Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures

Supplementary Materials

Jennifer F. Bobb[1,*], Linda Valeri[1], Birgit Claus Henn[2], David C. Christiani[3], Robert O. Wright[4], Maitreyi Mazumdar[3], John J. Godleski[2], and Brent A. Coull[1]

[1]*Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115*

[2]*Department of Environmental Health, Harvard School of Public Health, Landmark Center, 401 Park Drive Boston, Massachusetts 02215*

[3]*Department of Environmental Health, Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115*

[4]*Mount Sinai Hospital, 17 East 102 Street Floor 3, West Room D3-110, New York, NY 10029*

**jfbobb@hsph.harvard.edu*

A. Prior specification

Here we specify prior distributions for the parameters of the Bayesian kernel machine regression (BKMR) model described in Section 2 of the main text.

We assumed $\boldsymbol{\beta} \sim 1$ (flat prior) and $\sigma^{-2} \sim \text{Gamma}(a_\sigma, b_\sigma)$, where we set the shape parameter a_σ and scale parameter b_σ to each be 0.001. It is convenient to parameterize BKMR by $\lambda \equiv \tau\sigma^{-2}$, and we assumed a Gamma prior distribution for the variance component λ having mean and variance each set to 100 (Let a_λ, b_λ denote corresponding shape and rate parameters). For the distribution of the slab $f_1(r_m)$ of the variable selection prior in equation (3) of the main text we assumed $\rho_m = 1/r_m \sim \text{Unif}(a_r, b_r)$.

We assumed that the prior probability π that a mixture component z_m (group $\mathcal{S}_g$) was included in the model had a beta distribution with shape parameters a_π, b_π. We assumed that if a large number of mixture components (or groups) were under investigation, then a smaller subset of these components would likely be predictive of health. We selected $a_\pi = 2$ and $b_\pi = 6$ such that a priori we would expect 25% of the components (groups) to be included. Finally, for the hierarchical variable selection approach, for $\boldsymbol{\pi}_{\mathcal{S}_g}$ we assumed that each component of the same group was equally likely to be included in the model.

A.1. Hyperparameters for the simulation study

We set the following values for the hyperparameters in our simulation studies (described in Section 3.1 of the main text). For the uniform slab prior on $\rho_m = 1/r_m$ we set $a_r = 0$ and $b_r = 100$. For the beta prior on π we set a_π and b_π as described in Section A for the $M = 13$ component scenario and $a_\pi = b_\pi = 1$ for the $M = 3$ component scenario.

B. Estimation and prediction

Here we describe the Markov chain Monte Carlo (MCMC) sampler used to fit Bayesian kernel machine regression (BKMR) with component-wise and hierarchical variable selection described in Section 2 of the main text (see Box below for a summary of the model specification and notation).

To apply a standard Gibbs sampler in which samples are generated from the full conditional distributions of each of the parameters, the augmented kernel matrix $\mathbf{K}_{\mathbf{Z},\mathbf{r}}$ (Section 2.2 of the main text) must be inverted at each iteration of the sampler, which can lead to numerical instability if the kernel is nearly singular. This problem can be avoided by integrating out $\mathbf{h}$, and obtaining posterior samples from the marginal posterior distribution of the remaining parameters.

Summary of Bayesian kernel machine regression and variable selection

Model specification

$$\text{Likelihood} \begin{cases} \mathbf{y} \mid \mathbf{h}, \boldsymbol{\beta}, \sigma^2, \mathbf{X} & \sim \mathrm{N}(\mathbf{h} + \mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_n) \\ \mathbf{h} \mid \tau, \mathbf{r}, \mathbf{Z} & \sim \mathrm{N}(\mathbf{0}, \tau \mathbf{K}_{\mathbf{Z},\mathbf{r}}) \end{cases}$$

$$\text{Component-wise variable selection} \begin{cases} r_m \mid \delta_m & \sim \delta_m \mathrm{Unif}^{-1}(a_r, b_r) + (1-\delta_m) P_0, \\ \delta_m \mid \pi & \sim \text{Bernoulli}(\pi), \end{cases}$$

$$\text{Hierarchical variable selection} \begin{cases} r_m \mid \delta_m & \sim \delta_m \mathrm{Unif}^{-1}(a_r, b_r) + (1-\delta_m) P_0, \\ \boldsymbol{\delta}_{\mathcal{S}_g} \mid \omega_g & \sim \text{Multinomial}(\omega_g, \boldsymbol{\pi}_{\mathcal{S}_g}) \\ \omega_g \mid \pi & \sim \text{Bernoulli}(\pi), \end{cases}$$

$$\text{Priors} \begin{cases} \boldsymbol{\beta} & \sim 1 \\ \sigma^{-2} & \sim \text{Gamma}(a_\sigma, b_\sigma) \\ \lambda \equiv \tau\sigma^{-2} & \sim \text{Gamma}(a_\lambda, b_\lambda) \\ \pi & \sim \text{Beta}(a_\pi, b_\pi) \end{cases}$$

Notation

$$\text{Indices} \begin{cases} i = 1, \ldots, n & \text{subjects} \\ m = 1, \ldots, M & \text{mixture components} \\ g = 1, \ldots, G & \text{mixture groups} \end{cases}$$

$$\text{Data} \begin{cases} \mathbf{y} = (y_1, \ldots, y_n)^T & \text{health outcomes} \\ \mathbf{X} & \text{covariate design matrix with rows } \mathbf{x}_i^T \\ \mathbf{Z} & \text{exposure design matrix with rows } \mathbf{z}_i^T = (z_{i1}, \ldots, z_{iM}) \\ \{\mathcal{S}_g\}_{g=1,\ldots,G} & \text{partition of mixture components into groups} \end{cases}$$

$$\text{Parameters} \begin{cases} \mathbf{h} = (h_1, \ldots, h_n)^T & \text{subject-specific health effects } h_i = h(\mathbf{z}_i) \\ \mathbf{K}_{\mathbf{Z},\mathbf{r}} & n \times n \text{ kernel matrix for variable selection with} \\ & (i,j)\text{-element } \exp\left\{-\sum_{m=1}^M r_m (z_{im} - z_{jm})^2\right\} \\ \mathbf{r} = (r_1, \ldots, r_M)^T & \text{augmented variables in kernel matrix for} \\ & \text{variable selection, which controls smoothness of } h(\cdot) \\ \boldsymbol{\delta} = (\delta_1, \ldots, \delta_M)^T & \text{inclusion indicators for mixture components} \\ \boldsymbol{\delta}_{\mathcal{S}_g} = (\delta_m)_{z_m \in \mathcal{S}_g} & \text{inclusion indicators for the components in group } g \end{cases}$$

B.1. MCMC sampler

Integrating over π and $\mathbf{h}$ and applying the prior distributions specified in Section A of the Supplementary Material, the posterior is given by

$$f(\boldsymbol{\beta}, \sigma^2, \lambda, \mathbf{r}, \boldsymbol{\theta}, \mid \mathbf{y}) \propto \mathrm{N}(\mathbf{y} \mid \mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{V}_{\lambda,\mathbf{Z},\mathbf{r}}) \left\{ \prod_{m=1}^M f(r_m \mid \delta_m) \right\} f(\boldsymbol{\theta}) \tag{1}$$

$$\times \text{Gamma}(\sigma^{-2} \mid a_\sigma, b_\sigma) \text{Gamma}(\lambda \mid a_\lambda, b_\lambda),$$

where $\mathbf{y} = (y_1, \ldots, y_n)^T$, $\mathbf{X}$ is the covariate design matrix with rows $\mathbf{x}_i^T$, and $\mathbf{V}_{\lambda,\mathbf{Z},\mathbf{r}} = \mathbf{I}_n + \lambda \mathbf{K}_{\mathbf{Z},\mathbf{r}}$. The $\boldsymbol{\theta}$ corresponds to the parameter vector for variable selection, such that $\boldsymbol{\theta} = \boldsymbol{\delta}$ for component-wise selection and $\boldsymbol{\theta} = (\boldsymbol{\delta}, \boldsymbol{\omega})$ for hierarchical variable selection, where $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_M)^T$ is the vector of component indicators and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_G)^T$ is the vector of group indicators. The term $f(\boldsymbol{\theta})$ is equal to $\Gamma\left(\sum_m \delta_m + a_\pi\right) \Gamma\left(M - \sum_m \delta_m + b_\pi\right)$ under component-wise variable selection and is equal to $\Gamma\left(\sum_g \omega_g + a_\pi\right) \Gamma\left(G - \sum_g \omega_g + b_\pi\right) \prod_{g=1}^G \left\{\prod_{m:z_m \in \mathcal{S}_g} \mathrm{Card}(\mathcal{S}_g)^{-\delta_m}\right\}$ under hierarchical variable selection, where $\sum_{m:z_m \in \mathcal{S}_g} \delta_m = \omega_g$ for each g.

We updated $\boldsymbol{\beta}$ and σ^2 using separate Gibbs steps, with full conditionals given by

$$\boldsymbol{\beta} \mid \sigma^2, \lambda, \mathbf{r}, \mathbf{y} \sim \mathrm{N}(\boldsymbol{\beta} \mid \mathbf{V}_{\boldsymbol{\beta}} \mathbf{X}^T \mathbf{V}_{\lambda,\mathbf{Z},\mathbf{r}}^{-1} \mathbf{y}, \sigma^2 \mathbf{V}_{\boldsymbol{\beta}}), \quad \text{where } \mathbf{V}_{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{V}_{\lambda,\mathbf{Z},\mathbf{r}}^{-1} \mathbf{X})^{-1},$$
$$\sigma^{-2} \mid \boldsymbol{\beta}, \lambda, \mathbf{r}, \mathbf{y} \sim \mathrm{Gamma}\{\alpha_\sigma + n/2, b_\sigma + WSS_{\boldsymbol{\beta},\lambda,\mathbf{r}}/2\},$$

where $WSS_{\boldsymbol{\beta},\lambda,\mathbf{r}}$ is the weighted sum of squares $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{V}_{\lambda,\mathbf{Z},\mathbf{r}}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$. We updated λ using a Metropolis-Hastings step, with full conditional $f(\lambda \mid \boldsymbol{\beta}, \mathbf{r}, \boldsymbol{\delta}, \mathbf{y}, \mathbf{X}, \mathbf{Z}) \propto (\det \mathbf{V}_{\lambda,\mathbf{Z},\mathbf{r}})^{-1/2} \times \exp\{-WSS_{\boldsymbol{\beta},\lambda,\mathbf{r}}/(2\sigma^2)\} \times \mathrm{Gamma}(\lambda \mid a_\lambda, b_\lambda)$. We used a gamma proposal distribution with mean set to the value of λ from the previous iteration and variance tuned to produce a good acceptance rate.

Component-wise variable selection. Because sampling individually from the full conditionals of $\mathbf{r}$ and $\boldsymbol{\delta}$ leads to a reducible Markov chain, we instead sampled $(\mathbf{r}, \boldsymbol{\delta})$ jointly from their conditional distribution

$$f(\mathbf{r}, \boldsymbol{\delta}, \mid \boldsymbol{\beta}, \sigma^2, \lambda, \mathbf{y}) \propto \Gamma\left(\sum_m \delta_m + a_\pi\right) \Gamma\left(M - \sum_m \delta_m + b_\pi\right) \left\{\prod_{m=1}^M f(r_m \mid \delta_m)\right\}$$

by adapting the Metropolis-Hastings algorithm from Sha et al. (2004). To obtain a sample at the sth iteration of the MCMC, this procedure generates a proposal $(\mathbf{r}^*, \boldsymbol{\delta}^*)$ by randomly

selecting one of the following moves:

1. Randomly select $m \in \{1, \ldots, M\}$ and set $\delta_m^* = 1 - \delta_m^{(s-1)}$. If $\delta_m^* = 0$ set $r_m^* = 0$; else, generate the proposal r_m^* from a proposal distribution $Q_1(\cdot)$.

2. Among the components of $\boldsymbol{\delta}^{(s-1)}$ equal to one, randomly choose one (say δ_m^*) and generate the corresponding r_m^* from a proposal distribution $Q_2(\cdot \mid r_m^{(s-1)})$.

For Q_1 we sampled $\rho_m^* = 1/r_m^* \sim \text{Unif}(a_r, b_r)$, and we set Q_2 to be a truncated normal distribution with mean set to $r_m^{(s-1)}$, variance tuned to have a good acceptance rate for those iterations where move 2 was selected, and truncated to fall within the range (b_r^{-1}, a_r^{-1}).

Hierarchical variable selection. We generalized the previous Metropolis-Hastings sampling scheme to sample $(\mathbf{r}, \boldsymbol{\delta}, \boldsymbol{\omega})$ jointly from their conditional distribution

$$f(\mathbf{r}, \boldsymbol{\delta}, \boldsymbol{\omega}, \mid \boldsymbol{\beta}, \sigma^2, \lambda, \mathbf{y}) \propto \Gamma\left(\sum_g \omega_g + a_\pi\right) \Gamma\left(G - \sum_g \omega_g + b_\pi\right) \left\{\prod_{m=1}^{M} f(r_m \mid \delta_m)\right\} \times \prod_{g=1}^{G} \left\{ \prod_{m: z_m \in \mathcal{S}_g} \text{Card}(\mathcal{S}_g)^{-\delta_m} \right\}, \quad \text{where} \sum_{m: z_m \in \mathcal{S}_g} \delta_m = \omega_g.$$

At the sth iteration of the MCMC, we generated a proposal $(\mathbf{r}^*, \boldsymbol{\delta}^*, \boldsymbol{\omega}^*)$ by randomly selecting one of the following moves:

1. **Change the state of a group $\mathcal{S}_g$.** Randomly select $g \in \{1, \ldots, G\}$ and set $\omega_g^* = 1 - \omega_g^{(s-1)}$. If $\omega_g^* = 0$ set $\delta_{m'}^* = 0$ where m' is the component of group $\mathcal{S}_g$ with $\delta_{m'}^{(s-1)} = 1$; else, randomly select a component m^* from group $\mathcal{S}_g$, set $\delta_{m^*}^* = 1$ and generate the proposal $r_{m^*}^*$ from a proposal distribution $Q_1(\cdot)$.

2. **Switch components within a group $\mathcal{S}_g$.** Among the groups with $\text{Card}(\mathcal{S}_g) > 1$ and $\omega_g = 1$, randomly choose one, say $\mathcal{S}_g^*$. For the component m' of $\mathcal{S}_g^*$ with $\delta_{m'}^{(s-1)} = 1$, set

$\delta^*_{m'} = 0$ and $r^*_{m'} = 0$. Among the components of $\boldsymbol{\delta}_{\mathcal{S}^*_g}$ equal to zero, randomly select one, say m^*, set $\delta^*_{m*} = 1$ and generate the corresponding $r^*_{m^*}$ from a proposal distribution $Q_2(\cdot)$.

3. **Refinement step.** Among the components of $\boldsymbol{\delta}^{(s-1)}$ equal to one, randomly choose one (say δ^*_m) and generate the corresponding r^*_m from a proposal distribution $Q_3(\cdot \mid r^{(s-1)}_m)$.

For Q_1 and Q_2 we sampled $\rho^*_m = 1/r^*_m \sim \text{Unif}(a_r, b_r)$, and we set Q_3 to be a truncated normal distribution with mean set to $r^{(s-1)}_m$, variance tuned to have a good acceptance rate for those iterations where move 3 was selected, and truncated to fall within the range (b_r^{-1}, a_r^{-1}).

B.2. Estimating subject-specific health effects

To obtain posterior samples of h_i, which represents the subject-specific association between exposure to the environmental mixture on health, first note that the posterior density $f(\mathbf{h}, \boldsymbol{\beta}, \sigma^2, \lambda, \mathbf{r}, \boldsymbol{\theta} \mid \mathbf{y})$ can be decomposed in the usual way as $f(\mathbf{h} \mid \boldsymbol{\beta}, \sigma^2, \lambda, \mathbf{r}, \boldsymbol{\theta}, \mathbf{y}) \times f(\boldsymbol{\beta}, \sigma^2, \lambda, \mathbf{r}, \boldsymbol{\theta} \mid \mathbf{y})$, where the conditional distribution of $\mathbf{h}$ is given by

$$\mathbf{h} \mid \boldsymbol{\beta}, \sigma^2, \lambda, \mathbf{r}, \boldsymbol{\theta}, \mathbf{y} \sim \text{N}(\lambda \mathbf{K}_{\mathbf{Z},\mathbf{r}} \mathbf{V}^{-1}_{\lambda,\mathbf{Z},\mathbf{r}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \sigma^2 \lambda \mathbf{K}_{\mathbf{Z},\mathbf{r}} \mathbf{V}^{-1}_{\lambda,\mathbf{Z},\mathbf{r}}). \tag{2}$$

Therefore for each sample $(\boldsymbol{\beta}^{(s)}, \sigma^{2(s)}, \lambda^{(s)}, \mathbf{r}^{(s)}, \boldsymbol{\theta}^{(s)})$ generated from the marginal posterior in (1) with our MCMC sampling algorithm, we generated a sample $\mathbf{h}^{(s)}$ from (2).

B.3. Predicting health effects at new exposure profiles

A critical aim in analyzing the health effects of environmental mixtures is to estimate the (multivariate) exposure-response function. This entails not only estimating $h_i = h(\mathbf{z}_i)$ at the observed data points, but also predicting h at a collection of unobserved exposure profiles. Predicted health effects are also of importance to regulators, who may wish to evaluate the

health benefit attributable to proposed regulatory scenarios. For example, if $\bar{\mathbf{z}}$ denotes the vector of mean pollutant levels, and $\mathbf{z}_{reg}$ denotes the anticipated levels under a particular regulatory scenario, then the health impact could be estimated as $h(\mathbf{z}_{reg}) - h(\bar{\mathbf{z}})$.

Let $\mathbf{Z}_{new}$ be the $P \times M$ design matrix (with rows $\mathbf{z}_p^{new}$) of new exposure profiles, and let $\mathbf{h}_{new}^T = (h_1^{new}, \ldots, h_P^{new})$ denote the corresponding desired predictions. In the mixed model representation of KMR, we can consider the joint distribution of the multi-pollutant risks at the observed and unobserved exposure profiles,

$$\begin{pmatrix} \mathbf{h} \\ \mathbf{h}_{new} \end{pmatrix} \sim \mathrm{N}\left\{ \mathbf{0}, \tau \begin{pmatrix} \mathbf{K}_{\mathbf{Z},\mathbf{r}} & \mathbf{K}_{\mathbf{Z},\mathbf{Z}_{new},\mathbf{r}} \\ \mathbf{K}_{\mathbf{Z},\mathbf{Z}_{new},\mathbf{r}}^T & \mathbf{K}_{\mathbf{Z}_{new},\mathbf{r}} \end{pmatrix} \right\},$$

where $\mathbf{K}_{\mathbf{Z},\mathbf{r}}$ denotes the augmented kernel matrix (defined in Section 2.2 of the main text), $\mathbf{K}_{\mathbf{Z},\mathbf{Z}_{new},\mathbf{r}}$ is the $n \times P$ matrix with (i,j)-element $\exp\left\{-\sum_{m=1}^M r_m(z_{im} - z_{jm}^{new})^2\right\}$, and $\mathbf{K}_{\mathbf{Z}_{new},\mathbf{r}}$ is the $P \times P$ matrix with (i,j)-element $\exp\left\{-\sum_{m=1}^M r_m(z_{im}^{new} - z_{jm}^{new})^2\right\}$. Following routine calculations, the conditional posterior distribution of $\mathbf{h}_{new}$ is given by

$$\mathbf{h}^{new} \mid \boldsymbol{\beta}, \sigma^2, \lambda, \mathbf{r}, \boldsymbol{\theta}, \mathbf{y} \sim \mathrm{N}\{\lambda \mathbf{K}_{\mathbf{Z},\mathbf{Z}_{new},\mathbf{r}}^T \mathbf{V}_{\lambda,\mathbf{Z},\mathbf{r}}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \sigma^2\lambda(\mathbf{K}_{\mathbf{Z}_{new},\mathbf{r}} - \lambda \mathbf{K}_{\mathbf{Z},\mathbf{Z}_{new},\mathbf{r}}^T \mathbf{V}_{\lambda,\mathbf{Z},\mathbf{r}}^{-1} \mathbf{K}_{\mathbf{Z},\mathbf{Z}_{new},\mathbf{r}})\}. \tag{3}$$

In theory, we could obtain predictions by generating $\mathbf{h}^{new}$ from its conditional distribution at each iteration of the MCMC. However, because a large number of predictions are typically desired (e.g., to plot the exposure-response function on a grid of points), this posterior simulation can be very computationally expensive because it requires repeatedly simulating from a high-dimensional multivariate normal distribution. Therefore, we propose to approximate the posterior mean (variance) of $\mathbf{h}^{new}$ as its conditional posterior mean (variance) from equation (3) evaluated at the estimated posterior mean of the other parameters.

References

Sha, N., Vannucci, M., Tadesse, M., Brown, P., Dragoni, I., Davies, N., Roberts, T., Contestabile, A., Salmon, M., and Buckley, C. (2004). Bayesian variable selection in multinomial probit models to identify molecular signatures of disease stage. *Biometrics* **60,** 812–819.

C. Tables and figures

	Single-metal models			**Multi-metal model**		
	Est.	SE	p-value	Est.	SE	p-value
Pb	0.066	0.073	0.37	0.065	0.081	0.42
Mn	0.021	0.056	0.72	0.049	0.071	0.49
As	−0.066	0.058	0.25	−0.102	0.067	0.13

Table 1. Estimated coefficients, standard errors (SE) and p-values from linear models of the association between metal exposure and the motor composite score (MCS) in the Bangladesh application (Section 4 of the main text).
As = arsenic, Mn = manganese, Pb = lead.

Constituent	Single-constituent models			Multi-constituent model		
	Est.	SE	p-value	Est.	SE	p-value
Group $\mathcal{S}_1$						
Al	2.28	0.94	0.02	−6.20	6.20	0.32
Si	3.06	1.09	0.01	10.23	7.83	0.19
Ti	2.20	0.97	0.02	0.40	3.53	0.91
Ca	2.50	1.05	0.02	0.51	2.93	0.86
K	2.19	1.07	0.04	−3.32	3.30	0.32
Cu	1.99	1.03	0.06	−0.40	1.92	0.83
Mn	2.82	1.07	0.01	3.07	1.95	0.12
Group $\mathcal{S}_2$						
Ni	−0.28	0.88	0.75	0.69	1.43	0.63
V	−1.94	1.14	0.09	−3.34	1.78	0.06
Zn	0.22	2.12	0.92	−2.12	2.72	0.44
Group $\mathcal{S}_3$						
S	−0.39	0.66	0.56	0.27	0.90	0.77
Group $\mathcal{S}_4$						
Cl	−0.41	0.99	0.68	0.55	1.19	0.65
Group $\mathcal{S}_5$						
BC	0.34	0.91	0.71	1.86	1.69	0.28

Table 2. Estimated coefficients, standard errors (SE) and p-values from linear mixed models of the association between elemental air pollution constituents and heart rate in the toxicology application (Section 5 of the main text).
Al = aluminum, Si = silicon, Ti = titanium, Ca = calcium, Ni = nickel, V = vanadium, Zn = zinc, S = sulphur, bc = black carbon, Cu = copper, K = potassium, Cl = chlorine, Mn = manganese.

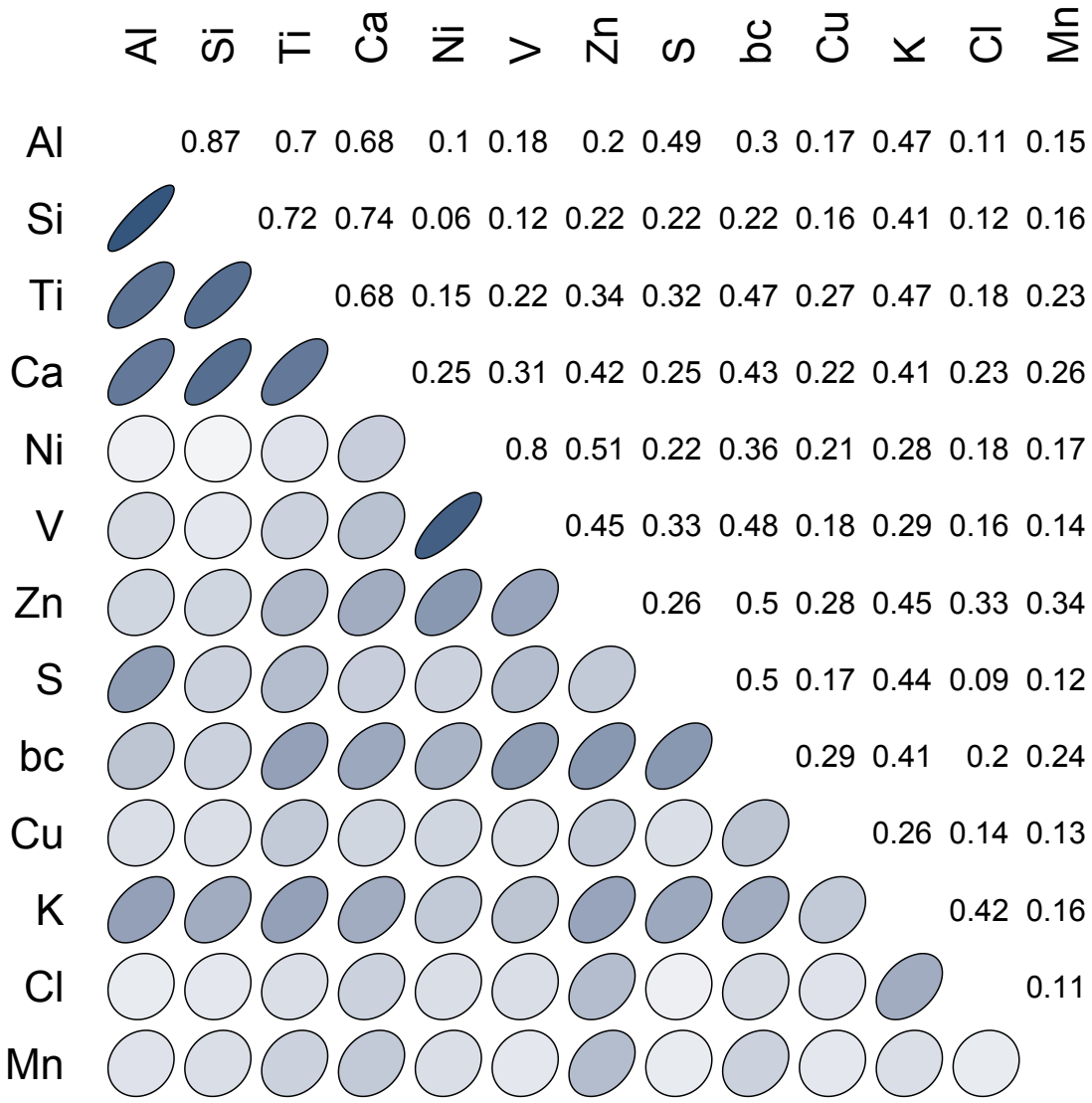

Figure 1. Correlation matrix of Boston air pollution data used for the $M = 13$ component scenario of the Simulation Study (Section 3 of the main text). The components are aluminum (Al), silicon (Si), titanium (Ti), calcium (Ca), nickel (Ni), vanadium (V), zinc (Zn), sulphur (S), black carbon (bc), copper (Cu), potassium (K), chlorine (Cl), and manganese (Mn).

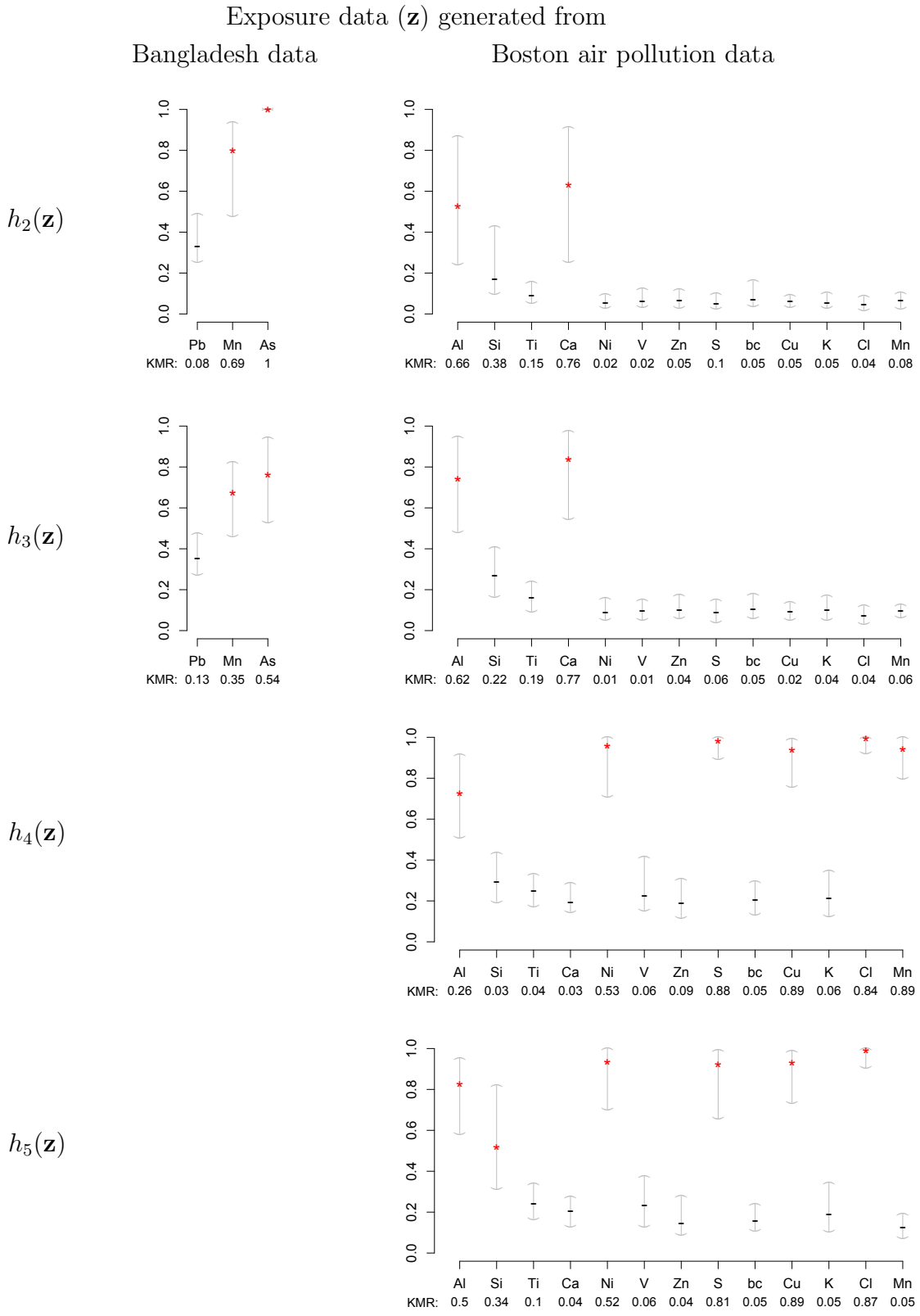

Figure 2. Median (25%, 75%) of the posterior inclusion probabilities from Bayesian kernel machine regression (BKMR) with component-wise variable selection, across 100 simulated datasets for each of four true $h(\mathbf{z})$ functions. The vector of exposure data $\mathbf{z}$ were generated either based on the Bangladesh data with $M = 3$ mixture components or on the Boston air pollution data with $M = 13$ mixture components. The truly associated components are shown in red. The proportion of simulation iterations for which each mixture component had p-value < 0.05 under the garrote test for Kernel machine regression (KMR) is printed below the x-axis.

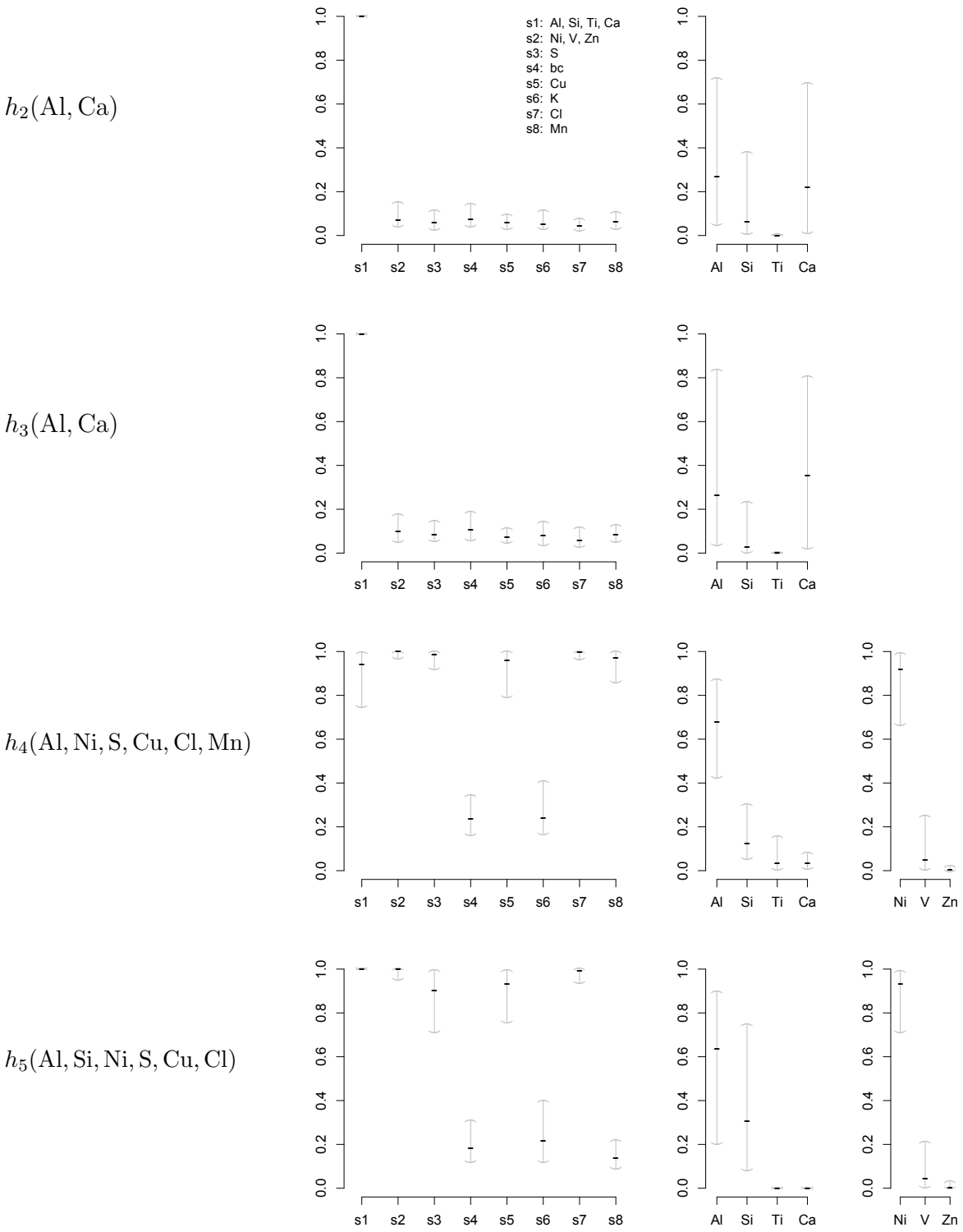

Figure 3. Median (25%, 75%) of the posterior inclusion probabilities from Bayesian kernel machine regression (BKMR) with hierarchical variable selection, across 100 simulated datasets for each of four true $h(\mathbf{z})$ functions. Exposure data $\mathbf{z}$ were generated based on the Boston air pollution data with $M = 13$ mixture components partitioned into 8 groups. Leftmost plots show the posterior inclusion probabilities for each group, and middle and right plots show the conditional posterior inclusion probabilities for the components in groups 1 and 2 given that the group was included in the model.

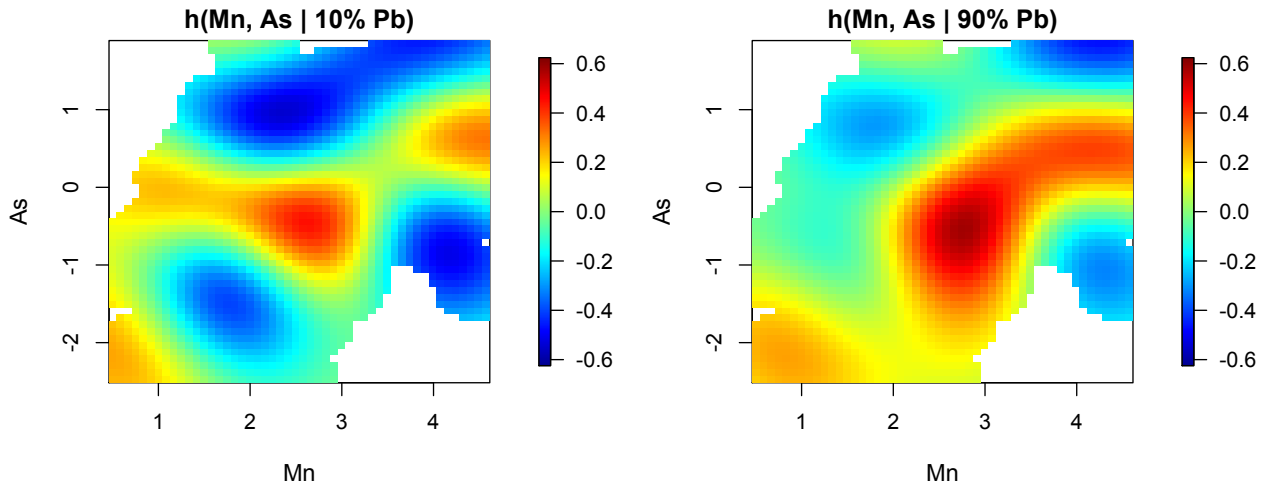

Figure 4. Relationship of manganese (Mn) and arsenic (As) with the motor composite score (MCS), for lead (Pb) fixed at its 10th (left panel) and 90th (right panel) percentile.

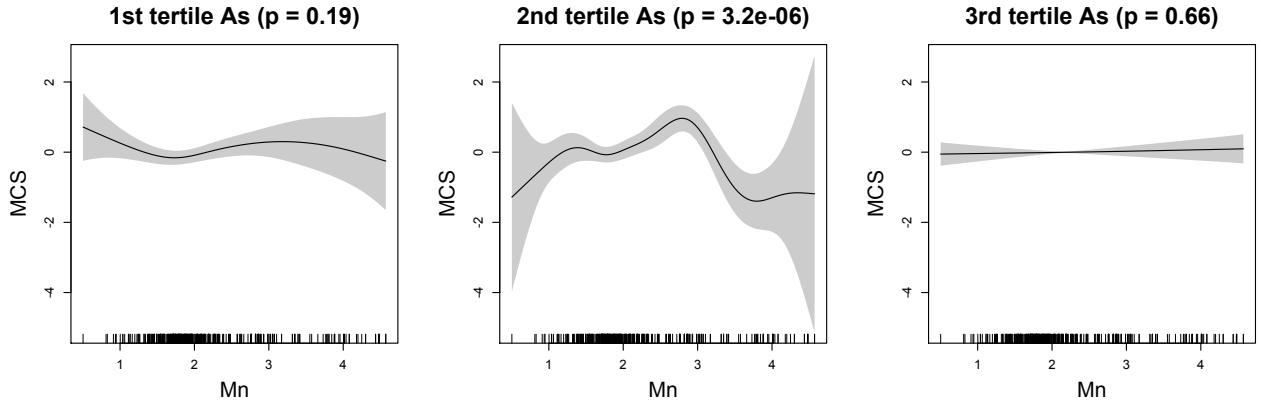

Figure 5. Estimated exposure response function of manganese (Mn) with the motor composite score (MCS) at each tertile of arsenic (As) exposure from fitting a generalized additive model to the Bangladesh application (Section 4 of the main text). P-values for the smooth terms are printed above each plot.