

UC Berkeley

UC Berkeley Previously Published Works

Title

Identifying Vulnerable Populations through an Examination of the Association Between Multipollutant Profiles and Poverty

Permalink

<https://escholarship.org/uc/item/49f6f7v9>

Journal

Environmental Science and Technology, 45(18)

ISSN

0013-936X

Authors

Molitor, John
Su, Jason G
Molitor, Nuoo-Ting
[et al.](#)

Publication Date

2011-09-15

DOI

10.1021/es104017x

Peer reviewed

Identifying Vulnerable Populations through an Examination of the Association Between Multipollutant Profiles and Poverty

John Molitor,^{*,†,‡} Jason G. Su,[§] Nuoo-Ting Molitor,[‡] Virgilio Gómez Rubio,^{||} Sylvia Richardson,[†] David Hastie,[†] Rachel Morello-Frosch,[§] and Michael Jerrett[§]

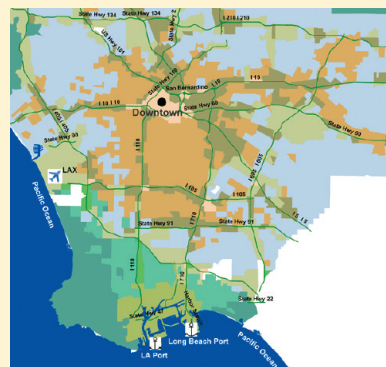
[†]Department of Epidemiology and Biostatistics, Imperial College, London, United Kingdom

[‡]Department of Public Health, Oregon State University, Corvallis, Oregon 97331, United States

[§]Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, California 94720, United States

^{||}Departamento de Matemáticas (Estadística e I.O.), Instituto de Matematica Aplicada a la Ciencia y la Ingenieria E.T.S.I. Industriales, Universidad de Castilla-La Mancha, Avda. Camilo José Cela, 313071 Ciudad Real Spain

ABSTRACT: Recently, concerns have centered on how to expand knowledge on the limited science related to the cumulative impact of multiple air pollution exposures and the potential vulnerability of poor communities to their toxic effects. The highly intercorrelated nature of exposures makes application of standard regression-based methods to these questions problematic due to well-known issues related to multicollinearity. Our paper addresses these problems by using, as its basic unit of inference, a profile consisting of a pattern of exposure values. These profiles are grouped into clusters and associated with a deprivation outcome. Specifically, we examine how profiles of NO₂, PM_{2.5}, and diesel- (road and off-road) based exposures are associated with the number of individuals living under poverty in census tracts (CT's) in Los Angeles County. Results indicate that higher levels of pollutants are generally associated with higher poverty counts, though the association is complex and nonlinear. Our approach is set in the Bayesian framework, and as such the entire model can be fit as a unit using modern Bayesian multilevel modeling techniques via the freely available WinBUGS software package,¹ though we have used custom-written C++ code (validated with WinBUGS) to improve computational speed. The modeling approach proposed thus goes beyond single-pollutant models in that it allows us to determine the association between entire multipollutant profiles of exposures with poverty levels in small geographic areas in Los Angeles County.



1. INTRODUCTION

Growing health disparities exist in the United States between persons with high socioeconomic position (SEP) and people of color who are of lower socioeconomic position. These disparities translate into higher rates for mortality, morbidity, and disability for the lower SEP groups and people of color.² Although these health disparities are frequently attributed to individual health behaviors such as smoking, individual factors account for only a fraction of the overall inequalities between these groups.³ As such, social epidemiology research has focused on the effects of SEP on many health outcomes, differential access to health and social services,⁴ and neighborhood or community characteristics that may promote or adversely affect health.⁵ Researchers and policy makers concerned about environmental justice also investigate whether disadvantaged groups experience higher environmental exposures. These studies generally report that areas with a greater portion of nonwhite, low-income, and poor residents face higher single and cumulative environmental exposures.^{2,6,7} Environmental justice researchers argue that socially disadvantaged groups bear a greater environmental exposure burden and are more susceptible to the effects of these exposures due to factors such as psychosocial stressors, underlying health

conditions, and occupational exposures.⁸ These disparities in environmental exposures are increasingly recognized as potential determinants of health inequities.^{9,10}

These health inequities have led researchers to recognize the need for knowledge about the cumulative health effects of multiple environmental hazards such as exposure to multiple air pollutants¹¹ and the potential vulnerability of people in poor communities who suffer from their toxic effects.^{11,12} As a result, environmental justice advocates have urged regulators to develop scientifically valid indicators from a multipollutant approach to examine environmental health inequities and guide decision making.^{13,14} Developing such measures is complicated due to issues of comparison and high levels of correlation among some pollutants. For example, the highly intercorrelated nature of air pollutants makes it difficult to examine their combined effects on health-related outcomes and measures of deprivation.¹⁵ As such, epidemiological studies have traditionally focused on

Received: November 30, 2010

Accepted: July 28, 2011

Revised: June 4, 2011

Published: July 28, 2011

single-pollutant models that use regression-based techniques to examine the marginal association between a pollutant and an outcome. Pollutants, however, occur in complex mixtures consisting of highly correlated combinations of individual exposures. For example, evidence for synergy among pollutants in causing health effects has been recently reviewed by Mauderly and Samet.¹⁶ Several studies have confirmed that synergisms exist between O₃ and other pollutants in laboratory studies involving humans and animals.¹⁷ Thus, the highly correlated nature of air pollution exposures and the potential for synergies raise questions about the adequacy of marginal, single-pollutant models.

In this paper, we propose a Bayesian statistical approach that examines association between several components of air pollution exposures and poverty. Our approach addresses these problems by using, as its basic unit of inference, a profile consisting of a sequence of exposure values. These profiles are grouped into clusters and associated with poverty levels in small geographic areas in Los Angeles County.

2. MATERIALS AND METHODS

Our overall approach is to cluster joint patterns of air pollution exposures, denoted as an air pollution *profile*, and relate these clusters to measure of deprivation. The multipollutant profile approach developed will be applied to estimates of air pollution concentrations for NO₂, PM_{2.5}, diesel (road), and diesel (off-road) exposures obtained using a recently published paper.¹⁸ Briefly, NO₂ estimates were obtained via a land-use regression. The variables used for NO₂ model selection included land-use information (e.g., commercial and industrial), road network, traffic, population distribution, physical properties, and remote sensing-derived greenness and soil brightness. We used a land-use regression model to estimate NO₂ (ppb) levels. Methods are described in detail elsewhere.¹⁵ In brief, the model was based on field measurements at 201 locations in Los Angeles. The measurements were obtained in summer for 2006 and winter for 2007, each for a 2-week period closest to the seasonal mean. These measurements were averaged to represent the annual mean. Some monitors were stolen or vandalized, leaving 181 sites for the analysis. Sixteen measurements were chosen at random to use as cross-validation sites. The measurements from the remaining 167 sites then were used as the dependent variables in a spatial land-use regression model with traffic, land use, population, and physical geography as predictors of pollution levels. The model was highly predictive; the R² between the field-measured and the predicted pollutant level was 86% with similar performance at the out-of-sample cross-validation sites predictions. To estimate PM_{2.5} exposure, we interpolated from 23 state and local district monitoring stations in the Los Angeles basin for year 2000 with a universal kriging algorithm.¹⁹

On-road and off-road diesel exposures were estimated through the Assessment System for Population Exposure Nationwide (ASPEN) with a Gaussian dispersion that accounted for meteorological conditions, wind speed, and atmospheric chemistry.²⁰ Exposure concentrations were obtained at the CT level, and measures of deprivation such as the number of people living under the poverty level and number of people of color were obtained for each CT from the US Census Bureau for year 2000.

The methods developed utilize recently developed Bayesian dimension-reduction and clustering techniques that will characterize the pollutant patterns. This overall approach has been used in a recent paper,²¹ which profiles health-related variables

associated with measures of mental health in children. The multipollutant profile approach adopts a global point of view, where inference is based on the joint pattern of pollution exposures. The methodology consists of the following two key components discussed in detail in the following sections. First, an exposure profile assignment submodel assigns individual multipollutant profiles to clusters via a normal mixture model. Clusters consist of exposures patterns that are similar to each other and are consistent with cluster-specific parameters, such as cluster means that help identify the kind of profiles that reside within each group. Second, a response submodel links clusters of exposure profiles to measures of deprivation via a regression model. The model is formulated in a Bayesian context, and as such all components of the modeling framework will be fitted jointly using Markov chain Monte Carlo methods.²² Though the models can be fit using the freely available software WinBUGS,²³ we have written custom C++ code to speed up computations. The models were run for a total of 100 000 iterations with 20 000 iterations discarded for burn in. Visual inspection of MCMC output revealed that the model mixed well, and shorter runs gave very similar results, indicating that convergence was not a problem.

Exposure Profile Assignment Submodel. Our basic data structure consists of, for each CT, i , a covariate profile, $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$, where each covariate, x_{ip} , $p = 1, \dots, P$, within each multipollutant profile denotes a measure of exposure for pollutant p in area i . We first construct an allocation submodel of the probability that an individual area is assigned to a particular cluster. The basic model we use to cluster profiles is a standard multivariate normal mixture model, which is described by Jain and Neal²⁴ and Neal.²⁵ Our mixture model incorporates a Dirichlet process prior to the mixing distribution. For further background information regarding mixture models with Dirichlet process priors, see Green and Richardson.²⁶ Profiles of areas are grouped into clusters, and an allocation variable, $Z_i = C$, indicates the C th cluster to which area, i , belongs. Our assignment submodel is then

$$f(x_i) = \sum_{c=1}^C \psi_c f(x_i | \mu_c, \Sigma_c) \quad (1)$$

where $f(x_i | \mu_c, \Sigma_c)$ denotes a multivariate normal distribution with location parameters $\mu_c = \mu_c^1, \dots, \mu_c^P$ and covariance matrix Σ_c . The mixture weights ψ_c , $c = 1, \dots, C$, are given a “stick-breaking prior”²⁷ with clustering parameter α . Since we have little a priori information regarding the specification of α , we assign it a vague exponential prior. As is often done in the Dirichlet process literature, we impose a finite maximum number of clusters; however, since clusters can be empty, the submodel provides flexibility in terms of the number of clusters actually required and used for the data set at hand. This is important as it incorporates cluster uncertainty into the modeling process both in terms of assignment of individual areas to clusters and in terms of number of clusters used. This flexibility also frees the analyst from prespecifying a fixed number of clusters based on an ad-hoc decision rule. Note that for all analyses in this paper, we set the maximum number of clusters allowed at $C = 20$, which we found to be sufficiently large for our applications.

Because it is possible that clusters will be empty, we cannot assign noninformative, “flat”, priors to cluster parameters. Therefore, we adopt an empirical Bayes approach and assign a prior for the mean of each pollutant across clusters as $\mu_c^p \approx N(\nu^p, \phi^p)$, where each ν^p is set to the observed empirical average, \bar{x}^p , but

each ϕ^p is set equal to the square of the empirical range squared as suggested in Richardson and Green.²⁸ Similarly, we assign a Wishart prior for the precision matrices as $\Sigma_c^{-1} \approx \text{Wish}(\rho R, \rho)$, where $\rho = P$. Since under this formulation the mean of the Wishart distribution is $E(\Sigma_c^{-1}) = R^{-1}$, we set R to the empirical variance, namely, $R = \Sigma$. Note in our model formulation cluster hyperparameters are assumed to come from distributions centered on empirical averages. Thus cluster-specific parameters are used to represent subgroups that deviate from a single empirically derived population.

Deprivation Submodel. This submodel uses the allocation variables defined for the exposure profile submodel above, namely, for each CT, i , $z_i = c, c = 1, \dots, C$, indicates the cluster to which individual i belongs. However, in this submodel, the c th cluster is assigned a random-effect parameter that measures the cluster's influence on the outcome (on the logistic scale) denoted as θ_c . Since it is possible for a particular θ_c to be associated with an empty cluster, these parameters must be assigned a proper prior. Therefore, we assign to each θ_c a proper t density function with 7 degrees of freedom and scale 2.5 as a prior, as discussed by Gelman et al.,²⁹ which corresponds to the baseline case of one-half of a success and one-half of a failure for a single binomial trial with probability $p = \text{logit}^{-1}(\theta_c)$. Our response model, which links the clusters with poverty counts, y_i for CT, i , is simply $y_i \approx \text{Bin}(n_i, p_i)$ with

$$\text{logit}(p_i) = \theta_{z_i} + \varepsilon_i \quad (2)$$

where $\varepsilon_i \approx N(0, \sigma^2)$ represents unexplained CT-level variation in the outcome not explained by air pollution exposures, n_i indicates the number of individuals in CT, i , and p_i indicates that probability that a randomly chosen individual in CT, i , is living under the poverty line.

At each iteration of the sampler, we define $V^\theta = \text{Var}(\theta_{z_i})$ and $V^\varepsilon = \text{Var}(\varepsilon_i)$ across all regions so we can then obtain a posterior distribution for the overall amount of variation in deprivation explained by air pollution clusters versus unexplained residual error defined as

$$\rho = V^\theta / (V^\theta + V^\varepsilon) \quad (3)$$

Note that the posterior distribution for ρ is *not* based on an optimal clustering referred to above but rather represents the ratio obtained by *model averaging* through the entire MCMC output, thus properly taking into account uncertainty regarding cluster assignment and the number of clusters used.

At each iteration of our algorithm CTs are grouped into a relatively small number of clusters with a set of parameters associated with each cluster. However, as cluster membership and the number of clusters used changes from iteration to iteration, the iterative process will create, for each CT, a unique posterior distribution for each parameter of interest, such as air pollution risk and modeled exposure parameters (NO_2 , $\text{PM}_{2.5}$, road and off-road diesel). This feature of using shared cluster parameters estimated at each iteration of the model fitting process to form unique posterior distributions at the “individual” level is well known as Bayesian partition models. For an overview of these models, see Denison et al.³⁰

Finding Clustering That Best Fits the Data. One important aspect of our flexible Bayesian modeling framework is that our model implementation allows the number of air pollution exposure clusters to change from iteration to iteration of the MCMC sampler, and this added flexibility leads to a rich output

that requires careful interpretation. The goal is to find the “typical” or “best” way in which the algorithm groups profiles into clusters and then process this best partition using modern Bayesian model-averaging techniques that utilize the entire output from the MCMC sampler. Consistent clustering throughout the sampling process will be associated with greater certainty regarding subgroup parameter estimates, such as disease risk, leading to narrower posterior credible intervals. This is important because while even noisy data will exhibit a “best” clustering, the Bayesian model-averaging techniques will reveal little confidence in the clustering of the best partition of the noisy data as it will not generally coincide with the way individual observations are clustered at each iteration from the sampler.

To determine this optimal clustering, we construct a correlation matrix with cells that correspond to the percentage of MCMC iterations in which two CTs are assigned to the same cluster and then find a single clustering that best represents this final matrix (see Molitor et al.³¹ and Dahl et al.³² for details). Despite the fact that a single “best” clustering is computed, uncertainty regarding cluster parameters is still determined by model averaging through all the different clustering found throughout the entire run. This model-averaging approach will generally produce smaller posterior errors for cluster parameters when the clustering is consistent from iteration to iteration and reveal larger errors in the presence of “noisy” data that produce inconsistent, haphazard clustering.

3. RESULTS

Since we are interested in the joint distribution of exposures we examine the optimal clustering obtained using the profile-based Bayesian modeling approach with the number of individuals living under the poverty line per CT as the outcome. The optimal partition is displayed in Figure 1, with mean values and posterior credible intervals listed in Table 1 and graphically displayed in Figure 3. In Table 1, clusters are sorted according to poverty risk. Note that in Table 1 we highlight at the top the value of $\rho = V^\theta / (V^\theta + V^\varepsilon)$, denoted in eq 3, which indicates the proportion of variance explained by the air pollution clusters relative to the residual error. CT clusters with statistically significant air pollution effects are displayed in Figure 2.

In Table 1 the value of $\rho = 0.79$ (0.47, 0.97) reveals that variation in air pollution exposures throughout Los Angeles County coincide with variation in poverty levels. If we examine the clusters significantly associated with poverty in Figure 2, we see that populations living in the port neighborhoods of Los Angeles and Long Beach mainly suffer from nonroad diesel impacts, probably from goods movement vessels (Cluster 9). Further, the roadways (Cluster 7) exhibit higher than average levels of NO_2 , $\text{PM}_{2.5}$, and road diesel, while the high-traffic area of downtown Los Angeles (Cluster 10) exhibits higher than average levels of all pollutants. These results reveal that people who live in the port neighborhoods of Los Angeles and Long Beach, the main artery closeby, the Los Angeles downtown core, and the central areas not only suffer from poverty but also face significant pollution impacts from multiple air pollutants.

In general, the results depicted in Table 1 and Figure 3 reveal that areas with higher levels of air pollution exposures are associated with higher levels of poverty. However, the association between air pollutants and poverty is not entirely linear. For example, Cluster 9 (L.A. ports) is more deprived than Cluster 7 (roads) as it has a higher poverty risk, 0.28 (0.24, 0.32) versus

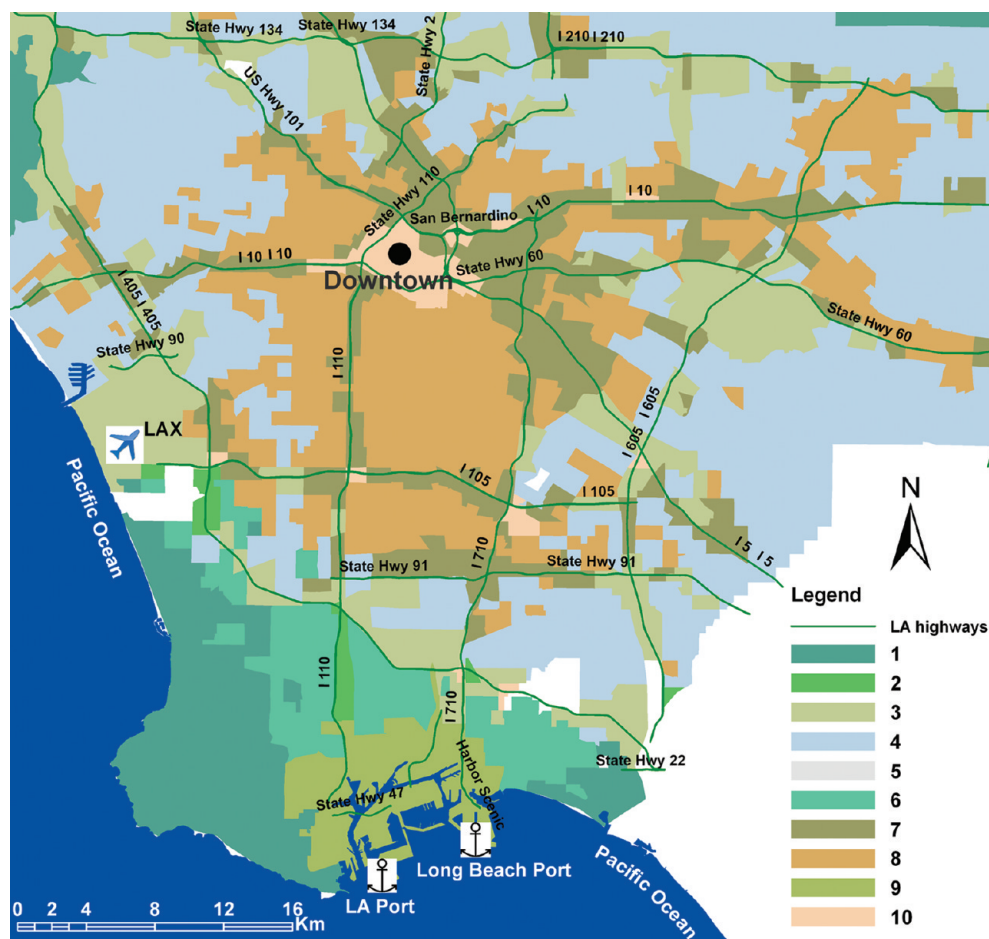


Figure 1. Optimal clusters as defined in Materials and Methods section.

0.23 (0.22, 0.25). While the marginal C.I.'s for cluster risks just barely overlap, the joint probability that the risk for Cluster 9 is greater than Cluster 7 is significant as $\Pr(p_9 > p_7) = 0.99$. As one might expect, the more-deprived Cluster 9 has much higher levels of off-road diesel emissions, 7.91 (6.53, 9.27), compared to the less-deprived Cluster 7, 1.29 (1.26, 1.33). However, the more-deprived Cluster 9 has lower levels of NO_2 , $\text{PM}_{2.5}$, and road diesel, with exposure values of 20.64 (19.42, 21.84), 16.67 (16.07, 17.30), and 0.90 (0.74, 1.08) compared to exposure values for Cluster 7 of 26.69 (26.22, 27.14), 21.68 (21.55, 21.81), and 1.21 (1.15, 1.26). Thus, what is different between Clusters 9 and 7 cannot be summarized by an additive effect of all pollutants but is instead related to a contrast between off-road diesel and other pollutants. Similar remarks can be made for other clusters. For example, the relatively more-deprived Cluster 10 (downtown) has higher levels of NO_2 and road and off-road diesel emissions, 32.60 (30.42, 34.18), 2.49 (2.10, 2.89), and 1.80 (1.60, 2.03), compared with lower emission levels corresponding to the relatively less-deprived Cluster 8 (central Los Angeles, off-roads), 24.18 (23.90, 24.47), 0.72 (0.70, 0.74), and 1.42 (1.39, 1.46). The levels of $\text{PM}_{2.5}$ however are nearly the same and not statistically different with values of 21.94 (21.58, 22.31) for Cluster 10 versus 21.70 (21.63, 21.77) for Cluster 8, despite differences in poverty levels. Therefore, while it might be generally true that increased poverty is associated with increased air pollution exposures, the nature of these associations in Los Angeles County is complex and nonlinear.

4. DISCUSSION

There has been increased interest in the air pollution literature in the examination of the combined effect of air pollution and poverty. In this paper, we examined the joint effects of air pollution mixtures to help identify vulnerable populations in Los Angeles County. The results showed a general relationship between elevated levels of air pollution exposures and poverty. The results also revealed that the relationship is complex in that poverty levels do not increase linearly with increased levels of exposure, as is assumed when such relationships are examined using linear additive regression models. The approach employed here examined the combined effects of several air pollutants on poverty, revealing vulnerable populations were not always subject to elevated levels of different exposures uniformly but rather different combinations of exposure levels were associated with different subgroups of poverty populations as explained in the Results section.

The approach used here clusters exposure profiles into risk groups that were then associated with poverty. The flexible MCMC-based parameter estimation techniques allowed the assignment of exposure profiles to risk groups and the number of risk groups to vary throughout the run of the sampler. The results displayed exploratory "best" clustering of profiles along with more robust results obtained from the model averaging through the clustering patterns obtained from the sampler. The approach identifies cumulative environmental hazard inequalities

Table 1. Modeled Values for Air Pollution/Poverty Clusters^a

cluster	percentage of poverty explained by air pollution clusters $\rho = 0.79$ (0.47, 0.97)				
	NO ₂ ($\bar{x} = 22.33$)	PM _{2.5} ($\bar{x} = 20.25$)	road diesel ($\bar{x} = 0.77$)	off-road diesel ($\bar{x} = 1.36$)	risk ($\sigma = 0.17$)
1 (n = 192)	15.50 (14.94, 16.13)	17.00 (16.63, 17.39)	0.45 (0.43, 0.49)	1.12 (1.04, 1.20)	0.05 (0.04, 0.06)
2 (n = 12)	22.08 (19.70, 24.19)	19.33 (18.16, 20.34)	1.37 (0.96, 1.81)	1.82 (1.22, 3.06)	0.08 (0.05, 0.13)
3 (n = 203)	22.18 (21.40, 22.90)	20.18 (19.89, 20.45)	0.96 (0.90, 1.01)	1.09 (1.04, 1.16)	0.10 (0.10, 0.11)
4 (n = 543)	21.82 (21.55, 22.09)	21.22 (21.10, 21.33)	0.60 (0.59, 0.62)	1.08 (1.06, 1.10)	0.11 (0.10, 0.11)
5 (n = 72)	16.77 (15.56, 18.02)	12.02 (10.92, 13.27)	0.33 (0.29, 0.38)	0.62 (0.53, 0.73)	0.13 (0.12, 0.16)
6 (n = 178)	19.95 (19.38, 20.61)	18.46 (18.16, 18.77)	0.59 (0.56, 0.65)	1.54 (1.41, 1.67)	0.16 (0.15, 0.18)
7 (n = 285)	26.69 (26.22, 27.14)	21.68 (21.55, 21.81)	1.21 (1.15, 1.26)	1.29 (1.26, 1.33)	0.23 (0.22, 0.25)
8 (n = 479)	24.18 (23.90, 24.47)	21.70 (21.63, 21.77)	0.72 (0.70, 0.74)	1.42 (1.39, 1.46)	0.25 (0.24, 0.26)
9 (n = 38)	20.64 (19.42, 21.84)	16.67 (16.07, 17.30)	0.90 (0.74, 1.08)	7.91 (6.53, 9.27)	0.28 (0.24, 0.32)
10 (n = 36)	32.60 (30.42, 34.81)	21.94 (21.58, 22.31)	2.49 (2.10, 2.89)	1.80 (1.60, 2.03)	0.34 (0.29, 0.38)

^a Bold rows indicate clusters with statistically significant poverty risks.

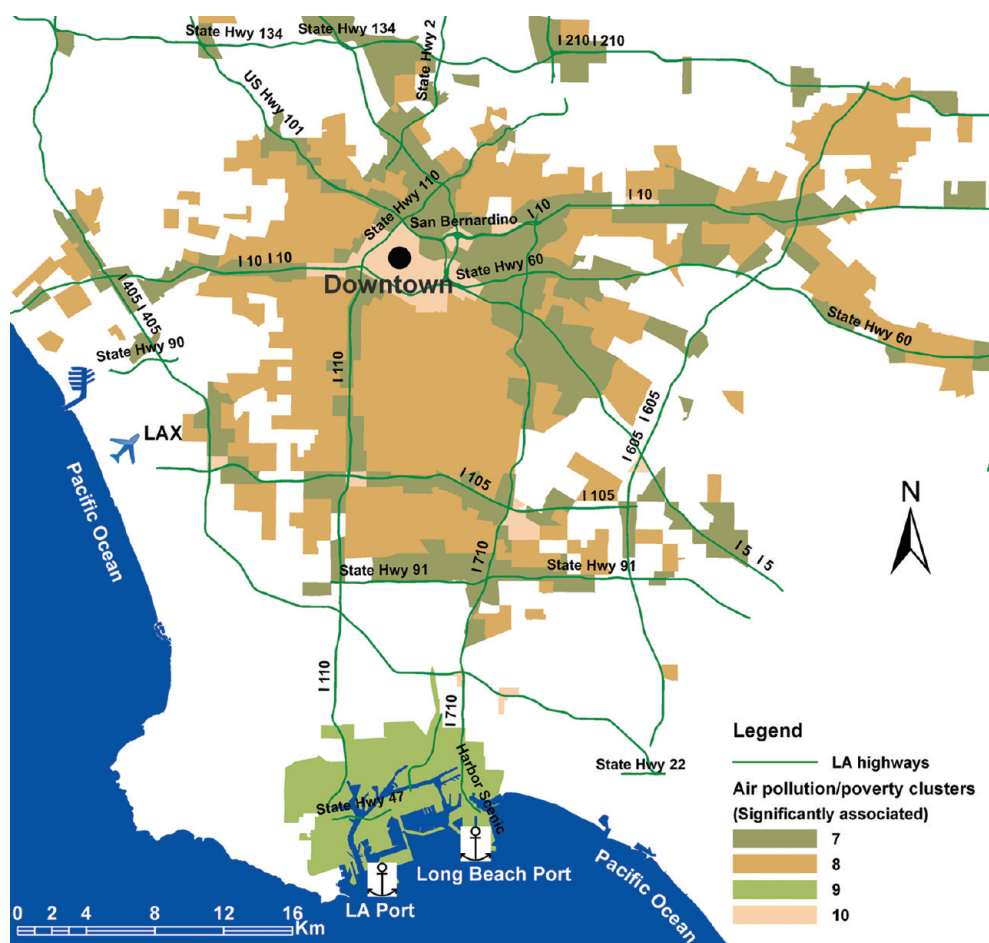


Figure 2. Clusters with statistically significant association with poverty.

within a region at the CT level. It further extends the framework that identified cumulative environmental risks at the regional level.¹⁸

Not surprisingly, the results often display a decidedly non-linear pattern, as some clusters display extremely high values for one pollutant but average or below average values for other pollutants. Unlike a conventional linear model approach, the clustering approach applied here allows one to examine the manner in which pollutants vary together.

The empirical results we observed here are broadly consistent with the literature on environmental hazard inequalities in the United States (see Morello-Frosch, Zuk et al.³³ for a recent review). Substantial evidence now suggests that numerous environmental hazards, including air pollution, are worse in poor neighborhoods and places with high proportions of racial or ethnic minority groups. In an international context, the findings from this study fit within a fairly consistent pattern, suggesting

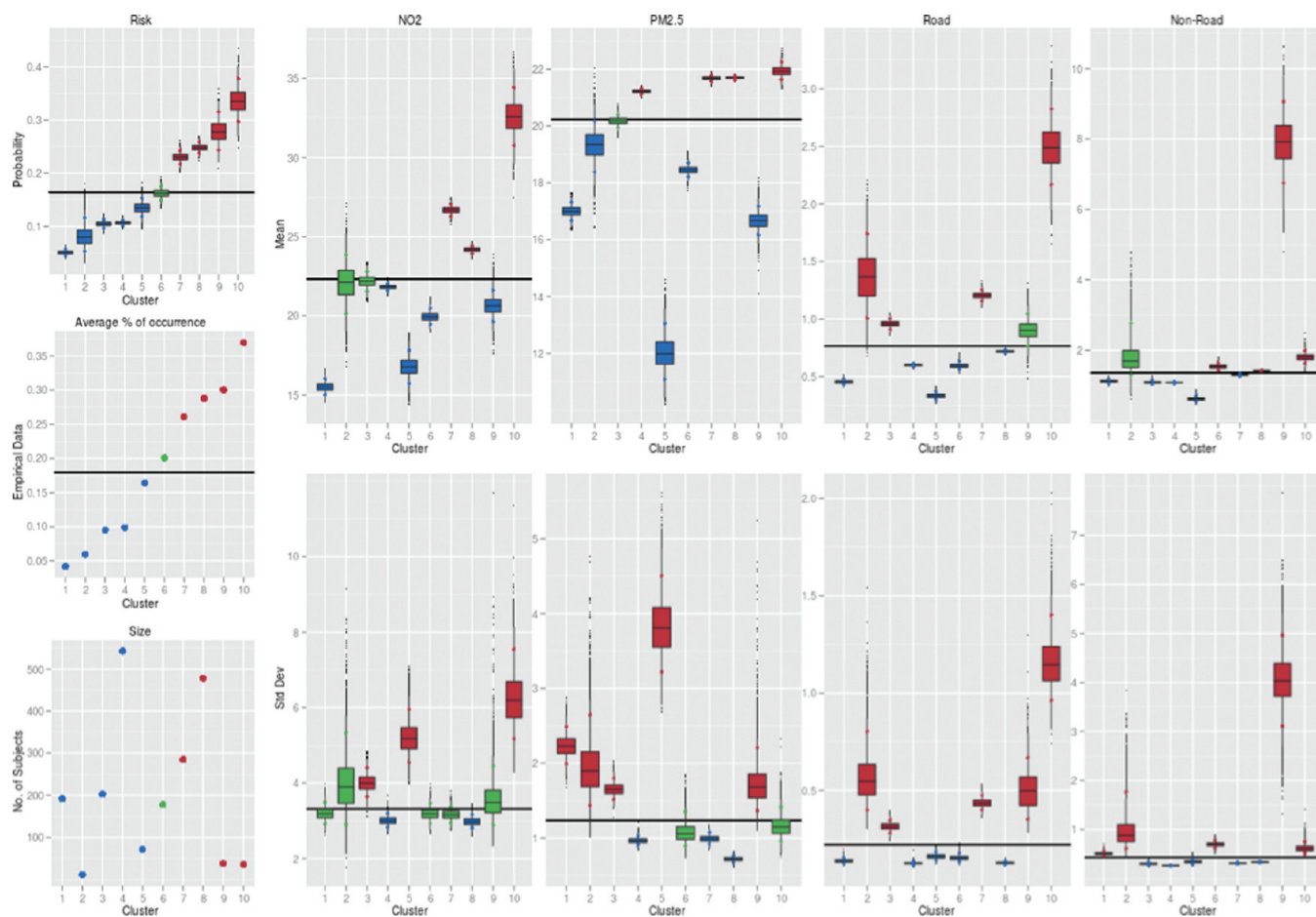


Figure 3. Cluster summaries. Left vertical panel: (a) Risk associated with each cluster, (b) average empirical proportion of individuals living under poverty averaged over CT's in the optimal clustering, (c) optimal cluster sizes. Right four vertical panels: (top four) Boxplots corresponding to cluster means for each pollutant/cluster, (bottom four) boxplots corresponding to cluster standard deviations for each pollutant/cluster.

that air pollution and other environmental risks remain unequally distributed with an inverse social gradient. Even in economically advanced countries with many income and social equalization programs, universal access to health care, and some of the highest life expectancies, air pollution and other environmental risks remain unequally distributed with an inverse social gradient. Unequal distributions of traffic pollution by race and SEP have been documented, albeit with more mixed results than in the United States. Pearce et al.³⁴ used atmospheric dispersion modeling to demonstrate a relationship between traffic pollution and disadvantaged social groups in New Zealand, finding higher levels of air pollution in areas of relatively high deprivation. In England, Brainard et al.³⁵ found that CO and NO₂, both markers of traffic pollution, related strongly to racial and ethnic minority status and to social deprivation. In Sweden, Chaix et al.³⁶ investigated the distribution of NO₂ in relation to young children. They reported higher levels of NO₂ for children living in poorer housing and neighborhoods. A Canadian study based on a land-use regression prediction of NO₂ in Toronto reported that lower SEP was related to higher air pollution exposures, but there were exceptions that contrasted with the U.S. literature.³⁷ For example, racial minority groups tended to be less exposed in Toronto than other groups, probably due to the city's role as a gateway city for highly educated immigrants. Dwelling values also took an unexpected positive sign, which may have been partly

explained by the dense urban structure of the downtown area and the relatively high traffic and land rents in this district. Similar diversions from the pattern were reported in an Italian study.³⁸ These subtle differences highlight the need not only to examine the specific intricacies of place but also to employ methods used in this paper, which may elucidate more subtle patterns and relationships.

The profiling method used in this paper can be extended to examine joint distributions of pollution, SEP, and health. For this initial demonstration we limited the investigation to exploring the joint distribution of poverty and several important air pollutants. In the future, we will extend this to examine health outcomes, which may reveal complex and subtle interactions between multiple or cumulative exposures, SEP, and health. Such investigations will have implications for public health policy because they may lead to better protection for vulnerable populations who experience high cumulative exposures, high susceptibilities, or both. At present, current public health protections do not take into account these cumulative exposures and susceptibilities, which may be significant contributing factors to observed health inequalities that follow social gradients.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: john.molitor@oregonstate.edu.

ACKNOWLEDGMENT

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. J.M. is also supported by Medical Research Council grant G0600609.

REFERENCES

- (1) Spiegelhalter, D.; Thomas, A.; Best, N. *WinBUGS version 1.4 user manual*; MRC Biostatistics Unit: Cambridge, 2003.
- (2) Brulle, R. J.; Pellow, D. N. Environmental justice: human health and environmental inequalities. *Annu. Rev. Public Health* **2006**, *27*, 103–24.
- (3) Lantz, P. M.; Lynch, J. W.; House, J. S.; Lepkowski, J. M.; Mero, R. P.; Musick, M. A.; Williams, D. R. Socioeconomic disparities in health change in a longitudinal study of US adults: the role of health-risk behaviors. *Soc. Sci. Med.* **2001**, *53* (1), 29–40.
- (4) Shumka, L.; Benoit, C. Social suffering and gaps in alternative health care for vulnerable women workers. *Res. Sociol. Health Care* **2007**, *25*, 253–275.
- (5) Phuong, D. D.; Finch, B. K.; Basurto-Davila, R.; Bird, C.; Escarce, J.; Lurie, N. Does place explain racial health disparities? Quantifying the contribution of residential context to the Black/white health gap in the United States. *Social Sci. Med.* **2008**, *67*, 1258–1268.
- (6) Morello-Frosch, R.; Pastor, M., Jr.; Porras, C.; Sadd, J. Environmental justice and regional inequality in southern California: implications for future research. *Environ. Health Perspect.* **2002**, *110* (Suppl 2), 149–54.
- (7) Evans, G. W.; Kantrowitz, E. Socioeconomic status and health: the potential role of environmental risk exposure. *Annu Rev Public Health* **2002**, *23*, 303–31.
- (8) *Toward Environmental Justice: Research, Education, and Health Policy Needs*; Committee on Environmental Justice, Health Sciences Policy Program, Health Sciences Section., Institute of Medicine: Washington, DC, 1999.
- (9) Morello-Frosch, R.; Jesdale, B. M. Separate and unequal: residential segregation and estimated cancer risks associated with ambient air toxics in U.S. metropolitan areas. *Environ. Health Perspect.* **2006**, *114* (3), 386–93.
- (10) Finkelstein, M. M.; Jerrett, M. A study of the relationships between Parkinson's disease and markers of traffic-derived and environmental manganese air pollution in two Canadian cities. *Environ. Res.* **2007**, *104* (3), 420–32.
- (11) National Research Council, Science and Decisions: Advancing Risk Assessment, 2008.
- (12) National Environmental Justice Advisory Council, Ensuring Risk Reduction in Communities with Multiple Stressors: Environmental Justice and Cumulative Risks/Impacts-Draft Report; NEJAC (National Environmental Justice Advisory Council: New Orleans, 2001.
- (13) Cal-EPA EJAP (Environmental Justice Action Plan), In 2004.
- (14) DeFur, P. L.; Evans, G. W.; Cohen Hubal, E. A.; Kyle, A. D.; Morello-Frosch, R. A.; Williams, D. R. Vulnerability as a function of individual and group resources in cumulative risk assessment. *Environ. Health Perspect.* **2007**, *115* (5), 817–24.
- (15) Su, J. G.; Jerrett, M.; Beckerman, B. A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. *Sci. Total Environ.* **2009**, *407* (12), 3890–8.
- (16) Mauderly, J. L.; Samet, J. M. Is there evidence for synergy among air pollutants in causing health effects? *Environ. Health Perspect.* **2009**, *117* (1), 1–6.
- (17) *National Air Toxics Assessment*; U.S. Environmental Protection Agency: Washington, DC, 2006.
- (18) Su, J. G.; Morello-Frosch, R.; Jesdale, B. M.; Kyle, A. D.; Shamasunder, B.; Jerrett, M. An Index for Assessing Demographic Inequalities in Cumulative Environmental Hazards with Application to Los Angeles, California. *Environ. Sci. Technol.* **2009**, *43* (20), 7626–7634.
- (19) Jerrett, M.; Burnett, R. T.; Ma, R.; Pope, C. A., 3rd; Krewski, D.; Newbold, K. B.; Thurston, G.; Shi, Y.; Finkelstein, N.; Calle, E. E.; Thun, M. J. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* **2005**, *16* (6), 727–36.
- (20) *National Air Toxics Assessment*; U.S. Environmental Protection Agency: Washington, DC, 2005.
- (21) Molitor, J.; Papathomas, M.; Jerrett, M.; Richardson, S. Bayesian profile regression with an application to the National Survey of Children's Health. *Biostatistics* **2010**, *11*, 484–498.
- (22) Gilks, W. R.; Richardson, S.; Spiegelhalter, D. J. *Markov chain Monte Carlo in practice*; Chapman & Hall: Boca Raton, FL, 1998; p xvii, 486 p.
- (23) Lunn, D. J.; Thomas, A.; Best, N.; Spiegelhalter, D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat. Comput.* **2000**, *10*, 325–337.
- (24) Jain, S.; Radford, N. A split-merge Markov chain Monte carlo procedure for the Dirichlet process mixture model. *J. Comput. Graphical Stat.* **2004**, *13*, 158–182.
- (25) Neal, R. Markov Chain Sampling Methods for Dirichlet Process Mixture Models. *J. Comput. Graph. Stat.* **2000**, *9*.
- (26) Green, P. J.; Richardson, S. Modelling Heterogeneity with and without the Dirichlet Process. *Scand. J. Stat.* **2001**, *28* (2), 355–375.
- (27) Ishwaran, H.; James, L. F. Gibbs sampling methods for stick-breaking priors. *J. Am. Statist. Assoc.* **2001**, *96*, 161–173.
- (28) Richardson, S.; Green, P. On Bayesian analysis of mixtures with an unknown number of components. *J. Roy. Statist. Soc. B* **1997**, *59*, 731–792.
- (29) Gelman, A.; Jakulin, A.; Pittau, M.; Su, Y.-S. A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* **2008**, *2* (4), 1360–1383.
- (30) Denison, D. G. T.; Holmes, C. C.; Mallick, B. K.; Smith, A. F. M. *Bayesian methods for nonlinear classification and regression*; John Wiley & Sons, Inc.: New York, 2002.
- (31) Brody, J. G.; Morello-Frosch, R.; Zota, A.; Brown, P.; Perez, C.; Rudel, R. A. Linking exposure assessment science with policy objectives for environmental justice and breast cancer advocacy: the northern California household exposure study. *Am. J. Public Health* **2009**, *99* (Suppl 3), S600–9.
- (32) Do, K.-A.; Müller, P.; Vannucci, M. In *Bayesian Inference for Gene Expression and Proteomics*; Dahl, D. B., Ed.; Cambridge Books Online. Cambridge University Press: Cambridge, U.K., 2006 (<http://dx.doi.org/10.1017/CBO9780511584589.011>, accessed 11 August 2011).
- (33) Morello-Frosch, R.; Zuk, M.; Jerrett, M.; Shamasunder, B.; Kyle, A. D. Understanding the cumulative impacts of inequalities in environmental health: implications for policy. *Health Aff (Millwood)* **2011**, *30* (5), 879–87.
- (34) Pearce, J.; Kingham, S.; Zavar-Reza, P. Every breath you take? Environmental justice and air pollution in Christchurch, New Zealand. *Environ. Planning A* **2006**, *38*, 919–938.
- (35) Brainard, J. S.; Jones, A. P.; Bateman, I. J.; Lovett, A. A.; Fallon, P. J. Modelling environmental equity: Access to air quality in Birmingham, England. *Environ. Planning A* **2002**, *34*, 695–716.
- (36) Chaix, B.; Gustafsson, S.; Jerrett, M.; Kristersson, H.; Lithman, T.; Boalt, A.; Merlo, J. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *J. Epidemiol. Community Health* **2006**, *60* (3), 234–241.
- (37) Buzzelli, M.; Jerrett, M. Geographies of susceptibility and exposure in the city: environmental inequity of traffic-related air pollution in Toronto. *Can. J. Regional Sci.* **2007**, *30* (2), 195–210.
- (38) Cesaroni, G.; Badaloni, C.; Romano, V.; Donato, E.; Perucci, C. A.; Forastiere, F. Socioeconomic position and health status of people who live near busy roads: the Rome Longitudinal Study (RoLS). *Environ. Health* **2010**, *9*, 41.