Environment International 91 (2016) 1–13

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Multi-pollutant exposure profiles associated with term low birth weight in Los Angeles County

Eric Coker [a,*], Silvia Liverani [b], Jo Kay Ghosh [c], Michael Jerrett [d], Bernardo Beckerman [d], Arthur Li [e], Beate Ritz [c], John Molitor [a]

[a] College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, United States
[b] Department of Mathematics, Brunel University, London, UK
[c] School of Public Health, University of California, Los Angeles, Los Angeles, CA, United States
[d] School of Public Health, University of California, Berkeley, Berkeley, CA, United States
[e] Department of Information Science, City of Hope National Cancer Center, Duarte, CA, United States

ARTICLE INFO

Article history:
Received 19 September 2015
Received in revised form 4 February 2016
Accepted 5 February 2016
Available online 15 February 2016

Keywords:
Air pollution
Bayesian
Clustering
Low birth weight
Pollutant profile
Profile regression

ABSTRACT

Research indicates that multiple outdoor air pollutants and adverse neighborhood conditions are spatially correlated. Yet health risks associated with concurrent exposure to air pollution mixtures and clustered neighborhood factors remain underexplored. Statistical models to assess the health effects from pollutant mixtures remain limited, due to problems of collinearity between pollutants and area-level covariates, and increases in covariate dimensionality. Here we identify pollutant exposure profiles and neighborhood contextual profiles within Los Angeles (LA) County. We then relate these profiles with term low birth weight (TLBW). We used land use regression to estimate NO_2, NO, and $PM_{2.5}$ concentrations averaged over census block groups to generate pollutant exposure profile clusters and census block group-level contextual profile clusters, using a Bayesian profile regression method. Pollutant profile cluster risk estimation was implemented using a multilevel hierarchical model, adjusting for individual-level covariates, contextual profile cluster random effects, and modeling of spatially structured and unstructured residual error. Our analysis found 13 clusters of pollutant exposure profiles. Correlations between study pollutants varied widely across the 13 pollutant clusters. Pollutant clusters with elevated NO_2, NO, and $PM_{2.5}$ concentrations exhibited increased log odds of TLBW, and those with low $PM_{2.5}$, NO_2, and NO concentrations showed lower log odds of TLBW. The spatial patterning of pollutant cluster effects on TLBW, combined with between-pollutant correlations within pollutant clusters, imply that traffic-related primary pollutants influence pollutant cluster TLBW risks. Furthermore, contextual clusters with the greatest log odds of TLBW had more adverse neighborhood socioeconomic, demographic, and housing conditions. Our data indicate that, while the spatial patterning of high-risk multiple pollutant clusters largely overlaps with adverse contextual neighborhood cluster, both contribute to TLBW while controlling for the other.

1. Introduction

Evidence has been accumulating that birth outcomes may be particularly sensitive to air pollution mixtures, specifically components related to traffic sources of airborne particulate matter (PM) (Bell et al., 2010; Laurent et al., 2014; Wilhelm et al., 2011a). While earlier research has linked increased prevalence of term low birth weight (TLBW) with various outdoor air pollutants including NO_2, NO, and $PM_{2.5}$ (Geer, 2014; Ritz and Wilhelm, 2008), most evidence relied on single pollutant modeling of exposures (Ritz and Wilhelm, 2008). A number of studies (Brauer and Tamburic, 2009; Ghosh et al., 2012; Gouveia et al., 2004; Laurent et al., 2014; Le et al., 2012; Morello-Frosch et al., 2010; Wilhelm et al., 2011a, 2011b) investigated exposures to multiple pollutants in relation to birth outcomes; however, these studies are limited in assessing which combination of pollutants are most hazardous or how multipollutant health effects vary spatially. Despite there being no single exposure-measure-framework to holistically address the health effects of multipollutant exposures (Oakes et al., 2014), investigating health effects of profiles of multiple pollutants using clustering techniques has recently shown promise (Gu et al., 2012; Molitor et al., 2014a; Papathomas et al., 2010; Pirani et al., 2015; Qian et al., 2004; Zanobetti et al., 2014).

1.1. Multipollutant exposures

Considerable intra-urban spatial variations in outdoor air pollution concentrations exists, and recent research indicates that between-pollutant correlations and $PM_{2.5}$ composition exhibit highly localized spatial patterns to create complex mixtures (Austin et al., 2013; Austin

* Corresponding author.

http://dx.doi.org/10.1016/j.envint.2016.02.011

et al., 2012; Bell et al., 2011; Geer, 2014; Hasheminassab et al., 2014; Houston et al., 2014; Janhäll et al., 2012; Laurent et al., 2014; Levy et al., 2013a; Molitor et al., 2011; Monn, 2001; Tsai et al., 2015). Spatially correlated air pollution mixtures correspond to localized sources, such as transportation-related emissions (Laurent et al., 2014), local industrial activities (Morello-Frosch et al., 2002; Zhu et al., 2011), or small-area commercial land uses (Morello-Frosch et al., 2002; Morello-Frosch et al., 2001). Factors that combine to determine exposure to spatially correlated pollutants from a particular source are complex and diverse, e.g., traffic-source driven exposures are influenced by traffic volumes and congestion, proximity to traffic, the types of fuel and engines, operating conditions of emitting sources, types of emitting sources, background air pollution levels, local meteorology, chemical reactions between pollutants, and local topographies (Austin et al., 2012; Boehmer et al., 2013; Cho et al., 2009; Greco et al., 2007; Hu et al., 2012; Janhäll et al., 2012; U.S. EPA, 2008; Zhang and Batterman, 2013).

Correlations across different pollutants hinder our ability to assess their individual or combined health effects, since estimates of effects may become unstable when adjusting for co-pollutants using regression techniques (Mauderly et al., 2010). Correlations between $PM_{2.5}$ concentrations and nitrogen oxides (NO_2 and NO [NO_x]) are typically weak to moderate (Ghosh et al., 2013; Laurent et al., 2014; Levy et al., 2013). However, such correlations can vary spatially (Levy et al., 2013; Tsai et al., 2015) based upon whether the particulates represent primary $PM_{2.5}$ (particles emitted directly from the source, e.g. fuel combustion (Fine et al., 2008)) or secondary $PM_{2.5}$ (particles formed in the atmosphere (Fine et al., 2008)). Therefore, since some $PM_{2.5}$ components represent "fresh" traffic emissions (i.e. ultrafine PM and black carbon), they can exhibit high correlations with outdoor concentrations of NO_x, as studies of urban air pollution from Asia, Europe, and North America demonstrate (Brauer et al., 2011; Dionisio et al., 2014; Janhäll et al., 2012; Levy et al., 2013; Tsai et al., 2015; Wang et al., 2014). Furthermore, the spatial variation in between-pollutant correlations also suggests a strong potential for a unique spatial pattern of multipollutant-related health risks, yet research on this question is lacking.

1.2. Contextual factors

Neighborhood-level "contextual" factors may also affect risk of birth outcomes (English et al., 2003; Morello-Frosch and Shenassa, 2006). Contextual effects are non-chemical stressors (Lewis et al., 2011) that arise when grouped neighborhood-level factors – such as socio-demographic or built environment (e.g. housing as one component of the built environment) factors – influence health outcomes across populations (Sheppard et al., 2012). Data from Southern California indicates that spatially clustered socioeconomic deprivation and racial segregation correlate with air pollution exposures, including pollutant mixtures (Molitor et al., 2011; Morello-Frosch et al., 2011; Morello-Frosch et al., 2002; Morello-Frosch and Shenassa, 2006; Su et al., 2012). In addition, compared to newer homes, older homes are shown to have higher indoor air pollution levels within LA County (Spengler et al., 1994). Older housing stock may further correlate with higher poverty, residential racial segregation patterns, substandard housing conditions, and a lack of compliance with building or sanitary codes (Shennassa et al, 2004). Moreover, research from LA County indicates that housing ventilation conditions may be associated with TLBW (Ghosh et al., 2013).

Air pollution and birth outcomes research studies, however, generally do not account for spatial clustering of multiple neighborhood-level vulnerabilities (i.e. race-ethnicity, poverty, and adverse housing conditions) related to exposure. This may confound multipollutant exposure-response relationships (Geer, 2014; Morello-Frosch et al., 2011; Morello-Frosch and Shenassa, 2006; Ponce, 2005; Ritz and Wilhelm, 2008). Given this gap in air pollution and health effects literature, our study includes contextual factors, as clustering covariates, to better control for highly correlated contextual factors known to influence differential exposures across socio-demographic groups and risk of TLBW.

1.3. Study objectives

The primary objectives of our study are to first identify profiles of exposure to multiple different air pollutants (pollutant profiles) for pregnant women within LA County, and secondly to assess whether and which pollutant profiles relate with elevated prevalence of TLBW. Additionally, our approach identifies pollutant profiles most likely related to primary traffic emissions, based on examination of the spatial patterning of pollutant exposure profiles and well established pollutant source emissions relationships.

2. Methods

2.1. Study population and birth outcomes

Electronic birth certificates from the California Department of Public Health provided the data on baby's birth weight and individual-level covariates for LA County births during the years 2000–2006. Individual data from the birth records included maternal characteristics (age, race and ethnicity, education, total number of previous maternal births, and residential address) and information on the infant and birth (date of birth, abnormalities, birth season, gestational age at birth [as determined by self-report of last menstruation], birth weight and baby's sex). The dataset was restricted to singleton births without apparent abnormalities, while births with extreme gestational days (<140 days or >320 days) and births with weight <500 g or >5000 g were excluded from the analysis. Such extreme values are likely attributable to recording errors. These data restrictions provided a sample size of 899,554. Finally, we defined TLBW as full-term (≥259 gestation days) infants with a birth weight < 2500 g, which further restricted the study population to term births, to provide a final sample size of 804,726 to assess the relationship between TLBW with neighborhood-level pollutant profile exposures. Human subjects research was approved through the University of California, Los Angeles' Office of the Human Subjects Protection Program, the California Committee for the Protection of Human Subjects, and the University of Southern California's Office for the Protection of Research Subjects. Geocoding of residential addresses are explained elsewhere (see Goldberg et al., 2008).

2.2. Exposure estimation

Two separate land use regression (LUR) models estimated individual-level exposures for $PM_{2.5}$, NO, and NO_2 (Beckerman et al., 2013a, 2013b; Su et al., 2009). LUR estimates were temporally adjusted to derive the entire pregnancy average exposures. LUR exposure predictions for NO and NO_2 were based on traffic volumes, truck routes, road networks, land use data, satellite-derived vegetation greenness and soil brightness, truck route slope gradients, and air monitoring data. NO_2 and NO data were collected during 2-week time periods in Summer of 2006 and Winter of 2007, from over 200 monitoring locations (Su et al., 2009). The $PM_{2.5}$ exposure estimates came from a LUR model that utilized long-term governmental monitoring data of $PM_{2.5}$ measurements collected between 1998 through 2002 (Beckerman et al., 2013b). A machine learning deletion/substitution technique (Beckerman et al., 2013a) assessed as many as 70 covariates to develop the final $PM_{2.5}$ LUR model, such as land use data (i.e. agricultural, barren, all developed land, high-density development, green space, water, and wetland), long-term traffic counts (1990–2001), and road networks from the year 2000 (Beckerman et al., 2013a; Jerrett et al., 2013).

We adjusted LUR exposure estimates temporally to derive "*seasonalized*" values that correspond to each pregnancy time span. For temporal adjustments, we first used daily air monitoring data from LA County between the years 1999–2006, for all monitors to

calculate an overall daily average for $PM_{2.5}$, NO_2, and NO_x. Pregnancy time period averages were then calculated for each pregnancy from these daily averages. As NO was not directly measured by these monitors, we subtracted the NO_2 pregnancy average from the NO_x pregnancy average to derive NO pregnancy time period estimates. Temporal adjustments of the LUR estimates for NO_2 and NO were achieved using the following equation, which is similar in approach to our earlier work (see (Ghosh et al., 2012)):

$$NO_2\ LUR\ Pregnancy\ Average = NO_2\ LUR \times \frac{monitor\ station\ pregnancy\ average\ NO_2}{2006\ NO_2\ average\ for\ all\ monitor\ stations}$$

$$NO\ LUR\ Pregnancy\ Average = NO\ LUR \times \frac{monitoring\ station\ pregnancy\ average\ NO}{2006\ NO\ average\ for\ all\ monitoring\ stations}.$$

Since the $PM_{2.5}$ LUR estimates represent long-term estimates spanning the study time period and some pregnancies began in 1999, we performed seasonal adjustments with the following equation:

$$PM_{2.5}\ LUR\ Pregnancy\ Average = PM_{2.5}LUR \times \frac{monitor\ station\ pregnancy\ average\ PM_{2.5}}{PM_{2.5}\ average\ for\ all\ monitoring\ stations\ (1999–2006)}.$$

Such temporal adjustment via region-wide monitoring station ratios has been validated for the purposes of estimating pregnancy exposures in birth outcomes studies (Ross et al., 2013).

All of the available data from the temporally adjusted LUR model estimates (N = 899,554) were then averaged over census block groups to develop air pollution exposure profiles at the census block group level. Data aggregation at the census block group-level for individual estimates was performed since we were interested in assessing between neighborhood multipollutant exposure-related TLBW risks. Moreover, implementation of the Bayesian profile regression using individual-level estimates with a dataset as large as ours is not feasible given the current computational limitations of the R PReMiuM package (described below).

2.3. Bayesian profile regression

We developed the profile clusters using a non-parametric dimension reduction technique known as Bayesian profile regression, based on commonly used Dirichlet process mixture model methods (Neal, 2000). Profile regression is set in a Bayesian framework using Markov chain Monte Carlo (MCMC) methods. Bayesian profile regression uses covariate values to observe joint patterns within the covariate data. This approach was used in recent studies (Hastie et al., 2013; Molitor et al., 2014a; Molitor et al., 2011, Molitor et al., 2010; Papathomas et al., 2012), including environmental epidemiology studies (Papathomas et al., 2010; Pirani et al., 2015; Vrijheid et al., 2014). This clustering approach is advantageous because it reduces the dimensionality of the covariate data and allows for examining health risks as they relate to joint patterns of exposure, while avoiding the pitfalls of exposure variables that are highly collinear. This approach is also quite flexible because it does not rely on setting a total number of allowable clusters, as seen with k-means clustering procedures (Austin et al., 2013; Austin et al., 2012; Gu et al., 2012). We implemented the profile regression using the PReMiuM package in R (Liverani et al., 2015). Since our interest is in obtaining clustering that best fits the data for sub-regions within the LA County area, we utilized a feature of the PReMiuM package that excludes the outcome variable from the profile regression model (Liverani et al., 2015). We relied on "hard clustering" (Fang et al., 2011) in the sense that a census block group's final allocation is to a single cluster. Cluster allocation is based on the "*best*" clustering derived from the Bayesian averaging process, rather than probabilistic allocation to several different clusters simultaneously (as in fuzzy [or soft] clustering). Briefly, for each census block group, j, a covariate profile is defined as, $x_j = (x_2, x_2, \ldots, x_p)$, where every covariate, x_p, $p = 1, \ldots, P$, within each profile signifies a level of exposure for covariate p in region j. The primary model for cluster profiles was defined by a multivariate normal mixture model (Jain and Neal, 2004) that further integrates a Dirichlet process prior into the mixing distribution. For greater details on this Bayesian profile regression approach, the reader is referred to other recent works (Hastie et al., 2013; Liverani et al., 2015; Molitor et al., 2014a; Molitor et al., 2011, Molitor et al., 2010; Papathomas et al., 2012; Papathomas et al., 2010).

We performed two separate profile regressions to develop a set of two unique profile clusters to fit in the TLBW risk model. The first clustering procedure developed pollutant-only profile clusters. The second clustering procedure developed contextual-only profile clusters. The co-pollutants for our pollutant-only profile regression included average census block group-level concentrations for NO_2, NO, and $PM_{2.5}$. Furthermore, since our LUR estimates were seasonalized and thus provide temporally resolved estimates of exposure for each pregnancy, we also performed pollutant profile regression across different birth seasons and across different birth years. We present the results of these seasonal and yearly pollutant clusters in the Supplemental materials (Figs. S1–S4), however, in this paper we focus on the overall pollutant profiles as described above since the spatial patterning and between-pollutant correlations were very stable across each of these different seasonal and yearly pollutant profile cluster analyses.

The contextual-only exposure profile regression utilized year 2000 U.S. census data and included census block group-level race-ethnicity (percent non-Hispanic White, percent non-Hispanic Black, and percent Hispanic), median household income, and percent of homes built prior to year 1950. Even though our multivariate risk model (described below) adjusts for individual-level maternal race-ethnicity, we included census block group-level racial/ethnic composition as a clustering contextual covariate under the rationale that area-level racial/ethnic composition may act as a contextual risk factor for TLBW separate from an individual's race-ethnicity (Debbink and Bader, 2011). Similarly, while our multivariate model adjusts for maternal education as a marker of individual-level SES, we included census block group level median household income as a contextual SES variable in the clustering procedure, under the same rationale that area-level SES acts as a contextual risk factor for TLBW independent of individual-level SES (Grady, 2011; Grady, 2006). We included the percentage of homes built before 1950 since disparities in housing quality and other housing characteristics correlated with older housing may act as an important contextual risk factor in TLBW risk (Ghosh et al., 2013; Grady, 2011). Individual mothers were then assigned to both a pollutant cluster and a contextual cluster as determined by which census block group the mother resided in according to their address at time of delivery.

2.4. Multilevel risk model

Our multilevel logistic regression model was set in a Bayesian framework with pollutant profile clusters and contextual profile clusters used as separate random effects variables in the regression equation, along with spatially structured and unstructured independent error terms fit as additional random effects. The model specification is detailed in turn:

$$logit(p_i) = \alpha + V\eta' + \gamma_{k[i]}^{pollutant-cluster} + \gamma_{c[i]}^{contextual-cluster} + S_j + \epsilon_j \qquad (1)$$

where p_i denotes the logit of TLBW ($y_i = 1$) for individual i, $V\eta'$ represents the *individual-level* covariate fixed effects, $\gamma_k^{pollutant-cluster} \sim N(0, \sigma^2_{pollutant-cluster})$, $k = 1, \ldots, 13$ represents the random effects for the pollutant-clusters and $\gamma_c^{contextual-cluster} \sim N(0, \sigma^2_{contextual-cluster})$, $c = 1, \ldots, 14$ represents the contextual cluster random effects. Following Gelman and Hill (2006), we use the notation $k[i]$ to denote the pollutant profile group k to which individual i belongs and $c[i]$ to denote the contextual profile group c to which individual i belongs. Thus, each pollutant random error term represents the variation in TLBW prevalence in the pollutant profile clusters and likewise each contextual random error term represents the variation in TLBW in the contextual clusters. In other words, the cluster random effect can be interpreted

as measuring the change in baseline log odds of TLBW for individual i in cluster k, when all other covariates in the model are set to zero (see Molitor et al., 2010).

Regarding the spatial and independent residual error terms, here S_j and ϵ_j denote spatial and independent residual error terms, respectively, with the restriction $\sum_j S_j = 0$ imposed for indefinability reasons. While the independent error term is defined in the standard way as $\epsilon_j \sim N(0, \sigma^2)$, the spatial error term is defined as,

$$S_{j|k\neq j} \sim N\left(\frac{\sum_{k\neq j} w_{jk} S_k}{\sum_{k\neq j} w_{jk}}, \frac{\tau^2}{\sum_{k\neq j} w_{jk}}\right),$$

where the weights $w_{j,k}$ are elements of the zero-one neighborhood adjacency matrix defined to be equal to one when census tracts i, k are adjacent and zero otherwise. This approach implements the Besag-York-Molly (BYM) model (Besag et al., 1991) and has been successfully employed in a variety of exposure/health association studies (Molitor et al., 2007). Given the large number of records in the dataset, we "pre-clustered" exposure profiles as described in our clustering section and then used the R-INLA (integrated nested Laplace approximations) package to implement the Bayesian multilevel random effects model described in equation (1) above. R-INLA estimates Bayesian posterior marginal distributions (Rue et al., 2014; Rue and Martino, n.d.) without relying on computationally intensive Markov chain Monte Carlo techniques (Gilks et al., 1998).

2.5. Assessing uncertainty in pollutant random effects

As our analysis is in a Bayesian framework, with random effects terms for each pollutant cluster and for each contextual cluster, we assessed the uncertainty with respect to the random effect for each cluster on the baseline log odds of TLBW. For instance, we calculated the posterior probability that a specific profile cluster's posterior distribution of baseline log odds for TLBW ($[\text{logit}[\Pr(y_i = 1)] = \theta_k]$) is above the overall baseline log odds for TLBW (θ) (Papathomas et al., 2010). Said another way, for each cluster we calculate the probability as $P(\theta_k > 0)$, with probability values close to 1 indicative of a high probability for a baseline log odds above zero for each cluster (evidence for adverse effect). Conversely, a posterior probability close to zero is indicative of a low probability for a baseline log odds above zero for each cluster (evidence for no adverse effect). The posterior probabilities for each pollutant cluster and for each contextual cluster were then mapped in ArcGIS V.10.1 (Redlands, CA) to investigate the spatial distribution of these clusters effects on the log odds of TLBW. These kinds of probability effect maps are commonly used in Bayesian modeling of spatial effects of exposure (Bivand et al., 2013; Coker et al., 2015).

2.6. Covariates

Individual-level covariates adjusted for were maternal factors including age at delivery (<20 years, 20–24 years, 25–29 years, 30–34 years, ≥35 years), race-ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian, and Other race), highest education level attained (< 9 years, 9–12 years, 13–15 years, and ≥16 years), parity, along with infant factors such as gestational days, gestational days squared, and infant sex (male/female).

2.7. Characterization of pollutant clusters

In order to infer which clusters are most likely affected by near highway traffic emission (or primary emissions) we characterized each pollutant cluster in terms of their respective pollutant ratios, between-pollutant correlations, and maternal residential distance to major highways. Such metrics have shown to be helpful in terms of assessing sources of emissions related to near road vehicle traffic (Austin et al., 2012; Janhäll et al., 2012; Laurent et al., 2014; Levy et al., 2013). Also, since both NO and NO_2 could be highly correlated with $PM_{2.5}$ under certain emissions scenarios, we normalized NO and NO_2 concentration to $PM_{2.5}$ as described in Austin et al. (2012). This normalization helps to indicate which pollutant clusters have elevated NO and NO_2 concentrations, after accounting for their overall relationship with $PM_{2.5}$. We obtained the normalized concentrations by calculating the cluster-specific ratio of NO or NO_2 to $PM_{2.5}$ and dividing by the overall study area ratio of NO or NO_2 to $PM_{2.5}$.

3. Results

We observed an overall TLBW prevalence of 2.07% (95% CI: 2.04–2.11, n = 16,694) for the study population. Our data also showed spatial autocorrelation at the census tract level with respect to prevalence of TLBW (Fig. S5, Supplemental materials). Average census block group concentrations of NO_2, NO, and $PM_{2.5}$ were 22.49 ppb (interquartile range [IQR]: 19.68, 25.30 ppb), 21.84 ppb (IQR: 16.05, 26.11), and 16.94 μg/m^3 (IQR: 15.96, 18.18), respectively (Table 1).

3.1. Pollutant and contextual variable correlations

As indicated in Fig. 1, there is evidence of strong correlation (*Pearson's r*) for between pollutant concentrations, most notably for NO_2 and NO. The positive correlations between $PM_{2.5}$ and NO_2 and between $PM_{2.5}$ and NO are considerably weaker (Fig. 1). However, we did find that the between-pollutant correlations for $PM_{2.5}$ and the NO_x species varied widely across LA County (Fig. S7, Supplemental materials). Correlations between our contextual variables and the study pollutants ranged from only weak to moderate (range: −0.53 to 0.54) (Fig. 1). Correlations between the contextual variables on the other hand were stronger (range: −0.66 to 0.64).

3.1.1. Pollutant clusters

The profile regression identified 13 pollutant profile clusters (P1–P13) from the 6280 census block groups from which we had complete air pollution data. Summary statistics for each pollutant overall, and stratified by cluster, are in Table 1. These data summaries are color coded to help indicate which exposure profile clusters have either elevated (red), typical (green), or lowered (blue) pollutant concentrations compared to the overall concentrations. According to pollutant summaries in Table 1, the elevated NO_2 clusters are pollutant clusters P3, P7, P9, P10, and P13. The elevated NO pollutant clusters are the same as for NO_2, plus P12. The elevated $PM_{2.5}$ pollutant clusters are clusters P6, P7, P9, P10and P13. Four of the pollutant clusters show elevated levels for all pollutants, including P7, P9, P10, and P13. Whereas pollutant profile clusters P1, P2, P4, P5, and P11 show low levels for all pollutants. We mapped the spatial distributions of pollutant clusters in Fig. 2. Clusters with high concentrations for all pollutants are mostly within the downtown/metro area of LA and South-Central LA (clusters P7, P9, and P10), and a relatively fewer number of census block groups in the eastern section of the county (P13).

Between pollutant correlations for each pollutant cluster are indicated in Table 2 (see Supplemental materials for the spatial pattern of pollutant correlations). In Fig. 3 we present the distributions of NO/NO_2 ratios and the normalized NO_2 and normalized NO values for each pollutant cluster, while the median residential distances to major highways throughout the county are shown in Fig. 4. According to these pollutant and near-highway exposure metrics, mothers residing in pollutant clusters P9 and P10 are most likely exposed to higher levels of primary traffic pollution since these two clusters are characterized by high NO/NO_2 ratios and high normalized NO and NO_2 concentration. Additionally, P9 and P10 are characterized by low correlations between NO and NO_2 (suggestive of near roadway emissions) and elevated $PM_{2.5}$

Table 1
Summary statistics of mean census block group-level pollutant concentrations for pollutant exposure profile clusters.

Pollutant cluster (number of census block groups)	$PM_{2.5}$ μg/m^3 (IQR)	NO_2 ppb (IQR)	NO ppb (IQR)
P1 (329)	13.20 (11.95, 14.55)	18.19 (15.58, 20.79)	10.99 (9.30, 12.87)
P2 (242)	12.83 (12.03, 13.69)	21.34 (19.28, 23.2)	19.17 (16.07, 21.60)
P3 (633)	14.93 (14.58, 15.30)	28.14 (26.12, 29.84)	29.10 (23.46, 33.90)
P4 (1399)	16.91 (16.41, 17.36)	23.87 (21.87, 26.17)	21.08 (17.35, 24.53)
P5 (624)	15.96 (15.27, 16.52)	21.94 (20.61, 23.33)	23.50 (21.84, 25.21)
P6 (500)	18.87 (18.31, 19.41)	17.84 (16.12, 19.66)	15.46 (13.24, 18.20)
P7 (1715)	18.16 (17.79, 18.50)	25.63 (24.00, 27.31)	29.36 (26.11, 32.57)
P8 (96)	16.90 (16.40, 17.18)	16.46 (13.23, 18.78)	22.34 (16.29, 30.91)
P9 (513)	17.66 (17.12, 18.15)	31.74 (30.14, 33.27)	42.31 (37.57, 46.37)
P10 (52)	17.23 (16.84, 17.67)	42.51 (39.91, 44.29)	74.08 (63.40, 80.79)
P11 (29)	10.01 (8.93, 10.83)	14.44 (7.44, 16.80)	18.67 (10.86, 23.07)
P12 (52)	16.87 (16.28, 17.47)	16.51 (15.78, 17.42)	59.62 (35.43, 83.08)
P13 (96)	22.77 (21.89, 23.75)	31.39 (29.55, 32.85)	30.69 (28.13, 33.10)
Overall (6280)	16.88 (15.84, 18.13)	24.41 (21.33, 27.45)	25.99 (19.12, 30.94)

concentrations. Moreover, clusters P9 and P10 have the shortest median residential distance to major highways (478 m and 230 m respectively). Even though pollutant cluster P13 exhibited the highest average $PM_{2.5}$ in addition to elevated NO_2 and NO, this cluster is emblematic of combined high levels of regional sources of $PM_{2.5}$ and secondary particulate formation from traffic emissions. For instance, compared to all other clusters, P13 shows some of the strongest positive correlation between NO_2 and $PM_{2.5}$, a low NO/NO_2 ratio and low normalized NO and NO_2 levels. The pollutant metrics including maternal distance to highways (Fig. 4), between-pollutant correlations (Table 2), normalized NO concentration, and NO/NO_2 ratio (Fig. 3), indicate that all other pollutant profile clusters are more characteristic of secondary traffic-related pollution and other non-local (or regional transport) $PM_{2.5}$ sources instead of primary traffic-related.

3.1.2. Pollutant cluster random effects

For each pollutant cluster, Table 3 presents the total number of births, the number of TLBW cases, percent prevalence of TLBW, and the posterior means and 95% credible intervals for the pollutant cluster random effects. The cluster-specific effects represent the variation in the baseline log odds of TLBW, after adjusting for individual-level covariates, contextual cluster random effects, and accounting for spatially structured and unstructured residual error. In Table 3 we also present the posterior probability that a pollutant cluster effect is above the overall baseline log odds for TLBW (i.e. probability effect > zero).

Pollutant clusters with the highest probabilities for a random effect above zero are clusters P9 and P10, with probabilities of 94.9% and 91.6%, respectively (Table 3). Pollutant cluster P13 showed the next highest probability (77.4%). All other pollutant clusters showed probabilities below 70% for posterior probabilities with effects above zero. Pollutant clusters P1, P2, and P3 showed substantially lower baseline log odds of TLBW (Table 3).

3.1.3. Spatial distribution of pollutant cluster effects

In Fig. 5 we mapped the posterior probabilities for the pollutant cluster random effects. This map indicates clustering within LA County's urban core of downtown/metro LA, South-Central LA, and parts of east LA County for the pollutant profile clusters associated with the highest probability for increased baseline log odds of TLBW. Furthermore, the census block groups with the largest certainty for elevated TLBW log odds are mostly confined to census block groups near major highways; suggesting that women exposed to air pollution mixtures near highways have the greatest probability of delivering a TLBW baby.

3.2. Contextual clusters

The profile regression determined 14 distinct contextual profiles clusters (C1–C14) and are summarized in Fig. 6. Since the contextual clusters were developed separately from the pollutant clusters, these contextual clusters are distinct from the pollutant clusters. Table 4 presents a summary of the contextual profile cluster random effects. Again, these random effects represent the variation in the baseline log odds of TLBW across contextual clusters in our multilevel model. Compared to all other clusters, contextual cluster C6 showed the largest posterior mean effect (0.124) and the highest probability for a baseline log odds above the overall baseline log odds (probability = 99.5%). The income distribution for cluster C6 is significantly below the overall median income for LA County and consists of a significantly lower percentage of non-Hispanic Whites. Additionally, we find that cluster C6 has above average percentage of homes older than 1950, percentage of Blacks, and percentage of Hispanics. The next two highest probability contextual clusters are clusters C11 and C14, with elevated baseline TLBW log odds probabilities of 90.2% and 81.6%, respectively. While contextual clusters C11 and C14 have elevated percentages for homes older than 1950 and elevated percentages for Black populations, only C11 has significantly lower median income levels. All other clusters fell below 80% for elevated baseline TLBW log odds probabilities.

We also mapped the spatial distribution of contextual profile cluster effect probabilities in Fig. 7. This map indicates that contextual profile clusters with the highest probabilities for an elevated baseline log odds of TLBW are mostly in the urban core of LA County (central LA and south central LA). As anticipated, we find a large degree of spatial overlap between pollutant profile clusters with elevated effects on TLBW and contextual profile clusters with elevated effects on TLBW. Despite this spatial overlap, the two types high risk clusters take on distinct spatial patterns from one another.

3.3. Fixed effects results

In Table 5 we summarize each fixed effects estimate and corresponding 95% credible intervals from the multilevel spatial model.

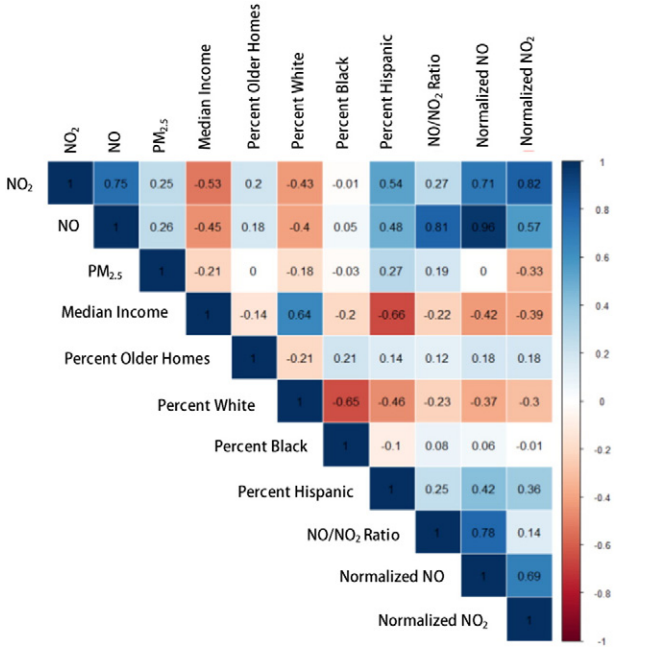

Fig. 1. Pearson correlation matrix of census block group level averages for air pollutants, contextual variables, and pollutant metrics.

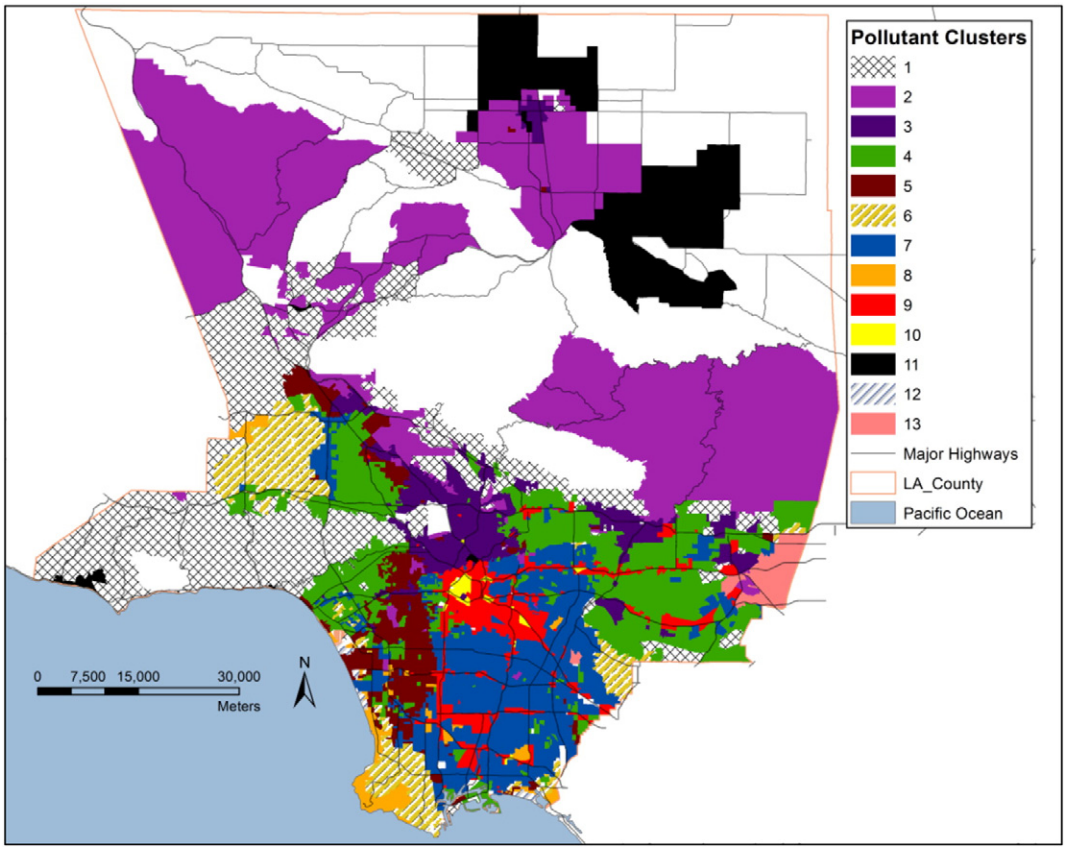

Fig. 2. Spatial distribution of pollutant exposure profile clusters.

Individual-level factors associated with odds of TLBW were female infant sex, gestational days, and gestational days squared, as well as maternal factors such as parity, age, race-ethnicity, and education level.

4. Discussion

Our Bayesian profiling approach highlights the varied and distinct spatial patterns of pollutant exposure profiles and how such exposures contribute to TLBW within the context of clustered indicators of socio-demographic and housing. While exposure profile clustering has been used in previous epidemiologic studies (Molitor et al., 2014b; Papathomas et al., 2010; Pirani et al., 2015; Zanobetti et al., 2014), no such studies have examined birth outcomes. Our clustering procedure and multilevel analysis provided concentration estimates for pollutants and TLBW risk estimates for place-based air pollution mixtures across LA County that take on a strong spatial structure.

4.1. Pollutant profile clusters and effects on TLBW

We find that census block groups adjacent to major highways in the downtown/metro LA, South-Central LA, and parts of eastern LA County show elevated $PM_{2.5}$, NO_2, and NO concentrations. We also find that census block groups in downtown/metro LA, South-Central LA, and east LA County, consistently had the most hazardous air pollution mixtures in relation to prevalence of TLBW.

We also identified profiles of clustered neighborhood contextual factors to show that sub-populations previously shown to be vulnerable to TLBW are concentrated in the downtown/metro area of LA and South-Central LA County. These high risk contextual clusters partially overlapped spatially with the most hazardous air pollution mixtures. Such co-occurrence of clustered indicators of disadvantage and hazardous air pollution mixtures reinforced the validity in our approach of adjusting for such correlated factors in examining multiple pollutant health effects.

Table 2
Between pollutant Pearson's correlations for the pollutant exposure profile clusters (P1–P13)[a].

Pollutants	Pollutant clusters												
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13
NO_2–NO	0.86	0.85	0.90	0.88	0.79	0.86	0.73	0.93	0.53	0.15	0.94	0.41	0.75
NO_2–$PM_{2.5}$	0.49	0.75	0.24	0.26	−0.42	0.37	0.12	−0.27	0.15	−0.23	0.58	0.40	0.55
NO–$PM_{2.5}$	0.45	0.67	0.33	0.10	−0.14	0.38	−0.11	−0.18	0.23	0.43	0.56	0.30	0.36
PM2.5 ~ NO/NO_2[b]	0.16	0.38	0.36	−0.06	0.32	0.26	−0.27	0.02	0.18	0.54	0.15	0.27	0.14

[a] Overall correlations: NO_2–NO = 0.76, NO_2–$PM_{2.5}$ = 0.25, NO–$PM_{2.5}$ = 0.26, $PM_{2.5}$–NO/NO_2 ratio = 0.19.
[b] Correlation between $PM_{2.5}$ concentration and the ratio of NO concentration to NO_2 concentration.

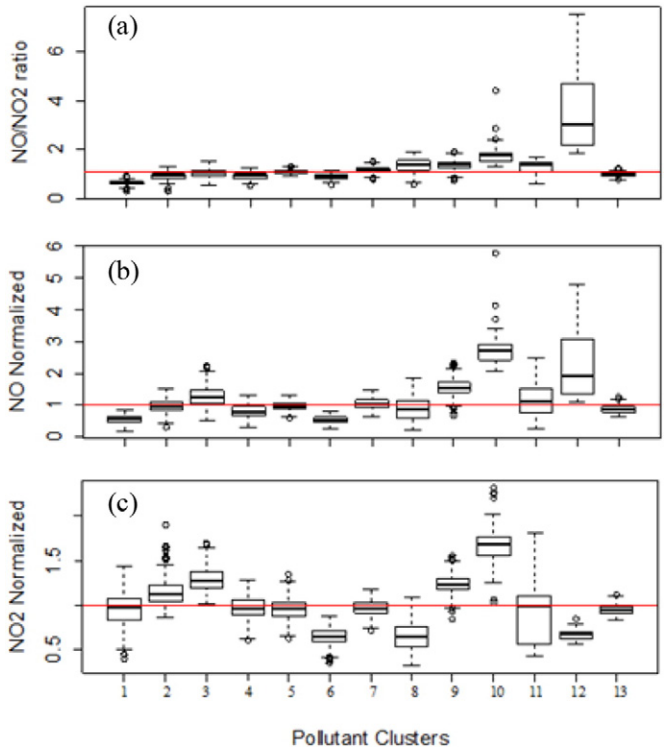

Fig. 3. Boxplots of (a) NO/NO_2 ratios, (b) normalized NO, and (c) normalized NO_2 for each pollutant cluster. The red line indicates the overall average for each of these pollutant metrics.

Spatially clustered emissions related to residential proximity to emissions sources may explain why our multilevel analysis of pollutant profile TLBW risks reveals such distinct spatial patterning. Spatial cluster inducing factors may include localized circumstances, such as traffic volumes and congestion, the vehicle fleet (e.g. heavy-duty trucks), and higher exposures to specific primary PM components that result from proximity to PM emissions sources. In our study, the spatial patterning across pollutant profiles of maternal distance to major highways, between pollutant correlations, elevated concentrations of NO and NO_2, and elevated NO/NO_2 ratios suggest that clusters reflective of primary traffic emissions tended to impart the greatest risk of TLBW (i.e. pollutant clusters P9 and P10). For instance our spatial proximity data show that the highest risk pollutant clusters (P9 and P10) are characterized by the smallest median maternal residential distances to major roadways, compared to all other cluster-specific residential distances to highways. In terms of primary traffic emission above background levels, the residential distances for P9 and P10 are consistent with the literature that suggest primary traffic emissions decay to background levels between 115 m to 570 m (Karner et al., 2010). In addition, there was clear spatial clustering of higher risk pollutant profiles (P9 and P10) at interchanges where LA's major interstate highways (I-5, I-110, I-710, and I-10) converge in central and south-central areas of LA County. This is suggestive of highly localized traffic and emission patterns that are germane to these major highways and their intersections.

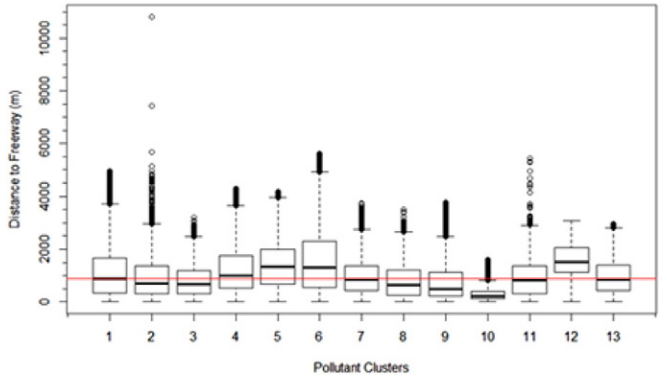

Fig. 4. Distribution of residential distance to highways among mothers within each pollutant cluster. Redline indicates the overall median for all maternal residential distances (median = 867 m).

Findings from our study are notably consistent with our earlier studies and other's conducted in LA County that found variation in estimated effects on birth outcomes between traffic-related sources of air pollution (e.g. traffic-related versus natural background sources) and proximity to major roadways (Laurent et al., 2014; Ritz et al., 2007; Wilhelm et al., 2011a; Wilhelm and Ritz, 2005). When juxtaposed with other recent studies (Bell et al., 2011; Coker et al., 2015; Laurent et al., 2014; Pirani et al., 2015; Wilhelm et al., 2011a; Zanobetti et al., 2014), our findings carry the implication that TLBW risks related to spatial patterns in exposure combined with the physical and chemical properties of $PM_{2.5}$ requires further investigation, and further suggests important spatially derived hypotheses. For instance, recent findings by our group showed that the exposure response relationship of $PM_{2.5}$ on TLBW varied spatially across LA County (Coker et al., 2015). Spatially varying effects suggests greater than additive health impacts influenced by (1) the sources of localized emissions, (2) proximity to PM sources (Buonocore et al., 2009; Cho et al., 2009; Greco et al., 2007; Kuhn et al., 2005; Laurent et al., 2014; Wagner et al., 2012), and (3) the varied pollutant profiles associated with proximity to different emissions sources of $PM_{2.5}$ (Laurent et al., 2014).

An important limitation of previous studies that attempted to find gradients in TLBW risk associated with various $PM_{2.5}$ components is the inability to pinpoint major sources or components contributing to TLBW risks. Instead, nearly all sources and components imparted a risk of exposure and are correlated. For example, the inherent dependencies between $PM_{2.5}$, $PM_{0.1}$, and various carbonaceous particulates (e.g. organic carbon, black carbon, and elemental carbon), or between $PM_{2.5}$ and sulfates in the exposure model used by Laurent et al. (2014), made it impossible to parse out which fraction sizes, components within $PM_{2.5}$- or combination thereof - are most likely to impart the greatest TLBW risk (Laurent et al., 2014). A single major source combined with certain spatially determined factors may produce a particular air pollution *mixture* that is more hazardous, yet multiple regression techniques struggle to distinguish between them.

The results from ours and a recent study by Pirani et al. (2015), show that Bayesian profile regression provides a tangible clustering procedure to develop profiles of exposure to multiple pollutants and simultaneously provide visualization tools. For instance, Pirani et al. (2015) studied variations in respiratory mortality across exposure profile clusters using a similar Bayesian profile regression. They found that days with high levels of secondary particulates (e.g. nitrates and sulfates) imparted the highest mortality risk in comparison to all other $PM_{2.5}$ component exposure profiles. Thus, rather than multiple regression models with pollutants and sources being highly correlated (Hampel et al., 2015; Laurent et al., 2014), our clustering approach could be applied to develop PM-exposure profiles using data on $PM_{2.5}$ components, sources, and size fractions. Furthermore, our spatially-based clustering approach enables identification and mapping of sub-regions that are characterized by the most hazardous PM-source components.

Rather than simply examining gradients in multipollutant health outcome risks devoid of spatial information, our study illustrates the importance in examining the spatial patterning of multipollutant health effects to help bring out the likely causes of apparent non-linear effects. For example, pollutant cluster P9 - reflective of primary traffic $PM_{2.5}$ pollution - was not the only cluster with elevated concentrations for all

Table 3
Prevalence of TLBW for multipollutant clusters and model results for multipollutant exposure profile cluster random effects (N = 804,726).

Cluster	No. births[a]	No. TLBW	% TLBW[b] (95% CI)	Cluster effect[c,d] (95% CI)	Probability effect > 0[e]
P1	29394	431	1.47* (1.33, 1.61)	−0.122 (−0.254, 0.008)	0.033
P2	34263	589	1.72* (1.58, 1.86)	−0.108 (−0.232, −0.015)	0.042
P3	79199	1555	1.96 (1.87, 2.06)	−0.026 (−0.133, 0.081)	0.315
P4	168145	3305	1.97 (1.89, 2.03)	0.006 (−0.093, 0.106)	0.548
P5	66035	1595	2.42* (2.30, 2.53)	0.017 (−0.092, 0.126)	0.619
P6	58164	1067	1.83* (1.73, 1.95)	0.003 (−0.109, 0.115)	0.519
P7	256817	5612	2.19* (2.13, 2.24)	0.019 (−0.080, 0.117)	0.642
P8	7452	113	1.52* (1.25, 1.82)	−0.096 (−0.275, 0.079)	0.141
P9	77336	1824	2.36* (2.25, 2.47)	0.089 (−0.018, 0.195)	0.949
P10	3171	82	2.59 (2.06, 3.20)	0.141 (−0.060, 0.338)	0.916
P11	3491	74	2.12 (1.67, 2.65)	0.043 (−0.168, 0.252)	0.656
P12	3357	51	1.52 (1.13, 1.99)	−0.012 (−0.232, 0.205)	0.459
P13	17902	396	2.21 (2.00, 2.44)	0.055 (−0.087, 0.196)	0.774
Overall	804,726	16,694	2.07 (2.04, 2.11)		

[a] Full term births (>259 days gestation).
[b] Percent prevalence of TLBW without model adjustment.
[c] Adjusted for maternal age, race, education, and parity, infant sex, gestation (days), gestation-squared, and contextual random effect clusters.
[d] Random effect presented on the log-odds scale.
[e] Probabilities were calculated utilizing the "*inla.pmarginal*" function in INLA.
* TLBW prevalence is significantly (p-value < 0.01, two-sided) different from the overall proportion of TLBW (unadjusted).

study pollutants. Cluster P9, however, displayed the highest probability for an estimated effect above zero and the second highest effect estimate size, despite other pollutant clusters displaying higher $PM_{2.5}$ or higher NO and NO_2 concentrations. The only other pollutant cluster with a larger estimated effect size was cluster P10, which was also characteristic of *primary* traffic pollution. Whereas pollutant cluster P13 – reflective of mostly secondary traffic $PM_{2.5}$ – has a $PM_{2.5}$ concentration 30% higher than P9 and P10. Despite this, the estimated effect size for cluster P13 is lower than P9 and P10. Our approach thus identifies patterns that help explain apparent non-linear effects, such as: (1) mapping of cluster effects that exhibited strong spatial patterns related to major roadways, (2) the variations in pollutant metrics such as correlations and pollutant ratios across clusters combined with the spatial patterning of these pollutant metrics, and (3) the influence of residential distance to major roadways. These spatial data provide strong evidence that *primary* traffic emissions uniquely impart the largest effect on TLBW.

4.2. Correlated pollutants and health effects research

Our approach to examining health effects of correlated exposures via exposure profile clusters offers several advantages compared to co-pollutant regression methods. Problems with collinearity within a multivariate regression include inflated variance in regression coefficients, unstable effect estimates, and causal inference challenges (Dormann et al., 2013; Lin, 2008; Schmidt and Muller, 1978). Several outdoor air pollution studies find that pollutants contribute to a health outcome in single pollutant models; however, mutual adjustment for correlated pollutants can result in no pollutant showing an association (Ebisu and Bell, 2012) or coefficients flipping sides i.e. opposite in direction

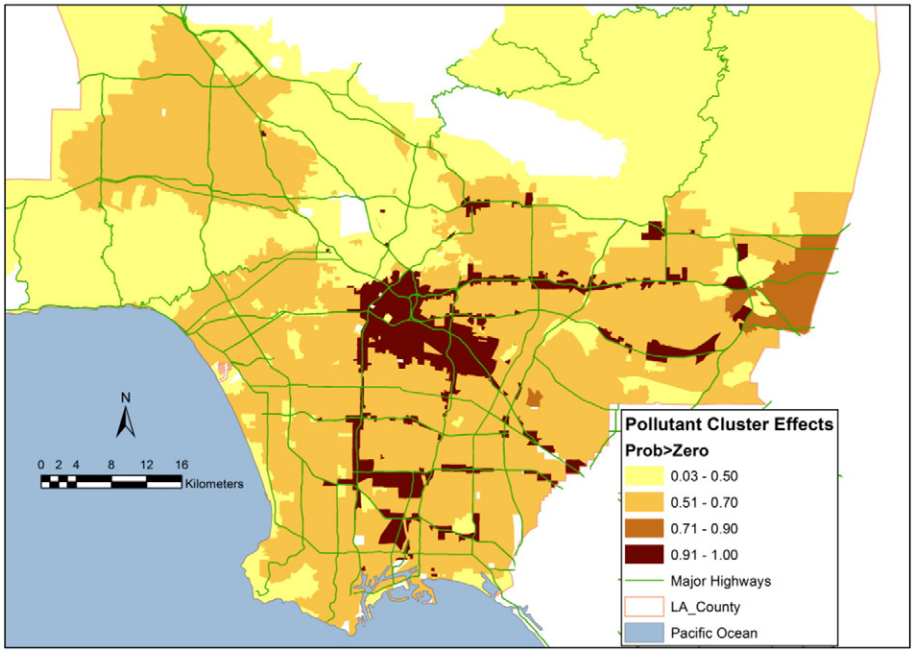

Fig. 5. Spatial distribution of pollutant profile cluster random effect posterior probabilities.

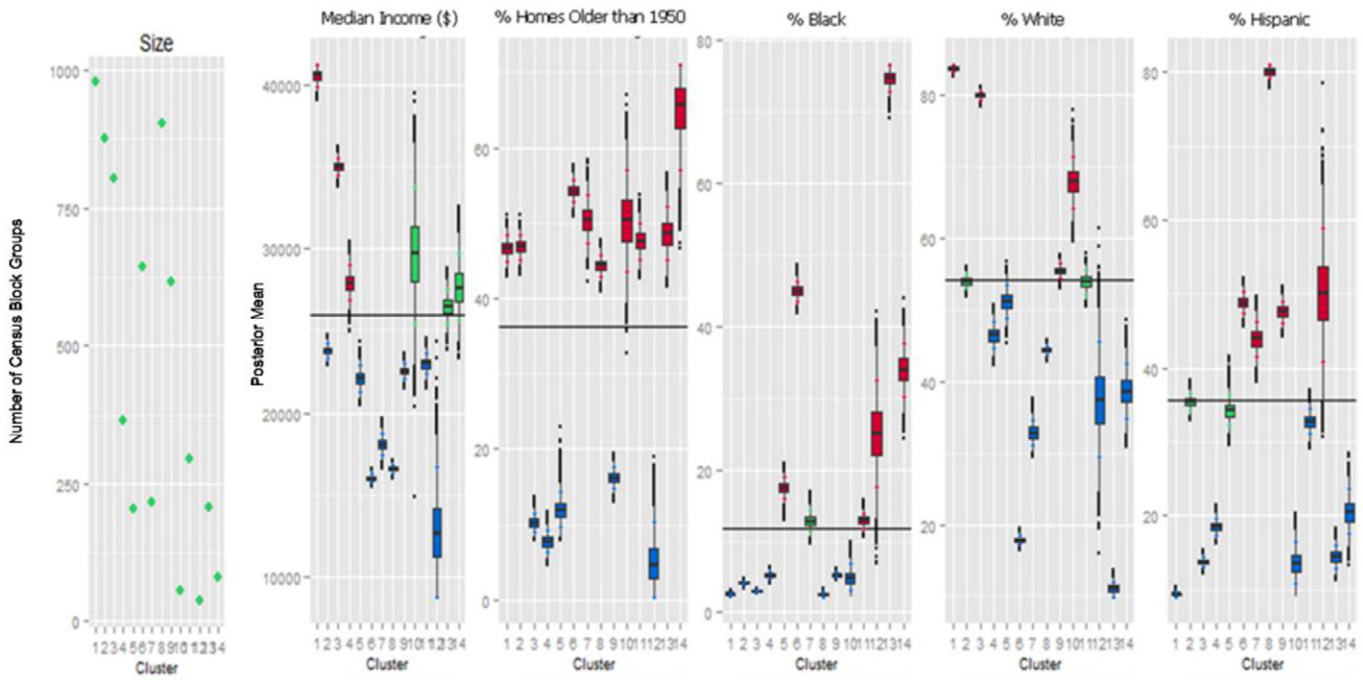

Fig. 6. Contextual profile clusters: cluster size (number of census block groups) and posterior distributions for median income, percent older homes, percent black, percent white, and percent Hispanic ($N_{clusters} = 14$). Black lines indicate the overall average exposure for a given covariate.

from single pollutant models (Kelsall et al., 1997). Furthermore, whether mutual adjustment is necessary – i.e. whether multiple pollutants actually confound each other – cannot be determined in models with highly correlated pollutants. When we analyze our data using a co-pollutant model that includes all pollutants in a multivariate regression model, we find instability of specific pollutant coefficient estimates and increased standard errors (see Table S2 in the Supplemental materials). Also, while we find suggestive evidence for a statistically significant interaction between NO and NO_2 in a co-pollutant model (Table S1, Supplemental materials), the interpretation is challenging since both pollutants tend to co-vary and are linked through conversion into one-another by atmospheric chemistry involving ozone. Bayesian profile regression on the other hand harnesses the collinearity of air pollutants to find meaningful patterns of joint exposure that are relevant for determining different health risks across pollutant clusters.

Another important problem with multipollutant modeling is that correlations between pollutants can vary over space (Dionisio et al., 2014; Levy et al., 2013; Snowden et al., 2015). Spatial variability in pollutant correlations between and within urban communities is challenging because it can lead to exposure measurement error and further calls into question estimating co-pollutant effects reliably. It is also unclear whether results from studies in one particular region are generalizable to others that have different spatial patterns of pollutant correlations. Consistent with previous studies (Dionisio et al., 2014; Levy et al., 2013), we found substantial within county variability in pollutant correlations between our pollutant clusters (Figs. S6 and S7, Supplemental materials). To some extent our approach overcomes spatial variation in pollutant correlations because we characterized the heterogeneity in pollutant relationships across space, and further relate these exposure profiles to a health outcome. Hence, employing our approach in environmental health studies may better inform policies designed to protect public health since policies can be tailored towards pollution mixtures relevant to a specific area.

Table 4
Prevalence of TLBW for contextual clusters and model results for contextual exposure profile cluster random effects (N = 804,726).

Cluster	No. births[a]	No. TLBW	% TLBW[b] (95% CI)*	Cluster effect[c,d] (95% CI)	Probability effect > 0[e]
C1	64018	965	1.51* (1.42, 1.60)	−0.09 (−0.19, 0.02)	0.05
C2	95559	1869	1.96 (1.87, 2.05)	−0.03 (−0.12, 0.07)	0.30
C3	76163	1218	1.60* (1.51, 1.69)	−0.08 (−0.18, 0.02)	0.07
C4	39105	787	2.01 (1.88, 2.16)	−0.03 (−0.14, 0.08)	0.31
C5	38084	809	2.12 (1.98, 2.27)	−0.002 (−0.11, 0.11)	0.49
C6	112919	2970	2.63* (2.54, 2.73)	0.12 (0.03, 0.22)	0.99
C7	39738	912	2.30* (2.15, 2.45)	0.03 (−0.08, 0.14)	0.71
C8	165328	3384	2.05 (1.98, 2.12)	0.03 (−0.07, 0.12)	0.71
C9	101455	2040	2.01 (1.93, 2.10)	0.02 (−0.07, 0.14)	0.67
C10	2624	42	1.60 (1.18, 2.16)	−0.08 (0.30, 0.13)	0.22
C11	43403	975	2.25 *(2.11, 2.39)	0.07 (−0.04, 0.17)	0.90
C12	963	18	1.87 (1.18, 2.95)	0.07 (−0.31, 0.18)	0.30
C13	18784	538	2.86*(2.64, 3.11)	0.03 (−0.09, 0.15)	0.68
C14	6583	167	2.54*(2.18, 2.95)	0.07 (−0.09, 0.23)	0.82

[a] Full term births (>259 days gestation).
[b] Percent prevalence of TLBW without model adjustment.
[c] Adjusted for maternal age, race, education, and parity, infant sex, gestation (days), gestation-squared, and pollutant random effect clusters.
[d] Random effect presented on the log-odds scale.
[e] Probabilities were calculated utilizing the "*inla.pmarginal*" function in INLA.
* TLBW prevalence is significantly (p-value < 0.01, two-sided) different from the overall proportion of TLBW (unadjusted).

4.3. Contextual neighborhood effects on TLBW

An important aspect to our pollutant clustering approach lies in simultaneously adjusting for clustered neighborhood indicators of disadvantage (i.e. income and race) and older housing. While other contextual factors related to TLBW could have been included in our clustering procedure, it is clear that our clustering variables are highly correlated with other adverse contextual factors (i.e. education, housing values, low social support, neighborhood greenness, violent crimes, etc.) in southern California, and thus likely account for these other contextual factors (Boggess and Hipp, 2010; Conway et al., 2010; English et al., 2003; Ghosh et al., 2010). Furthermore, since our contextual variables correlate with one another and correlate with pollutants, it was important to separate out the contextual area-level effects from the

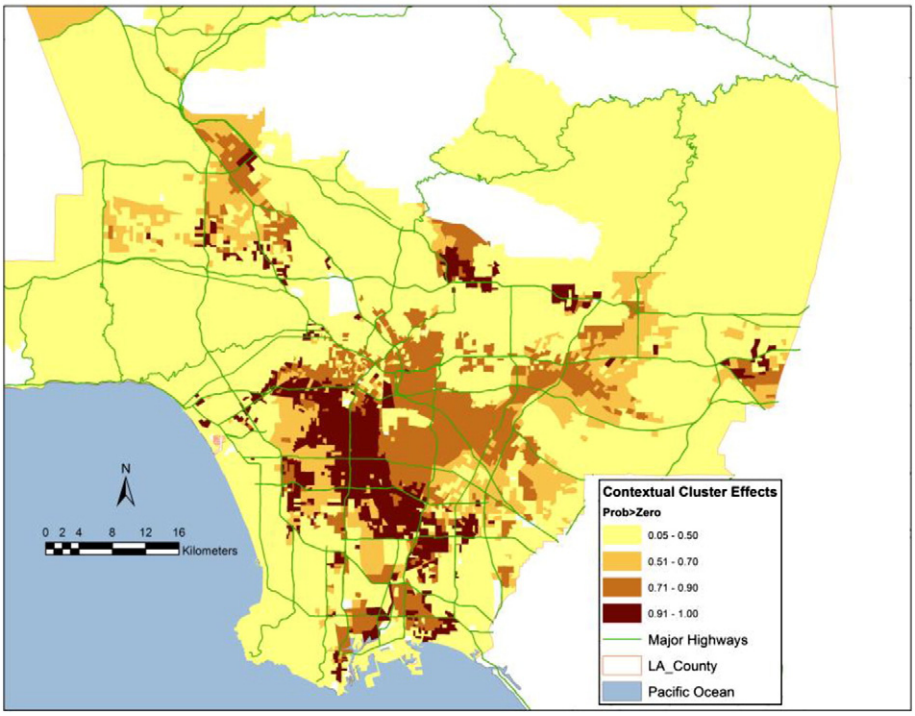

Fig. 7. Spatial distribution of contextual profile cluster random effect posterior probabilities.

pollutant profile cluster effects. Despite the spatial similarity between the two different types of clusters in our study, the high risk contextual clusters display a spatial pattern that is distinct from the high risk pollutant clusters, suggesting that these two separate exposure profiles measured different spatial patterns of risk related to their respective variables.

Table 5
Fixed effects odds ratios of TLBW for individual-level covariates[a] (N = 804,726).

Fixed effects covariates	Odds ratio	2.5% quantile	97.5% quantile
Infant			
Male (reference)	1		
Female	1.45	1.40	1.49
Parity			
No previous births (reference)	1		
≥1 previous birth	0.59	0.57	0.61
Maternal age			
<20 years (reference)	1		
20–24 years	0.97	0.92	1.03
25–29 years	0.90	0.85	0.96
30–34 years	0.91	0.86	0.97
≥35	1.06	1.00	1.14
Maternal education			
0–8 years (reference)	1		
9–12 years	0.90	0.86	0.94
13–15 years	0.75	0.70	0.79
≥16 years	0.66	0.62	0.71
Race/ethnicity			
Non-Hispanic White (reference)	1		
Hispanic	1.08	1.02	1.15
Non-Hispanic Black	2.16	2.01	2.32
Non-Hispanic Asian	1.42	1.32	1.53
Other	1.81	1.67	1.97
Gestation (days)	0.32	0.30	0.33
Gestation-squared	1.0019	1.0018	1.002

[a] ORs represent covariate fixed effects estimated from the multilevel model. Random effects in this model were pollutant clusters, contextual clusters, a spatial random error term based on adjacent census tracts, and a spatially unstructured random error term.

4.4. Study limitations

We lacked data on speciation and tracers for specific sources of $PM_{2.5}$, which limits our ability to attribute a particular air pollution source to effects on TLBW. Despite this limitation, the spatial patterning of our results, supplemented with metrics such as cluster-specific NO/NO_2 ratios and between pollutant correlations, offers strong evidence implicating primary traffic pollution. Another limitation lies in the lack of fine-scale spatial data for other air toxics (i.e. benzene, ozone and carbon monoxide). Since we lack data on other air toxics that correlate with the pollutants considered in our study (Fujita et al., 2011; Ghosh et al., 2012; Laurent et al., 2014; Protano et al., 2012; Salam et al., 2005; Wilhelm et al., 2011a), we cannot say whether and how these other pollutants may contribute to the observed spatial patterning of TLBW risks. Other limitations include a lack of information on indoor air pollution exposure and time-activity patterns that may influence air pollution exposures, such as information about whether the women worked outside the home (Ritz et al., 2007) or commuted daily (Zuurbier et al., 2010); all of which can contribute further to exposure misclassification. However, recent findings suggest that maternal outdoor air pollution estimates at the home address are unlikely affected by a lack of time-activity patterns during pregnancy (Ouidir et al., 2015). Finally, while our approach offers several important advantages over previous air pollution profile studies, our Bayesian approach is currently limited in regards to handling a dataset with high dimensionality, which can only be overcome via future gains in computational efficiencies.

4.5. Study strengths

The primary strength of our study is that we were able to examine exposures to multiple correlated air pollutants and a health outcome, mitigating some of the typical problems encountered with correlated exposures. Our study also had a large sample size and used population-wide data for exposure, thus avoiding selection of a study

population based on proximity to major sources of air pollution or proximity to central site monitors (Dionisio et al., 2014; Kumar, 2012). Another strength is adjustment for individual-level covariates and contextual factors associated with TLBW, and adjustment for spatial residual confounding at the census tract-level. Another important strength is that we were able to fit a multilevel/hierarchical random effects model for the clustered pollutant profiles and contextual profiles enabling us to look at multiple profile-specific risks, thus avoiding some of the issues related to multiple testing of myriad exposure/SES effects on health (Gelman and Hill, 2006).

5. Conclusion

Our Bayesian clustering procedures allowed us to go beyond simple one-at-a-time analyses usually employed to examine marginal effects of individual pollutants on health. Further, this spatially distributed *mixtures* approach provides information on the spatial distribution of exposure/SES profiles that pertain to the *levels* of various pollutants and SES factors. Policy analysts can use this information to determine which exposure/SES profiles dominate a particular sub-region of L.A. County, as a starting point for regulatory considerations. In our analyses, we found that neighborhood-level $PM_{2.5}$, NO_2, and NO concentrations were correlated with census block group-level contextual factors throughout LA County; and the nature of these relationships was quite complex and highly spatially variable across the County. Moreover, the pollutant profile clusters showed a strong spatial contrast with respect to exposure-related TLBW risks. LA County's urban core, south-central urban region, and parts of the eastern-most region of the county exhibited the largest exposures for $PM_{2.5}$, NO_2, and NO, which decreased with distance from major highways. Moreover, the highest concentration pollutant profile clusters imparted the greatest TLBW risks, especially those closest to major highways, which suggests near roadway emissions are more important in terms of risk of adverse birth outcomes related to these air pollution profiles.

Acknowledgments

Research described in this article was conducted under contract to the Health Effects Institute (HEI), an organization jointly funded by the United States Environmental Protection Agency (EPA) (Assistance Award No. R-82811201) and certain motor vehicle and engine manufacturers. The contents of this article do not necessarily reflect the views of HEI, or its sponsors, nor do they necessarily reflect the views and policies of the EPA or motor vehicle and engine manufacturers.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.envint.2016.02.011.

References

Austin, E., Coull, B., Thomas, D., Koutrakis, P., 2012. A framework for identifying distinct multipollutant profiles in air pollution data. Environ. Int. 45, 112–121. http://dx.doi.org/10.1016/j.envint.2012.04.003.

Austin, E., Coull, B.A., Zanobetti, A., Koutrakis, P., 2013. A framework to spatially cluster air pollution monitoring sites in US based on the PM2.5 composition. Environ. Int. 59, 244–254. http://dx.doi.org/10.1016/j.envint.2013.06.003.

Beckerman, B.S., Jerrett, M., Martin, R.V., van Donkelaar, A., Ross, Z., Burnett, R.T., 2013a. Application of the deletion/substitution/addition algorithm to selecting land use regression models for interpolating air pollution measurements in California. Atmos. Environ. 77, 172–177. http://dx.doi.org/10.1016/j.atmosenv.2013.04.024.

Beckerman, B.S., Jerrett, M., Serre, M., Martin, R.V., Lee, S.-J., van Donkelaar, A., Ross, Z., Su, J., Burnett, R.T., 2013b. A hybrid approach to estimating national scale spatiotemporal variability of PM $_{2.5}$ in the contiguous United States. Environ. Sci. Technol., 130617085617008 http://dx.doi.org/10.1021/es400039u.

Bell, M.L., Belanger, K., Ebisu, K., Gent, J.F., Lee, H.J., Koutrakis, P., Leaderer, B.P., 2010. Prenatal exposure to fine particulate matter and birth weight: variations by particulate constituents and sources. Epidemiology 21, 884–891. http://dx.doi.org/10.1097/EDE.0b013e3181f2f405.

Bell, M.L., Ebisu, K., Peng, R.D., 2011. Community-level spatial heterogeneity of chemical constituent levels of fine particulates and implications for epidemiological research. J. Expo. Sci. Environ. Epidemiol. 21, 372–384. http://dx.doi.org/10.1038/jes.2010.24.

Besag, J., York, J.C., Molli'e, A., 1991. Bayesian image restoration, with two applications in spatial statistics (with discussion). Ann Stat Math 43, 1–59.

Bivand, R.S., Pebesma, E., Gómez-Rubio, V., 2013. Applied Spatial Data Analysis With R, 2. ed. ed, Use R! Springer, New York, NY.

Boehmer, T.K., Foster, S.L., Henry, J.R., Woghiren-Akinnifesi, E.L., Yip, F.Y., Centers for Disease Control and Prevention (CDC), 2013. Residential proximity to major highways — United States, 2010. Morb. Mortal. Wkly. Rep. Surveill. Summ. Wash. DC 2002 (62 Suppl 3), 46–50.

Boggess, L.N., Hipp, J.R., 2010. Violent Crime, Residential Instability and Mobility: Does the Relationship Differ in Minority Neighborhoods? J. Quant. Criminol. 26, 351–370. http://dx.doi.org/10.1007/s10940-010-9093-7.

Brauer, M., Tamburic, L., 2009. Multi-pollutant analysis of reproductive outcomes and air pollution using the CMAQ model. Epidemiology 20, S72–S73. http://dx.doi.org/10.1097/01.ede.0000362926.16562.96.

Brauer, M., Hystad, P., Poplawski, K., 2011. Assessing the spatial representativeness of PM2.5 and O3 measurements from the national air pollutant surveillance system. Envrionment Canada.

Buonocore, J.J., Lee, H.J., Levy, J.I., 2009. The influence of traffic on air quality in an Urban neighborhood: a community–university partnership. Am. J. Public Health 99, S629–S635. http://dx.doi.org/10.2105/AJPH.2008.149138.

Cho, S.-H., Tong, H., McGee, J., Baldauf, R., Krantz, Q., Gilmour, M., 2009. Comparative toxicity of size-fractionated airborne particulate matter collected at different distances from an urban highway. Environ. Health Perspect. http://dx.doi.org/10.1289/ehp.0900730.

Coker, E., Ghosh, J., Jerrett, M., Gomez-Rubio, V., Beckerman, B., Cockburn, M., Liverani, S., Su, J., Li, A., Kile, M.L., Ritz, B., Molitor, J., 2015. Modeling spatial effects of PM2.5 on term low birth weight in Los Angeles County. Environ. Res. 142, 354–364. http://dx.doi.org/10.1016/j.envres.2015.06.044.

Conway, D., Li, C.Q., Wolch, J., Kahle, C., Jerrett, M., 2010. A Spatial Autocorrelation Approach for Examining the Effects of Urban Greenspace on Residential Property Values. J. Real Estate Financ. Econ. 41, 150–169. http://dx.doi.org/10.1007/s11146-008-9159-6.

Debbink, M.P., Bader, M.D.M., 2011. Racial residential segregation and low birth weight in Michigan's metropolitan areas. Am. J. Public Health 101, 1714–1720. http://dx.doi.org/10.2105/AJPH.2011.300152.

Dionisio, K.L., Baxter, L.K., Chang, H.H., 2014. An empirical assessment of exposure measurement error and effect attenuation in bipollutant epidemiologic models. Environ. Health Perspect. http://dx.doi.org/10.1289/ehp.1307772.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36, 27–46. http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x.

Ebisu, K., Bell, M.L., 2012. Airborne PM2.5 chemical components and low birth weight in the northeastern and mid-Atlantic regions of the United States. Environ. Health Perspect. http://dx.doi.org/10.1289/ehp.1104763.

English, P.B., Kharrazi, M., Davies, S., Scalf, R., Waller, L., Neutra, R., 2003. Changes in the spatial pattern of low birth weight in a southern California county: the role of individual and neighborhood level factors. Soc. Sci. Med. 56, 2073–2088. http://dx.doi.org/10.1016/S0277-9536(02)00202-2.

EPA, U.S., 2008. Idling Vehicle Emissions for Passenger Cars, Light-Duty Trucks, and Heavy-Duty Trucks (No. EPA420-F-08-025). U.S. Environmental Protection Agency, Washington, D.C.

Fang, H., Johnson, C., Stopp, C., Espy, K.A., 2011. A new look at quantifying tobacco exposure during pregnancy using fuzzy clustering. Neurotoxicol. Teratol. 33, 155–165. http://dx.doi.org/10.1016/j.ntt.2010.08.003.

Fine, P.M., Sioutas, C., Solomon, P.A., 2008. Secondary particulate matter in the United States: insights from the particulate matter supersites program and related studies. J. Air Waste Manage. Assoc. 58, 234–253. http://dx.doi.org/10.3155/1047-3289.58.2.234.

Fujita, E.M., Campbell, D.E., Zielinska, B., Arnott, W.P., Chow, J.C., 2011. Concentrations of air toxics in motor vehicle-dominated environments. Res. Rep. Health Eff. Inst. 3–77.

Geer, L.A., 2014. Identifying exposure disparities in air pollution epidemiology specific to adverse birth outcomes. Environ. Res. Lett. 9, 101001. http://dx.doi.org/10.1088/1748-9326/9/10/101001.

Gelman, A., Hill, J., 2006. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Leiden.

Ghosh, J.K.C., Wilhelm, M.H., Dunkel-Schetter, C., Lombardi, C.A., Ritz, B.R., 2010. Paternal support and preterm birth, and the moderation of effects of chronic stress: a study in Los Angeles County mothers. Arch. Womens Ment. Health 13, 327–338. http://dx.doi.org/10.1007/s00737-009-0135-9.

Ghosh, J.K.C., Wilhelm, M., Su, J., Goldberg, D., Cockburn, M., Jerrett, M., Ritz, B., 2012. Assessing the influence of traffic-related Air pollution on risk of term low birth weight on the basis of land-use-based regression models and measures of air toxics. Am. J. Epidemiol. 175, 1262–1274. http://dx.doi.org/10.1093/aje/kwr469.

Ghosh, J.K.C., Wilhelm, M., Ritz, B., 2013. Effects of residential indoor air quality and household ventilation on preterm birth and term low birth weight in Los Angeles County, California. Public Health 103, 686–694. http://dx.doi.org/10.2105/AJPH.2012.300987.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), 1998. Markov Chain Monte Carlo in Practice. Chapman & Hall, Boca Raton, Fla.

Goldberg, D.W., Wilson, J.P., Knoblock, C.A., Ritz, B., Cockburn, M.G., 2008. An effective and efficient approach for manually improving geocoded data. Int. J. Health Geogr. 7, 60. http://dx.doi.org/10.1186/1476-072X-7-60.

Gouveia, N., Bremner, S.A., Novaes, H.M.D., 2004. Association between ambient air pollution and birth weight in São Paulo, Brazil. J. Epidemiol. Community Health 58, 11–17.
Grady, S.C., 2006. Racial disparities in low birthweight and the contribution of residential segregation: a multilevel analysis. Soc. Sci. Med. 1982 (63), 3013–3029. http://dx.doi.org/10.1016/j.socscimed.2006.08.017.
Grady, S.C., 2011. Housing quality and racial disparities in low birth weight: a GIS assessment. In: Maantay, J.A., McLafferty, S. (Eds.), Geospatial Analysis of Environmental Health. Springer Netherlands, Dordrecht, pp. 303–318.
Greco, S.L., Wilson, A.M., Spengler, J.D., Levy, J.I., 2007. Spatial patterns of mobile source particulate matter emissions-to-exposure relationships across the United States. Atmos. Environ. 41, 1011–1025. http://dx.doi.org/10.1016/j.atmosenv.2006.09.025.
Gu, J., Pitz, M., Breitner, S., Birmili, W., von Klot, S., Schneider, A., Soentgen, J., Reller, A., Peters, A., Cyrys, J., 2012. Selection of key ambient particulate variables for epidemiological studies — applying cluster and heatmap analyses as tools for data reduction. Sci. Total Environ. 435-436, 541–550. http://dx.doi.org/10.1016/j.scitotenv.2012.07.040.
Hampel, R., Peters, A., Beelen, R., Brunekreef, B., Cyrys, J., de Faire, U., de Hoogh, K., Fuks, K., Hoffmann, B., Hüls, A., Imboden, M., Jedynska, A., Kooter, I., Koenig, W., Künzli, N., Leander, K., Magnusson, P., Männistö, S., Penell, J., Pershagen, G., Phuleria, H., Probst-Hensch, N., Pundt, N., Schaffner, E., Schikowski, T., Sugiri, D., Tiittanen, P., Tsai, M.-Y., Wang, M., Wolf, K., Lanki, T., 2015. Long-term effects of elemental composition of particulate matter on inflammatory blood markers in European cohorts. Environ. Int. 82, 76–84. http://dx.doi.org/10.1016/j.envint.2015.05.008.
Hasheminassab, S., Daher, N., Saffari, A., Wang, D., Ostro, B.D., Sioutas, C., 2014. Spatial and temporal variability of sources of ambient fine particulate matter (PM2.5) in California. Atmos. Chem. Phys. 14, 12085–12097. http://dx.doi.org/10.5194/acp-14-12085-2014.
Hastie, D.I., Liverani, S., Azizi, L., Richardson, S., Stücker, I., 2013. A semi-parametric approach to estimate risk functions associated with multi-dimensional exposure profiles: application to smoking and lung cancer. BMC Med. Res. Methodol. 13, 129. http://dx.doi.org/10.1186/1471-2288-13-129.
Houston, D., Li, W., Wu, J., 2014. Disparities in exposure to automobile and truck traffic and vehicle emissions near the Los Angeles-Long Beach port complex. Am. J. Public Health 104, 156–164. http://dx.doi.org/10.2105/AJPH.2012.301120.
Hu, S., Paulson, S.E., Fruin, S., Kozawa, K., Mara, S., Winer, A.M., 2012. Observation of elevated air pollutant concentrations in a residential neighborhood of Los Angeles California using a mobile platform. Atmos. Environ. 51, 311–319. http://dx.doi.org/10.1016/j.atmosenv.2011.12.055.
Jain, S., Neal, R.M., 2004. A split-merge Markov chain Monte Carlo Procedure for the dirichlet process mixture model. J. Comput. Graph. Stat. 13, 158–182. http://dx.doi.org/10.1198/1061860043001.
Janhäll, S., Molnar, P., Hallquist, M., 2012. Traffic emission factors of ultrafine particles: effects from ambient air. J. Environ. Monit. 14, 2488. http://dx.doi.org/10.1039/c2em30235g.
Jerrett, M., Burnett, R.T., Beckerman, B.S., Turner, M.C., Krewski, D., Thurston, G., Martin, R.V., van Donkelaar, A., Hughes, E., Shi, Y., Gapstur, S.M., Thun, M.J., Pope, C.A., 2013. Spatial analysis of air pollution and mortality in California. Am. J. Respir. Crit. Care Med. 188, 593–599. http://dx.doi.org/10.1164/rccm.201303-0609OC.
Karner, A.A., Eisinger, D.S., Niemeier, D.A., 2010. Near-roadway air quality: synthesizing the findings from real-world data. Environ. Sci. Technol. 44, 5334–5344. http://dx.doi.org/10.1021/es100008x.
Kelsall, J.E., Samet, J.M., Zeger, S.L., Xu, J., 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am. J. Epidemiol. 146, 750–762.
Kuhn, T., Biswas, S., Fine, P.M., Geller, M., Sioutas, C., 2005. Physical and chemical characteristics and volatility of PM in the proximity of a light-duty vehicle freeway. Aerosol Sci. Technol. 39, 347–357. http://dx.doi.org/10.1080/027868290930024.
Kumar, N., 2012. Uncertainty in the relationship between criteria pollutants and low birth weight in Chicago. Atmos. Environ. 49, 171–179. http://dx.doi.org/10.1016/j.atmosenv.2011.12.001.
Laurent, O., Hu, J., Li, L., Cockburn, M., Escobedo, L., Kleeman, M.J., Wu, J., 2014. Sources and contents of air pollution affecting term low birth weight in Los Angeles County, California, 2001–2008. Environ. Res. 134, 488–495. http://dx.doi.org/10.1016/j.envres.2014.05.003.
Le, H.Q., Batterman, S.A., Wirth, J.J., Wahl, R.L., Hoggatt, K.J., Sadeghnejad, A., Hultin, M.L., Depa, M., 2012. Air pollutant exposure and preterm and term small-for-gestational-age births in Detroit, Michigan: long-term trends and associations. Environ. Int. 44, 7–17. http://dx.doi.org/10.1016/j.envint.2012.01.003.
Levy, I., Mihele, C., Lu, G., Narayan, J., Brook, J.R., 2013. Evaluating multipollutant exposure and urban air quality: pollutant interrelationships, neighborhood variability, and nitrogen dioxide as a proxy pollutant. Environ. Health Perspect. http://dx.doi.org/10.1289/ehp.1306518.
Lewis, A.S., Sax, S.N., Wason, S.C., Campleman, S.L., 2011. Non-chemical stressors and cumulative risk assessment: an overview of current initiatives and potential air pollutant interactions. Int. J. Environ. Res. Public. Health 8, 2020–2073. http://dx.doi.org/10.3390/ijerph8062020.
Lin, F.-J., 2008. Solving multicollinearity in the process of fitting regression model using the nested estimate procedure. Qual. Quant. 42, 417–426. http://dx.doi.org/10.1007/s11135-006-9055-1.
Liverani, S., Hastie, D., Azizi, L., Papathomas, M., Richardson, S., 2015. PReMiuM : an R package for profile regression mixture models using dirichlet processes. J. Stat. Softw. 3, 1–29.
Mauderly, J.L., Burnett, R.T., Castillejos, M., Ozkaynak, H., Samet, J.M., Stieb, D.M., Vedal, S., Wyzga, R.E., 2010. Is the air pollution health research community prepared to support a multipollutant air quality management framework? Inhal. Toxicol. 22 (Suppl. 1), 1–19. http://dx.doi.org/10.3109/08958371003793846.
Molitor, J., Jerrett, M., Chang, C.C., Molitor, N.T., Gauderman, J., Berhane, K., McConnell, R., Lurmann, F., Wu, J., Winer, A., Thomas, D., 2007. Assessing uncertainty in spatial exposure models for air pollution health effects assessment. Environ. Health Perspect. 115, 1147–1153.
Molitor, J., Papathomas, M., Jerrett, M., Richardson, S., 2010. Bayesian profile regression with an application to the National Survey of Children's health. Biostat. Oxf. Engl. 11, 484–498. http://dx.doi.org/10.1093/biostatistics/kxq013.
Molitor, J., Su, J.G., Molitor, N.-T., Rubio, V.G., Richardson, S., Hastie, D., Morello-Frosch, R., Jerrett, M., 2011. Identifying vulnerable populations through an examination of the association between multipollutant profiles and poverty. Environ. Sci. Technol. 45, 7754–7760. http://dx.doi.org/10.1021/es104017x.
Molitor, J., Brown, I.J., Chan, Q., Papathomas, M., Liverani, S., Molitor, N., Richardson, S., Van Horn, L., Daviglus, M.L., Dyer, A., Stamler, J., Elliott, P., for the INTERMAP Research Group, 2014a. Blood pressure differences associated with optimal macronutrient intake trial for heart health (OMNIHEART)-like diet compared with a typical American diet. Hypertension 64, 1198–1204. http://dx.doi.org/10.1161/HYPERTENSIONAHA.114.03799.
Molitor, J., Brown, I.J., Chan, Q., Papathomas, M., Liverani, S., Molitor, N., Richardson, S., Van Horn, L., Daviglus, M.L., Dyer, A., Stamler, J., Elliott, P., for the INTERMAP Research Group, 2014b. Blood pressure differences associated with optimal macronutrient intake trial for heart health (OMNIHEART)-like diet compared with a typical American diet. Hypertension 64, 1198–1204. http://dx.doi.org/10.1161/HYPERTENSIONAHA.114.03799.
Monn, C., 2001. Exposure assessment of air pollutants: a review on spatial heterogeneity and indoor/outdoor/personal exposure to suspended particulate matter, nitrogen dioxide and ozone. Atmos. Environ. 35, 1–32. http://dx.doi.org/10.1016/S1352-2310(00)00330-7.
Morello-Frosch, R., Shenassa, E.D., 2006. The environmental "riskscape" and social inequality: implications for explaining maternal and child health disparities. Environ. Health Perspect. 114, 1150–1153. http://dx.doi.org/10.1289/ehp.8930.
Morello-Frosch, R., Pastor, M., Sadd, J., 2001. Environmental justice and Southern California's "riskscape": the distribution of air toxics exposures and health risks among diverse communities. Urban Aff. Rev. 36, 551–578. http://dx.doi.org/10.1177/10780870122184993.
Morello-Frosch, R., Pastor, M., Porras, C., Sadd, J., 2002. Environmental justice and regional inequality in southern California: implications for future research. Environ. Health Perspect. 110 (Suppl. 2), 149–154.
Morello-Frosch, R., Jesdale, B.M., Sadd, J.L., Pastor, M., 2010. Ambient air pollution exposure and full-term birth weight in California. Environ. Heal. 9, 44. http://dx.doi.org/10.1186/1476-069X-9-44.
Morello-Frosch, R., Zuk, M., Jerrett, M., Shamasunder, B., Kyle, A.D., 2011. Understanding the cumulative impacts of inequalities in environmental health: implications for policy. Health Aff. (Millwood) 30, 879–887. http://dx.doi.org/10.1377/hlthaff.2011.0153.
Neal, R.M., 2000. Markov chain sampling methods for dirichlet process mixture models. J. Comput. Graph. Stat. 9, 249. http://dx.doi.org/10.2307/1390653.
Oakes, M., Baxter, L., Long, T.C., 2014. Evaluating the application of multipollutant exposure metrics in air pollution health studies. Environ. Int. 69, 90–99. http://dx.doi.org/10.1016/j.envint.2014.03.030.
Ouidir, M., Giorgis-Allemand, L., Lyon-Caen, S., Morelli, X., Cracowski, C., Pontet, S., Pin, I., Lepeule, J., Siroux, V., Slama, R., 2015. Estimation of exposure to atmospheric pollutants during pregnancy integrating space–time activity and indoor air levels: does it make a difference? Environ. Int. 84, 161–173. http://dx.doi.org/10.1016/j.envint.2015.07.021.
Papathomas, M., Molitor, J., Richardson, S., Riboli, E., Vineis, P., 2010. Examining the joint effect of multiple risk factors using exposure risk profiles: lung cancer in nonsmokers. Environ. Health Perspect. 119, 84–91. http://dx.doi.org/10.1289/ehp.1002118.
Papathomas, M., Molitor, J., Hoggart, C., Hastie, D., Richardson, S., 2012. Exploring data from genetic association studies using Bayesian variable selection and the dirichlet process: application to searching for gene × gene patterns: searching for gene × gene patterns. Genet. Epidemiol. 36, 663–674. http://dx.doi.org/10.1002/gepi.21661.
Pirani, M., Best, N., Blangiardo, M., Liverani, S., Atkinson, R.W., Fuller, G.W., 2015. Analysing the health effects of simultaneous exposure to physical and chemical properties of airborne particles. Environ. Int. 79, 56–64. http://dx.doi.org/10.1016/j.envint.2015.02.010.
Ponce, N.A., 2005. Preterm birth: the interaction of traffic-related air pollution with economic hardship in Los Angeles neighborhoods. Am. J. Epidemiol. 162, 140–148. http://dx.doi.org/10.1093/aje/kwi173.
Protano, C., Scalise, T., Orsi, G.B., Vitali, M., 2012. A systematic review of benzene exposure during pregnancy and adverse outcomes on intrauterine development and birth: still far from scientific evidence. Ann. Ig. Med. Prev. E Comunità 24, 451–463.
Qian, Z., Chapman, R.S., Hu, W., Wei, F., Korn, L.R., Zhang, J. (Jim), 2004. Using air pollution based community clusters to explore air pollution health effects in children. Environ. Int. 30, 611–620. http://dx.doi.org/10.1016/j.envint.2003.11.003.
Ritz, B., Wilhelm, M., 2008. Ambient air pollution and adverse birth outcomes: methodologic issues in an emerging field: ambient air pollution and adverse birth outcomes. Basic Clin. Pharmacol. Toxicol. 102, 182–190. http://dx.doi.org/10.1111/j.1742-7843.2007.00161.x.
Ritz, B., Wilhelm, M., Hoggatt, K.J., Ghosh, J.K.C., 2007. Ambient air pollution and preterm birth in the environment and pregnancy outcomes study at the University of California, Los Angeles. Am. J. Epidemiol. 166, 1045–1052. http://dx.doi.org/10.1093/aje/kwm181.
Ross, Z., Ito, K., Johnson, S., Yee, M., Pezeshki, G., Clougherty, J.E., Savitz, D., Matte, T., 2013. Spatial and temporal estimation of air pollutants in New York City: exposure assignment for use in a birth outcomes study. Environ. Health Glob. Access Sci. Source 12, 51. http://dx.doi.org/10.1186/1476-069X-12-51.
Rue, H., Martino, S., n.d. Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with discussion). J. R. Stat. Soc..

Rue, H., Martino, S., Lindgren, F., Simpson, D., Riebler, A., Krainski, E., 2014. Functions which Allow to Perform Full Bayesian Analysis of Latent Gaussian Models Using Integrated Nested Laplace Approximaxion [WWW Document]. URL http://inla.googlecode.com/hg-history/default/rinla/DESCRIPTION.

Salam, M.T., Millstein, J., Li, Y.-F., Lurmann, F.W., Margolis, H.G., Gilliland, F.D., 2005. Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: results from the children's health study. Environ. Health Perspect. 113, 1638–1644.

Schmidt, P., Muller, E., 1978. The problem of multicollinearity in multistage causal alienation model: a comparison of ordinary least squares, maximum likelihood and Ridge estimation. Qual. Quant. 12, 267–297.

Shenassa, E.D., Stubbendick, A., Brown, M.J., 2004. Social disparities in housing and related pediatric injury: a multilevel study. Am. J. Public Health 94, 633–639.

Sheppard, L., Burnett, R.T., Szpiro, A.A., Kim, S.-Y., Jerrett, M., Pope, C.A., Brunekreef, B., 2012. Confounding and exposure measurement error in air pollution epidemiology. Air Qual. Atmos. Health 5, 203–216. http://dx.doi.org/10.1007/s11869-011-0140-9.

Snowden, J.M., Reid, C.E., Tager, I.B., 2015. Framing Air Pollution Epidemiology in Terms of Population Interventions, With Applications to Multipollutant Modeling: Epidemiology 26, 271–279. http://dx.doi.org/10.1097/EDE.0000000000000236.

Spengler, J., Schwab, M., Ryan, P.B., Colome, S., Wilson, A., Billick, I., Becker, E., 1994. Personal exposure to nitrogen dioxide in the Los Angeles Basin. Air Waste 44, 39–47. http://dx.doi.org/10.1080/1073161X.1994.10467236.

Su, J.G., Jerrett, M., Beckerman, B., Wilhelm, M., Ghosh, J.K., Ritz, B., 2009. Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. Environ. Res. 109, 657–670. http://dx.doi.org/10.1016/j.envres.2009.06.001.

Su, J.G., Jerrett, M., Morello-Frosch, R., Jesdale, B.M., Kyle, A.D., 2012. Inequalities in cumulative environmental burdens among three urbanized counties in California. Environ. Int. 40, 79–87. http://dx.doi.org/10.1016/j.envint.2011.11.003.

Tsai, M.-Y., Hoek, G., Eeftens, M., de Hoogh, K., Beelen, R., Beregszászi, T., Cesaroni, G., Cirach, M., Cyrys, J., De Nazelle, A., de Vocht, F., Ducret-Stich, R., Eriksen, K., Galassi, C., Gražuleviciene, R., Gražulevicius, T., Grivas, G., Gryparis, A., Heinrich, J., Hoffmann, B., Iakovides, M., Keuken, M., Krämer, U., Künzli, N., Lanki, T., Madsen, C., Meliefste, K., Merritt, A.-S., Mölter, A., Mosler, G., Nieuwenhuijsen, M.J., Pershagen, G., Phuleria, H., Quass, U., Ranzi, A., Schaffner, E., Sokhi, R., Stempfelet, M., Stephanou, E., Sugiri, D., Taimisto, P., Tewis, M., Udvardy, O., Wang, M., Brunekreef, B., 2015. Spatial variation of PM elemental composition between and within 20 European study areas — results of the ESCAPE project. Environ. Int. 84, 181–192. http://dx.doi.org/10.1016/j.envint.2015.04.015.

Vrijheid, M., Slama, R., Robinson, O., Chatzi, L., Coen, M., van den Hazel, P., Thomsen, C., Wright, J., Athersuch, T.J., Avellana, N., Basagaña, X., Brochot, C., Bucchini, L., Bustamante, M., Carracedo, A., Casas, M., Estivill, X., Fairley, L., van Gent, D., Gonzalez, J.R., Granum, B., Gražulevičienė, R., Gutzkow, K., Julvez, J., Keun, H.C., Kogevinas, M., McEachan, R.R.C., Meltzer, H.M., Sabidó, E., Schwarze, P.E., Siroux, V., Sunyer, J., Want, E., Zeman, F., Nieuwenhuijsen, M.J., 2014. The human early-life exposome (HELIX): project rationale and design. Environ. Health Perspect. http://dx.doi.org/10.1289/ehp.1307204.

Wagner, J.G., Morishita, M., Keeler, G.J., Harkema, J.R., 2012. Divergent effects of urban particulate air pollution on allergic airway responses in experimental asthma: a comparison of field exposure studies. Environ. Heal. 11, 45. http://dx.doi.org/10.1186/1476-069X-11-45.

Wang, Y., Ying, Q., Hu, J., Zhang, H., 2014. Spatial and temporal variations of six criteria air pollutants in 31 provincial capital cities in China during 2013–2014. Environ. Int. 73, 413–422. http://dx.doi.org/10.1016/j.envint.2014.08.016.

Wilhelm, M., Ritz, B., 2005. Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California. USA. Environ. Health Perspect. 113, 1212–1221.

Wilhelm, M., Ghosh, J.K., Su, J., Cockburn, M., Jerrett, M., Ritz, B., 2011a. Traffic-related air toxics and term low birth weight in Los Angeles County. California. Environ. Health Perspect. 120, 132–138. http://dx.doi.org/10.1289/ehp.1103408.

Wilhelm, M., Ghosh, J.K., Su, J., Cockburn, M., Jerrett, M., Ritz, B., 2011b. Traffic-related air toxics and preterm birth: a population-based case-control study in Los Angeles county. California. Environ. Health 10, 89. http://dx.doi.org/10.1186/1476-069X-10-89.

Zanobetti, A., Austin, E., Coull, B.A., Schwartz, J., Koutrakis, P., 2014. Health effects of multi-pollutant profiles. Environ. Int. 71, 13–19. http://dx.doi.org/10.1016/j.envint.2014.05.023.

Zhang, K., Batterman, S., 2013. Air pollution and health risks due to vehicle traffic. Sci. Total Environ. 450-451, 307–316. http://dx.doi.org/10.1016/j.scitotenv.2013.01.074.

Zhu, X., Fan, Z. (tina), Wu, X., Jung, K.H., Ohman-Strickland, P., Bonanno, L.J., Lioy, P.J., 2011. Ambient concentrations and personal exposure to polycyclic aromatic hydrocarbons (PAH) in an urban community with mixed sources of air pollution. J. Expo. Sci. Environ. Epidemiol. 21, 437–449. http://dx.doi.org/10.1038/jes.2011.2.

Zuurbier, M., Hoek, G., Oldenwening, M., Lenters, V., Meliefste, K., van den Hazel, P., Brunekreef, B., 2010. Commuters' exposure to particulate matter air pollution is affected by mode of transport, fuel type, and route. Environ. Health Perspect. 118, 783–789. http://dx.doi.org/10.1289/ehp.0901622.