal Size}_k) + \beta_{22}(\text{P\&D}_k) + \beta_{23}(\text{Shop}_k) + \beta_{24}(\text{Ventilation}_{jk}) + \beta_{25}\log(\text{YardEC}_k) + \beta_{26}(\text{Job}_{jk}) + \gamma_{ijk})$ Work area exposure to EC is predicted by terminal characteristics, work location (dock or shop), indoor ventilation, and background concentrations observed in the yard ($R^2=0.64$).

The choice of the three variables representing terminal characteristics (terminal size, P&D, and Shop) were the result of a principal components analysis (PCA) that used the larger set of terminal characteristics data to define three factors that could more efficiently describe the variability of the dataset (Discussion A-2, Table A-3, Supporting Information). These PCA factors were used both as covariates (Table A-6, Supporting Information), as well as to identify important individual variables for inclusion in the model. The PCA results identified terminal size and equipment, P&D activities, and shop activities, as the primary factors driving the complete set of terminal characteristics, and yard size, number of P&D drivers, and number of mechanics were chosen as representative of these factors and incorporated into the current equation describing work area particle exposures. The results were similar for both the PCA factors and three specific terminal characteristics variables. A one standard deviation rise in yard size increases work area EC concentrations by 12.6%, while the effect of P&D drivers is much smaller at 3.7% (number of mechanics was not statistically significant).

Another factor in the determination of work area concentrations are the ventilation rates or the degree of closure of the buildings where the source activity is occurring, as measured by an interaction term between job location and outdoor temperature levels. This covariate for the winter closure effect significantly predicts that work area EC is 57.1% lower in the warmer versus colder outdoor temperatures in the shop, with a negligible effect on the dock.

The indoor work area particle concentrations are also affected by background levels in and around the terminal locations, as captured by the yard monitors placed at the perimeter of the properties. Based on the model, a one standard deviation rise in background concentrations detected in the yard significantly increases the predicted work area EC levels by 78.8%. Also, the large differential between shop and dock exposures are in line with expectations, and the results show that area exposures are over 700% higher in the shop than the dock.

Background Exposure

$(\log(\text{YardOC}) = \beta_{30} + \beta_{31}(\text{Relative Humidity}) + \beta_{32}(\text{Temperature}) + \beta_{33}(\text{Windspeed}) + \beta_{34}(\text{Interstate}) + \beta_{35}(\text{Industrial}) + \beta_{3(6-9)}(4 \text{ Regional Dummies}) + \eta_{ijk})$ Background particle levels at the terminal locations were observed in the yard, which is located away from and upwind of regular work activities. These background concentrations were consistently less than the work area and personal levels. They are significantly predicted by a number of weather conditions, as well as location-specific factors such as the distance to the road and the percent of industrial activity nearby ($R^2=0.51$).

The percent of land characterized as industrial or commercial within a one kilometer radius is significant in the model, where a one standard deviation increase above the mean is followed by a 28.1% increase in the amount of EC observed in the terminal yard. A dichotomous variable representing a cutoff distance of 500 meters from an interstate is a significant predictor of background EC levels, and yard concentrations are 35.1% higher at terminals located closer to an interstate.

Relative humidity, temperature, and wind speed are included in the final model to control for weather effects on pollution dispersal and dilution, all of which have a negative impact on predicted background exposure levels. In particular, a one standard deviation rise in relative humidity decreases predicted background EC by 7.9%, while the same effect represents a decline of 9.2% for temperature, and 45.5% for wind speed. These weather relationships are likely to be strongly effected by the unique composition of EC source activity in our work environments, and for that reason differ from previous studies of small particles at roadways (29). Precipitation was excluded from the model due to lack of variability – 80% of sampling sessions had no precipitation and over 90% experienced less than a tenth of an inch.

Finally, the Census designations show significant differences in background EC levels across regions of the US. The constant term is excluded from the background equation because all four regional dummies are represented (p-values test the hypothesis stated as “the constant for region x is zero”), with the result that each regional coefficient is a multiplicative constant elevating background exposures to various degrees above the hypothetical zero scenario. The biggest regional impact on background EC exposure levels at trucking terminals is the South (2.94), followed by the Northeast (2.32), the West (1.97) and finally the Midwest (1.67).

Organic Carbon and PM2.5

Although the EC component of PM2.5 is used as the primary measure of combustion particle exposure at the terminal locations, PM2.5 and the OC component were also evaluated. The concentrations for these two measures were analyzed using the same type of model developed to predict EC and results are provided in Supporting Information, Tables A-4 and A-5. The SEM outcomes are similar for most of the predictors, although the coefficient estimates changed somewhat. The effect of temperature on OC was opposite that for EC, likely reflecting the increased rate of chemical reactions for organic vapors in the atmosphere, which lead to more formation of condensation aerosol (30,31). Also, the smoking effect is much higher for OC and PM2.5, where smoker exposures are 70% and 72% higher than nonsmoker exposures, respectively (compared to 19% for EC). This is consistent with the large OC content in cigarette smoke, whereas EC is only a small fraction of cigarette smoke (18).

Another major difference between the particle exposure models is the percent of variability explained by each equation. In particular, the equations for personal and work area exposures explain more of the variability in these exposure measurements (R^2 equals 0.64 for both), with less variability explained in the background equation ($R^2=0.51$). The opposite is true for OC and PM2.5, with a large percent of variability left unexplained by the personal and work area equations (Personal R^2 : OC=0.32, PM2.5=0.43; Area R^2 : OC=0.25, PM2.5=0.27), with little variability left unexplained in the background equation (Background R^2 : OC=0.92, PM2.5=0.91). In other words, personal and work area exposure to EC is more highly predicted by on-site work related exposures to vehicle emissions, while it is more difficult to predict personal exposure to OC and PM2.5 based on the same modeling approach. However, the variability in background OC and PM2.5 is almost completely described by the model, in contrast to the identical background model for EC. Although the background R^2 values for the SEMs are very high for OC and PM2.5, it would be useful to further refine the model beyond the regional dummy variables in order describe individual characteristics within regions that drive these differences.

Discussion

The statistical modeling approach presented here provides a new alternative to modeling environmental and occupational health data. The complicated three-tiered sampling scheme – personal measurements from individuals, indoor measurements in work areas, and background measurements in a nearby outdoor location – requires a statistical approach that can handle high dimensional data and complex covariance structures. This type of study design of multiple and simultaneous within site measurements at a given study location (indoor and outdoor) is common in environmental health research, and the SEM approach presented in this paper should provide a useful alternative to exposure modeling in these settings.

The modeling results suggest a number of interesting trends. Overall, personal exposures are found to depend primarily on the work area concentrations of the individual monitored, and to a lesser extent their smoking status. They are also indirectly determined by terminal characteristics and activities, ventilation in the work areas, job location, and the background particle conditions. These background conditions will impact observed personal concentrations via increasing or decreasing work area exposure levels, and can be accounted for in large part by characteristics specific to local areas such as nearby pollution sources and weather conditions.

The model works quite well in explaining extreme values that are still within the relevant range of the collected exposure data. In the personal EC regression equation, the largest and smallest predicted values from the model are within the range of observed values at the high and low ends of the distribution. The same is the case for the other equations, including those for OC and PM_{2.5}. These results provide further evidence in favor of the SEM approach, as well as support the use of lognormal transformations of the exposure variables.

If we instead assumed limited correlation across sampling locations (personal, work area, and background) and no directionality in the effect of covariates on exposure (Figure A-4, Supporting Information), these different sampling components could be broken down into their individual effects and a standard additive linear model applied to predict personal exposures. An examination of the differences between these two approaches provides much support for the SEM method, and the results from ordinary least squares (OLS) estimation of EC personal concentrations (including all covariates) are provided in Supporting Information, Table A-7. Although this approach explains a high percentage of variability among personal EC exposures ($R^2=0.71$), all of the coefficients are smaller (with the exception of smoking), dampening their estimated impact on personal exposure levels predicted by the SEM approach. This underestimation of effect sizes would be expected given that the OLS approach does not account for either the correlation among exposure measurements or the directionality of the covariate effects on the individual equations (i.e., weather and background conditions impact personal exposure indirectly thru elevating work area particle concentrations). Furthermore, in the OLS model, temperature, relative humidity, distance to road, and regional differences are no longer significant predictors of exposure, while the coefficient for the number of P&D drivers changes sign. These types of changes in regression output are common when there is a high degree of correlation among model covariates, with some of the covariates endogenous, or predicted within the system. The use of the OLS approach in this scenario would under predict personal exposure (with important implications for the epidemiologic model relying on these estimates), the results of which lend strong support to the use of SEM methods over OLS to estimate exposures in this type of setting. We are currently working to address the theoretical and methodological implications of SEMs in exposure assessment modeling, and plan to further test this approach in settings such as the railroad industry.

An important implication of the model concerns the size of the cigarette smoking effect, since we will have limited data on the smoking habits of each individual in the cohort for the epidemiologic study. The results from the EC analysis suggest that both work area concentrations and smoking status are significant predictors of personal exposure to particles. However, although the coefficient for smoking is statistically significant, its impact on personal exposure to EC is small by comparison with the impact of work area EC levels. Furthermore, excluding smoking status from the model does not significantly change the predictive value of the personal EC equation – the R^2 declines slightly to 0.62 (from 0.64) and all other variables remain significant predictors. This is not the case for OC or PM_{2.5}, where the smoking effect explains a much higher percentage of the variability in personal exposure to these particles. Therefore, as long as EC remains the primary marker of diesel and other combustion particle exposures in this study, information on smoking status is not pivotal to obtaining accurate estimates of historical exposure levels.

Another interesting result of the model concerns the differences in predictability across exposure equations and particle components. A high percentage of variability in personal EC exposure was predicted by individual and work area characteristics (personal $R^2=0.64$; work area $R^2=0.64$), while less of the variability in PM_{2.5} and OC exposures were explained by work related exposures compared to background exposures (work area R^2 : OC=0.25, PM_{2.5}=0.27; yard background R^2 : OC=0.93, PM_{2.5}=0.91). These results would appear to reinforce the choice of EC as our primary marker of on-site personal exposure to diesel and vehicular particles, since the equations related to diesel exposure in the work environment were more important in explaining EC than either OC or PM_{2.5}. In other words, the model suggests that EC is more highly predicted by the work area environments at each trucking terminal, while PM_{2.5} and OC are better explained by background conditions at each location. Since PM_{2.5} is a predictor of cancer risk from urban air pollution, this difference between PM_{2.5} and EC will also allow us to determine the relative contributions to cancer risk for PM_{2.5} and EC, which has important implications for prevention strategies. Note that the high OC levels relative to EC suggests that idling and lightly loaded diesels and perhaps some car emissions were the predominant sources of EC in the terminal work environments, which is also supported by source apportionment data that is being reported elsewhere.

The current model of exposure will set the stage for developing a model for the historical extrapolation of exposures for the epidemiologic study, linking cumulative lifetime exposures with lung cancer outcomes in the cohort of trucking industry employees. However, knowledge of the current work environment will need to be supplemented with historical data on policy and regulatory changes in the industry, as well as economic fluctuations in business activity levels that impact exposure levels across time. For example, since the late 1980s when the last large exposure assessment of the industry occurred (18), the EPA has strengthened heavy-duty diesel truck engine standards multiple times, from 0.60 g/bhp-hr in 1990 to 0.10 g/bhp-hr in 1998. These changes have certainly impacted exposure levels at these locations and a thorough review of their effects is essential to an accurate historical extrapolation of particle exposures. We are also working on establishing a link between macroeconomic trends and particle exposures at these locations that will aid in estimating historical exposure levels. The trucking industry is highly susceptible to economy-wide fluctuations in business activity, and the current model identifies a significant and positive association between terminal-specific activity levels (terminal size and number of drivers) and particle exposure. We will use economic trend data to estimate historical exposure levels to particles, with the assumption that economic recessions and booms will have a direct impact on activity levels at the terminals (trucks coming in and out, forklift activity, maintenance activity, nearby traffic patterns, etc.), and a subsequent indirect effect on personal exposure to combustion particles.

Acknowledgements

A special thanks to Drew Blicharz, Paul Reaser, Kevin Lane, and Jonathan Natkin for excellent field work and technical data support. We would also like to acknowledge the contribution of the other members of the Trucking Industry Particle Study: Drs. Douglas W. Dockery and Frank E. Speizer. This work was supported by NIH/NCI R01 CA90792 and HEI 4705-RFA03-1/04-1.

Literature Cited

1. Schenker MB. Diesel exhaust - an occupational carcinogen? *Journal of Occupational Medicine* 1980;22:41–46. [PubMed: 6153403]
2. Bhatia R, Lopipero P, Smith AH. Diesel exhaust exposure and lung cancer. *Epidemiology* 1998;9:84–91. [PubMed: 9430274]
3. Lipsett M, Campleman S. Occupational exposure to diesel exhaust and lung cancer: a meta-analysis. *Am J Public Health* 1999;89:1009–1017. [PubMed: 10394308]
4. Stayner L, Bailer AJ, Smith R, Gilbert S, Rice F, Kuempel E. Sources of uncertainty in dose-response modeling of epidemiological data for cancer risk assessment. *Ann N Y Acad Sci* 1999;895:212–22. [PubMed: 10676419]
5. IARC. Diesel and gasoline engine exhausts and some nitroarenes. Lyon, France: International Agency for Research on Cancer; 1989. Online source: <http://www-cie.iarc.fr/htdocs/monographs/vol46/46-01.htm>
6. HEI. Diesel exhaust: a critical analysis of emissions, exposure and health effects. Cambridge, MA: Health Effects Institute; 1995. Online source: <http://www.healtheffects.org/Pubs/diesum.htm>
7. WHO. Diesel fuel and exhaust emissions. Geneva, Switzerland: World Health Organization; 1996.
8. U.S. EPA. Health Assessment Document for Diesel Engine Exhaust. Washington, D.C.: U.S. Environmental Protection Agency, 2002; Online source: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=29060>
9. Diesel Epidemiology Working Group. Research Directions to Improve Estimates of Human Exposure and Risk from Diesel Exhaust. Boston, MA: Health Effects Institute; 2002. Online source: <http://www.healtheffects.org/Pubs/DieselSpecialReport02.pdf>
10. Garshick, E.; Smith, T.J.; Laden, F. Diesel Epidemiology Working Group. Research Directions for Improve Estimates of Human Exposure and Risk from Diesel Exhaust. Boston, MA: Health Effects Institute; 2002. Quantitative Assessment of Lung Cancer Risk from Diesel Exhaust Exposure in the US Trucking Industry: A Feasibility Study. Online source: <http://www.healtheffects.org/Pubs/DieselSpecialReport02.pdf>
11. Schauer JJ. Evaluation of elemental carbon as a marker for diesel particulate matter. *J Expo Anal Environ Epidemiol* 2003;13(6):443–53. [PubMed: 14603345]
12. NIOSH. Carcinogenic effects of exposure to diesel exhaust. U.S. Department of Health and Human Services; 1988. p. 30
13. Olson DA, Norris GA. Sampling artifacts in measurement of elemental and organic carbon: Low-volume sampling in indoor and outdoor environments. *Atmospheric Environment* 2005;39(30):5437–5445.
14. Turpin BJ, Huntzicker JJ, Hering SV. Investigation of organic aerosol sampling artifacts in the Los Angeles basin. *Atmospheric Environment* 1994;28(19):3061–3071.
15. NIOSH. Elemental Carbon (Diesel Particulate) 5040. In: Cassinelli, ME.; O'Connor, PF., editors. NIOSH Manual of Analytical Methods. 4. NIOSH; 1998.
16. Groves J, Cain JR. A Survey of Exposure to Diesel Engine Exhaust Emissions in the Workplace. *Annals of Occupational Hygiene* 2000;44:435–447. [PubMed: 10963708]
17. Sauvain JJ, Vu Duc T, Guillemin M. Exposure to carcinogenic polycyclic aromatic compounds and health risk assessment for diesel-exhaust exposed workers. *Int Arch Occup Environ Health* 2003;76:443–455. [PubMed: 12750904]
18. Zaubst DD, Clapp DE, Blade LM, Marlow DA, Steenland K, Hornung RW, et al. Quantitative Determination of Trucking Industry Workers' Exposures to Diesel Exhaust Particles. *American Industrial Hygiene Association Journal* 1991;52:529–541. [PubMed: 1723577]

19. Chu SH, Paisie JW, Jang BW. PM data analysis - a comparison of two urban areas: Fresno and Atlanta. *Atmos Environ* 2004;38:3155–3164.
20. Weather Underground, Inc. 2005. <http://www.weatherunderground.com/>
21. Hitchins J, Morawska L, Wolff R, Gilbert D. Concentrations of submicrometer particles from vehicle emissions near a major road. *Atmos Environ* 2000;34:51–69.
22. Morawska L, Thomas S, Gilbert D, Greeaway C, Rijnders E. A study of the horizontal and vertical profile of submicrometer particles in relation to a busy road. *Atmos Environ* 1999;33:1261–1274.
23. Wrobel A, Rokita E, Maenhaut W. Transport of traffic-related aerosols in urban areas. *Sci Total Environ* 2000;257:199–211. [PubMed: 10989929]
24. Zhu Y, Hinds WC, 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.
25. Zhu Y, Hinds WC, Kim S, Sioutas C. Concentration and size distribution of ultrafine particles near a major highway. *J Air Waste Manag Assoc* 2002;52:1032–1042. [PubMed: 12269664]
26. Sanchez BN, Budtz-Joergensen E, Ryan LM, Hu H. Structural Equation Models: A Review with Applications to Environmental Epidemiology. *Journal of the American Statistical Association* 2005;100:1443–1455.
27. Zellner A, Theil H. Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations. *Econometrica* 1962;30:54–78.
28. Goldberger A. Structural Equation Methods in the Social Science. *Econometrica* 1972;40:979–1001.
29. Sioutas C, Delfino RJ, Singh M. Exposure assessment for atmospheric ultrafine particles (UFPs) and implications in epidemiologic research. *Environ Health Perspect* 2005;113(8):947–55. [PubMed: 16079062]
30. Schauer JJ, Mader BT, Deminter JT, Heidemann G, Bae MS, Seinfeld JH, et al. ACE-Asia Intercomparison of a Thermal-Optical Method for the Determination of Particle-Phase Organic and Elemental Carbon. *Environ Sci Technol* 2003;37:993–1001. [PubMed: 12666931]
31. Lim HJ, Turpin BJ, Russell LM, Bates TS. Organic and Elemental Carbon Measurements during ACE-Asia Suggest a Longer Atmospheric Lifetime for Elemental Carbon. *Environ Sci Technol* 2003;37:3055–3061. [PubMed: 12901650]

Table 1

Summary Statistics of Particles ($\mu\text{g}/\text{m}^3$) by Job Title and Location

	EC			OC			PM2.5			GM	SD	GSD
	Obs	Avg	SD	GM	SD	GSD	Obs	Avg	SD			
Dockworker												
Non-smoker	398	1.0	0.8	0.8	2.1	2.1	398	14.7	6.8	13.8	13.6	1.7
Smoker	124	1.4	1.0	1.0	1.9	1.9	124	29.1	22.6	24.0	43.4	2.0
Dock Area	483	1.0	0.8	0.7	2.8	2.8	483	8.1	4.1	7.4	20.7	1.9
Mechanic												
Non-smoker	107	4.4	9.1	2.0	3.8	3.8	107	19.4	13.8	16.9	38.8	2.0
Smoker	32	3.5	2.6	2.4	2.3	2.3	32	29.6	16.3	24.4	41.1	1.9
Shop Area	214	3.0	4.0	1.5	3.5	3.5	214	13.0	9.6	10.4	27.6	2.4
Yard	348	0.8	1.0	0.5	3.3	3.3	348	5.3	3.0	4.6	8.7	2.4

* The geometric mean (GM) is an estimator of the median, and the geometric standard deviation (GSD) is a multiplicative factor, e.g. one standard deviation above the GM is GM*GSD

Table 2

Regression Results for Elemental Carbon¹

Equation 1: $\log(\text{PersonalEC}) = \beta_{10} + \beta_{11}\log(\text{WorkAreaEC}) + \beta_{12}(\text{Smoking}) + \epsilon_{ijk}$	
Work Area	0.99 (0.04)**
Smoking	0.17 (0.06)**
Constant	0.05 (0.03)
Equation R²	0.64
Equation 2: $\log(\text{WorkAreaEC}) = \beta_{20} + \beta_{21}(\text{Terminal Size}) + \beta_{22}(\text{P\&D}) + \beta_{23}(\text{Shop}) + \beta_{24}(\text{Ventilation}) + \beta_{25}\log(\text{YardEC}) + \beta_{26}(\text{Job}) + \gamma_{ijk}$	
Terminal Size	0.01 (0.002)**
P&D Drivers	0.002 (0.001)*
Mechanics	0.002 (0.001)**
Ventilation	-0.09 (0.01)**
Yard Background	0.71 (0.05)**
Job	2.11 (0.09)**
Constant	-0.21 (0.08)**
Equation R²	0.64
Equation 3: $\log(\text{YardEC}) = \beta_{30} + \beta_{31}(\text{Relative Humidity}) + \beta_{32}(\text{Temperature}) + \beta_{33}(\text{Windspeed}) + \beta_{34}(\text{Interstate}) + \beta_{35}(\text{Industrial}) + \beta_{3(6-9)}(4 \text{ Regional Dummies}) + \eta_{ijk}$	
Relative Humidity	-0.004 (0.002)*
Temperature	-0.01 (0.004)**
Windspeed	-0.11 (0.01)**
Interstate Distance	-0.30 (0.07)**
Industrial Land Uses	0.01 (0.001)**
Region 1 – Midwest	0.51 (0.20)**
Region 2 – Northeast	0.84 (0.19)**
Region 3 – South	1.08 (0.20)**
Region 4 – West	0.68 (0.18)**
Constant	N/A
Equation R²	0.51

Coefficients listed with standard errors in parentheses;

** significant at 1% level,

* significant at 5% level

¹ All particulate concentrations are log transformations measured in $\mu\text{g}/\text{m}^3$