Biostatistics (2010), **0**, 0, *pp.* 1–17
doi:10.1093/biostatistics/"cep naaqs supplementary"

Supplementary Material to Estimating Causal Effects of Air Quality Regulations Using Principal Stratification for Spatially-Correlated Multivariate Intermediate Outcomes

CORWIN M. ZIGLER*, FRANCESCA DOMINICI, YUN WANG

Department of Biostatistics, Harvard University, Harvard School of Public Health, 655 Huntington Avenue, Boston, MA 02115

czigler@hsph.harvard.edu

1. APPENDIX A: DETAILS OF THE SPATIAL HIERARCHICAL MODEL AND MCMC STRATEGY

For $f(X_0(s), X_1(s)|Z(s))$, we propose the following spatial hierarchical model for potential air-pollution concentrations:

$$X(s) = Z^T(s)\beta + W(s) + \epsilon(s), \tag{1.1}$$

where $X(s) = (X_0^T(s), X_1^T(s))^T$ is the $2q-$dimensional vector of potential pollution concentrations (q pollutants under each of two regulations), $W(s)$ is a vector of spatially-varying random intercepts, and $\epsilon(s)$ represents nonspatial "nugget" error (e.g., measurement error). $Z^T(s)$ is a $2q \times p$ matrix of time-invariant covariates, where $p = \Sigma_k p_k$ and p_k is the number of covariates pertaining to the k^{th} pollutant at location s (including an intercept). The vector β represents regression coefficients relating air pollution to differences in covariates.

The spatial correlation structure follows from specifying $W(s)$ as a realization from a multivariate Gaussian Process (MVGP) with cross-covariance function $K(s_i, s_j; \nu)$ being the $2q \times 2q$ matrix of covari-

*To whom correspondence should be addressed.

ances between the $2q$ potential pollution concentrations measured at locations s_i and s_j. The parameter $\nu = (\nu_1, \ldots \nu_{2q})$ is included because we model the cross covariances as functions of ν that represent the spatial decay of correlations between pollution measurements across space.

When $q = 2$, $K(s_i, s_j; \nu)$ is a 4×4 matrix where, for example, $[K(s_i, s_j; \nu)]_{1,1}$ denotes the covariance between PM_{10} under $\mathbf{A} = \mathbf{0}$ at location s_i and PM_{10} under $\mathbf{A} = \mathbf{0}$ at location s_j, and $[K(s_i, s_j; \nu)]_{2,3}$ denotes the covariance between O_3 under $\mathbf{A} = \mathbf{0}$ at location s_i and PM_{10} under $\mathbf{A} = \mathbf{1}$ at location s_j. Note that when $i = j$, $K(s_i, s_j; \nu) = K(s, s)$ is in fact a covariance matrix characterizing the relationships among the $2q$ potential pollution concentrations within a a location.

For this model, $W = [W(s_i)]_{i=1}^n \sim MVN(0, \Sigma_W)$, where Σ_W is a $2qn \times 2qn$ covariance matrix with $(i, j)^{th}$ block equal to the cross covariance between locations s_i and s_j, $\Sigma_W = [K(s_i, s_j)]_{i,j=1}^n$. We assume independent nugget errors, with $\epsilon(s) \sim MVN(0, \Psi)$, and Ψ diagonal.

The mechanics of model (2.3) rely on two key features: the relationship among the $2q$ potential pollution concentrations within a location, and the decay of their correlations across space. For the relationship among pollutants within a location, note that the cross covariance within a location, $K(s, s)$, is in fact a covariance matrix of the $2q$ random effects corresponding to pollution measurements at a common site. We write $K(s, s) = LL^T$, where L is the lower-triangular Cholesky square root of this covariance matrix, and assume that $K(s, s)$ is the same for all s, that is, that the process is stationary.

For the spatial decay, we define a simpler MVGP, $\tilde{W}(s)$, such that $Var(\tilde{W}_k(s)) = 1$ and the cross covariance is diagonal: $\tilde{K}(s_i, s_j; \nu) = diag\{\rho_k(s_i, s_j; \nu_k)\}$, where $\rho_k(s_i, s_j; \nu_k)$ represents a function for the spatial decay of the correlation between the k^{th} element of $\tilde{W}(s)$ across space. We assume isotropic exponential covariance functions that depend only on the Euclidean distance between locations s_i and s_j ($||s_i - s_j||$), with $\rho_k(s_i, s_j) = e^{-\nu_k ||s_i - s_j||}$. The covariance matrix of $\tilde{W} = [\tilde{W}(s_i)]_{i=1}^n$ can be written as $\Sigma_{\tilde{W}} = [\tilde{K}(s_i, s_j)]_{i,j=1}^n$.

Rather than model $K(s_i, s_j; \nu)$ directly, we separately specify $\tilde{K}(s_i, s_j; \nu)$ and LL^T, and define $W(s) = L\tilde{W}(s)$, which implies that the spatial random effects in (2.3) are a linear transformation of

the simpler MVGP, with transformation defined by the relationships among the pollutants. With this specification, $K(s_i, s_j; \nu) = L\tilde{K}(s_i, s_j; \nu)L^T$.

Let $\mathbf{X}^T = [(X_0(s_i)^T, X_1(s_i)^T)]_{i=1}^n$ be the $2qn \times 1$ pollution vector and $\mathcal{Z}$ be the $2qn \times p$ matrix of regressors. The above model can, after marginalization over $\tilde{W}$, be equivalently stated as

$$\mathbf{X} \sim MVN(\mathcal{Z}\beta, \mathcal{L}\Sigma_{\tilde{W}}\mathcal{L}^T + I_n \otimes \Psi) \tag{1.2}$$

where $\mathcal{L} = I_n \otimes L$, and $\otimes$ is the Kronecker product. Details for this model formulation as well as generalizations can be found in Wackernagel (2003), Finley *and others* (2007), and Banerjee *and others*. (2008).

For the MCMC, $K(s, s)$ is updated via updates of $L_{\mathbf{a}}$, which are the lower-triangular Cholesky roots of the $q \times q$ diagonal blocks of $K(s, s)$ that are informed by the observed data. The off-diagonal blocks of $K(s, s)$ are updated according to the pre-specified value of ω and the values of $L_{\mathbf{a}}$, subject to a positive-definiteness constraint. Under our prior specification (detailed below), the posterior distribution for β is multivariate normal, with samples drawn using a fully-conditional Gibbs step. All other parameters and missing data are updated with a Metropolis step using normal proposal distributions, with appropriate transformations for all variables having restricted support. Each diagonal element of Ψ, each ν, and each missing quantity are updated individually, with block updating carried out for β, $\alpha^0 = (\alpha_0^0, \alpha_1^0, \alpha_2^0)$, $\alpha^1 = (\alpha_0^1, \alpha_1^1, \alpha_2^1)$, and (L_0, L_1). Proposal variances for each parameter and for the missing quantities are tuned during an adaptive "pre-burn-in" stage, after which proposal variances are fixed and the chain is run for 20,000 iterations. After discarding the first 1,000 iterations as burn in, inference is based on every 20^{th} posterior sample.

As pointed out in Finley *and others* (2007), values of ν are only weakly identifiable and require reasonably informative priors for satisfactory MCMC behavior, but the model decomposition described above entails adequate structure to identify $\Sigma_{\tilde{W}}, L, \Sigma_W$, and Ψ. We treat the parameters $\beta, \Psi, L_0, L_1, \nu, \alpha^0$, and α^1, as *a priori* independent. We specify flat priors for β, α^0 and α^1. For the diagonal elements of Ψ, we specify independent inverse-gamma distributions with shape parameters set to 3.0 and scale parameters set

to 0.12 for both PM_{10} and O_3 . For ν_k, we specify uniform prior distributions on the interval $(1.0, 6.0)$ for both PM_{10} and O_3 . Parameters for the prior distributions of Ψ and ν are meant to reflect diffuse prior information within the range of plausible parameter values. For each 2×2 block on the diagonal of $K(s, s)$, we specify an inverse-Wishart prior distribution with 3 degrees of freedom and scale matrix equal to the covariance structure of PM_{10} and O_3 observed among regulated and unregulated areas with co-located monitors.

2. APPENDIX B: ILLUSTRATION USING SIMULATED DATA

To illustrate that our method can distinguish between settings where a regulation effect on health is and is not associated with effects on pollution, we simulate data reflecting two scenarios: Scenario A where the regulation improves health the most when it also improves air pollution (high associative effect relative to dissociative effect) and Scenario B where the regulation improves health through other pathways (associative effect $\approx$ dissociative effect). Monitor locations, number of Medicare beneficiaries living at each location, and baseline covariates ($Z(s)$) were those from the observed data used in the analysis of the CAAA.

For both scenarios, we simulate $X(s) = (x_1(s), x_2(s), x_3(s), x_4(s))$ from:

$$X(s) = Z^T(s)\beta + W(s) + \epsilon(s), \tag{2.3}$$

where (x_1, x_2) represent, respectively, the logarithm of potential PM_{10} , O_3 concentrations under no regulation ($\mathbf{A} = \mathbf{0}$) and (x_3, x_4) represent the logarithm of PM_{10} , O_3 in the presence of regulation ($\mathbf{A} = \mathbf{1}$). The values of β were set to maximum likelihood estimates from simple linear regressions of actual CAAA post-regulation pollution concentrations on $Z(s)$. We used the following covariance matrices for $(x_1(s), x_2(s), x_3(s), x_4(s))$ to simulate the pollution data:

$$K(s,s) = \begin{pmatrix} 0.059 & 0.024 & 0.035 & 0.014 \\ 0.024 & 0.031 & 0.014 & 0.019 \\ 0.035 & 0.014 & 0.059 & 0.024 \\ 0.014 & 0.019 & 0.024 & 0.031 \end{pmatrix}, \Psi = \begin{pmatrix} 0.034 & 0 & 0 & 0 \\ 0 & 0.016 & 0 & 0 \\ 0 & 0 & 0.034 & 0 \\ 0 & 0 & 0 & 0.016 \end{pmatrix}, \tag{2.4}$$

and set $\nu = (3, 3, 3, 3)$. This specification implies a value of $\omega = 0.6$, which we assume known for

illustrative purposes. To simulate a causal effect of the regulation on pollution, we reduced the potential regulated pollution concentrations (x_3, x_4) by one quarter of one standard deviation, that is, we set $x_k = x_k - 0.25\sigma_k$ for $k = 3, 4$, where σ_k is the standard deviation of the k^{th} simulated pollutant.

We simulated mortality outcomes from Poisson distributions with means determined by the following log-linear models:

$$log(E[Y_{\mathbf{a}}(s)]) = \alpha_0^{\mathbf{a}} + Z^T(s)\alpha_1^{\mathbf{a}} + X_{\mathbf{a}}(s)\alpha_2^{\mathbf{a}} + log(N(s)). \tag{2.5}$$

To obtain values of α parameters for the simulation, we fit a Poisson regression predicting actual CAAA deaths in unregulated areas from $Z(s)$. We set covariate relative risks under both regulation programs (α_1^0, α_1^1) to maximum likelihood estimates from this Poisson regression. The intercept parameters $(\alpha_0^{\mathbf{a}})$ and relative risk parameters associated with post-regulation pollution $(\alpha_2^{\mathbf{a}})$ are varied in Scenarios A and B. The values used for each simulated scenario appear in Table 1. Note that for Scenario A, the relative risks associated with PM_{10} are stronger than those associated with O_3, reflecting a scenario where causal effects on PM_{10} have a larger impact on mortality than equally-sized causal effects on O_3.

Table 1. Values of $\alpha_0^{\mathbf{a}}$ and $\alpha_2^{\mathbf{a}}$ $(\mathbf{a} = 0, 1)$ for the two simulated scenarios.

		Scenario A (high assoc., low dissoc.)		Scenario B (assoc. ≈ dissoc.)	
Parameter		$\mathbf{a} = 0$	$\mathbf{a} = 1$	$\mathbf{a} = 0$	$\mathbf{a} = 1$
$\alpha_0^{\mathbf{a}}$	Intercept	-5.72	-5.74	-5.47	-5.50
$\alpha_2^{\mathbf{a}}$	log(PM_{10} RR)	0.14	0.14	0.005	0.005
$\alpha_2^{\mathbf{a}}$	log(O_3 RR)	0.05	0.05	0.005	0.005

2.1 *Simulation Results*

The overall causal effects of the regulation on mortality were similar in the two simulated scenarios. For Scenario A, the estimated overall average causal effect of the regulated program (adjusted for $Z(s)$ using a Poisson regression model without post-regulation pollution) was 1.68 fewer mortalities per 1,000 Medicare beneficiaries (posterior mean (sd) deaths/1,000: 63.59(0.29) vs. 65.27 (0.34)). For Scenario B, the

analogous estimated overall average causal effect of the regulated program was 1.31 fewer mortalities per 1,000 Medicare beneficiaries (posterior mean (sd) deaths/1,000: 62.19 (0.29) vs. 63.50 (0.33)). However, our proposed methodology to estimate associative and dissociative effects yields very different results in the two scenarios.

The $h(\mathbf{x}, \mathbf{y}), C^D_{\mathcal{K}}$, and $C^A_{\mathcal{K}}$ used for $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ were as in the main text. Figures 1 and 2, examine estimated causal effects on mortality as a function of estimated causal effects on air quality. As in the main text, the shaded region represents a causal effect on air pollution below the threshold $C^D_{\mathcal{K}}$, and the area below the shaded region represents a causal effect on pollution in excess of the threshold $C^A_{\mathcal{K}}$.

In Figure 1 for Scenario A, note that for PM_{10} , mortality is estimated to decrease for almost all values of $[X_0(s)]_{\mathcal{K}}, [X_1(s)]_{\mathcal{K}}$, but that the decreases below the shaded area tend to be larger than the decreases within the shaded area. This is evidence of a larger associative effect than dissociative effect. Notice that this pattern is much less apparent for O_3 .

In contrast, note that for Scenario B (Figure 2), decreases in mortality are distributed more evenly around shaded region for both PM_{10} and O_3 , suggesting that decreases in mortality tend to be the same regardless of changes in pollution.

Figures 3 presents posterior distributions of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ under Scenarios A and B, for $\mathcal{K} = \{$ PM_{10} $\}, \{$ O_3 $\}$, and $\{$ PM_{10} , O_3 $\}$. $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ are estimated to be negative for all $\mathcal{K}$ and in both scenarios, reflecting a regulation-induced decrease in mortality. For Scenario A, estimates of $EAE_{\mathcal{K}}$ always indicate more pronounced mortality decreases than estimates of $EDE_{\mathcal{K}}$. In contrast, for Scenario B, estimates of $EAE_{\mathcal{K}}$ and $EDE_{\mathcal{K}}$ are always estimated to be similar. Note that, the contrast between Scenarios A and B is most stark for $\mathcal{K} = \{$ PM_{10} $\}$ and $\mathcal{K} = \{$ PM_{10} , O_3 $\}$, and that the largest mortality decrease is estimated in Scenario A when the regulation causally affects both PM_{10} and O_3 .

3. APPENDIX C: POSTERIOR SUMMARIES OF ALL PARAMETERS IN MODELS FOR THE CAAA

Figures 4 – 12 summarize posterior distributions of all model parameters from our analysis of the CAAA. In these figures, solid squares represent posterior medians, open circles represent posterior means, and error bars are 95% posterior intervals.

4. APPENDIX D: ASSESSMENT OF AGIA

For $\omega = 0.0, 0.3, 0.6,$ and 0.9, posterior mean estimates of ν for PM_{10} were 3.46, 3.48, 3.32, and 3.13, and for O_3 were 3.01, 3.01, 2.68, and 3.42, respectively. To investigate violations of AGIA we determine, for each location, the distance to the nearest location in the opposite regulation group. Denote this distance with d_s. For each pollutant, Figure 13 presents a histogram of $e^{-\hat{\nu}_k d_s}$, representing the estimated correlation between concentrations of the k^{th} pollutant measured at locations d_s apart, where $\hat{\nu}_k$ is 3.13 for PM_{10} and 2.68 for O_3 (the lowest posterior mean estimates across all values of ω). For PM_{10} (O_3), 111 (127) regulated and 67 (79) unregulated locations lie within distances implying estimated correlations > 0.05, and 17 (31) regulated and 13 (23) unregulated locations lie within distances implying estimated correlations > 0.25. Thus, AGIA appears violated for many locations, but relatively few of these locations exhibit correlations greater than 0.25.

AGIA implies that potential outcomes for locations in $\mathcal{R}^{\mathbf{a^{obs}}}$ are the same under $\mathbf{A} = \mathbf{a^{obs}}$ and $\mathbf{A} = \mathbf{1}$ and that potential outcomes for locations in $\mathcal{U}^{\mathbf{a^{obs}}}$ are the same under $\mathbf{A} = \mathbf{a^{obs}}$ and $\mathbf{A} = \mathbf{0}$. However, estimates of ν suggest a violation of this assumption for some locations. Rather than being the same (as under AGIA), we argue that pollution under the hypothetical program $\mathbf{A} = \mathbf{1}$ would likely be *lower* than that observed in $\mathcal{R}^{\mathbf{a^{obs}}}$, because adding regulations to nearby areas should decrease overall pollution. Similarly, we argue that pollution under the hypothetical program $\mathbf{A} = \mathbf{0}$ would likely be *higher* than that observed in $\mathcal{U}^{\mathbf{a^{obs}}}$, because removing regulations from nearby areas should increase overall pollution. If all other model assumptions held, this would lead pollution estimates in our analysis to be *overestimates* of potential pollution under $\mathbf{A} = \mathbf{1}$ and *underestimates* of potential pollution under $\mathbf{A} = \mathbf{0}$. Thus, it is

feasible that, assuming all other modeling assumptions hold, a violation of AGIA leads to estimates of the causal effect of $\mathbf{A} = \mathbf{1}$ vs. $\mathbf{A} = \mathbf{0}$ that are conservatively approximate in that they underestimate the true causal effect.

5. APPENDIX E: MISSING DATA IN THE ANALYSIS OF CAAA

5.1 *Missing outcome data*

In practice, we are confronted with two types of missing outcome data. First, we have the missing post-regulation potential outcomes in 1999-2001, that is, the potential pollution levels and mortality counts under the regulation program that would imply the opposite designation of that which was observed. Second, we have missing pollution outcomes in 1999-2001 for locations where not all of the q pollutants are measured, which we assume to be missing at random. In our analysis of the CAAA, co-located monitors for both PM_{10} and O_3 were present in 108 locations during 1999-2001, but air pollution concentrations were missing for 135 locations having monitors only for PM_{10} , 70 locations having monitors only for O_3 , and 49 locations having no monitor for either pollutant during this period. For the MCMC, both types of missing outcome data are treated the same and are sampled at each iteration conditional on observed data, current values for parameters, and imputed values for other missing outcomes.

5.2 *Missing covariate data*

A distinct type of missing data in our application is the missing pre-regulation PM_{10} and O_3 concentrations during 1987-1989, which are considered covariates in our analysis. This issue arises because monitor locations in the EPA monitoring network are not static over time, and there are many monitor locations in operation during 1999-2001 that were not in operation during the pre-regulation period. The issue of missing covariates is different from the issue of missing outcomes. If observations were to have *both* missing covariates *and* missing outcomes, then there would be limited information with which to ground imputations for these quantities absent any model for the covariates. This is particularly problematic in

our case because *every* observation has missing outcomes (unobserved potential outcomes).

Rather than exclude observations with missing pre-regulation pollution, we employ an additional spatial model to impute the missing pre-regulation pollution and use these imputations in our analysis of the CAAA. Specifically, we fit the same type of hierarchical model as detailed in Appendix A, with the outcome specified as the 2-dimensional vector consisting of log-transformed PM_{10} and O_3 during the years 1987-1989. Because this imputation model is not designed to evaluate causal effects, it does not entail the complications of potential outcomes and nonidentifiable associations. Covariates included in this model were the same as those in the model of the main text (See Table 1 of the main text), with the addition of a covariate denoting whether a location lies in an attainment or nonattainment area. In total, 179 pre-regulation PM_{10} and 228 pre-regulation O_3 values were imputed using posterior-predictive means from this model and treated as fixed covariates in the final analysis of the CAAA.

To assess the predictive ability of our baseline imputation model, we re-fit the imputation model to a data set that withheld 100 known baseline pollution values (50 PM_{10} and 50 O_3) selected at random. Figure 14 plots the true known values of these 100 pollution concentrations against predicted values from the baseline imputation model. We see that the baseline model produces reasonable predictions for these pre-regulation pollution values. While our final analysis of the CAAA does not reflect uncertainty in these imputations, the out-of-sample prediction ability displayed in Figure 14 suggests that our imputations are not likely to introduce substantial bias.

To further investigate the possible implications of imputing missing pre-regulation pollution values, Figure 15 redisplays box plots of posterior distributions of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ (same as those in Figure 4 of the main text), overlaid with the same quantities calculated using only locations without imputed baseline data. That is, we used the same model fit to the entire data set (with imputed baseline values) to obtain posterior distributions of potential outcomes for all locations, but calculated $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ using only posterior samples for locations that had completely observed baseline pollution. For all $\mathcal{K}$, we see more variability when using only observations with completely observed baseline pollution. In Figure

15(a), estimates for $\mathcal{K} = \{ \text{PM}_{10} \}$ using only locations with completely observed baseline pollution are largely consistent with estimates using all locations. For $\mathcal{K} = \{ \text{O}_3 \}$ in Figure 15(b) we see that using only locations with completely observed baseline pollution suggests more pronounced health effects than estimates using all locations. A similar pattern is seen in Figure 15(c).

6. APPENDIX F: SENSITIVITY TO THE CHOICE OF $C^D_{\mathcal{K}}$ AND $C^A_{\mathcal{K}}$

As defined in the main text, $EAE_{\mathcal{K}}$ represents the average causal regulation effect on mortality among locations where the regulation causally decreases pollution by at least $C^A_{\mathcal{K}}$, and $EDE_{\mathcal{K}}$ represents average causal effects on mortality among locations where the regulation effect on pollution is less than $C^D_{\mathcal{K}}$. That is, these threshold values distinguish locations that exhibit a "meaningful" decrease in pollution from areas where pollution was not meaningfully affected. One benefit of our approach is that estimates of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ can be obtained using different threshold values of $C^D_{\mathcal{K}}$ and $C^A_{\mathcal{K}}$ without re-fitting the model. This is because posterior samples of the missing potential outcomes involved in the definition of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ do not depend on the choice of $C^D_{\mathcal{K}}$ and $C^A_{\mathcal{K}}$. Thus, posterior samples of missing potential outcomes can be obtained from fitting the model once, and these posterior samples can be used to calculate average causal effects in any combination of principal strata. Figures 16, 17, and 18 display posterior boxplots $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ calculated using two different sets of threshold values $C^D_{\mathcal{K}}$ and $C^A_{\mathcal{K}}$: one set entailing a more lenient definition of "meaningful" causal effect on pollution ($C^D_{\mathcal{K}} = C^A_{\mathcal{K}} = 2, 0.0025$ for PM_{10} , O_3) and one entailing a more stringent definition of "meaningful" causal effect on pollution ($C^D_{\mathcal{K}} = C^A_{\mathcal{K}} = 8, 0.01$ for PM_{10} , O_3). We see that qualitative conclusions regarding the relative magnitudes of associative and dissociative effects do not change with different values of $C^D_{\mathcal{K}}$ and $C^A_{\mathcal{K}}$, but that uncertainty regarding $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ depends on these thresholds. This is to be expected, because the threshold values determine how many locations contribute to estimates of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$. For example, with more stringent definitions of meaningful causal effects ($C^D_{\mathcal{K}} = C^A_{\mathcal{K}} = 8, 0.01$ for PM_{10} , O_3), more locations are estimated to have no meaningful causal effect, and $EDE_{\mathcal{K}}$ is estimated with

more precision.

Note that definitions of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ need not rely on principal strata specified by differences (or absolute values of differences) between $[X_0(s)]_{\mathcal{K}}$ and $[X_1(s)]_{\mathcal{K}}$. Our formulations are one specific instance of more general formulations that rely on a (possibly vector-valued) contrast function $h(\mathbf{w}, \mathbf{v})$, where $h(\mathbf{w}, \mathbf{w}) = \mathbf{0}$, and $\mathbf{0}$ denotes a vector with every element equal to 0. For example, $h(\mathbf{w}, \mathbf{v})$ could denote the Euclidean distance between multivariate pollution vectors. $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ could then be defined as:

$$EAE_{\mathcal{K}} = E[Y_1^r(s) - Y_0^r(s) | h([X_0(s)]_{\mathcal{K}}, [X_1(s)]_{\mathcal{K}}) > C_{\mathcal{K}}^A] \tag{6.6}$$

$$EDE_{\mathcal{K}} = E[Y_1^r(s) - Y_0^r(s) | h([X_0(s)]_{\mathcal{K}}, [X_1(s)]_{\mathcal{K}}) < C_{\mathcal{K}}^D]. \tag{6.7}$$

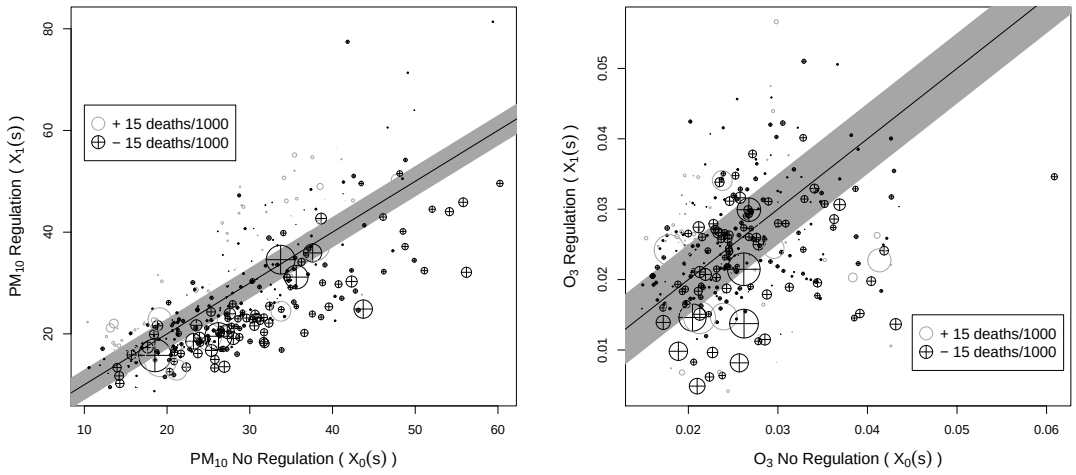
REFERENCES

BANERJEE, SUDIPTO AND GELFAND, ALAN E. AND FINLEY, ANDREW O. AND SANG, HUIYAN (2008), Gaussian predictive process models for large spatial data sets, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **70**, 825–848.

FINLEY, A. O AND BANERJEE, S. AND CARLIN, B. P, (2007), spBayes: an R package for univariate and multivariate hierarchical point-referenced spatial models, *Journal of Statistical Software*, **19**, 4–24.

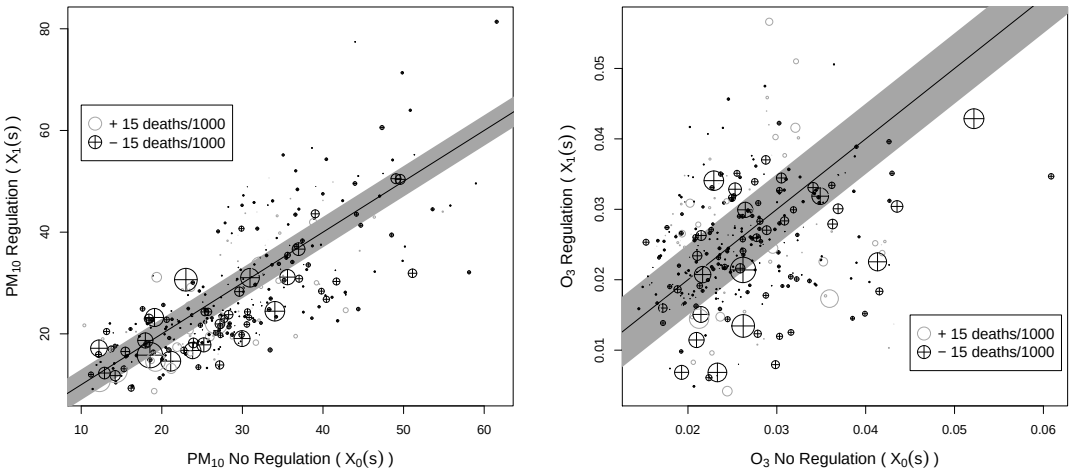
WACKERNAGEL, HANS, (2003), *Multivariate geostatistics: an introduction with applications*, Springer.

Fig. 1. Detailed depiction of the causal effect surface for $\mathcal{K} = \text{PM}_{10}$ and $\mathcal{K} = \text{O}_3$ for simulated Scenario A with large associative effects relative to dissociative effects. Values of $X_{\mathbf{a}}(s)$ represent observed or posterior predictive mean values. Size and plotting symbol of point indicates the posterior mean causal effect on mortality for that location. Points in shaded area represent areas with $[X_0(s)]_{\mathcal{K}} - [X_1(s)]_{\mathcal{K}} < C^D_{\mathcal{K}}$.



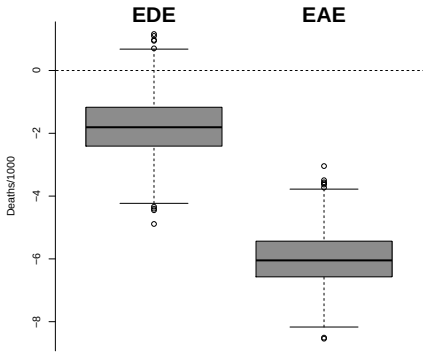
(a) PM_{10} (b) O_3

Fig. 2. Detailed depiction of the causal effect surface for $\mathcal{K} = P$ and $\mathcal{K} = 0$ for simulated Scenario B with associative effects $\approx$ dissociative effects. Values of $X_{\mathbf{a}}(s)$ represent observed or posterior predictive mean values. Size and plotting symbol of point indicates the posterior mean causal effect on mortality for that location. Points in shaded area represent areas with $[X_0(s)]_{\mathcal{K}} - [X_1(s)]_{\mathcal{K}} < C^D_{\mathcal{K}}$.

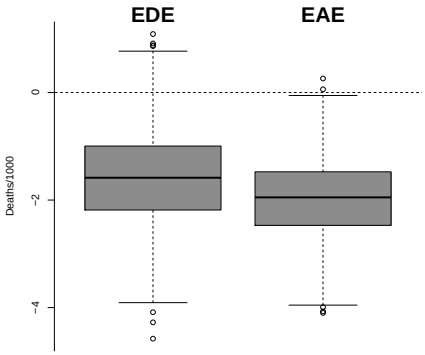


(a) PM_{10} (b) O_3

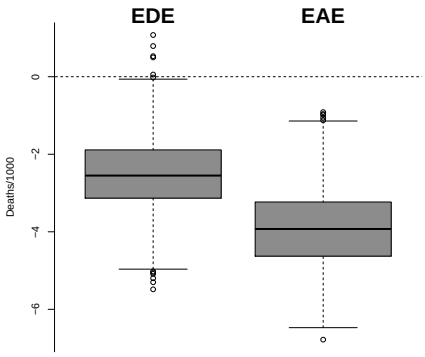
Fig. 3. Posterior summaries of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ for Scenarios A and B.



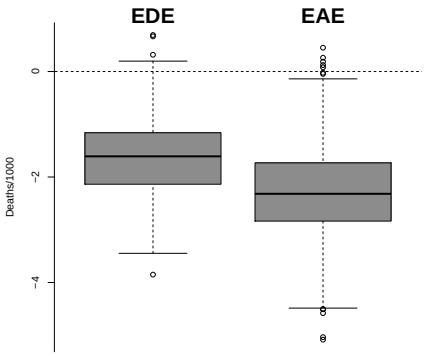
(a) $\mathcal{K} = \{ \text{PM}_{10} \}$, Scenario A



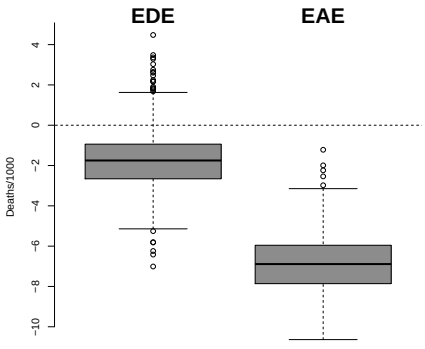
(b) $\mathcal{K} = \{ \text{PM}_{10} \}$, Scenario B



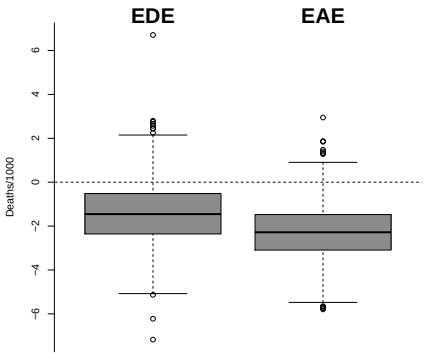
(c) $\mathcal{K} = \{ \text{O}_3 \}$, Scenario A



(d) $\mathcal{K} = \{ \text{O}_3 \}$, Scenario B



(e) $\mathcal{K} = \{ \text{PM}_{10}, \text{O}_3 \}$, Scenario A



(f) $\mathcal{K} = \{ \text{PM}_{10}, \text{O}_3 \}$, Scenario B

Fig. 4. Posterior distributions for $K(s, s)$.

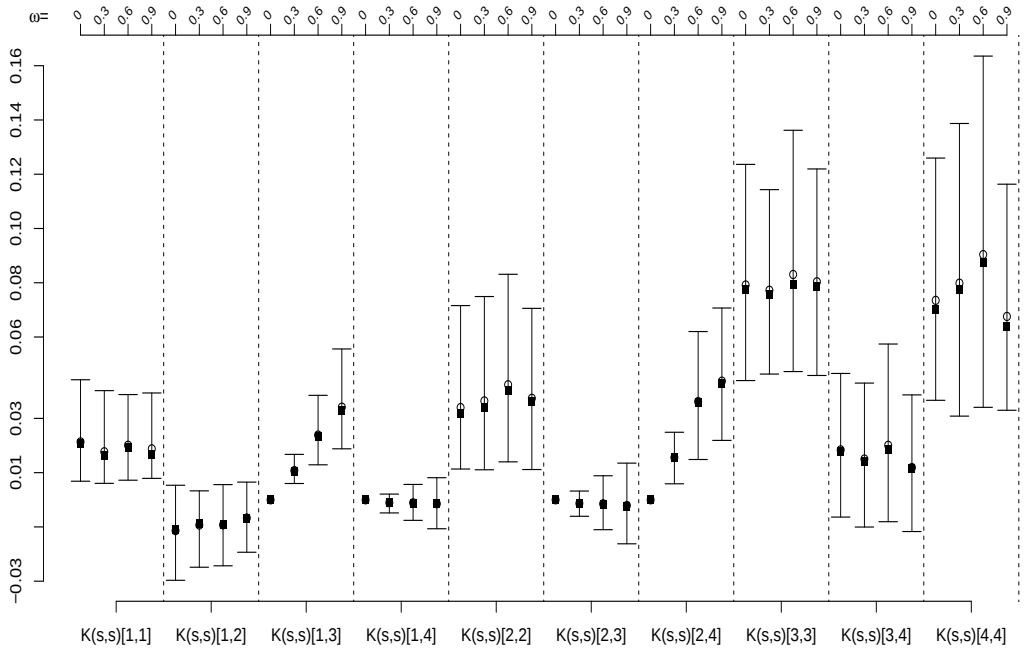


Fig. 5. Posterior distributions for Ψ.

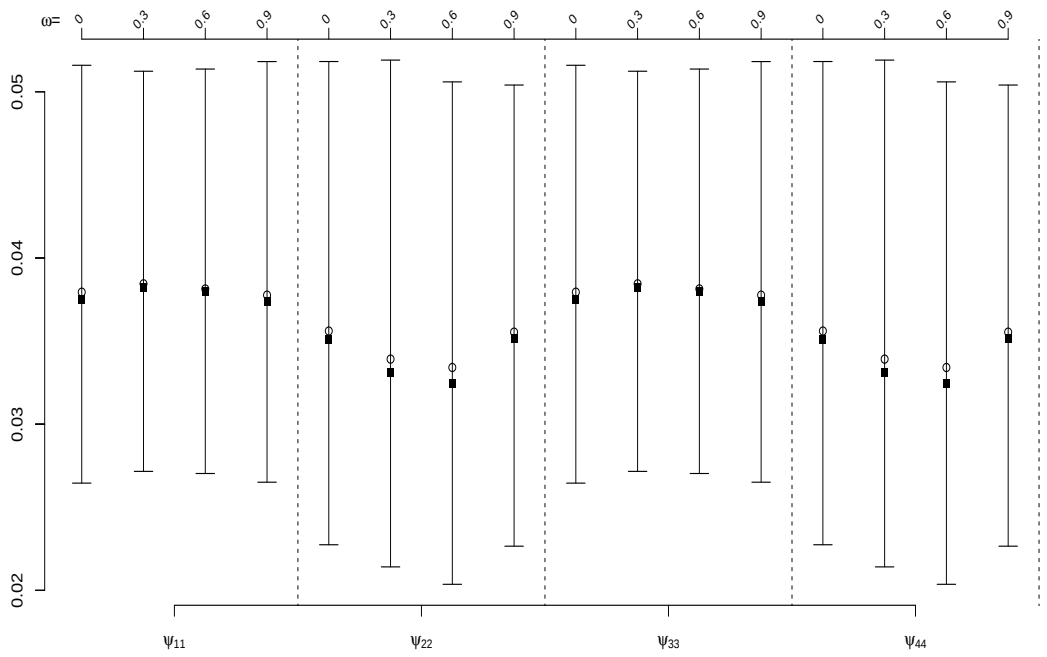


Fig. 6. Posterior distributions for ν.

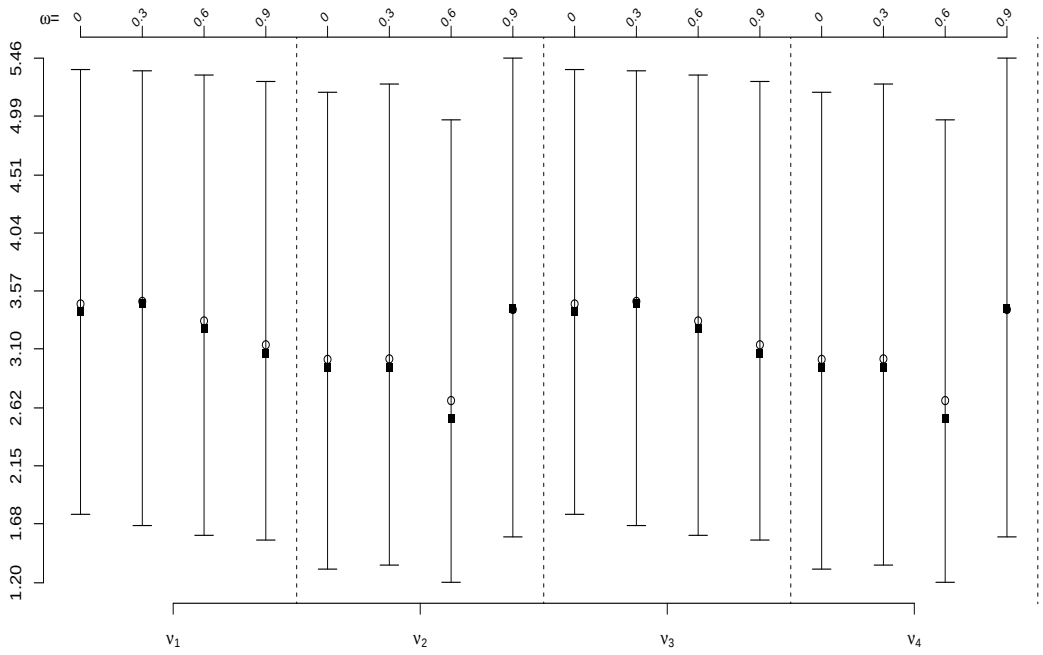


Fig. 7. Posterior distributions for β for PM_{10} under $\mathbf{A} = \mathbf{0}$.

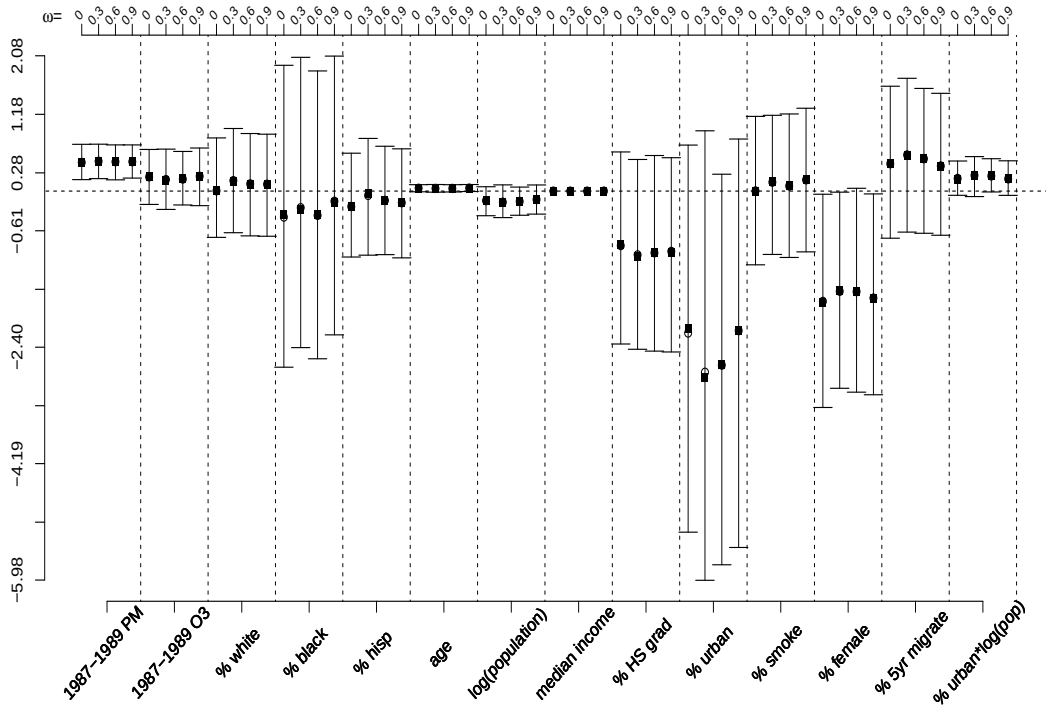


Fig. 8. Posterior distributions for β for O_3 under $\mathbf{A} = \mathbf{0}$.

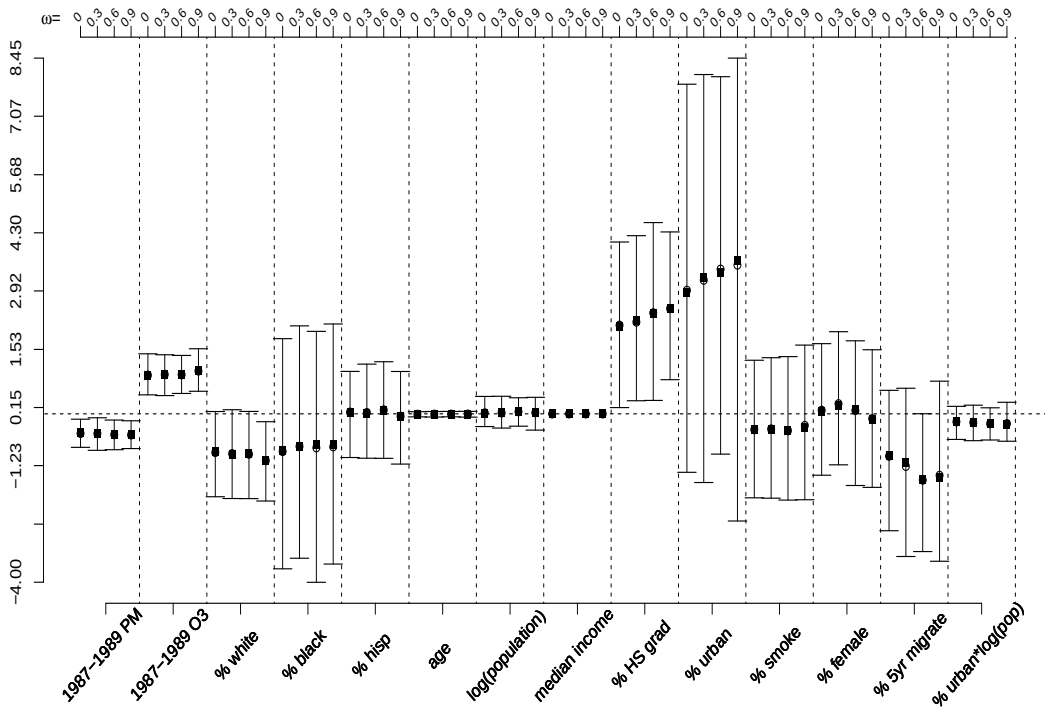


Fig. 9. Posterior distributions for β for PM_{10} under $\mathbf{A} = \mathbf{1}$.

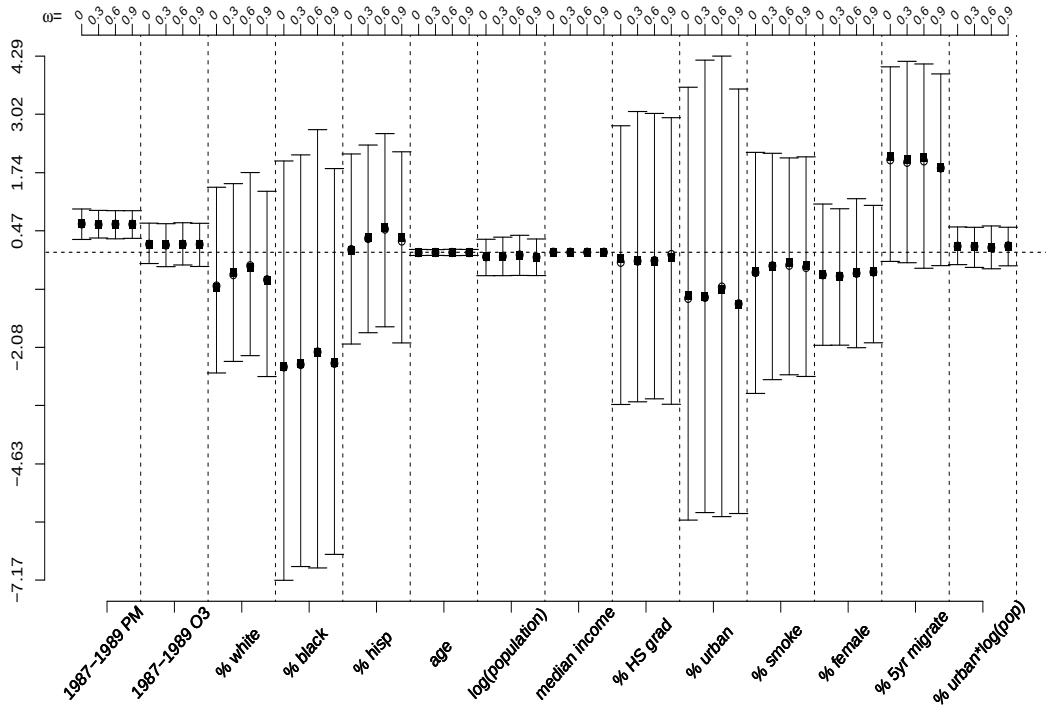


Fig. 10. Posterior distributions for β for O_3 under $\mathbf{A} = \mathbf{0}$.

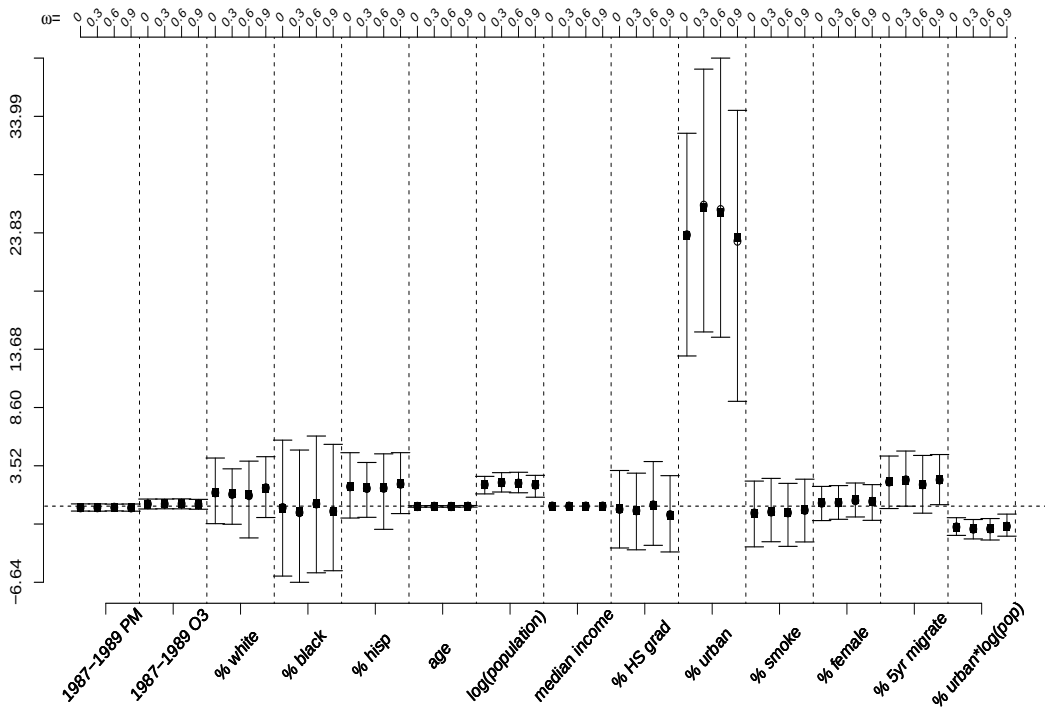


Fig. 11. Posterior distributions for α^0.

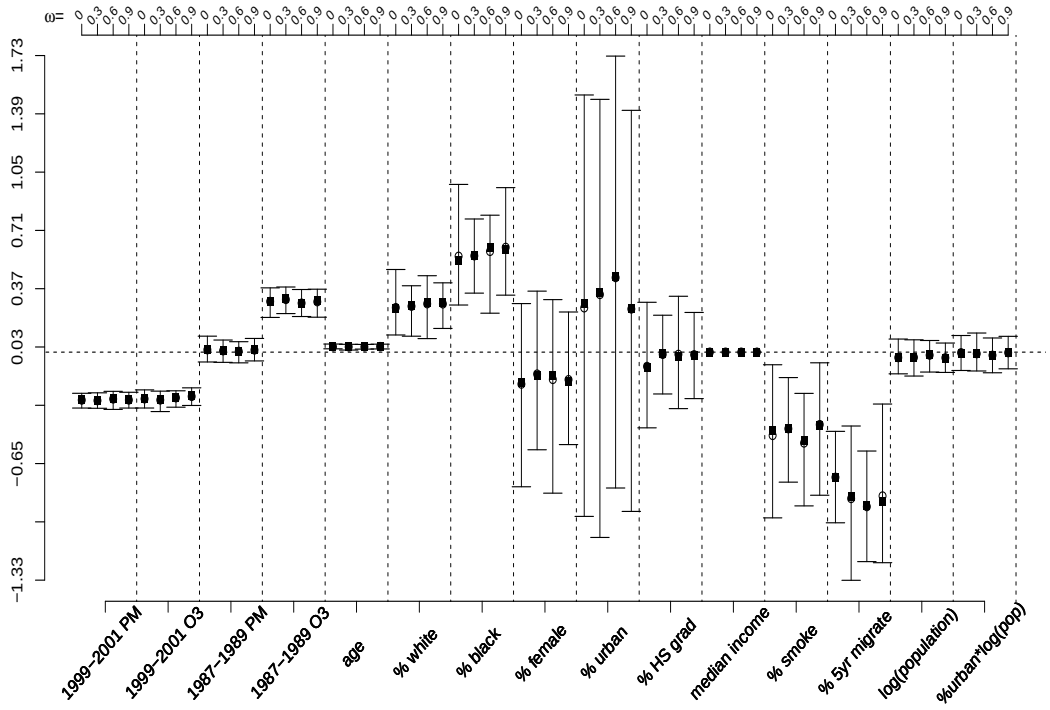


Fig. 12. Posterior distributions for α^1.

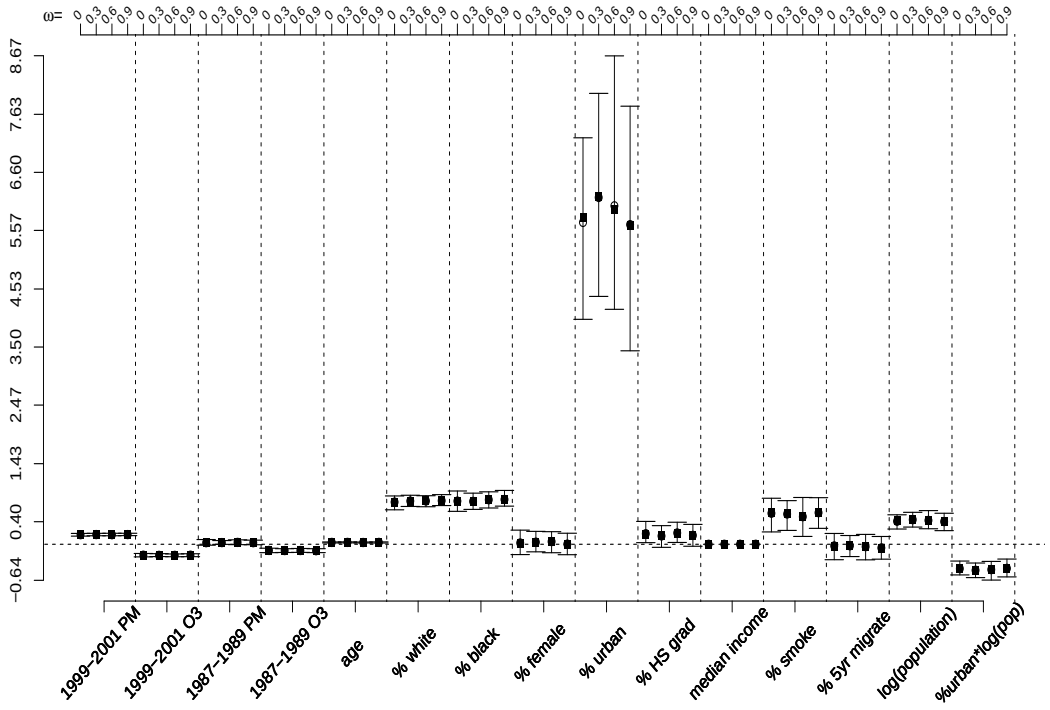
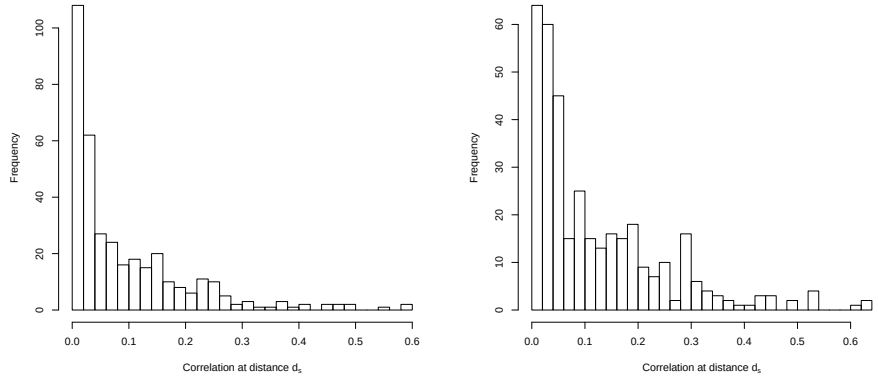


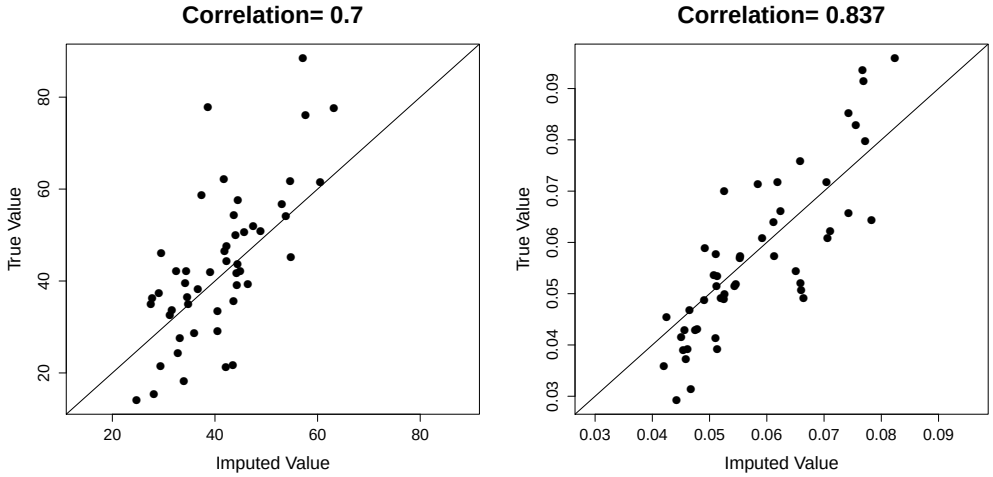
Fig. 13. Histograms of the correlation ($e^{-\hat{\nu}_k d_s}$) of pollution concentrations for each location and the nearest location with opposite observed regulation (d_s). Values of $\hat{\nu}_k$ are the minimum posterior mean value for each pollutant across all values of the sensitivity parameter, ω.



(a) PM_{10} , $\hat{\nu}_k = 3.13$

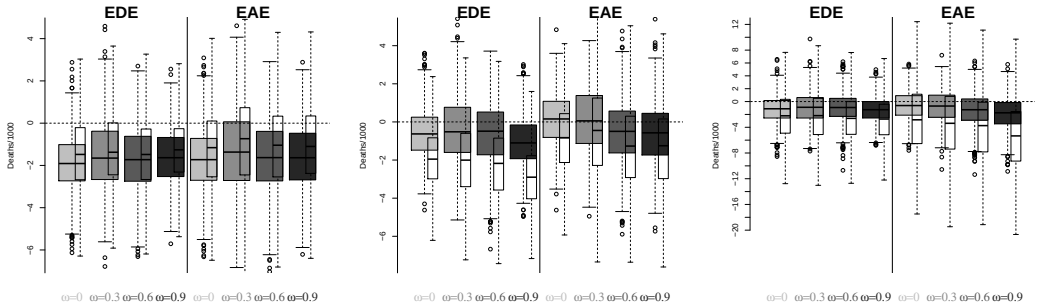
(b) O_3 , $\hat{\nu}_k = 2.68$

Fig. 14. True baseline pollution concentration vs. posterior-predictive mean pollution concentration from the baseline imputation model fit to a data set that withheld 100 known pollution values during 1987-1989.



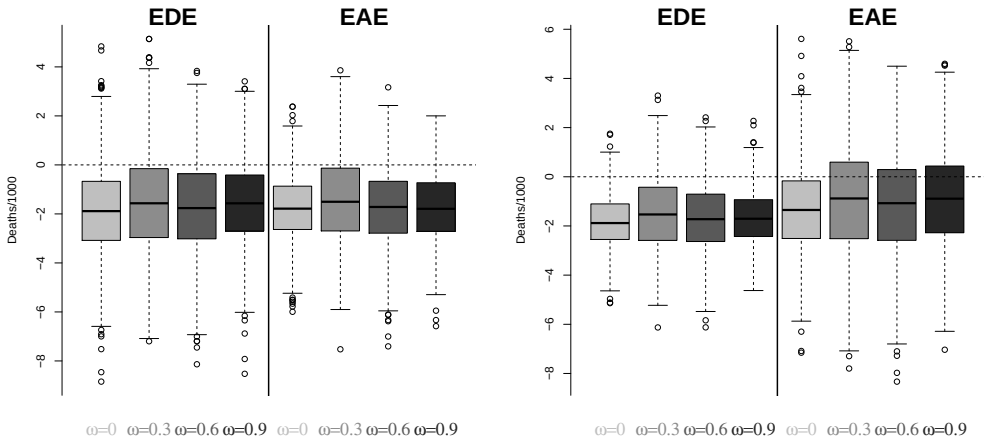
(a) PM_{10} (b) O_3

Fig. 15. Posterior summaries of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$. Shaded box plots use all locations, hollow box plots use only observations with observed baseline pollution.



(a) $\mathcal{K} = \{ PM_{10} \}$ (b) $\mathcal{K} = \{ O_3 \}$ (c) $\mathcal{K} = \{ PM_{10}, O_3 \}$

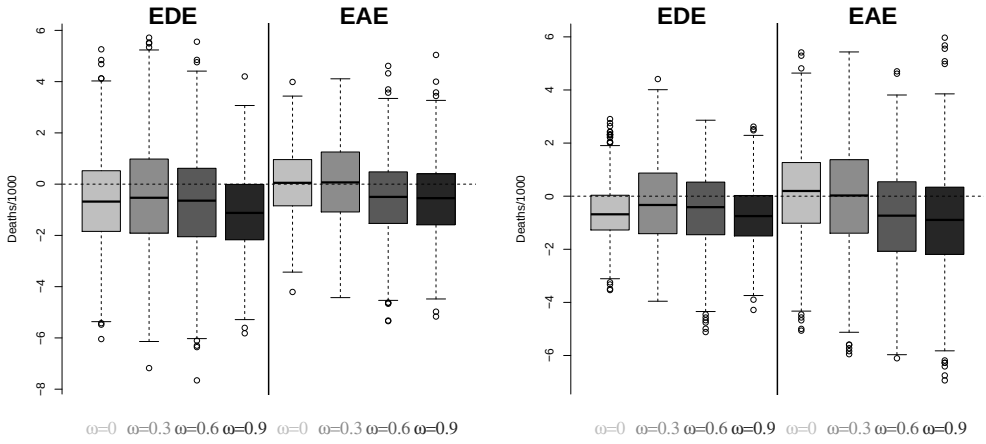
Fig. 16. Posterior summaries of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ for different values of $C^D_{\mathcal{K}}$ and $C^A_{\mathcal{K}}$. $\mathcal{K} = \{ \text{PM}_{10} \}$.



(a) $C^D_{\mathcal{K}} = 2, C^A_{\mathcal{K}} = 2$

(b) $C^D_{\mathcal{K}} = 8, C^A_{\mathcal{K}} = 8$

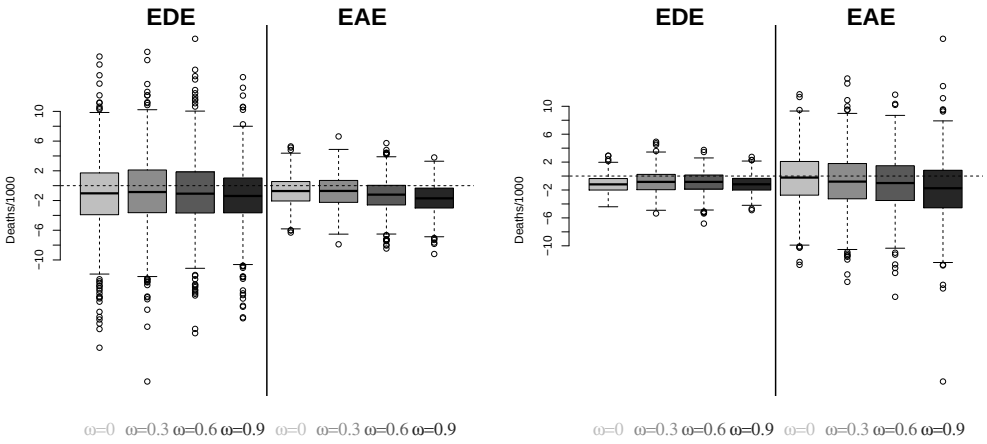
Fig. 17. Posterior summaries of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ for different values of $C^D_{\mathcal{K}}$ and $C^A_{\mathcal{K}}$. $\mathcal{K} = \{ \text{O}_3 \}$.



(a) $C^D_{\mathcal{K}} = 0.0025, C^A_{\mathcal{K}} = 0.0025$

(b) $C^D_{\mathcal{K}} = 0.01, C^A_{\mathcal{K}} = 0.01$

Fig. 18. Posterior summaries of $EDE_{\mathcal{K}}$ and $EAE_{\mathcal{K}}$ for different values of $C^D_{\mathcal{K}}$ and $C^A_{\mathcal{K}}$. $\mathcal{K} = \{ \text{PM}_{10} , \text{O}_3 \}$.



(a) $C^D_{\mathcal{K}} = \{2, 0.0025\}, C^A_{\mathcal{K}} = \{2, 0.0025\}$

(b) $C^D_{\mathcal{K}} = \{8, 0.01\}, C^A_{\mathcal{K}} = \{8, 0.01\}$