RESEARCH ARTICLE

Frequentist Grouped Weighted Quantile Sum Regression for Correlated Chemical Mixtures

Daniel Rud[1] | Md Mostafijur Rahman[1,2] | Anny H. Xiang[3] | Rob McConnell[1] | Fred Lurmann[4] | Michael J. Kleeman[5] |
Joel Schwartz[6] | Zhanghua Chen[1] | Sandy Eckel[1] | Juan Pablo Lewinger[1]

[1]Department of Population and Public Health Sciences, University of Southern California, Los Angeles, CA, USA | [2]Department of Environmental Health Sciences, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA | [3]Department of Research & Evaluation, Kaiser Permanente Southern California, Pasadena, CA, USA | [4]Sonoma Technology, Inc., Petaluma, CA, USA | [5]Department of Civil and Environmental Engineering, University of California, Davis, CA, USA | [6]Department of Environmental Health and Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Correspondence: Juan Pablo Lewinger (lewinger@usc.edu)

ABSTRACT

As individuals are exposed to a myriad of potentially harmful pollutants every day, it is important to determine which actors have the greatest influence on health outcomes. However, jointly modeling the associations of multiple pollutant exposures is often hindered by the presence of highly correlated chemicals originating from a common source. A popular approach to analyzing associations between a disease outcome and several highly correlated exposures is Weighted Quantile Sum Regression (WQSR) modeling. WQSR provides increased stability in estimating model parameters but requires data splitting to estimate individual and group effects of chemicals, which reduces the power of the approach. A recent Bayesian implementation of WQSR regression provides a model fitting procedure that avoids data splitting at the cost of high computational expense on large data. In this paper, we introduce a Frequentist Grouped Weighted Quantile Sum Regression (FGWQSR) model that can be fitted efficiently to large datasets without requiring data splitting. FGWQSR produces estimates of the joint effect of mixture groups and of individual chemicals, and likelihood-ratio-based tests that account for FGWQSR's non-standard asymptotics. We demonstrate that FGWQSR is well calibrated for type-I errors while outperforming both Bayesian Grouped Weighted Quantile Sum Regression and Quantile Logistic Regression in terms of statistical power to detect the effects of mixture groups and individual chemicals. In addition, we show that FGWQSR is robust to model misspecification and can be fitted on large datasets in a fraction of the time required for BGWQSR. We apply FGWQSR to a dataset of 317 767 mother-child pairs with exposure profiles generated by chemical transport models to study the associations between several components found in particulate matter with an aerodynamic diameter smaller than 2.5 μm (PM$_{2.5}$) and child Autism Spectrum Disorder (ASD) diagnosis before age 5. PM$_{2.5}$ copper and PM$_{2.5}$ crustal material are found to be statistically significantly associated with ASD diagnosis by five years of age.

1 | Introduction

Every day, people are exposed to a variety of potentially harmful chemicals through the air they breathe, the food and water they consume, and the items they touch. Chemicals can originate from many sources, including vehicle emissions, industrial emissions, and consumer products [1, 2]. Recent environmental epidemiology studies have attempted to uncover underlying associations between several chemical exposure levels and diverse health outcomes [3–7]. Many of these studies focus on the link between environmental pollution and the risk of different cancers, pulmonary illnesses (asthma), diabetes, heart disease, and stroke. In these studies, researchers have utilized various statistical methods to study both the joint and individual effects of highly correlated chemical exposures.

Models for complex chemical mixtures can be described as statistical models equipped to handle the underlying correlation of chemical mixtures and provide means for inference on the relative importance of individual mixture components. We define a mixture as any collection of exposures that may be grouped by having a common source. For example, one might consider jointly investigating exposure to hydrocarbons, carbon monoxide, nitrogen oxides, and particulate matter (PM) components that arise as the products of combustion from diesel fuels in vehicles [8]. Because chemical mixtures often originate from common sources, their individual chemical components can be highly correlated. In smaller studies, traditional modeling techniques including linear and logistic regression with highly correlated exposures (i.e., multicollinearity), often lead to inflated standard error estimates, which can in turn result in a lack of power to detect important effects. Additionally, the inclusion of highly multicollinear covariates may lead to the *reversal paradox*, a situation in which the directionality of effect estimates are reversed in the presence of co-varying variables [9]. This occurs when the correlation between exposures induces a negative correlation between estimated model parameters, which may in turn "flip" a parameter's direction of association. Thus, a framework for modeling mixtures must be able to handle a non-negligible degree of underlying correlation within chemical mixtures.

A traditional approach for dealing with covariate multicollinearity is to employ shrinkage (penalization/regularization) techniques [10, 11]. Penalized regression including Lasso and Ridge regression, most commonly used in high-dimensional $p >> n$ scenarios, include a model complexity penalty in their loss functions that controls the bias-variance trade-off and adds stability to estimated effects in the presence of multicollinearity. Lasso models in particular benefit from variable selection, as its L_1 penalty allows estimated effects to be set to exactly zero. However, Lasso models tend to select an arbitrary covariate among a group of correlated covariates with fairly low specificity [10, 12, 13]. By contrast, Ridge regression does not perform variable selection as estimated parameter effects are never exactly zero. In the presence of highly correlated covariates, Ridge regression tends to "spread" the estimated group effect across all covariates within a correlated group [10]. Although either behavior may not be of concern when the goal is a prediction of the outcome, in our epidemiological context, both the lasso and ridge properties described above would be undesirable since we would like to identify important individual chemicals within correlated mixtures. The use of penalized methods also precludes performing inference on parameters, such as computing standard errors, confidence intervals, and performing hypothesis tests, which are critical for quantifying the association between a chemical and the disease outcome of interest [14].

To address the limitations of penalized regression for modeling mixtures, Carrico et al. proposed the Weighted Quantile Sum Regression (WQSR) model [15]. In the last several years, extensions and improvements of WQSR have been developed [16–18]. To handle nonlinearities, all of the quantile-sum approaches first transform individual chemical values (e.g., concentrations) into quantiles. WQSR was introduced for a single group of chemicals. Individual chemical weights, which are positive and add to one, quantify the importance of a chemical within a group (see Equation (2) and methods section below for a summary of the WQSR model). The group effect parameter represents the effect on the outcome corresponding to a unit increase in the weighted sum of the chemicals within the mixture group. A two-step model fitting procedure is used to fit the WQSR model. This requires a training/testing split of the data set to separately estimate the chemical weights and group effect. The individual chemical weights are first estimated through an importance-weighted bootstrap sampling procedure, and a group effect is subsequently estimated through a regression on the linearly weight-combined chemical exposure using the testing dataset. The estimation procedure of the group effect yields inferential statistics, which quantify the significance of the group of chemicals. Chemical weight estimates can be analyzed to infer the relative importance of components within the mixture.

Wheeler et al. extended WQSR to mixture problems that require joint analysis of multiple chemical groups [16]. The Grouped Weighted Quantile Sum Regression (GWQSR) approach estimates a set of group effects and accompanying chemical weight sets for each mixture group. This framework enables the estimation of mixture group effects that may have differing directions of association with the outcome. For example, a potential chemical mixture analysis could comprise both a group of chemicals previously found to have a deleterious health effect and a group of chemicals known to protect against disease. The model for which GWQSR performs parameter estimation can be seen in Equation (2) with $G \geq 1$. Similar to WQSR, GWQSR also uses a two-step model fitting procedure that requires a training/testing data split and provides inferential measures for the estimated group effects. As the weight estimation procedure relies on numerical nonlinear optimization performed on bootstrap samples, GWQSR is computationally intensive and slows down quickly as a function of sample size and mixture covariate space.

With the goal of applying GWQSR without the need for data splitting, Wheeler et al. proposed a Bayesian formulation of the model for estimating the parameters in Equation (2) [17]. Instead of a two-step fitting procedure, Bayesian Grouped Weighted Quantile Sum Regression (BGWQSR) relies on Markov Chain Monte Carlo (MCMC) to sample from the posterior distribution of chemical mixture parameters from which tests and posterior credible intervals can be derived. Because BGWQSR avoids data splitting, the model detects individual and group effects with greater power than GWQSR.

Although BGWQSR is an elegant solution to the problem of information loss due to data splitting, the approach has limitations that hinder its general applicability in epidemiological studies. To begin, the model fitting procedure of BGWQSR becomes computationally intensive as a function of sample size and eventually becomes infeasible for large sample size studies. In larger studies, utilizing BGWQSR requires condensing the dataset through a subsampling procedure to run models that can be fitted in reasonable amounts of time, highlighting BGWQSR's inability to handle large datasets. Furthermore, BGWQSR tends to be conservative in its power to detect significant group effects and individual chemical effects, which does not come as a surprise since Bayesian methods in general are biased toward their prior distributions. Although BGWQSR models can reduce said bias by incorporating priors that contain relevant information, the default configuration for BGWQSR utilizes uninformative priors, where many users of the method often would not modify the predefined model settings. Finally, some researchers may simply be unfamiliar with inferential procedures in a Bayesian context and can have difficulty interpreting BGWQSR models. To appropriately analyze a BGWQSR model, one should be familiar with extracting posterior estimates and credible intervals from MCMC results, verifying model convergence, and examining both traceplots and autocorrelation plots.

To overcome the limitations of BGWQSR, we introduce a Frequentist Grouped Weighted Quantile Sum Regression (FGWQSR) model. FGWQSR employs a Maximum Likelihood (ML) fitting procedure to estimate the parameters found in Equation (2) with corresponding inferential measures for both weights and group effects in addition to adjusting covariates. In contrast to GWQSR, which also calls itself a Frequentist method, FGWQSR estimates model parameters in a single ML procedure without requiring any data splitting. Inferences for group effects and weights are performed using Likelihood Ratio tests (LRT). In addition, FGWQSR can handle large datasets in a fraction of the time it would take BGWQSR to fit, enabling epidemiologists to use weighted quantile sum mixture models in large studies. We showcase FGWQSR's ability to handle large datasets in Section 2.4, where we fit an FGWQSR model on 317767 observations.

2 | Materials and Methods

2.1 | Model Specification

We first describe the WQSR model and the approaches that researchers have taken to fit it in the past. Afterward, we introduce our proposal for a simpler Frequentist fit of the model. In the context of chemical mixture modeling, we are generally interested in studying the relationship between $G \geq 1$ non-overlapping chemical mixtures and an outcome Y, where each chemical mixture group is composed of $c_g \geq 1$ distinct chemicals. For concreteness, we describe the weighted-quantile sum regression for a binary disease outcome (presence vs. absence of disease); however, the model can accommodate any generalized linear model outcome (e.g., continuous, binary, count). In addition to chemical exposures, we assume there are r covariates of interest the researcher wants to adjust for. The relation between the chemical mixture exposures, covariates, and

risk of disease is described through the following logistic regression model:

$$y_i \sim \text{Bernoulli}(\pi_i) \tag{1}$$

$$\text{logit}(\pi_i) = c_0 + \sum_{g=1}^{G} \gamma_g \left(\sum_{k=1}^{c_g} w_{g,k} \cdot q_{g,k,i} \right) + \sum_{r=1}^{R} \phi_r z_{r,i} \ i = 1, \ldots, n \tag{2}$$

where, for subject i, y_i represents the observed disease outcome, π_i the probability of disease and $q_{g,k,i}$ the exposure to chemical k in mixture group g. As described above, the chemical exposure is quantized to handle potential nonlinearities in exposure-outcome relations. Given a number of prespecified quantiles q, the chemical exposure data are quantized so that the first quantile of the exposure is coded as 0 and the qth quantile as $q - 1$. With this coding, the intercept term c_0 can be interpreted as the baseline odds of the outcome for individuals in the first quantile of exposure across all chemical exposures and baseline levels of the covariates. The parameter γ_g captures the overall effect of chemical mixture g, while $w_{g,k}$ represents the fraction of the effect due to the kth chemical in the mixture such that $\sum_{k=1}^{c_g} w_{g,k} = 1$ and $w_{g,k} \in [0, 1]$. Finally, $z_{r,i}$ denotes the rth adjustment covariate value for subject i, and ϕ_r the effect of covariate r on the outcome.

Importantly, all chemicals within a particular group are modeled to affect the outcome in the same direction—either positive or negative. We refer to this as the *directionality assumption*. It is noted that (2) can be viewed as a reparameterization of standard logistic regression under the constraint that all coefficients within a mixture group are in the same direction [18]. By forcing all effect estimates within a mixture group to be in the same direction, the reversal paradox is avoided, and increased stability of estimated effects in the model is achieved. The directionality constraints are akin to shrinkage methods in the manner that, by model optimization over a subset of $\mathbb{R}^p$, we can expect a decrease in model variability.

Before discussing our proposal for a Frequentist fit of the model, we give a brief overview of the fitting procedures of both the GWQSR and the BGWQSR approaches. The fitting procedure of the WQSR model is not discussed since it is very similar to the algorithm used in the GWQSR model.

GWQSR model fitting requires the splitting of the observations into training and validation data sets. Estimation of the group weights begins by taking bootstrapped resamples from the training data set. Each bootstrap resample dataset is fitted by constrained ML using nonlinear optimization to enforce the constraints of the group weights. The final estimates for the sets of group weights are calculated as a test statistic weighted average, with test statistics and sets of weight estimates resulting from each bootstrap fit on the training data. After the group weights have been estimated, each chemical group is linearly combined into a single variable per group. A logistic regression is then fitted using the validation data set on the new set of linearly combined group variables. From this regression, the estimated group effects are derived with corresponding inferential measures (e.g., p-values, confidence intervals). However, the GWQSR model does not yield inferential estimates for individual chemicals [16].

Wheeler et al. proposed a Bayesian formulation to estimate the parameters of the WQSR problem (2) [17]. Group effects and adjusting covariates are assigned noninformative Normal priors by default, but the user may change the specification of the prior distributions. Group weights are given Dirichlet priors, where

$$(w_1, \ldots, w_k) \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_k) \Rightarrow \sum_{i=1}^{k} w_k = 1, \ w_k \in (0,1)$$

BGWQSR uses Markov Chain Monte Carlo (MCMC) to estimate the posterior distributions of the model parameters. Unlike GWQSR, BGWQSR can perform inference on the group weights of chemicals through posterior 95% credible intervals.

2.2 | The Frequentist Grouped Weighted Quantile Sum Regression

We approach fitting the model in Equation (2) within a fully Frequentist ML framework. Unlike GWQSR, our proposal does not require data splitting, which yields a model more inferentially efficient. In addition, FGWQSR does not require the use of computationally costly MCMC, which allows its applicability to large data sets.

There are two main challenges in a fully Frequentist approach to model (2). First, the model must perform computationally efficient estimation of the model parameters subject to constraints. Second, after parameter estimation, statistical inference on these constrained parameters must be handled with care as constrained ML does not follow standard asymptotics. We first discuss parameter estimation in FGWQSR.

2.2.1 | FGWQSR Parameter Estimation

The challenge in FGWQSR model optimization comes from the estimation of the group weight parameters that are constrained to satisfy the directionality assumption. To perform constrained ML estimation, we considered two approaches: Repeated constrained optimization in the original WQSR parameterization, and reparameterization to an unconstrained model. We discuss both these approaches and their benefits/drawbacks. The final procedure utilized for estimation in FGWQSR hybridizes the two techniques according to their strengths.

If we knew the signs of the group effects in the model, the constrained MLE could be obtained using standard "box" constrained optimization. For example, suppose $G = 2$, $\beta_i = \{\beta_{i,j} \mid j \in I\}$ where $I = \{1, \ldots, c_i\}$ and c_1, c_2 denote the number of chemicals in groups 1 and 2 respectively. If we knew that all group effects were positive, we would have the constraints: $(\beta_1, \beta_2) \in [0, \infty)^{c_1} \times [0, \infty)^{c_2}$. Since these are in the form of intervals or "boxes," we could solve this optimization problem using a "box" constrained optimization routine like a quasi-newton method (e.g., limited memory BFGS algorithm [19]. In R, this can be implemented, for example, in the *optim* function. In general, we can partition the parameter space into 2^G separate regions represented by box constraints. In the example above we would have the following four regions: $(\beta_1, \beta_2) \in [0, \infty)^{c_1} \times [0, \infty)^{c_2}$, $(\beta_1, \beta_2) \in [0, \infty)^{c_1} \times (-\infty, 0]^{c_2}$, $(\beta_1, \beta_2) \in (-\infty, 0]^{c_1} \times [0, \infty)^{c_2}$,

and $(\beta_1, \beta_2) \in (-\infty, 0]^{c_1} \times (-\infty, 0]^{c_2}$. Within each region, we can run "box" constrained optimization to obtain the within-region maximum. The overall maximum will be the largest of the within-region maxima. Coefficients corresponding to confounding variables are unconstrained to avoid affecting partition into regions. For each group, we can derive both the group effect and pollutant weights as follows:

$$\gamma_g = \sum_{k=1}^{c_g} \beta_{g,k} \tag{3}$$

$$w_{g,k} = \frac{\beta_{g,k}}{\sum_{k=1}^{c_g} \beta_{g,k}} = \frac{\beta_{g,k}}{\gamma_g} \tag{4}$$

An immediate consequence of this model-fitting approach is that computational complexity increases quickly as a function of the number of groups, which becomes especially apparent in cases when $G \geq 4$. Although more computationally intensive, the box-constrained model fitting approach allows more precise estimation of parameters that are projected, under the constraints, onto the boundary of the parameter space, an advantage that we will later leverage in FGWQSR's hybrid fitting procedure.

As a computationally less intensive approach, we also consider fitting FGWQSR by transforming our constrained problem into an unconstrained one through reparameterization. Specifically, we propose a reparameterized formulation for FGWQSR as follows:

$$y_i \sim \text{Bernoulli}(\pi_i)$$

$$\begin{aligned}
\text{logit}(\pi_i) = c_0 + \sum_{g=1}^{G} \gamma_g \left(\sum_{k=1}^{c_g-1} \frac{e^{\alpha_{g,k}}}{1 + \sum_{j=1}^{c_g-1} e^{\alpha_{g,j}}} q_{g,k,i} + \frac{1}{1 + \sum_{j=1}^{c_g-1} e^{\alpha_{g,j}}} q_{g,c_g,i} \right) \\
+ \sum_{r=1}^{R} \phi_r z_{r,i} \\
= c_0 + \sum_{g=1}^{G} \frac{\gamma_g}{1 + \sum_{j=1}^{c_g-1} e^{\alpha_{g,j}}} \left(\sum_{k=1}^{c_g-1} e^{\alpha_{g,k}} q_{g,k,i} + q_{g,c_g,i} \right) + \sum_{r=1}^{R} \phi_r z_{r,i}
\end{aligned} \tag{5}$$

The reparameterization in Equation (5) is essentially the parameterization used in multinomial logistic regression applied separately to each set of group weights. For a particular group g, the chemical weights can be recovered as

$$(w_{g,1}, \ldots, w_{g,c_{g-1}}, w_{g,c_g})$$

$$= \left(\frac{e^{\alpha_{g,1}}}{1 + \sum_{j=1}^{c_g-1} e^{\alpha_{g,j}}}, \ldots, \frac{e^{\alpha_{g,c_g-1}}}{1 + \sum_{j=1}^{c_g-1} e^{\alpha_{g,j}}}, \frac{1}{1 + \sum_{j=1}^{c_g-1} e^{\alpha_{g,j}}} \right)$$

The parameterization in Equation (5) has unconstrained parameters, eliminating the need to optimize over multiple different subregions. Although more computationally efficient, this parameterization yields less precise estimates when chemical weights are close to the boundary (either 0 or 1). This is because 0 and 1 correspond to $-\infty$ and $+\infty$ respectively in the reparameterized scale. Numerically, this means that weights with true underlying values of close to 0 or 1 would be estimated as a large (in magnitude) negative or positive value in the reparameterized scale. These estimates would be necessarily far from the

boundaries ($-\infty$ or $+\infty$), resulting in a loss of precision when converted back to the original [0,1] scale. Although these errors may be small, they can affect the asymptotic evaluation of our estimators.

Given that the box-constrained optimization is more precise yet slower and the reparameterized optimization is less precise but faster, we propose a hybrid approach that tailors to the strengths of each. First, the constrained regression problem is optimized using the reparameterized fitting procedure to get an initial but less precise estimate. In particular, we can determine which sub-region of the constrained parameter space our ML solution falls in by looking at the signs of the initial estimated group effects. We then use box-constrained optimization over the derived sub-region hot-started with initial values set to the initial estimates. Chemical weights and group effects are finally computed from the final box-constrained fit.

2.2.2 | Inference for FGWQSR

FGWQSR allows inference on both group effects and individual chemical weight estimates. Specifically, in addition to the estimates, we can construct tests of the hypothesis that a group (omnibus test) or individual effect is zero. For testing a group or an individual effect, we compute a likelihood ratio test statistic by fitting models with and without (i.e., under H_0) the group or individual chemical of interest from the model. A challenge is that, because of the constrained ML estimation, these likelihood ratio statistics do not have standard asymptotic chi-square distributions [20]. We now briefly describe the non-standard asymptotics of these likelihood ratio statistics. The interested reader is referred to the Supporting Informations for details.

It can be shown that for a constrained model, the likelihood ratio statistic has an asymptotic distribution equal to that of the likelihood ratio test statistic for testing that the mean of a single multivariate normal realization equals zero under certain constraints [20]. More precisely, let $Z \sim N_p\left(\theta_0, I^{-1}(\theta_0)\right)$, where θ_0 denotes the true parameter values for the FGWQSR model $\{\theta = \beta_{g,k} | g \in \{1, \ldots, G\}, k \in \{1, \ldots, c_g\}\}$, and $I^{-1}(\theta_0)$ denote the inverse of the Fisher Information evaluated at θ_0. We can translate our likelihood ratio testing problem in FGWQSR into a likelihood ratio testing scenario where we perform ML estimation for the mean vector θ of Z over *approximating cone* regions C_{Ω_0} and C_{Ω_1} that *approximate* the constrained regions Ω_0 and Ω_1 corresponding to hypotheses H_0 and H_1, respectively. For example, the likelihood ratio statistic for testing that a particular group g has no effect on the outcome (i.e., $\beta_{g,k} = 0, k \in \{1, \ldots, c_g\}$) has asymptotic distribution equal to that of the likelihood ratio statistic for testing $\theta \in C_{\Omega_0}$ vs. $\theta \in C_{\Omega_1}$, where C_{Ω_0} and C_{Ω_1} are approximating cones to Ω_0 and Ω_1, the original null and alternative hypotheses H_0 and H_1. For a discussion on approximating cones and further derivation of the likelihood testing procedure for FGWQSR, we refer the reader to the Supporting Informations and Self and Liang [20].

Derivation of the asymptotic distribution requires knowledge of the true parameter values θ_0, which of course we do not know. However, Self and Liang propose using the ML estimates of the parameters of interest $\hat{\theta}$ as an approximation of the location of θ_0 [20]. We adopted their proposal for FGWQSR, using $\hat{\theta} =$ $\{\hat{\beta}_{g,k} | g \in \{1, \ldots, G\}, k \in \{1, \ldots, c_g\}\}$ as a consistent estimator of θ_0, where each $\hat{\beta}_{g,k}$ is derived from the fitted FGWQSR model. In general, the asymptotic distribution of the LRT above is analytically complex but is very straightforward to simulate. In our implementation, we approximate the null distribution of the LRT using 10 000 Monte Carlo simulations.

2.2.3 | Scalability of FGWQSR

Accounting for both the initial and subsequent nested fits for likelihood ratio testing, FGWQSR estimates $G + p + 1$ constrained regression models, where $p = \sum_{g=1}^{G} c_g$, and computes (by default) $10{,}000 \times (G + p)$ one sample multivariate normal constrained optimizations. For a fixed chemical mixture setting, the primary impact of increasing data sample size is longer optimization runtimes for fitting the $G + p + 1$ constrained regression models. Interestingly, FGWQSR's inferential procedure generally benefits from larger sample sizes, as having more mixture components estimated with greater precision potentially reduces the number of regions of optimization in the Monte Carlo simulations. Model fitting and one sample multivariate normal optimizations are all independent procedures and are completely parallelizable to multiple computing cores. The FGWQSR function includes an option *cores* that allows the user to specify how many computing cores to parallelize on in their machine.

2.3 | Simulation Study Design

2.3.1 | Simulations Under Correct Specification

We evaluated FGWQSR through a simulation study, comparing its performance to that of competing approaches: Quantile logistic regression (QLR) and BGWQSR. GWQSR is not included as a competitor in the simulation set as Wheeler et al. have already shown that BGWQSR outperforms GWQSR in terms of both statistical power and mean squared error [17]. We followed a simulation design similar to scenario set D in Wheeler et al. [17] QLR models were fitted with the *glm* function in R using quantized $PM_{2.5}$ components and BGWQSR models were fitted using an implementation of BGWQSR included in the *fgwqsr* package that utilizes parallel MCMC from the *runjags* package. QLR was included as the baseline comparison model since it can be viewed as the unconstrained analog to FGWQSR. In each simulation scenario, we considered three groups denoted as groups A, B, and C, with sizes 5, 4, and 5 chemicals each, respectively. Within each of the three groups, the individual chemicals were assigned the following weight sets: $\mathbf{w}_A = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$; $\mathbf{w}_B = (\frac{1}{2}, \frac{1}{2}, 0, 0)$; and $\mathbf{w}_C = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$. We included null chemicals in all scenarios to assess how well the nominal type 1 error rate was preserved.

We considered 5 configurations for the strengths of association (SOA) $SOA = (\exp(\gamma_1), \exp(\gamma_2), \exp(\gamma_3))$, where $\exp(\gamma_g)$ corresponds to the odds ratio for group g: $SOA_1 = (1,1,1)$; $SOA_2 = (0.9, 1.1, 1.1)$; $SOA_3 = (0.85, 1.15, 1.15)$; $SOA_4 = (0.8, 1.2, 1.2)$; and $SOA_5 = (0.7, 1.3, 1.3)$. Finally, we specified three different configurations of within– and between–group chemical correlation structures (CCS) $CCS = (\rho_w, \rho_b)$, where all pairwise correlations between chemicals within a group g are set to ρ_w, and all

pairwise correlations between chemicals in different groups are set to ρ_b: $CCS_1 = (0.5, 0.1)$; $CCS_2 = (0.7, 0.3)$; and $CCS_3 = (0.9, 0.5)$. These levels of within and between correlations are the same as in Wheeler et al. [17] aiming to capture realistic levels of correlation for chemicals within a mixture (weak, moderate, strong). The combination of SOA and CCS configurations yields $5 \times 3 = 15$ total simulation scenarios.

For each simulated dataset the sample size was set to $n = 500$ and the number of simulation replicates to $n.rep = 1,000$ datasets. Exposure data was generated by sampling from a multivariate normal distribution with mean set to $\mathbf{0}_p$ and covariance matrix defined by each simulation scenario CCS. The binary outcome was generated as Bernoulli random variables with probability depending on each scenario's SOA based on the exposure levels quantized into quintiles. The intercept was chosen to maintain a 1:1 case-control ratio for each simulation scenario. When fitting FGWQSR, BGWQSR, and QLR models, quintiles were used to quantize the multivariate normal exposure data. For inference using FGWQSR, we set the number of multivariate normal simulations to 10 000 to be performed within each of the $n.rep$ simulation replicates. For BGWQSR, we set the number of posterior samples to 10 000 per each of the two chains with 5 000 burn-in samples, 1 000 adaptive iterations, and a thinning interval of size 1.

Several metrics were computed from the fitted models in the simulation including power for detecting group and individual chemical effects, as well as bias and mean squared error (MSE) for estimating the model parameters. To perform power analysis using BGWQSR, we checked whether the highest posterior density 95% credible interval for the group effects and individual chemicals included 0. Since a value of zero is on the boundary of the support of the posterior distribution for each weight, we defined a cutpoint $b = 1 \times 10^{-8}$ to serve as a "tolerance". Any estimated weight 95% credible interval with left endpoint less than or equal to b was considered "statistically insignificant". For QLR, we computed likelihood ratio tests for group effects and used Wald tests for individual chemical tests at the $\alpha = 0.05$ level. To compare parameter estimation bias and MSE between the models, we parameterized FGWQSR and BGWQSR in the constrained regression parametrization

$$(\gamma_1 \cdot w_{1,1}, \ldots, \gamma_1 \cdot w_{1,c_1}, \ldots, \gamma_G \cdot w_{G,1}, \ldots, \gamma_G \cdot w_{G,c_g}) =$$
$$(\beta_{1,1}, \ldots, \beta_{1,c_1}, \ldots, \beta_{G,1}, \ldots, \beta_{G,c_g})$$

Bias and MSE calculations were computed using posterior mean estimates for BGWQSR, whereas point estimates were derived from ML estimation in FGWQSR and QLR.

2.3.2 | Simulations Under Misspecification

In practice, it may sometimes be the case that the directionality assumption (recall that this asserts all effects within a mixture group are in the same direction) of particular exposure variable groupings does not hold. To evaluate the performance of FGWQSR when the directionality assumption does not hold, we performed a similar set of simulations as discussed previously, with the addition of three different levels of model misspecification. Misspecification was introduced by the inclusion of negative

weights within weight group B, while groups A and C remain correctly specified. A negative weight $-w_{B,k}$ in group B corresponds to having an individual chemical effect of $-\gamma_B \cdot w_{B,k}$ included in group B, which directly models having chemical(s) in the model that violates the directionality assumption. In this simulation scenario, the total number of chemicals was set to 16, with group sizes of 5, 6, and 5 chemicals. We analyze three levels of model misspecification:

1. (Minor Violation):

$$\mathbf{w}_A = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0\right),$$
$$\mathbf{w}_B = \left(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, -\tfrac{1}{6}\right),$$
$$\mathbf{w}_C = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0\right)$$

2. (Moderate Violation):

$$\mathbf{w}_A = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0\right),$$
$$\mathbf{w}_B = \left(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, -\tfrac{1}{6}, -\tfrac{1}{6}\right),$$
$$\mathbf{w}_C = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0\right)$$

3. (Extreme Violation):

$$\mathbf{w}_A = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0\right),$$
$$\mathbf{w}_B = \left(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, -\tfrac{1}{6}, -\tfrac{1}{6}, -\tfrac{1}{6}\right),$$
$$\mathbf{w}_C = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0\right)$$

We iterate over the aforementioned 5 (SOA) $\times$ 3 (CCS) = 15 simulation sets and record power and precision metrics for the three models with the same options for FGWQSR and BGWQSR from the previous simulation set.

2.3.3 | Runtime Comparison Between FGWQSR and BGWQSR

Since the ability to analyze large datasets was a central motivation for developing FGWQSR, we performed simulations to compare the runtimes of FGWQSR and BGWQSR on datasets of varying sample sizes. Additionally, two different mixture group configurations were studied in our simulations to further understand the computational burden of increasing the number of mixture groups in FGWQSR and BGWQSR models. First, we use a mixture group configuration with 14 total chemicals distributed into group sizes of 5,4 and 5. We set SOA = (0.8, 1, 1.2), CCS = (0.7, 0.3), and the weight distribution as follows: $\mathbf{w}_1 = (\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0)$; $\mathbf{w}_2 = (\tfrac{1}{2}, \tfrac{1}{2}, 0, 0)$; and $\mathbf{w}_3 = (\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0)$. The second mixture group configuration consists of 28 chemicals distributed into groups of sizes 5, 4, 5, 5, 4, and 5. In these larger mixture simulations, SOA is set to SOA = (0.8, 1, 1.2, 0.8, 1, 1.2), CCS remains at CCS = (0.7, 0.3), and the chemical weight distributions are as follows: $\mathbf{w}_1 = (\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0)$; $\mathbf{w}_2 = (\tfrac{1}{2}, \tfrac{1}{2}, 0, 0)$; and $\mathbf{w}_3 = (\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0)$, $\mathbf{w}_4 = (\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0)$; $\mathbf{w}_5 = (\tfrac{1}{2}, \tfrac{1}{2}, 0, 0)$; and $\mathbf{w}_6 = (\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 0, 0)$ Simulations were performed on simulated

data with sample sizes ranging from 100 to 100 000, with 10 replications performed per sample size. The same arguments for FGWQSR and BGWQSR from the first simulation set were used here, with the addition of setting the number of cores for FGWQSR to 10 and fitting the two independent MCMC chains for BGWQSR on separate cores with 5000 posterior samples per chain (10 000 total, the default for BGWQSR).

Correctly specified and misspecified simulations were run on the University of Southern California's Center for Advanced Research Computing (CARC) resource. However, the runtime simulations for FGWQSR and BGWQSR were performed locally on a 2021 Macbook Pro with an M1 Max processor and 10 computing cores to reflect more common computing resources available to analysts.

2.4 | Application

We applied FGWQSR to a large dataset comprised of 317 767 mother-child pairs with data derived from electronic medical records from Kaiser Permanente of Southern California (KPSC) between the years of 2001–2014 and $PM_{2.5}$ exposure fields generated by researchers at the University of California, Davis. The goal was to study the associations between in-utero averaged exposure to components of particulate matter with aerodynamic diameter $< 2.5 \mu m$ ($PM_{2.5}$) and Autism Spectrum Disorder (ASD) diagnosis in children. Exposure fields were estimated using Source Oriented Chemical Transport models that describe emissions, transport, deposition, and chemical reaction to predict airborne $PM_{2.5}$ concentrations at maternal residential addresses geocoded using ArcGIS [21–23]. In particular, we focused on the associations between ASD and the following $PM_{2.5}$ averaged component concentrations: Iron (Fe), copper (Cu), manganese (Mn), nitrate (N), sulfate (S), secondary organic aerosols (PM-SOA), elemental carbons (EC), organic carbon (OC), other metals (Metl), and residual $PM_{2.5}$ mass (RM).

Several maternal- and child-level covariates were selected a priori to include in our model based on expert knowledge and existing literature to control for potential confounding, including child birth year, seasonality (wet = November–March, dry = April–October), maternal age at delivery, child parity, maternal race, maternal educational level, maternal comorbidity (0 = no comorbidity, ≥ 1 = heart, lung, kidney, and/or liver disease; cancer), median family income in census tract of residence, child sex, maternal obesity status, and maternal diabetes diagnosis, and total $PM_{2.5}$. In addition, we also adjusted for trimester averages of ozone and nitrogen dioxide levels, as these exposures may confound $PM_{2.5}$ component associations. A penalized spline with 4 degrees of freedom (guided by AIC) was used for the child birth year variable due to historical changes in incidence rates of ASD [24]. Child birth weight and gestational age were not adjusted for in the analysis, as the two variables are suspected to be on the causal pathway between air pollution and ASD [25, 26]. Child ASD diagnoses in the dataset are determined through ICD-9 codes s 299.0, 299.1, 299.8, 299.9 from at least 2 separate visits before five years of age [27]. Children were followed from birth through electronic medical records until either clinical diagnosis of ASD, loss to follow-up, or five years of age, as consistent with prior studies from this cohort [22, 23].

Due to the rarity of ASD, the dataset contains only 4522 ASD cases.

$PM_{2.5}$ components were organized into mixture groups based on their correlation structure. Specifically, we assigned components to a particular mixture group based on an apriori hierarchical clustering and non-negative matrix factorization of the correlation matrix C (for hierarchical clustering, the matrix of $1 - |C|$) of the components. Both methods agreed on the potential presence of 3 underlying latent mixture group clusters, where component group members agreed between the two methods. The $PM_{2.5}$ mixture group assignments are as follows: $G_1 = \{Cu, Fe, Mn\}$; $G_2 = \{N, S, PM\text{-}SOA\}$; and $G_3 = \{EC, OC, RM, Metl\}$. The cluster configurations in G_1, G_2, and G_3 can naturally be referred to as trace metals, secondary PM, and other primary $PM_{2.5}$.

As a baseline comparison with a traditional method, we analyzed the KPSC data using QLR with $PM_{2.5}$ components quantized into quintiles and the same set of adjusting covariates included in the analysis with FGWQSR. BGWQSR was not included as a comparison model due to the overly long fitting times required for this size dataset. In contrast, FGWQSR was fitted on all 317 767 mother-child pairs in 49 min across 24 cores on a Windows Server 2022×86-64 architecture.

3 | Results

3.1 | Simulation Study

3.1.1 | Correctly Specified Simulation

Unless explicitly mentioned, the results discussed here apply to all simulation scenarios. However, we only show the results of the simulation scenarios with the highest levels of within and between correlations, $CCS_3 = (0.9, 0.7)$, as these are the most challenging for traditional methods (Figure A1). Results for $CCS_1 = (0.5, 0.1)$ and $CCS_2 = (0.7, 0.3)$ are included in Figures S1 and S2. FGWQSR demonstrated the greatest power to detect nonzero group effects while achieving the nominal significance level for truly null group effects. Across all simulation sets, FGWQSR had the largest power to detect nonzero group effects, followed by BGWQSR and QLR, respectively. Additionally, as shown previously by Wheeler et al. [17] BGWQSR had greater power than QLR to detect group effects due to exploiting the directionality assumption when it holds.

In addition to the power to detect group effects, we also compared each model's ability to detect individual exposure effects. Generally, FGWQSR had the highest power for truly non-zero individual exposures. BGWQSR's 95% credible interval for individual exposure inference was underpowered, where many individual weights had particularly wide credible intervals. For a majority of nonzero individual exposure effects, QLR had higher power than BGWQSR across the 5 SOAs, which may be due to overly conservative Bayesian credible intervals.

Among the three models, BGWQSR generally had the lowest estimated MSE, followed by FGWQSR and QLR. We expected QLR to be the least precise, as BGWQSR and FGWQSR's constrained estimation reduces variability when the

directionality assumption holds. The reduction in model variability for the two constrained methods comes at the expense of increased bias. Generally QLR produces individual exposure effect estimates with the least amount of absolute bias. However, for simulations $SOA_1 = (1,1,1)$, BGWQSR had the smallest amount of absolute model bias because of the model's implicit bias toward the null prior. For a majority of the scenarios, FGWQSR and BGWQSR had comparable levels and directions of bias, with FGWQSR often showing slightly smaller absolute bias. Individual chemicals from groups with positive effects (groups B and C) had very small negative bias for nonzero effects and very small positive bias for null effects, and vice versa for members of group A with a negative group effect. This pattern of bias arises from the model constraints, where individual chemicals with null effects belonging to groups with non-null group effects have their constrained estimates in the direction of their group.

3.1.2 | Misspecified Simulations

Figure A2 presents the results of the moderate model misspecification and correlation structure $CCS_3 = (0.9, 0.5)$. All other results are included in the Supporting Information as Figures S3–S10. In these simulations, the misspecified group B is of the greatest interest. Overall, FGWQSR remains the highest powered for group and individual chemical effects for correctly specified groups A and C, while being well calibrated for type 1 error in those groups (for chemicals with null weights).

In the case of minor model misspecification (Figures S3–S5), FGWQSR and BGWQSR maintain their power gains over QLR to detect the misspecified group B, as well as for the correctly specified groups A and C. Individual chemical effects were still detected with higher power in FGWQSR models compared to QLR, except for the misspecified chemical 6 in group B, which QLR had higher power to detect than FGWQSR and BGWQSR. Chemical 6 in group B also had a much larger bias than other pollutants within group B. This is not surprising as QLR fits unconstrained models, and therefore, it is not misspecified. Nevertheless, BGWQSR and FGWQSR maintained the lowest levels of individual chemical effect MSEs.

The moderate level of misspecification introduces greater model violation to FGWQSR and BGWQSR. Across the range of effects explored, FGWQSR and QLR generally had similar power levels to detect a non-null effect for group B, with BGWQSR showing the lowest power (Figures S2, 2, and S7). At this level of model misspecification, FGWQSR, and QLR had comparable power to detect individual chemical effects, with the exception of QLR having had much higher power for detecting the misspecified chemicals 5 and 6 within group B. Bias for these two chemicals were accordingly higher in FGWQSR and BGWQSR models compared to QLR. The less flexible constrained models had lower MSE for individual exposure effects in comparison to QLR.

The extreme misspecification scenarios do result in considerably biased inference for FGWQSR and BGWQSR (Figures S8–S10). When half of the individual chemical effects in group B are positive and half negative, the constrained methods generally estimate a null effect for the group. In all the simulations under this scenario, QLR had the highest power to detect group B, while

power for FGWQSR and BGWQSR tended to remain at their null nominal levels. In addition, QLR also had the highest power to detect associations with individual chemicals compared to the constrained methods. FGWQSR and BGWQSR both exhibited positive biases for individual negative effects when the group B effect was estimated to be positive and negative bias for individual effects when the group B effect was estimated to be negative. QLR had a low bias for individual chemical effects. Again, this is not surprising as the QLR model is not misspecified under this scenario.

3.1.3 | Runtime Comparison Between FGWQSR and BGWQSR

For small sample sizes, BGWQSR fits quicker than FGWQSR (Figure A3). However, for moderate and large sample sizes FGWQSR proves to be faster than BGWQSR. For example, with $n = 100,000$, BGWQSR takes 1.14 and 1.66 h for three and size group mixture scenarios respectively, while FGWQSR fits the same models in 1.35 and 7.55 min respectively. In contrast to the computational expense of MCMC methods, as sample size increases, FGWQSR's inference has a fixed cost, with the $G + p + 1$ model optimizations required incurring only a slight increase in time as n grows. FGWQSR is affected by the size and number of mixture groups, as its Monte-Carlo-based inference needs to be run separately (or in parallel) for each mixture group and individual chemical. That being said, FGWQSR models can comfortably accommodate reasonable mixture group configurations for analyses commonly performed in environmental epidemiology.

3.1.4 | Data Application

In the KPSC cohort, 4522 children were diagnosed with ASD by age 5, of which 3673 (81.2%) were boys and 849 (18.8%) were girls. Children were more likely to be diagnosed with ASD by age 5 from mothers who were older, nulliparous, diabetic, obese, and who had health comorbidities (Table S1).

Among the $PM_{2.5}$ components, nitrate had the largest median concentration, with copper and manganese having much smaller concentrations (Figure A4). The copper concentration had the most skewed distribution, while the others remained fairly symmetric. The most correlated pairs of $PM_{2.5}$ components were iron and copper ($\rho = 0.95$) and iron and manganese ($\rho = 0.89$) (Figure A5).

The trace metals and other primary $PM_{2.5}$ mixture groups were both associated with ASD diagnosis by age 5, with estimated odds ratios (OR) of 1.13 and 1.11 per one unit increase in exposure respectively ($p < 0.001$, $p = 0.01$) (Table A1). A one-unit increase in exposure for a chemical mixture group corresponds to an increase of one quintile for all chemicals in the group. Secondary $PM_{2.5}$ was not significantly associated with ASD (OR=1.02, $p = 0.7$).

Within the trace metals group, the largest individual weight was for copper (0.65), while iron and manganese had weight estimates of 0.13 and 0.22 respectively. Copper was the only component within the trace mixture group that was statistically significant at

a conventional 5% significance level ($p = 0.02$). In the secondary $PM_{2.5}$ group, the weight for sulfate was estimated to be 1, and both nitrate and secondary organic aerosols had estimated weights of 0. Consistent with the group inference, none of the individual pollutants in the secondary $PM_{2.5}$ group were statistically significant. Finally, for the other primary $PM_{2.5}$ group, the individual weights were estimated as 0.46, 0.37, 0.17, and 0 for residual mass, other metals, elemental carbon, and organic carbon respectively. Among the three components with nonzero weight, residual mass was the only statistically significant pollutant ($p = 0.039$).

Estimates of individual pollutant ORs were comparable between FGWQSR and QLR for deleterious effects (Table A2). QLR estimated slightly protective but not significant effects for nitrate, secondary organic aerosols, and organic carbon ($p = 0.40$, $p = 0.78$, and $p = 0.29$). Both FGWQSR and QLR identified copper as an associated pollutant ($p = 0.019$ and $p = 0.036$ respectively). Residual mass was identified as significantly associated in the FGWQSR model ($p = 0.039$), whereas it was not identified as statistically associated within the QLR model ($p = 0.097$).

4 | Discussion

We have introduced FGWQSR, a Frequentist weighted quantile sum regression for chemical mixtures. Model parameters in FGWQSR are estimated by ML using a full dataset, in contrast to its non-Bayesian predecessor GWQSR, which requires data splitting. Similar to BGWQSR, FGWQSR yields estimates for both individual chemical and chemical mixture effects. Our simulations show that FGWQSR almost always had greater power to detect both groups of chemicals and individual chemicals in comparison to QLR and BGWQSR. In particular, BGWQSR yielded inferences that were often conservatively biased toward its null priors. Critically, FGWQSR can be fitted efficiently. As a result, it can be used in applications with large datasets as we have demonstrated in our analysis of the KPSC cohort with 317 767 observations.

FGWQSR's model parameterization, similar to that of BGWQSR, assumes that the effects on the outcome of chemicals within a mixture group are all in the same direction. By utilizing the extra information the directionality assumption provides, FGWQSR can detect significant group and individual chemical effects with higher power over unconstrained methods. This increased power is especially valuable in contexts where the effect of chemicals may be small. Since chemical exposures, either individual or in mixtures, are almost never expected to have protective effects, the directionality assumption is reasonably robust and unlikely to be violated in real studies. However, even in cases of moderate levels of misspecification, FGWQSR remains robust and maintains its ability to detect significant group and individual chemical effects. When the directionality assumption is violated, type-I errors for detecting chemicals with misspecified effects are often conservative. FGWQSR suffers the greatest in scenarios of extreme model misspecification, which occurs when the members of a mixture group have their effects in both positive and negative directions, with chemical effects in opposing directions having similar magnitudes. In this situation, although significant group and individual chemical associations exist, FGWQSR tends to estimate these effects toward the null, with a slight bias in a certain direction incurred by the particular correlation structure of the data.

FGWQSR, like other WQS models, is most effective when analyzing groups of highly correlated covariates, enabling researchers to jointly model disease risk based on grouped exposures. However, FGWQSR differentiates itself from WQS competitors as it is the first fully Frequentist WQS model that avoids data splitting to estimate both group effects and weights, and can be applied to large datasets. Previously, the high computational complexity of many WQS methods, including BGWQSR, has limited their applicability in large-scale datasets. FGWQSR was specifically designed to efficiently handle the data volume of the KPSC dataset and scales well to even larger datasets.

FGWQSR is most comparable to BGWQSR in terms of model output. However, while BGWQSR relies on posterior credible intervals for inference, FGWQSR uses a Monte Carlo-based approach centered on likelihood ratio test statistics. This distinction is particularly important when interpreting inference on individual constituents within a group. When credible intervals are used for inference on constrained weight parameters, they never numerically include the true null value of a zero weight. This is expected, as the null lies on the boundary of the parameter space. To address this, our simulations adopt a lenient threshold of 10^{-8} for BGWQSR credible intervals as the decision boundary for whether they include the null. In contrast, FGWQSR's inference procedure inherently respects the model's parameter constraints, allowing p-values to be directly used for statistical significance testing. As a result, FGWQSR provides more interpretable inferential estimates that align with the imposed model constraints. For the group effect, which is an unconstrained parameter, both p-value and credible interval estimation are equally valid for detecting signals across groups of chemicals. However, if interval estimation is the primary concern, BGWQSR is better suited, as FGWQSR in its current form outputs effect estimates and p-values for statistical inference.

When configuring an FGWQSR model, a researcher should select mixture groups a priori based on substantive knowledge. Mixture groups should be constructed to maximize within-group correlation levels while also obeying the directionality assumption. In the absence of prior knowledge, we suggest possible ways of constructing mixture groups solely based on correlation. For example, one may perform hierarchical clustering on the distance matrix constructed by evaluating one of $1 - C$, $1 - |C|$, or $1 - C^2$, where C is the correlation matrix corresponding to the data of our chemical mixture components of interest and C^2 denotes element-wise squaring of C. If there is suggestive clustering, one may accordingly lump chemicals into mixture groups for FGWQSR analysis. Ideally, the plausibility of the directionality assumption should still be evaluated by consulting with existing literature. Techniques other than clustering, for example, factor analysis or a non-negative matrix factorization on C, depending on whether there are negative correlations, can be similarly used to define mixture groups.

When fitting model models with FGWQSR, it may be the case that certain chemicals from particular mixture groups are estimated to the boundary value of 0. One may be interested in whether such boundary-estimated chemicals are misspecified within their

mixture groups and potentially violate the directionality assumption. A possible sensitivity analysis that can be performed is to remove the suspected chemicals from their respective mixture groups and include them as an (appropriately scaled) adjusting covariate in the model formula. By doing this, such chemical is instead fitted in a traditional unconstrained manner, and one may examine the direction of association for these suspected chemicals. If a misspecified chemical is identified, one can invert the functional association of the exposure variable by applying a monotonically decreasing function. For non-negative exposures, one such transformation could be $f(x) = x^{-1}$.

While FGWQSR offers significant advantages in computational efficiency and statistical inference, several limitations should be addressed. Like other WQS models, FGWQSR benefits from increased power over traditional linear models by imposing the directionality assumption model constraints. However, in cases where mixture groups are constructed with elements that have opposing directions of effect, any WQS method may struggle to identify important groups or individual actors. Additionally, as a Frequentist approach, FGWQSR relies on asymptotic evaluations that require adequate sample sizes for reliable inference. Through simulation, we have observed that FGWQSR can accommodate smaller datasets (e.g., 150 observations with 16 chemical constituents, not included in the manuscript) and still outperform traditional linear regression in terms of power due to its model constraints. However, for small datasets, we recommend comparing FGWQSR results with those from BGWQSR, as the Bayesian framework in BGWQSR uses exact posterior distributions and can leverage informative priors to achieve more robust results in such settings. In addition, FGWQSR performs similarly to standard logistic regression when handling varying case-control ratios, with estimation remaining robust except in cases of extreme imbalance. On a different note, FGWQSR requires complete observations without missing data. Therefore, the performance of FGWQSR when applied to an imputed dataset would rely on the quality of the data imputation.

FGWQSR can accommodate interaction variables as unconstrained parameters in the model. However, in its present form, FGWQSR does not handle interactions between mixture constituents. Extensions of FGWQSR to include constrained interactions would be nontrivial. To more effectively examine interaction in pollutant mixture models, the Bayesian Kernel Machine Regression (BKMR) flexibly measures interaction in pollutant mixture models using Gaussian process regression [11, 28]. However, because BKMR is computationally expensive, the method can only be applied to moderately sized datasets. In a recent paper, Lee et al. introduce a WQS model that investigates interaction by estimating stratum-specific weights of exposure in mixtures, which may be of interest to researchers who wish to perform WQS interaction analysis in mixture model settings [29].

We applied FGWQSR to a large longitudinal birth cohort comprised of 317 767 mother-child pairs to study the associations between several $PM_{2.5}$ components and child ASD diagnosis by age 5. From the results of the FGWQSR model fitting, it was found that the trace metal and other primary $PM_{2.5}$ groups were associated with ASD, where the two groups were comprised of the components {Cu, Fe, Mn} and {EC, OC, RM, OM}. Within the two mixture groups, Copper and Residual $PM_{2.5}$ mass were both

significant and can be viewed as the drivers of association within their mixture groups. Particles belonging to the Residual $PM_{2.5}$ mass component often derive from crustal material that arises from windblown dust and unpaved road dust resuspension. Cu $PM_{2.5}$ in ambient air primarily originates from vehicular brake wear and is most concentrated in high-traffic areas [23, 30, 31].

Our FGWQSR mixture analysis aligns with existing literature on the developmental neurotoxicity of prenatal $PM_{2.5}$ exposure on autism risk [21, 23, 32]. Additionally, several studies have documented the adverse effects of fine particulate metal exposure on children's developmental health, including cognitive impairments, respiratory issues, and asthma [33–35] Notably, the $PM_{2.5}$ constituents identified as statistically significant in our analysis primarily stem from non-tailpipe emissions. While stringent regulatory efforts have successfully reduced tailpipe-related $PM_{2.5}$ in Southern California over the past two decades, our findings highlight the need for increased attention to non-tailpipe sources, particularly brake wear, which remains unregulated. The contribution of non-tailpipe emissions is expected to rise with the widespread adoption of electric vehicles [36, 37]. Therefore, updating emissions standards to address brake-related particulate pollution is essential to further mitigate the health risks associated with vehicular emissions.

It should be noted that the spatial resolution of the exposure fields used in this analysis was 4 km, which precludes the analysis of near-roadway effects. Additionally, each exposure variable has potential measurement error and uncertainty that have not been accounted for in the present analysis. Future studies will use updated exposure fields that combine higher spatial resolution and more accurate concentration estimates to consider implications for FGWQSR analysis.

In our previous work with this cohort, we reported that Cu, Fe, and Mn were individually associated with an increased risk of ASD [21]. These results were derived from single-constituent models, that is, separate models for each constituent. We did not include all three constituents simultaneously in our previous work due to very high (ranging from 0.88 to 0.96) correlations among them. Consequently, it is possible that Fe and Mn showed significant associations in our previous study due to their high correlation with Cu. In this application, the multi-constituent FGWQSR model handled all constituents simultaneously and showed that Cu was the significant factor independent of other constituents including Fe and Mn, whereas Fe and Mn were not independently associated with child ASD.

Our proposed FGWQSR framework allows researchers to employ a mixture model framework in situations where the use of other mixture modeling tools were previously unfeasible due to computational limitations. We have demonstrated that FGWQSR outperforms its unconstrained analog QLR in terms of power to detect significant group and individual chemical effects in correctly and minorly-moderately misspecified scenarios. FGWQSR is also often higher powered than its Bayesian counterpart BGWQSR while maintaining nominal type 1 error levels. Although we have applied FGWQSR to pollutant exposure data, the method can be used in any situation where covariates of interest are highly correlated and an inherent grouping structure exists among the covariates. In our R package *fgwqsr*, we have

implemented FGWQSR for binary, continuous, and count outcomes; extension to Cox regression for survival outcomes should be straightforward. Future implementations of FGWQSR may also consider handling measurement errors on collected mixture constituents. We believe FGWQSR will enable researchers in the field to further understand underlying relationships between modifiable environmental exposures and disease outcomes.

5 | Software

Our R package fgwqsr implementing FGWQSR can be found at https://github.com/Daniel-Rud/fgwqsr with instructions for installation and examples of use.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. D. O. Carpenter, K. Arcaro, and D. C. Spink, "Understanding the Human Health Effects of Chemical Mixtures," *Environmental Health Perspectives* 110 (2002): 25–42.

2. I. Silins and J. Högberg, "Combined Toxic Exposures and Human Health: Biomarkers of Exposure and Effect," *International Journal of Environmental Research and Public Health* 8, no. 3 (2011): 629–647.

3. J. S. Colt, R. K. Severson, J. Lubin, et al., "Organochlorines in Carpet Dust and Non-Hodgkin Lymphoma," *Epidemiology* 16, no. 4 (2005): 516–525.

4. Y. Li, L. Xu, Z. Shan, W. Teng, and C. Han, "Association Between Air Pollution and Type 2 Diabetes: An Updated Review of the Literature," *Therapeutic Advances in Endocrinology and Metabolism* 10 (2019): 1–15.

5. J. Czarnota, C. Gennings, J. S. Colt, et al., "Analysis of Environmental Chemical Mixtures and Non-Hodgkin Lymphoma Risk in the NCI-SEER NHL Study," *Environmental Health Perspectives* 123, no. 10 (2015): 965–970, https://doi.org/10.1289/ehp.1408630.

6. d J. Bont, Y. Díaz, d M. Castro, et al., "Ambient Air Pollution and the Development of Overweight and Obesity in Children: A Large Longitudinal Study," *International Journal of Obesity* 45, no. 5 (2021): 1124–1132, https://doi.org/10.1038/s41366-021-00783-9.

7. E. A. Gibson, Y. Nunez, A. Abuawad, et al., "An Overview of Methods to Address Distinct Research Questions on Environmental Mixtures: An Application to Persistent Organic Pollutants and Leukocyte Telomere Length," *Environmental Health* 18, no. 76 (2019).

8. S. Mohankumar and P. Senthilkumar, "Particulate Matter Formation and Its Control Methodologies for Diesel Engine: A Comprehensive Review," *Renewable and Sustainable Energy Reviews* 80 (2017): 1227–1238, https://doi.org/10.1016/j.rser.2017.05.133.

9. R. K. Jain, "Ridge Regression and Its Application to Medical Data," *Computers and Biomedical Research* 18, no. 4 (1984): 363–368.

10. R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society* 58, no. 1 (1996): 267–288.

11. B. R. Joubert, M. A. Kioumourtzoglou, T. Chamberlain, et al., "Powering Research Through Innovative Methods for Mixtures in Epidemiology (PRIME) Program: Novel and Expanded Statistical Methods," *International Journal of Environmental Research and Public Health* 19, no. 3 (2022): 1378.

12. H. Xu, C. Caramanis, and S. Mannor, "Sparse Algorithms Are Not Stable: A no-Free-Lunch Theorem," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, no. 1 (2012): 187–193.

13. H. Wang, B. J. Lengerich, B. Aragam, and E. P. Xing, "Precision Lasso: Accounting for Correlations and Linear Dependencies in High-Dimensional Genomic Data," *Bioinformatics* 35, no. 7 (2019): 1181–1187.

14. R. Lockhart, J. Taylor, R. J. Tibshirani, and R. Tibshirani, "A Significance Test for the Lasso," *Annals of Statistics* 42, no. 2 (2014): 413–468, https://doi.org/10.1214/13-AOS1175.

15. C. Carrico, C. Gennings, D. C. Wheeler, and P. Factor-Litvak, "Characterization of Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting," *Journal of Agricultural, Biological, and Environmental Statistics* 20, no. 1 (2015): 100–120.

16. D. C. Wheeler, S. Rustom, M. Carli, T. P. Whitehead, M. H. Ward, and C. Metayer, "Assessment of Grouped Weighted Quantile Sum Regression for Modeling Chemical Mixtures and Cancer Risk," *International Journal of Environmental Research and Public Health* 18, no. 2 (2021): 504.

17. D. C. Wheeler, S. Rustom, M. Carli, T. P. Whitehead, M. H. Ward, and C. Metayer, "Bayesian Group Index Regression for Modeling Chemical Mixtures and Cancer Risk," *International Journal of Environmental Research and Public Health* 18, no. 7 (2021): 3486.

18. A. P. Keil, J. P. Buckley, K. M. O'Brien, K. K. Ferguson, S. Zhao, and A. J. White, "A Quantile-Based g-Computation Approach to Addressing the Effects of Exposure Mixtures," *Environmental Health Perspectives* 128, no. 4 (2020): 047004, https://doi.org/10.1289/EHP5838.

19. R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific Computing* 16, no. 5 (1995): 1190–1208, https://doi.org/10.1137/0916069.

20. S. G. Self and K. Y. Liang, "Asymptotic Properites of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions," *Journal of the American Statistical Association* 82, no. 398 (1987): 605–610.

21. M. M. Rahman, S. A. Carter, J. C. Lin, et al., "Prenatal Exposure to Tailpipe and Non-tailpipe Tracers of Particulate Matter Pollution and Autism Spectrum Disorders," *Environment International* 171 (2023): 107736.

22. S. A. Carter, M. M. Rahman, J. C. Lin, et al., "Maternal Exposure to Aircraft Emitted Ultrafine Particles During Pregnancy and Likelihood of ASD in Children," *Environment International* 178 (2023): 108061, https://doi.org/10.1016/j.envint.2023.108061.

23. M. M. Rahman, S. A. Carter, J. C. Lin, et al., "Associations of Autism Spectrum Disorder With PM2.5 Components: A Comparative Study Using Two Different Exposure Models," *Environmental Science & Technology* 57, no. 1 (2023): 405–414, https://doi.org/10.1021/acs.est.2c05197.

24. S. M. Meyers, R. G. Voigt, R. C. Colligan, et al., "Autism Spectrum Disorder: Incidence and Time Trends Over Two Decades in a Population-Based Birth Cohort," *Journal of Autism and Developmental Disorders* 49, no. 4 (2018): 1455–1474.

25. B. Bekkar, S. Pacheco, R. Basu, and N. DeNicola, "Association of air Pollution and Heat Exposure With Preterm Birth, Low Birth Weight, and Stillbirth in the US A Systematic Review," *JAMA Network Open* 3, no. 6 (2020): e208243, https://doi.org/10.1001/jamanetworkopen.2020.8243.

26. T. A. Becerra, M. Wilhelm, J. Olsen, M. Cockburn, and B. Ritz, "Ambient Air Pollution and Autism in Los Angeles County, California," *Environmental Health Perspectives* 121, no. 3 (2013): 380–386.

27. A. H. Xiang, X. Wang, M. P. Martinez, et al., "Association of Maternal Diabetes With Autism in Offspring," *JAMA* 313, no. 14 (2015): 1425–1434.

28. J. F. Bobb, L. Valeri, B. Claus Henn, et al., "Bayesian Kernel Machine Regression for Estimating the Health Effects of Multi-Pollutant Mixtures," *Biostatistics* 16, no. 3 (2014): 493–508, https://doi.org/10.1093/biostatistics/kxu058.

29. M. Lee, M. H. Rahbar, M. Samms-Vaughan, et al., "A Generalized Weighted Quantile Sum Approach for Analyzing Correlated Data in the Presence of Interactions," *Biometrical Journal* 61, no. 4 (2019): 934–954, https://doi.org/10.1002/bimj.201800259.

30. F. Oroumiyeh, M. Jerrett, I. Del Rosario, et al., "Elemental Composition of Fine and Coarse Particles Across the Greater Los Angeles Area: Spatial Variation and Contributing Sources," *Environmental Pollution* 292 (2022): 118356, https://doi.org/10.1016/j.envpol.2021.118356.

31. J. K. Gietl, R. Lawrence, A. J. Thorpe, and R. M. Harrison, "Identification of Brake Wear Particles and Derivation of a Quantitative Tracer for Brake Dust at a Major Road," *Atmospheric Environment* 44, no. 2 (2010): 141–146, https://doi.org/10.1016/j.atmosenv.2009.10.016.

32. H. Chun, C. Leung, S. W. Wen, J. McDonald, and H. H. Shin, "Maternal Exposure to Air Pollution and Risk of Autism in Children: A Systematic Review and Meta-Analysis," *Environmental Pollution* 256 (2020): 113307, https://doi.org/10.1016/j.envpol.2019.113307.

33. J. Pujol, R. Fenoll, D. Macià, et al., "Airborne Copper Exposure in School Environments Associated With Poorer Motor Performance and Altered Basal Ganglia," *Brain and Behavior: A Cognitive Neuroscience Perspective* 6, no. 6 (2016): e00467, https://doi.org/10.1002/brb3.467.

34. Y. Zheng, S. Chen, Y. Chen, et al., "Association Between PM2.5-Bound Metals and Pediatric Respiratory Health in Guangzhou: An Ecological Study Investigating Source, Health Risk, and Effect," *Frontiers in Public Health* 11 (2023): 1137933, https://doi.org/10.3389/fpubh.2023.1137933.

35. M. Kono, T. Y. Su, Y. Y. Chang, et al., "Assessing the Impact of Specific PM2.5-Bound Metallic Elements on Asthma Emergency Department Visits: A Case-Crossover Study in Taiwan," *Environmental Research* 255 (2024): 119130, https://doi.org/10.1016/j.envres.2024.119130.

36. R. M. Harrison, J. Allan, D. Carruthers, et al., "Non-exhaust Vehicle Emissions of Particulate Matter and VOC From Road Traffic: A Review," *Atmospheric Environment* 262 (2021): 118592, https://doi.org/10.1016/j.atmosenv.2021.118592.

37. V. R. Timmers and P. A. Achten, "Non-Exhaust PM Emissions From Electric Vehicles," *Atmospheric Environment* 134 (2016): 10–17, https://doi.org/10.1016/j.atmosenv.2016.03.017.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.
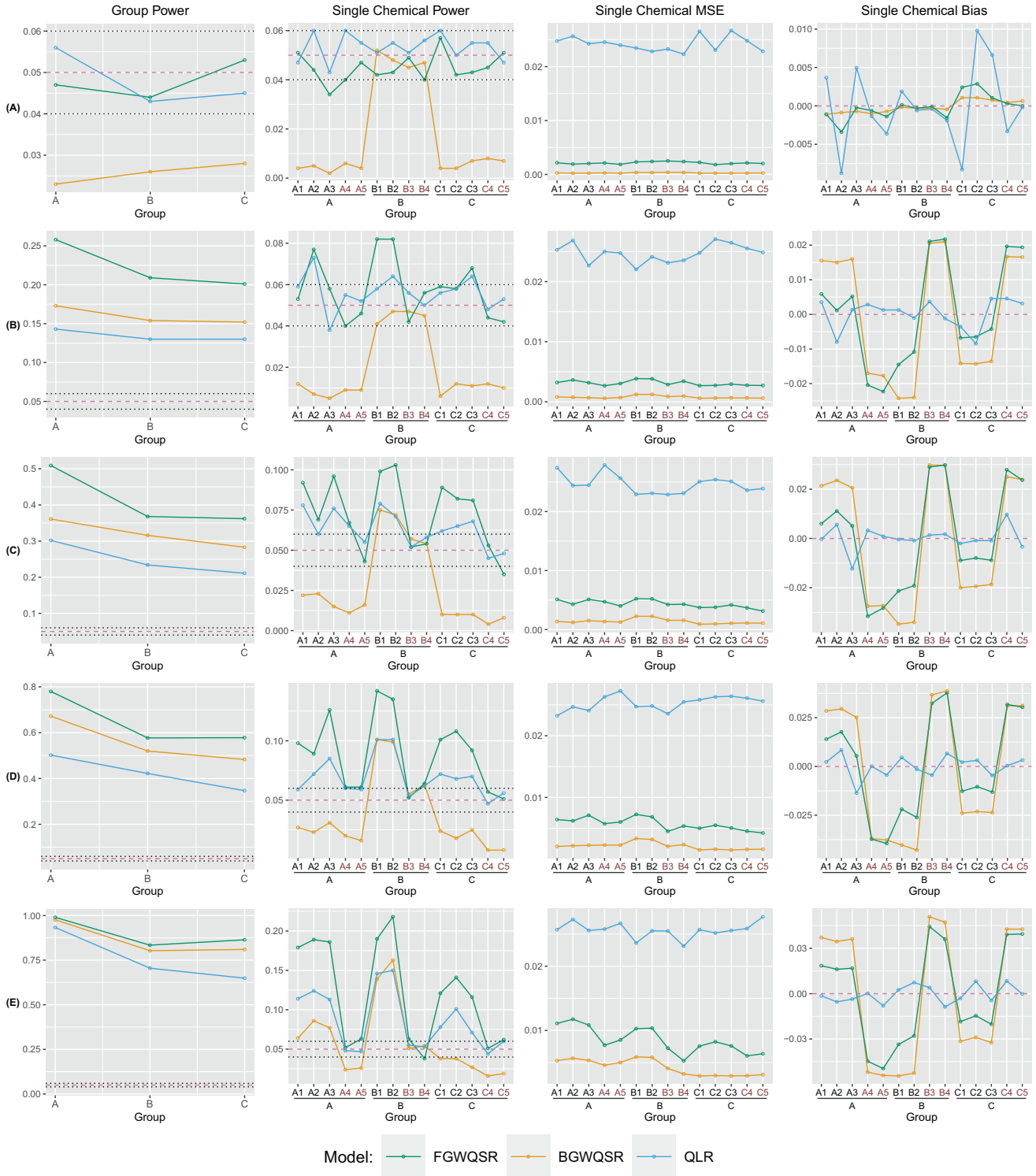
FIGURE A1 | Results of correctly specified simulations comparing the performance of FGWQSR, BGWQSR, and QLR over 1 000 replicates of simulated datasets with size $n = 500$ with a case-control ratio of approximately 1:1. Columns 1–4 correspond to mixture group effect power, single chemical power, individual chemical MSE, and individual chemical bias, respectively. In the presented simulations, CCS is fixed to (0.9,0.5) and the weight distribution of chemicals within groups is set as follows: $\mathbf{w}_A = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$; $\mathbf{w}_B = (\frac{1}{2}, \frac{1}{2}, 0, 0)$; and $\mathbf{w}_C = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$. Each row denoted (A)–(E) corresponds to five levels of SOA: $SOA_1 = (1, 1, 1)$; $SOA_2 = (0.9, 1.1, 1.1)$; $SOA_3 = (0.85, 1.15, 1.15)$; $SOA_4 = (0.8, 1.2, 1.2)$; and $SOA_5 = (0.7, 1.3, 1.3)$. Lines and points denoted by green, mustard, and light blue colors correspond to metrics computed from FGWQSR, BGWQSR, and QLR models, respectively. Dashed horizontal purple and black lines in power plots denote powers of 0.04, 0.05, and 0.06 as a reference for type 1 errors in the presence of null mixture group effects and null individual chemicals effects. Individual chemical effects colored in dark red denote null individual chemicals with weights of 0. Note that row (A) corresponds to the case where all group effects are null, which imposes that all individual chemical effects are also null.

FIGURE A2 | Results of moderate misspecification simulations comparing the performance of FGWQSR, BGWQSR, and QLR over 1 000 replicates of simulated datasets with size $n = 500$ with a case-control ratio of approximately 1:1. Columns 1–4 correspond to mixture group effect power, individual chemical power, individual chemical MSE, and individual chemical bias, respectively. In the presented simulations, CCS is fixed to (0.9,0.5) and the weight distribution of chemicals within groups is set as follows: $\mathbf{w}_A = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$; $\mathbf{w}_B = (\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, -\frac{1}{6}, -\frac{1}{6})$; and $\mathbf{w}_C = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$. Each row denoted (A)–(E) corresponds to one of five levels of SOA: $SOA_1 = (1, 1, 1)$; $SOA_2 = (0.9, 1.1, 1.1)$; $SOA_3 = (0.85, 1.15, 1.15)$; $SOA_4 = (0.8, 1.2, 1.2)$; and $SOA_5 = (0.7, 1.3, 1.3)$. Lines and points denoted by green, mustard, and light blue colors correspond to metrics computed from FGWQSR, BGWQSR, and QLR models, respectively. Dashed horizontal purple and black lines in power plots denote powers of 0.04, 0.05, and 0.06 as a reference for type 1 errors in the presence of null mixture group effects and null individual chemicals effects. Individual chemical effects colored in dark red denote null individual chemicals with weights of 0. Individual chemicals in group B with a dark blue shading (B1–B4) have a positive weight of size $\frac{1}{6}$ while chemicals with an orange shading (B5 and B6) have a negative weight of size $-\frac{1}{6}$. Note that row (A) corresponds to the case where all group effects are null, which imposes that all individual chemical effects are also null.
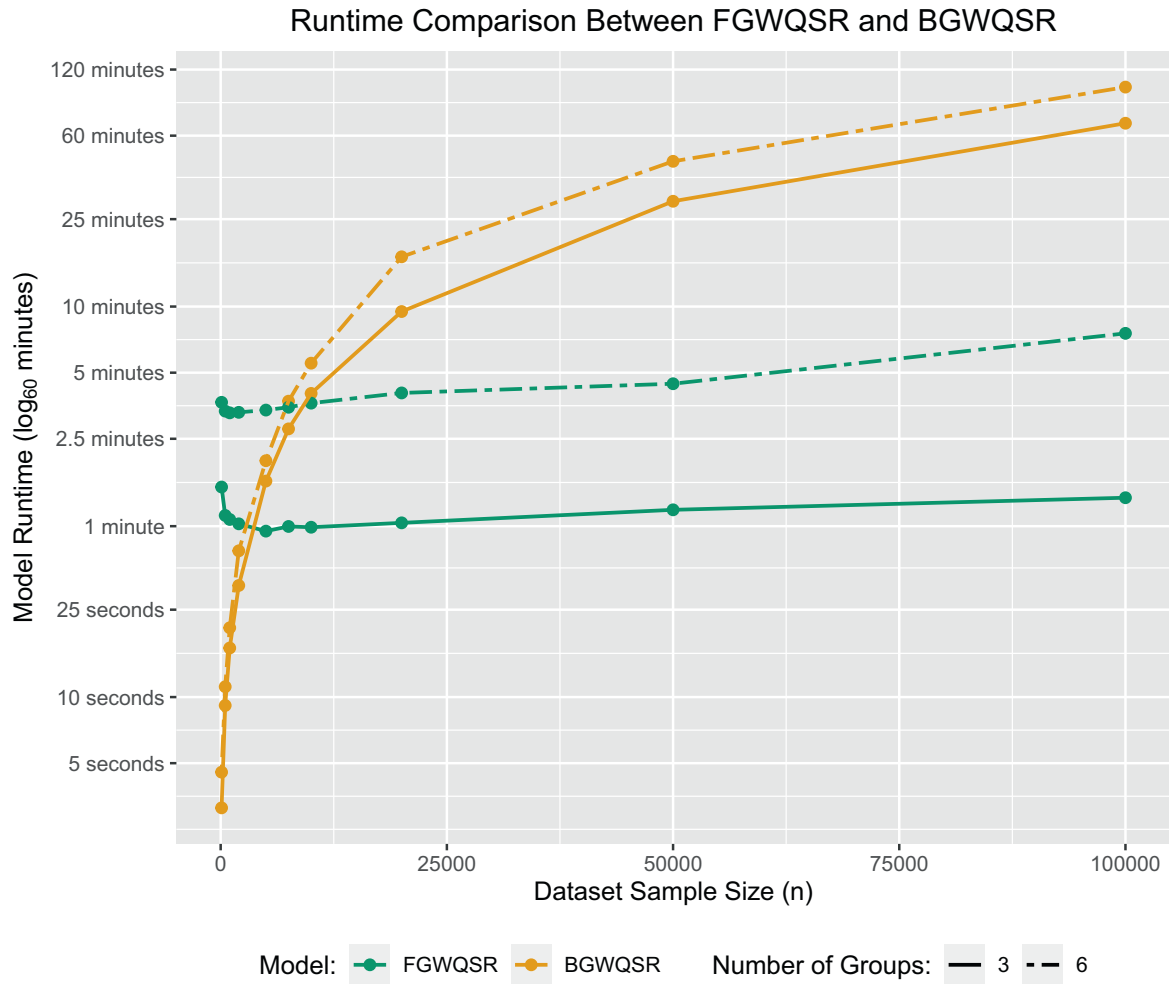
FIGURE A3 | Results of runtime comparison simulations between FGWQSR and BGWQSR. Runtimes are resolved at dataset sizes of $n = \{100,$ $500, 1\,000, 2\,000, 5\,000, 7\,500, 10\,000, 20\,000, 50\,000, 100\,000\}$ with a case-control ratio of approximately 1:1, and CCS fixed to $(0.7, 0.3)$. Simulations are run for 3-group and 6-group mixture configurations with $SOA_{3\text{-group}} = (0.8, 1, 1.2)$ and $SOA_{6\text{-group}} = (0.8, 1, 1.2, 0.8, 1, 1.2)$ with weight distributions as follows: $\mathbf{w}_1 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$, $\mathbf{w}_2 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$, and $\mathbf{w}_3 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$; $\mathbf{w}_1 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$, $\mathbf{w}_2 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$, and $\mathbf{w}_3 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$, $\mathbf{w}_4 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$, $\mathbf{w}_5 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$, and $\mathbf{w}_6 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$. Lines and points denoted by green and orange colors correspond to the averaged runtimes of FGWQSR and BGWQSR models, respectively, over 10 replicates. Solid lines correspond to runtime simulations under a 3-group configuration, while two-dash lines correspond to runtime simulations under a 6-group configuration. Model runtimes are plotted in the scale of $\log_{60}(\text{minutes})$ with labels at selected runtimes.
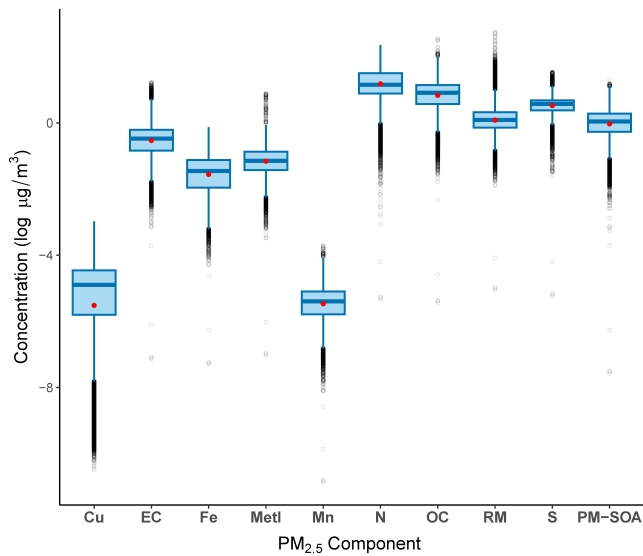
FIGURE A4 | Distribution of pregnancy averaged PM$_{2.5}$ component exposure generated from Source Oriented Chemical Transport models. Red points denote mean PM$_{2.5}$ component exposure levels. Pollutant concentrations are displayed in $\log(\mu g/m^3)$ units.
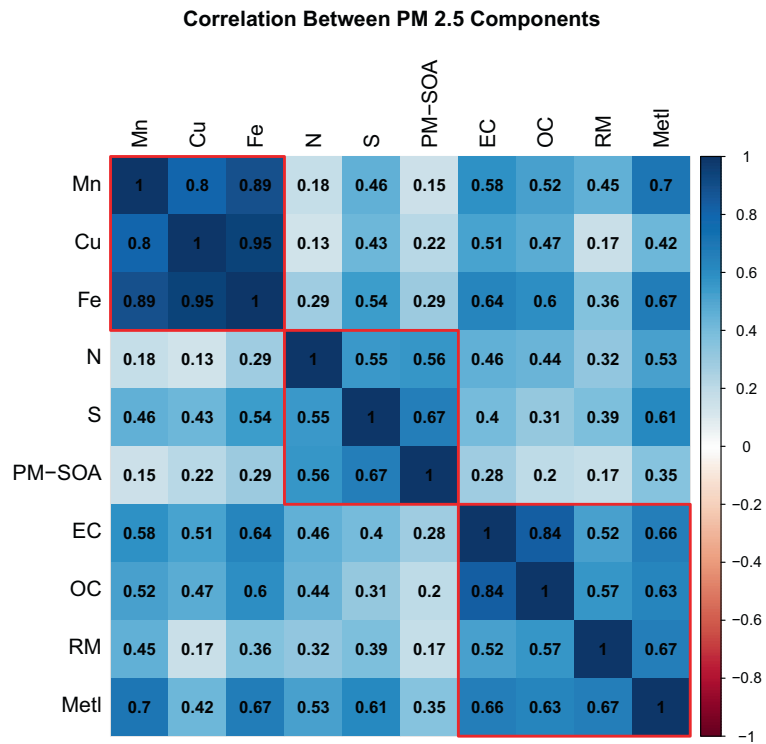
FIGURE A5 | Correlation plot of PM$_{2.5}$ components. Red boxes denote resulting clusters from hierarchical clustering on $1 - C$, where C denotes the correlation matrix, using a complete linkage and a cutpoint value to establish 3 clusters.

TABLE A1 | FGWQSR summary output for the association between pregnancy averaged maternal $PM_{2.5}$ component exposure levels on child ASD diagnosis by age 5. Included in this analysis are adjustments for the following potential confounders: Maternal ozone exposure, maternal nitrogen dioxide exposure, total $PM_{2.5}$ mass, child birth year, seasonality, maternal age at delivery, maternal race, maternal education level, maternal comorbidity, median family income in census tract of residence, child sex, maternal obesity status, and maternal diabetes diagnosis. Odds ratio (OR) estimates for mixture group effects can be interpreted per one unit increase in weighted mixture $PM_{2.5}$ exposure.

Group	Group effect OR	LRT	P	
Trace metals	1.127	25.830	< 0.001	***
Secondary $PM_{2.5}$	1.019	0.959	0.748	
Other primary $PM_{2.5}$	1.106	10.938	0.011	*

Group	Chemical name	Weight estimate	LRT	P	
Trace metals	Copper	0.648	4.215	0.019	*
	Iron	0.129	0.126	0.364	
	Manganese	0.223	0.812	0.208	
Secondary $PM_{2.5}$	Nitrate	0	0	1	
	Sulfate	1	0.471	0.336	
	Secondary organic aerosols	0	0	1	
Other primary $PM_{2.5}$	Elemental carbon	0.172	0.571	0.169	
	Organic carbon	0	0	1	
	Residual mass	0.457	3.385	0.039	*
	Other metals	0.371	1.538	0.116	

TABLE A2 | Comparison of odds ratio (OR) estimates and p-values for individual pollutant effects from FGWQSR and Quantile Logistic Regression (QLR) models. ORs are interpreted per one-quintile increase in pollutant exposure level. Individual chemical ORs for FGWQSR are computed as the exponential of the product between the chemical's group effect multiplied by the chemical's weight. Included in both models are adjustments for the following potential confounders: Maternal ozone exposure, maternal nitrogen dioxide exposure, total $PM_{2.5}$ mass, child birth year, seasonality, maternal age at delivery, maternal race, maternal education level, maternal comorbidity, median family income in census tract of residence, child sex, maternal obesity status, and maternal diabetes diagnosis. Bolded p-values indicate statistical significance at threshold $\alpha = 0.05$.

		Individual chemical OR estimate	P
Group	Chemical name	FGWQSR/QLR	FGWQSR/QLR
Trace metals	Copper	**1.081**/1.083	**0.019/0.036**
	Iron	1.016/1.017	**0.364**/0.702
	Manganese	1.027/1.022	0.208/0.462
Secondary $PM_{2.5}$	Nitrate	1/0.975	1/0.39
	Sulfate	1.019/1.012	0.336/0.564
	Secondary organic aerosols	1/0.993	1/0.783
Other primary $PM_{2.5}$	Elemental carbon	1.017/1.027	0.169/0.28
	Organic carbon	1/0.972	1/0.294
	Residual mass	1.047/1.043	**0.039**/0.097
	Other metals	1.038/1.037	0.116/0.227