**Materials and Methods**

*Satellite AOD Data.*

Daily spectral AOD data was obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Aqua satellite for the year 2003. Further details about MODIS satellite aerosol data retrieval and validation have been published previously.(Remer et al. 2005; Levy et al. 2007. A new algorithm MAIAC (Lyapustin et al. 2011)

Lyapustin, Alexei, et al. "Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look‐up tables." Journal of Geophysical Research: Atmospheres (1984–2012) 116.D3 (2011).

has been developed to process MODIS data. MAIAC retrieves aerosol parameters over land at 1 km resolution simultaneously with parameters of a surface bidirectional reflectance distribution function (BRDF). This is accomplished by using the time series of MODIS measurements and simultaneous processing of groups of pixels. The MAIAC algorithm ensures that the number of measurements exceeds the number of unknowns, a necessary condition for solving an inverse problem without empirical assumptions typically used by current operational algorithms. The MODIS time series accumulation also provides multi-angle coverage for every surface grid cell, which is required for the BRDF retrievals from MODIS data. The improved accuracy of MAIAC results from using the explicit surface characterization method in contrast to the empirical surface parameterization approach, which is utilized in the MYD04 algorithm. Further, MAIAC incorporates a cloud mask (CM) algorithm based on spatio-temporal analysis which augments traditional pixel-level cloud detection techniques. Daily values of AOD were assigned to the grid cell where the AOD retrieval centroid was located.

One feature of the AOD data is that some of the grid-specific AOD values are missing on some days due to cloud cover or snow cover(Kloog et al. 2011). Thus, the spatial coverage of the AOD data varies considerably by day.

*Air Pollution Monitors.*

Data for daily PM2.5 mass concentrations across New England (see Figure 1) for the year 2003 were obtained from the U.S. Environmental Protection Agency (EPA) Air Quality System (AQS) database as well as the IMPROVE (Interagency Monitoring of Protected Visual Environments) network. IMPROVE monitor sites are located in national parks and wilderness areas while EPA monitoring sites are located across New England including urban areas such as downtown Boston. There were 71 monitors with unique locations operating in New England during the study period.

*Spatial and Temporal Covariates.*

Spatial covariates included major roads, point emissions and area emissions.

Data on the density of major roads was based on A1 roads (hard surface highways including Interstate and U.S. numbered highways, primary State routes, and all controlled access highways) data obtained through the US census 2000 topologically integrated geographic encoding and referencing system. Because the distributions of major roads were highly right-skewed, they were log transformed.

Temporal covariates included wind speed, humidity, visibility, height of the planetary boundary layer. All meteorological variables (temperature, wind speed, humidity, visibility) were obtained through the national climatic data center (NCDC). height of the planetary boundary layer data was obtained from the North American Regional Reanalysis (NARR). Further details on spatial and temporal covariates are given in Kloog et al. 2011 and Kloog et al. 2012.

*Calibration of AOD.*

A description of the method used to calibrate the AOD values to represent PM2.5 concentrations is given in Kloog et al. 2011 and Kloog et al. 2012. Briefly, the relationship between PM2.5 and AOD at the monitoring sites was modeled using a mixed-effects regression model where PM2.5 was the dependent variable and AOD was the main explanatory predictor. The model included spatial covariates for major roads, point emissions and area emissions, and temporal covariates for wind speed, visibility, height of the planetary boundary layer, with interactions between AOD and random intercepts for each day.

Kloog et al. 2011 also includes a third stage of modeling which imputes PM2.5 at the missing AOD locations. In this study, we restricted to only days with ample AOD present to leverage the observed spatial variability in the data and minimize the use of known land-use regression models.

*Simulation setup.*

A simulation study was conducted to assess the performance of kriging and land use regression methods on a realistic representation of an air pollution surface. Separate simulation studies were conducted to consider studies of chronic health effects due to long-term air pollution exposures and acute health effects due to short-term air pollution exposures.

We considered two types of health effect models: a binary health outcome and a continuous health outcome. A linear regression health model was assumed for the continuous health outcome, where the outcome depends linearly on the exposure. For the binary health outcome, a logistic regression health model was assumed, where the outcome depends linearly on the exposure through a logit link function. No other confounding variables were included in the health model. We explored exposure models with four different covariance models (Matern function with different levels of smoothness, indexed by κ). We also contrasted two settings for the number of monitors where *m=100* is the realistic setting (although still higher than the actual number of monitors in this region during the study period), and *m=500* to represent a unrealistic ``best case scenario'' with much more spatial coverage to help highlight the problems due to sample size vs the problems due to model misspecification.

We restricted our simulation studies to the 32 days such that at least 50,000 grid-cells of AOD data were available from the satellite.

*Chronic Effects Simulation.*

To emulate the setting of a health study of the chronic effects of particulate matter, we generated a chronic exposure surface by averaging the calibrated PM2.5 data at each grid-cell over the 32 days of exposure. In this scenario, all subjects' exposures were sampled from this one common exposure surface. Thus, the spatial variability of the surface provided the only variability in the exposures of different subjects.

For each simulation, we generated 500 subjects' exposure and outcome measurements. To assign the exposure, we first generated each subjects' residential location by population density. Population density sampling was approximated by using the geocoded locations of births during 2003 from a previous study.(Kloog et al. 2012) We then assigned the corresponding calibrated PM2.5 value at the subjects' residential location as the exposure. The health outcome was generated to depend on the assigned exposure using the chosen health model type with no confounders. The monitor locations were chosen by a random uniform distribution across the exposure surface, and the corresponding calibrated PM2.5 value at the monitor location was used as the observed exposure. Using the measured exposure at the monitor locations, the kriging or land use model was fit to the data and chronic exposure predictions were generated at the residential locations of the subjects. The predicted exposures were then fit to the health outcomes to estimate the association.

*Acute Effects Simulation.*

We designed our acute effects simulation to mimic the setting of a health study of the short-term effects of particulate matter. Using the 32 days of calibrated PM2.5, we considered the exposure period of interest to be one day of PM2.5 exposure. For each simulation, we generated 1,000 subjects' residential location by randomly sampling the day of the exposure and then sampling the health locations by population density, as in the chronic simulation. Once the date and grid-cell were randomly chosen, we assigned the corresponding calibrated PM2.5 exposure at the grid-cell. The health outcomes were generated to depend on the assigned exposure using the chosen health model type with no confounders. We simulated 1000 subjects per simulation so that there were approximately 30 subjects sampled from each of the 32 days. The monitor locations were chosen by a random uniform distribution across the exposure surface, and the corresponding daily calibrated PM2.5 value at the monitor location was used as the observed exposure for each day. Using the measured exposure at the monitor locations, the kriging or land use model was fit to the data by day and exposure predictions were generated for each day at the residential locations of the subjects. The predicted exposures were then fit to the health outcomes to estimate the association.