The Annals of Applied Statistics
2024, Vol. 18, No. 1, 642–663
https://doi.org/10.1214/23-AOAS1805

BAYESIAN ADAPTIVE AND INTERPRETABLE FUNCTIONAL REGRESSION FOR EXPOSURE PROFILES

BY YUNAN GAO[a] AND DANIEL R. KOWAL[b]

Department of Statistics, Rice University, [a]*yunan.gao@rice.edu,* [b]*daniel.kowal@rice.edu*

Pollutant exposure during gestation is a known and adverse factor for birth and health outcomes. However, the links between prenatal air pollution exposures and educational outcomes are less clear, in particular, the critical windows of susceptibility during pregnancy. Using a large cohort of students in North Carolina, we study the link between prenatal daily $PM_{2.5}$ exposure and fourth end-of-grade reading scores. We develop and apply a locally adaptive and highly scalable Bayesian regression model for scalar responses with functional and scalar predictors. The proposed model pairs a B-spline basis expansion with dynamic shrinkage priors to capture both smooth and rapidly-changing features in the regression surface. The model is accompanied by a new decision analysis approach for functional regression that extracts the critical windows of susceptibility and guides the model interpretations. These tools help to identify and address broad limitations with the interpretability of functional regression models. Simulation studies demonstrate more accurate point estimation, more precise uncertainty quantification, and far superior window selection than existing approaches. Leveraging the proposed modeling, computational, and decision analysis framework, we conclude that prenatal $PM_{2.5}$ exposure during early and late pregnancy is most adverse for fourth end-of-grade reading scores.

1. Introduction. Scalar-on-function regression (SOFR) has emerged as a critical modeling tool for elucidating the relationship between a scalar response variable and data collected repeatedly over a continuous domain, such as time or space (Morris (2015)). The SOFR model is formally expressed as

$$y_i = \mu + \boldsymbol{z}_i'\boldsymbol{\alpha} + \int_{\mathcal{T}_i} X_i(t)\beta(t)\,dt + \epsilon_i, \quad \epsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2), i = 1, \ldots, n. \tag{1}$$

In this model the scalar response $y_i \in \mathbb{R}$ is linked to the functional covariate $X_i : \mathcal{T}_i \to \mathbb{R}$ and other scalar covariates $\boldsymbol{z}_i \in \mathbb{R}^p$. The compact domains $\mathcal{T}_i \subseteq \mathcal{T} \subset \mathbb{R}$ are subject-specific, which is important for our motivating application introduced below, and $\mathcal{T}$ denotes the maximal domain over which the functional covariates X_i may be observed. The integral term $\int_{\mathcal{T}_i} X_i(t)\beta(t)\,dt$ represents the cumulative effect of X_i over its domain $\mathcal{T}_i$, while the linear term $\boldsymbol{z}_i'\boldsymbol{\alpha}$ (or additive generalizations, see Section 5) accounts for other scalar covariates that may influence the response y_i.

The main goal of fitting a SOFR model (1) is to estimate and characterize the unknown regression function β, which describes the effects of the functional covariates X_i on the scalar response y_i. Further, we aim to identify any regions within the domain $\mathcal{T}$—referred to as *critical windows*—that are predictive of y_i, while adjusting for important confounding variables $\boldsymbol{z}_i$. Given that the functional covariates X_i are usually high dimensional, highly correlated, and often collected on irregularly-spaced locations over the domain, regularization of β is a central focus in SOFR. Regularization is enforced via penalties or priors to guard against

Received October 2022; revised May 2023.
Key words and phrases. Decision analysis, functional data analysis, nonparametric regression, shrinkage, spline.

both overfitting β and the multicollinearities induced by the within-function correlations of $\{X_i\}$. Classical approaches expand β, using a known basis expansion, and introduce a prior or penalty that encourages smoothness, such as splines with penalties on the differenced coefficients (James (2002), Marx and Eilers (1999)) or wavelets with sparsity priors or penalties (Brown, Vannucci and Fearn (1998), Morris et al. (2008), Morris and Carroll (2006)). Principal components analysis can be applied directly to (1) but does not account for the ordering within the X_i curves (Cardot, Ferraty and Sarda (1999), Müller and Stadtmüller (2005)).

This paper highlights and addresses two fundamental and significant challenges for SOFR. First, the performance and utility of model (1) hinges on the ability to estimate β. If the model for β fails to capture the shape of the true regression function, which may vary smoothly or exhibit rapid changes, then the estimates of β will be biased, the uncertainty quantification for β will be poorly calibrated, and the predictions of y will be suboptimal. Existing methods for SOFR commonly produce interval estimates for β that are far too conservative, which limits the power to detect important covariate effects (see Section 4). Thus, it is critical to produce estimation and inference tools for β that adapt to both smooth and rapid changes and provide precise yet well-calibrated uncertainty quantification. At the same time, modern high-dimensional datasets demand scalability in both the number of observations n and the number of observation points along $\mathcal{T}$. Our modeling and computing strategies emphasize both *adaptability* and *scalability*.

The second and more subtle challenge is that of interpretability: given an estimate and inference of the regression coefficient function β, how does one interpret the results? More concretely, consider $\beta(t^*)$ at a specific point $t^* \in \mathcal{T}$. In the context of (1), we may be tempted to interpret the coefficient function as

$$\beta(t^*) \approx \int_{N(t^*)} \{X(t) + 1\}\beta(t)\,dt - \int_{N(t^*)} X(t)\beta(t)\,dt \tag{2}$$

$$= \mathbb{E}\big[y \mid z_i, \{X(t) + 1\}_{t\in N(t^*)}\big] - \mathbb{E}\big[y \mid z_i, \{X(t)\}_{t\in N(t^*)}\big], \tag{3}$$

where $N(t^*)$ is a small neighborhood around t^* and the expectations also condition on the parameters $(\mu, \boldsymbol{\alpha}, \beta)$. Informally, (2)–(3) suggest that the regression function β at time t^* corresponds to the change in the expected response variable for a one-unit increase of X in a neighborhood of t^*, *all else equal*. Yet for functional covariates, this latter qualification is usually not meaningful: given a trajectory $\{X(t)\}_{t\in\mathcal{T}}$, it is difficult to envision that same trajectory but with $X(t)$ replaced by $X(t) + 1$ *only in a small neighborhood of* t^*. Such an abrupt and localized perturbation of the trajectory is typically not consistent with the data-generating process, especially when the curves X_i are modeled as smooth functions. These difficulties propagate more broadly, including effect directions and selection of critical windows. In particular, the interpretation of $\beta(t) > 0$ for $t \in \mathcal{T}^+$ and $\beta(t) < 0$ for $t \in \mathcal{T}^-$ for subdomains $\mathcal{T}^+, \mathcal{T}^- \subset \mathcal{T}$ is nontrivial, especially when the curves X_i exhibit structured (e.g., seasonal) correlations; this issue is discussed in detail in Section 5. Similar warnings regarding interpretability were issued by Dziak et al. (2019), although they did not suggest general purpose tools to resolve these challenges. As such, we are motivated to produce *model summarization* techniques that enable both interpretable estimation and powerful window selection for SOFR.

In response to these challenges, we propose a new Bayesian adaptive scalar-on-function regression (BASOFR) model paired with a decision analysis strategy to select critical windows and deliver more interpretable model summaries. The BASOFR specifies a B-spline basis expansion for β and a dynamic shrinkage prior (Kowal, Matteson and Ruppert (2019)) on the (second differenced) basis coefficients. Crucially, this local and adaptive shrinkage prior encourages smoothness yet can capture rapid changes in β, which produces better point estimates and more precise uncertainty quantification, especially in the presence of both smooth

and rapidly-changing features (see Section 4). Importantly, the proposed modeling structure admits a highly scalable Gibbs sampling algorithm, which is essential for handling high-dimensional datasets.

Leveraging the BASOFR output, we develop a decision analysis approach to select the critical windows of $\mathcal{T}$ and provide interpretable model summaries. A crucial observation is that estimation and uncertainty quantification for β is *not* sufficient for selecting critical windows: some decision analysis or other selection criteria (see below) are required. We propose to extract locally constant point estimates from the BASOFR model—or more generally, any Bayesian SOFR model—which feature estimates of the form $\hat{\beta}(t) = \hat{\delta}_k$ for $t \in \mathcal{T}_k$ and $\{\mathcal{T}_k\}$ a learned partition of $\mathcal{T}$. In conjunction, the estimated coefficients and partition $\{\hat{\delta}_k, \mathcal{T}_k\}$ identify effect sizes, effect directions, and critical windows. These locally constant estimates also provide a partial resolution to the challenges raised by (2)–(3): namely, $\hat{\delta}_k$ estimates the change in the expected response variable for a one-unit increase in the aggregated trajectory $X(\mathcal{T}_k) := \int_{\mathcal{T}_k} X(t)\,dt$ while holding $\{X(\mathcal{T}_j)\}_{j\neq k}$ constant. Here, the notion of *all else equal* is more plausible and less restrictive: it refers to distinct regions of the domain—rather than neighboring time points—and only requires the aggregated trajectories $\{X(\mathcal{T}_j)\}_{j\neq k}$—rather than the entire $\{X(t)\}_{t\in\mathcal{T}}$ paths—to be held constant outside of $\mathcal{T}_k$. These tools contribute minimal computational cost and complement more traditional posterior summaries, such as expectations and credible intervals.

The proposed decision analysis approach deviates from the vast majority of methods for critical window selection, which rely on pointwise credible intervals for β (Bose et al. (2017), Lee et al. (2018), Leon Hsu et al. (2015), Warren et al. (2012), Wilson et al. (2017)) or other marginal criteria (Warren et al. (2020)) under a Bayesian SOFR. However, it has been shown that *variable* selection based on credible interval—that is, variables are selected if the credible intervals exclude zero—is severely underpowered and overconservative, especially compared to recent decision analysis strategies (Kowal (2022a), Kowal (2022b), Kowal and Bourgeois (2020)). We confirm this effect for *window* selection (Section 4.2), which implicitly warns that popular existing methods may be erroneously omitting key windows of susceptibility.

Our decision analysis approach continues a line of research on posterior summarization of Bayesian models, which has been directed primarily for variable selection, including linear regression (Hahn and Carvalho (2015)), graphical models (Bashir et al. (2019)), seemingly-unrelated regressions (Puelz, Hahn and Carvalho (2017)), and function-on-scalars regression (Kowal and Bourgeois (2020)). The window selection problem is more closely related to change point detection than variable selection but specific to the regression coefficient function β in (1) rather than observed data.

An intuitive and alternative Bayesian approach, BLISS (Grollemund et al. (2019)), places a prior on β that constrains the regression function to be locally constant with unknown levels and partitions. However, BLISS faces a substantial computational burden and does not scale to moderate or large datasets such as ours (see Figure 3). In addition, BLISS requires careful specification of multiple hyperparameters, including a fixed choice for the number of partitions. Grollemund et al. (2019) proposed to fit separate models for each specified number of partitions and then to compare them using BIC. Naturally, this exercise further increases the computational cost. Regardless, we emphasize that the prior alone cannot *select* windows: whether using BASOFR, BLISS, or any other Bayesian SOFR model, some decision analysis or selection criterion is still required. Our two step procedure—fitting the scalable and adaptive BASOFR model and summarizing the output using customized decision analysis—circumvents these computational challenges yet still provides posterior uncertainty quantification, locally constant point estimates, and powerful window selection.

Among frequentist methods, Picheny, Servien and Villa-Vialaneix (2019) introduced a semiparametric approach for window selection, which does not require the linearity in (1)

TABLE 1

Variables in the NC dataset. Data are restricted to individuals with 30–42 *weeks of gestation, zero to* 104 *weeks of age-within-cohort, mother's age* 15-44, `Blood_Lead` $\leq$ 10, *birth order* $\leq$ 4, *no current limited English proficiency, and residence in NC at the time of birth and the time of fourth EOG test*

	Air quality during gestation
`Prenatal_PM2.5`	Daily $PM_{2.5}$ level estimated at the 2010 census tract of the mother's home address (with length `Gestation`)
	Birth information
`mEdu`	Mother's education group at the time of birth (NoHS = no high school diploma, HS = high school diploma, higherHS = some college/associates or higher)
`mRace`	Mother's race/ethnicity group (Non-Hispanic (NH) White, NH Black, Hispanic)
`mAge`	Mother's age at the time of birth
`Male`	Male infant? (1 = Yes)
`Smoker`	Mother smoked? (1 = Yes)
`Gestation`	Clinical estimate of the gestation length (days)
`BirthMonth`	Birth month of the student
	Education/End-of-grade (EOG) test information
`Reading_Score`	Standardized score for the (chronologically first) fourth EOG reading test
`Age_w_cohort`	Age-within-cohort: the relative age of each student within their cohort, see the online Supplementary Material for details and summary statistics (Gao and Kowal (2024)).
	Blood lead surveillance
`Blood_lead`	Blood lead level (micrograms per deciliter)
	Social/Economic status
`EconDisadvantage`	Economically disadvantaged students are indicated by participation in the free/reduced price lunch program (1 = Participation in the program)

and has been applied for datasets with a moderate to large number of observation points along $\mathcal{T}$. However, this approach requires specification of four tuning parameters, does not readily adjust for scalar covariates z_i, and does not provide estimation or inference for β, which is essential to quantify the direction and magnitude of the effects within these critical intervals. James, Wang and Zhu (2009) introduced an approach to SOFR that imposes sparsity on the derivatives of β, but similarly, it requires multiple tuning parameters and does not provide uncertainty quantification for β.

Our methods are motivated by a real-world study examining the impact of daily exposure to $PM_{2.5}$ during pregnancy on standardized fourth end-of-grade (EOG) reading test scores in a large cohort of students in North Carolina (NC) (Section 5). The dataset is created by linking multiple administrative datasets in NC that include birth and demographic information, blood lead level measurements, socioeconomic status, and fourth EOG reading test scores on $n = 98{,}159$ mother-child pairs; see Table 1 for the primary variables and Bravo and Miranda (2021) and Feldman and Kowal (2022) for additional details. Daily $PM_{2.5}$ exposure at the mother's home address is computed using the Fused Air Quality Surface Using Downscaling (FAQSD) data provided by the United States Environmental Protection Agency.

The goal of our analysis is to estimate and characterize the effects of prenatal $PM_{2.5}$ exposure on educational outcomes and, in particular, to identify the time periods during gestation—if any—that are predictive of adverse educational outcomes, while adjusting for important confounding variables. Since $PM_{2.5}$ is recorded as a function of days-in-gestation and EOG reading scores are scalar outcomes, SOFR offers a natural modeling framework. This analysis necessitates statistical methodologies capable of: (i) adapting to both smooth and abrupt changes in the coefficient function β, (ii) scaling to handle $n = 98{,}159$ observations and approximately 300 observation points per function, and (iii) providing tools for in-

ference to describe, interpret, and visualize the model-based results. Our proposed BASOFR model and accompanying decision analysis effectively cater to these needs.

In the context of estimating the effects of cumulative exposures, the distributed lag model (DLM) is a widely-used variant of (1) that replaces the integral with a multiple regression equation featuring lagged exposure measurements (Schwartz (2000)). As in SOFR, the DLM emphasizes flexible modeling of the regression surface along with regularization of the coefficients and has been generalized for spatiotemporal data (Warren et al. (2012)) and tree-based regression models (Mork and Wilson (2022)). Despite the similarities between SOFR and DLMs, we prefer the representation in (1) because it does not require the exposures $X_i(t)$ to be observed at the same time points for all subjects. In particular, gestational length ranges from 30 to 42 weeks in our dataset, which requires careful consideration of the subject-specific domain $\mathcal{T}_i$ in (1). Nonetheless, the proposed modeling, computational, and decision-analysis strategies remain relevant for DLMs.

This article is outlined as followed. Section 2 presents the proposed BASOFR model. Section 3 contains the decision analysis approach for window selection and model summarization. Section 4 contains a detailed simulation study. The methods are applied to the NC data in Section 5. Section 6 concludes. The online Supplementary Material includes computational details, additional simulation results, and supporting information about the NC data (Gao and Kowal (2024)). An `R` package is available at https://GitHub.com/YunanGao/BaiSOFR.

2. Bayesian adaptive scalar-on-function regression. The core task in fitting the SOFR model (1) is to learn the unknown regression coefficient function β, given observations $\{(X_i, y_i)\}_{i=1}^n$; here we omit the scalar covariates $\boldsymbol{z}_i$ for simplicity. With real data the functions are observed at discrete points: $\boldsymbol{x}_i = (X_i(t_{i,1}), \ldots, X_i(t_{i,m_i}))'$, where the m_i observation points may be unequally-spaced or differ from subject to subject. Although it is tempting to apply (nonfunctional) linear regression models to $\{(\boldsymbol{x}_i, y_i)\}$, there are several drawbacks to this approach. First, the resulting model will be high dimensional with highly-correlated covariates and, therefore, requires regularization. Second, such a model fails to account for the ordering among the covariates with respect to the domain $t_{i,j} \in \mathcal{T}$, which is crucial information for both regularization and interpretation. Lastly, it is unclear how to apply this approach when the functional covariates are not observed on a common grid, which occurs for our application (Section 5).

We instead pursue a basis expansion strategy for both the functional predictors and the regression coefficient function (Ramsay and Silverman (2005)). By expanding $X_i(t) = \sum_{k=1}^{K_X} X_{ik}^* \phi_k(t)$ and $\beta(t) = \sum_{k=1}^{K_B} B_k^* \psi_k(t)$ for known basis functions $\{\phi_k(\cdot)\}_{k=1}^{K_X}$ and $\{\psi_k(\cdot)\}_{k=1}^{K_B}$ and unknown coefficients $\{X_{ik}^*\}_{k=1}^{K_X}$ and $\{B_k^*\}_{k=1}^{K_B}$, the key term in (1) simplifies to

$$\int_{\mathcal{T}_i} X_i(t)\beta(t)\,dt = \mathbf{X}_i^* \mathbf{J}_i^{\phi,\psi} \boldsymbol{B}^* = \mathbf{X}_i^{**} \boldsymbol{B}^*, \tag{4}$$

where $\mathbf{X}_i^* = (X_{i1}^*, \ldots, X_{iK_X}^*)$, $\boldsymbol{B}^* = (B_1^*, \ldots, B_{K_B}^*)$, and $\mathbf{J}_i^{\phi,\psi} = [\int_{\mathcal{T}_i} \phi_j(t)\psi_k(t)\,dt]_{jk}$. The basis expansions resolve the difficulties with unequally-spaced or noncommon observation points for the functional predictors $\{X_i\}$, since we instead work with the coefficients $\boldsymbol{X}_i^*$. In addition, the basis expansion of each X_i serves as a regularization tool to smooth over the measurement errors associated with the direct observations $\boldsymbol{x}_i$. Lastly, the representation in (4) shows that fitting the SOFR (1) can be made equivalent to fitting a multiple linear regression model with covariates $\mathbf{X}_i^{**} = \mathbf{X}_i^* \mathbf{J}_i^{\phi,\psi}$, which is known. Additional details regarding the basis expansions are provided in the online Supplementary Material (Gao and Kowal (2024)).

Under these basis expansions, estimation and inference on the basis coefficients $\boldsymbol{B}^*$ are sufficient for estimation and inference on the regression function β. Thus, a prior on $\boldsymbol{B}^*$ implies a prior on β. Yet despite the promise of the multiple linear regression interpretation

of (4), the prior on $\boldsymbol{B}^*$ must be specified carefully. Common shrinkage priors for regression are designed to shrink redundant linear coefficients to zero. However, sparsity in $\boldsymbol{B}^*$ does not guarantee smoothness or other desirable properties of β. More specifically, the choice of prior on $\boldsymbol{B}^*$ cannot be decoupled from the choice of basis. For instance, when $\boldsymbol{B}^*$ is assigned a Gaussian prior, the resulting coefficient function is a Gaussian process with covariance function $\mathrm{Cov}\{\beta(t), \beta(s)\} = \sum_{k,\ell} \psi_k(t)\psi_\ell(s)\mathrm{Cov}(B_k^*, B_\ell^*)$. Thus, the prior on β inherits key properties from both the basis functions and the prior on $\boldsymbol{B}^*$.

Our strategy marries a particular choice of basis functions with a locally adaptive shrinkage prior. Specifically, we select a B-spline basis with a moderate number of equally-spaced knots. B-splines are numerically stable with convenient computational properties, in part due to the local compactness and ordering among the basis functions. These properties further motivate and enable our prior specification for the basis coefficients. As an illustrative example, consider the B-spline basis coefficients for the nonlinear function in Figure 1. The function is smooth yet features two regions with rapid changes, which are highlighted by the plot of β''. The B-spline basis coefficients $\boldsymbol{B}^*$ (determined via least squares for this illustration) offer several suggestions for an ideal prior. First, the coefficients inherit an ordering similar to that in the original domain $\mathcal{T}$. Thus, neighboring coefficients should be shrunk together to encourage smoothness. Next, the second differences of the coefficients, $\Delta^2 B_k^* = \Delta B_k^* - \Delta B_{k-1}^*$ with $\Delta B_k^* = B_{k+1}^* - B_k^*$, closely resemble the second derivatives β''. This observation has motivated P-splines (Marx and Eilers (1999)), which imitate the familiar roughness penalty $\int_{\mathcal{T}}\{\beta''(t)\}^2\,dt$ with the coefficient analog $\sum_k(\Delta^2 B_k^*)^2$. However, this *global* penalty ignores the final critical observation in Figure 1: the smooth periods of β correspond to zeros in $\Delta^2 B_k^*$, while the rapidly-changing periods exhibit *volatility clustering*, that is, consecutive sequences of large absolute values. More specifically, the shrinkage in the peaked regions (around the 50–80th and the 120–150th coefficients) should not be as aggressive as in the flat regions. Classical P-splines cannot capture this behavior: the rate of shrinkage is global across all k and $\mathcal{T}$. Thus, adequate B-spline modeling of functions with both smooth and rapidly-changing features requires a prior that: (i) encourages smoothness via near-sparsity of $\Delta^2 B_k^*$ and (ii) admits local adaptivity via dynamic volatility modeling.

Motivated by these considerations, we propose the following locally adaptive shrinkage prior on the regression coefficient function:

$$\beta(t) = \sum_{k=1}^{K_B} B_k^* \psi_k(t), \tag{5}$$

$$\Delta^2 B_k^* \mid \lambda_k \overset{\text{indep}}{\sim} \mathcal{N}(0, \lambda_k^2), \quad \{\lambda_k\} \sim \mathrm{DHS}, \tag{6}$$

where $\{\psi_k\}_{k=1}^{K_B}$ is a collection of equally-spaced B-splines and DHS refers to the dynamic horseshoe prior (Kowal, Matteson and Ruppert (2019)). The local scale parameters λ_k determine the smoothness of the function over the (compact) support of ψ_k: when λ_k is small, the function β is approximately locally linear; when λ_k is large, the function β admits large changes in the slope. This local adaptivity is enabled by the basis-specific scales λ_k. By comparison, classical P-splines apply global smoothness via a common scale $\lambda_k = \lambda$.

Relative to the horseshoe prior (Carvalho, Polson and Scott (2010)), the dynamic horseshoe prior offers key advantages for adaptive function estimation. The horseshoe prior assumes *independent* half-Cauchy priors for λ_k, which does not account for the volatility clustering observed in Figure 1. Informally, nondynamic shrinkage priors do not incorporate information regarding the shrinkage behavior of neighboring regions, which produces inferior estimates and uncertainty quantification (see Figure 2). The dynamic horseshoe prior resolves these issues using a volatility model,

$$h_k := \log(\lambda_k^2), \qquad h_{k+1} = \mu_h + \phi(h_k - \mu_h) + \eta_{k+1}, \qquad \eta_k \overset{\text{iid}}{\sim} Z(1/2, 1/2, 0, 1), \tag{7}$$

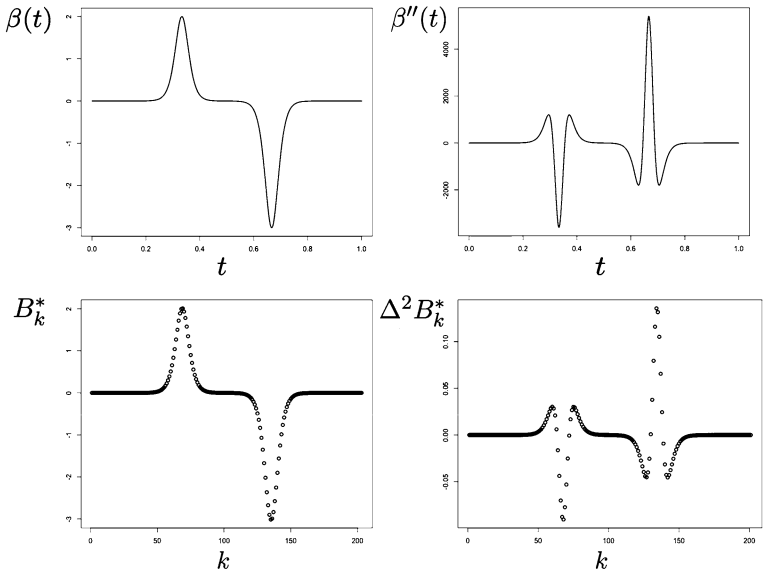

FIG. 1. *The nonlinear function* $\beta(t) = 8\{2 + \exp(20 - 60x) + \exp(60t - 20)\}^{-1} - 12\{2 + \exp(40 - 60t) + \exp(60t - 40)\}^{-1}$ (top left), *the second derivatives* β'' (top right), *the* (*ordinary least squares*) *basis coefficients using a dense B-spline basis with equally-spaced knots* (bottom left), *and the second differences of the basis coefficients* (bottom right).

where $Z(a, b, 0, 1)$ denotes the Z-distribution with density $[z] = \{B(a, b)\}^{-1} \exp\{z\}^a [1 + \exp\{z\}]^{-(a+b)}$ and $B(\cdot, \cdot)$ is the beta function. The dynamic horseshoe prior models the log-variances of the second-differenced basis coefficients with an autoregressive model of order one, which resembles classical Bayesian volatility models for time series analysis (Kim, Shephard and Chib (1998)). The key distinctions here are: (i) the presence of the Z-distribution and (ii) the role of the coefficient indices k. First, Kowal, Matteson and Ruppert (2019) showed that many common shrinkage priors expressed via λ_k can be represented on the log-scale with a Z-distribution; see Table 2. When there is no autoregressive behavior $\phi = 0$, the prior (7) with $a = b = 1/2$ exactly reproduces the horseshoe prior. As such, the dynamic horseshoe is capable of providing both aggressive shrinkage and persistence of large signals, which here corresponds to local smoothness and rapidly-changing features in β, respectively.

Second, the "time" index for the volatility model is k, which corresponds to the second differenced basis coefficient $\Delta^2 B_k^*$. This modeling structure is appropriate due to the local compactness and ordering of the (equally-spaced) B-spline basis functions (Figure 1).

To demonstrate the importance of both the *dynamic* and *shrinkage* aspects of the prior for SOFR, we consider a brief example with simulated data. Data from the SOFR model (1) are generated for $n = 500$ observations with a moderate signal-to-noise ratio (SNR $= 5$) using the nonlinear function from Figure 1 for the true regression coefficient function; additional details are provided in Section 4. To compare with the proposed approach, we consider a vari-

TABLE 2
Each of these priors is reproduced by a Z-distribution on the log-variance

$a = b = 1/2$	Horseshoe Prior (Carvalho, Polson and Scott (2010))
$a = 1/2, b = 1$	Strawderman–Berger Prior (Berger (1980), Strawderman (1971))
$a = 1, b = c - 2, c > 0$	Normal-Exponential-Gamma Prior (Griffin and Brown (2005))
$a = b \to 0$	(Improper) Normal-Jeffreys' Prior (Figueiredo (2003))

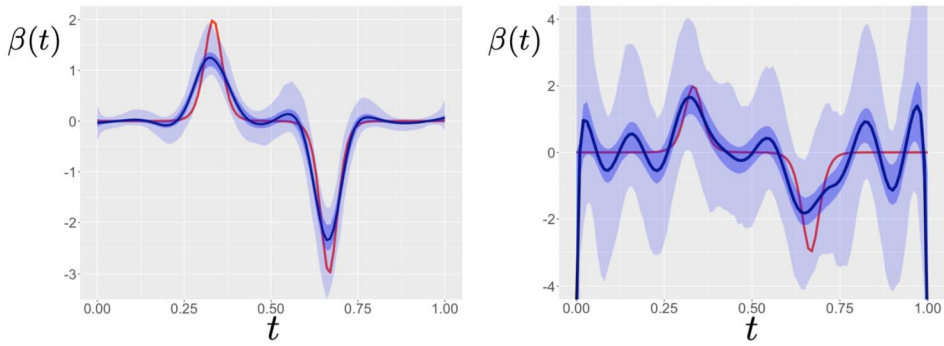

FIG. 2. *True curve* (*thin line*) *with the posterior mean* (*dark line*) *and* 50% (*dark shade*) *and* 95% (*light shade*) *credible intervals for the proposed method* (left) *and the local P-spline competitor* (right). *Both the* dynamic *and* shrinkage *properties of the proposed prior are necessary for accurate point estimation and precise uncertainty quantification.*

ation of (5)–(6) that instead uses independent and diffuse inverse-Gamma priors on $\{\lambda_k^2\}$. This *local P-spline* competitor includes local scale parameters but fails to provide either the aggressive shrinkage or the dynamics of the proposed approach. The posterior means and 50% and 95% pointwise credible intervals for β are presented in Figure 2. Clearly, this example is highly challenging: the nonlinear regression function includes both flat and rapidly-changing features and is not directly observable and must be inferred via the regression model (1). Most striking, the dynamic shrinkage provides better point estimation, especially in the smooth regions, and significantly more narrow interval estimates. By comparison, the local P-spline incorrectly estimates oscillations that are not present in the true function and produces credible intervals that are far too wide to be useful in practice.

The proposed *Bayesian adaptive scalar-on-function regression* (BASOFR) model (1) and (4)–(7) is completed by specifying priors on the remaining parameters in (1) and (7). By default, we assume the diffuse priors $[\mu] \propto 1$ and $[\sigma^{-2}] \sim \text{Gamma}(0.01, 0.01)$. For the boundary coefficients in (6), we include the prior $[B_1^*, B_{K_B}^* \mid \lambda_0] \sim \mathcal{N}(0, \lambda_0^2)$ indepenently with $\lambda_0^{-2} \sim \text{Gamma}(0.01, 0.01)$, which guards against excessively wide interval estimates of β near the boundaries of $\mathcal{T}$. For the dynamic horseshoe parameters, we follow Kowal, Matteson and Ruppert (2019) and assume $[\exp(\mu_h/2)] \sim C^+(0, 1)$, which corresponds to the global scale parameter in the nondynamic horseshoe ($\phi = 0$) and $[(\phi + 1)/2] \sim \text{Beta}(10, 2)$, which encourages persistence in the log-volatility but maintains stationarity via $|\phi| < 1$.

Posterior inference under this model is available using an efficient Gibbs sampler that cycles through the basis coefficients $\{B_k^*\}$ in (5), the log-volatilities $\{h_k\}$ in (7), the autoregressive parameters $\{\mu_h, \phi\}$, and the variance component σ^2 in (1). The crucial features of our sampling algorithm are: (i) the full conditional distribution of $\{B_k^*\}$ is K_B-dimensional Gaussian, which can be sampled efficiently and used to update $\beta(t) = \sum_{k=1}^{K_B} B_k^* \psi_k(t)$ for any $t \in \mathcal{T}$, and (ii) the log-volatilities $\{h_k\}$ can be sampled using a fast $\mathcal{O}(K_B)$ algorithm that relies on Gaussian parameter expansions and banded precision matrices; the details are provided in the online Supplementary Material (Gao and Kowal (2024)).

To highlight the computational scalability, we compare the empirical computing time for BASOFR against BLISS (Grollemund et al. (2019)) in Figure 3. We vary the sample sizes n and use the true regression function from Figure 2. To ensure favorable conditions for BLISS, we fix the number of local levels at two so that the computation time for BIC model selection is not included; thus, these computing times *underestimate* the usual computational burden of BLISS. Nonetheless, it is clear that BLISS does not scale to even moderate sample sizes ($n > 1000$), while the proposed algorithm scales approximately linearly in n. Since our application features $n \approx 100{,}000$, such scalability is essential.

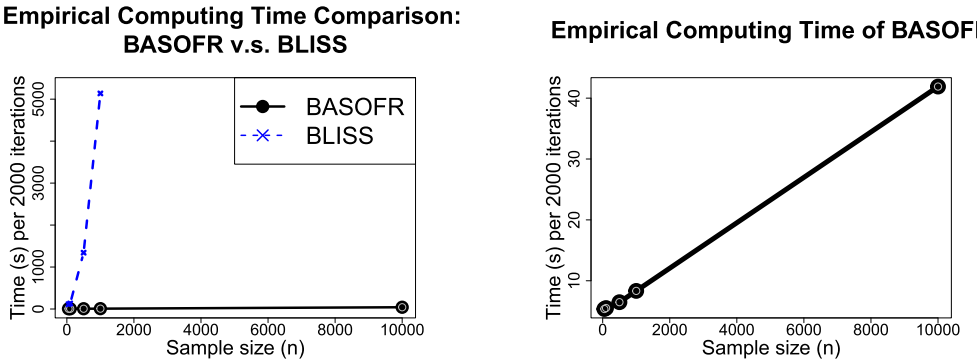

FIG. 3. *The empirical computational time for BASOFR and BLISS* (left) *and BASOFR* (right) *for* $n \in \{50, 100, 500, 1000, 10{,}000\}$ *in seconds per* 2000 *iterations* (*using R on a Windows PC*, 2.79 *GHz AMD EPYC-Rome*).

3. Decision analysis for window selection in SOFR. The BASOFR proposed in (1) and (4)–(7) offers several key benefits for Bayesian SOFR, including more adaptive point estimation, more precise uncertainty quantification, and scalable computing capabilities (see Section 4). Despite these advantages the posterior distribution of β alone cannot select critical windows of susceptibility: selection must be considered carefully and requires a decision analysis. Further, the local adaptivity induced by the prior does *not* resolve the interpretability issues noted in (2)–(3). This challenge permeates Bayesian SOFR and Bayesian inference more broadly: models may produce accurate data-generating processes, yet the parameters of those models do not necessarily offer the most convenient interpretations.

To address these challenges, we introduce *posterior summarization* tools for SOFR. Informally, the strategy adopts a two-stage approach: first, we estimate an adequate Bayesian SOFR model, such as BASOFR; second, we extract more interpretable point summaries from the model. These summaries are designed to replace more common point estimates, such as the posterior mean of $\beta(t)$, but customized to provide window selection and to mitigate the challenges observed in (2)–(3). Specifically, we target *locally constant* (or stepwise) point estimates of the form $\hat{\beta}(t) = \hat{\delta}_k$ for $t \in \mathcal{T}_k$ and $\{\mathcal{T}_k\}_{k=1}^K$ a partition of $\mathcal{T}$. Importantly, the key SOFR term simplifies

$$\int_{\mathcal{T}} X_i(t)\hat{\beta}(t)\,dt = \sum_{k=1}^{K} X_i(\mathcal{T}_k)\hat{\delta}_k, \tag{8}$$

where $X(\mathcal{T}_k) := \int_{\mathcal{T}_k} X(t)\,dt$ is the aggregated trajectory over $\mathcal{T}_k$. As a technical note, we define $X_i(t) = 0$ for $t \notin \mathcal{T}_i$, which allows the subject-specific domain in (1) to be expressed more conveniently using the maximal domain $\mathcal{T}$. The partition $\{\mathcal{T}_k\}_{k=1}^K$ provides window selection, that is, the regions of $\mathcal{T}$ that are most important for predicting y. Hence, it directly targets our goal of identifying the susceptible windows of $PM_{2.5}$ exposure during gestation, specifically linked to educational outcomes. These coefficients are also arguably more interpretable than $\beta(t)$: $\hat{\delta}_k$ estimates the change in the expectation of y for a one-unit increase in the *aggregated* trajectory $X(\mathcal{T}_k)$ while holding the remaining *aggregated* trajectories $\{X(\mathcal{T}_j)\}_{j\neq k}$ constant. As an added benefit, this representation significantly reduces storage requirements: point predictions can be computed using the aggregated trajectories $\{X(\mathcal{T}_k)\}_{k=1}^K$ instead of the entire trajectories $\{X(t)\}_{t\in\mathcal{T}}$.

To extract these summaries, we adopt a decision analysis approach and use the *acceptable families* of Kowal (2021) to compare partitions. Specifically, consider the following predictive

loss function for δ, omitting the scalar covariates z_i for now:

$$\tilde{L}_\lambda(\delta) = n^{-1}\sum_{i=1}^{n}\left\|(\tilde{y}_i-\mu)-\sum_{k=1}^{K}X_i(\mathcal{T}_k)\delta_k\right\|_2^2+\lambda\sum_{k=2}^{K}|\delta_k-\delta_{k-1}|, \tag{9}$$

where each $\tilde{y}_i$ is a posterior predictive variable at X_i under the BASOFR model. The loss function $\tilde{L}_\lambda(\delta)$ combines a "goodness-of-fit" component with an ℓ_1-penalty on the increments $\delta_k - \delta_{k-1}$ to encourage fewer change points in δ_k. The loss function may be further augmented with an ℓ_1-penalty on δ_k or other thresholding to encourage additional sparsity.

Since this loss inherits a posterior (predictive) distribution via $\{\tilde{y}_i\}_{i=1}^n$ and μ, Bayesian decision analysis proceeds by integrating over the posterior (predictive) distribution and minimizing the resulting quantity: $\hat{\delta}_\lambda := \arg\min_\delta \mathbb{E}_{[\tilde{y},\mu|y]}\tilde{L}_\lambda(\delta)$, which simplifies to

$$\hat{\delta}_\lambda = \arg\min_{\delta}\left\{n^{-1}\sum_{i=1}^{n}\left\|(\hat{y}_i-\hat{\mu})-\sum_{k=1}^{K}X_i(\mathcal{T}_k)\delta_k\right\|_2^2+\lambda\sum_{k=2}^{K}|\delta_k-\delta_{k-1}|\right\} \tag{10}$$

(assuming $\mathbb{E}_{[\tilde{y},\mu|y]}\|\tilde{y}_i-\mu\|^2<\infty$; Kowal (2021)), where $\hat{y}_i := \mathbb{E}_{[\tilde{y}_i|y]}\tilde{y}_i$ and $\hat{\mu} := \mathbb{E}_{[\mu|y]}\mu$ are posterior predictive expectations under the BASOFR model. Crucially, the *optimal Bayes action* $\hat{\delta}_\lambda$ is a "fit-to-the-fit" using pseudo-data $\hat{y}_i - \hat{\mu} = \int_{\mathcal{T}} X_i(t)\hat{\beta}(t)\,dt$ for $\hat{\beta}(t) = \mathbb{E}_{[\beta|y]}\beta(t)$ and covariates $\{X_i(\mathcal{T}_k)\}_{k=1}^K$. As such, $\hat{\delta}_\lambda$ seeks to simplify point estimation not by targeting $\hat{\beta}(t)$ directly but rather optimizing for the point predictions generated by $\hat{\beta}(t)$ under the SOFR model. Given these point predictions, the solution in (10) is readily computed using existing software, such as the `R` package `genlasso` (Tibshirani and Taylor (2011)).

The decision-analytic optimality of $\hat{\delta}_\lambda$ is valid only for a fixed λ, which controls the number of partitions (or steps) in the locally constant estimator $\{\hat{\delta}_k\}$. Hence, further comparisons are required across the path of λ values. We consider two metrics for each $\hat{\delta}$: the *empirical* mean squared error

$$\mathcal{E}_\lambda = \frac{1}{n}\sum_{i=1}^{n}\left\{(y_i-\hat{\mu})-\int_{\mathcal{T}_i}X_i(t)\hat{\delta}_\lambda(t)\,dt\right\}^2 \tag{11}$$

and the *predictive* mean squared error $\widetilde{\mathcal{E}}_\lambda := \tilde{L}_0(\hat{\delta}_\lambda)$ via (9), which replaces y_i with $\tilde{y}_i$ and $\hat{\mu}$ with μ in (11). Both metrics are important: $\mathcal{E}_\lambda$ provides an empirical point summary of the predictive accuracy, while $\widetilde{\mathcal{E}}_\lambda$ inherits a posterior predictive distribution under the BASOFR model via $\tilde{y}$ and μ.

The uncertainty quantification provided by the predictive version $\widetilde{\mathcal{E}}_\lambda$ is valuable for comparing across approximations of varying complexities λ. In particular, a primary drawback of the locally constant representation (8) is the potential for instability, that is, distinct partitions $\{\mathcal{T}_k\}$ and $\{\mathcal{T}'_k\}$ that produce similar predictive performance yet differ in their identification of the important windows of $\mathcal{T}$. This issue is not unique to our posterior summarization strategy but persists more broadly for estimators of the form (8). To address this instability, we leverage the uncertainty quantification from $\widetilde{\mathcal{E}}_\lambda$ to construct the *acceptable family* (Kowal (2021)), which collects the approximations $\hat{\delta}_\lambda$ that offer "near-optimal" predictive performance,

$$\mathcal{A}_\varepsilon := \{\lambda : \mathbb{P}_{\mathcal{M}}(\widetilde{\mathcal{D}}_\lambda < 0) \geq \varepsilon\}, \quad \varepsilon \in [0,1], \tag{12}$$

where $\widetilde{\mathcal{D}}_\lambda := 100\times(\widetilde{\mathcal{E}}_\lambda - \widetilde{\mathcal{E}}_{\lambda_{\min}})/\widetilde{\mathcal{E}}_{\lambda_{\min}}$ is the percent increase in predictive mean squared error relative to the empirical loss minimizer $\lambda_{\min} := \arg\min_\lambda \mathcal{E}_\lambda$. Informally, $\mathcal{A}_\varepsilon$ collects all approximations $\hat{\delta}_\lambda$ for which the predictive performance matches or exceeds that of $\hat{\delta}_{\lambda_{\min}}$ with at least ε probability under the BASOFR model $\mathcal{M}$. Equivalently, $\lambda \in \mathcal{A}_\varepsilon$ if and only if a lower $(1-\varepsilon)$ posterior prediction interval for $\widetilde{\mathcal{D}}_\lambda$ includes zero (Kowal (2021)). The acceptable family has been applied for targeted prediction (Kowal (2021)), variable selection

(Kowal et al. (2021)), subset selection (Kowal (2022a)), and selection in mixed effects models (Kowal (2022b)). By default, we select $\varepsilon = 0.10$; smaller values expand the acceptable family, but results are generally robust to moderate changes in ε (Kowal (2021), Kowal (2022a), Kowal et al. (2021)). We focus on the *simplest* member of the acceptable family, that is, the locally constant point estimate with the fewest changes in the local level (yet still satisfies (12)).

To incorporate the scalar covariates $\boldsymbol{z}_i$ in (1), we replace $(\tilde{y}_i - \mu)$ with $(\tilde{y}_i - \mu - \boldsymbol{z}_i'\boldsymbol{\alpha})$ in (9) and $\widetilde{\mathcal{E}}_\lambda$ and $(\hat{y}_i - \hat{\mu})$ with $(\hat{y}_i - \hat{\mu} - \boldsymbol{z}_i'\hat{\boldsymbol{\alpha}})$ for $\hat{\boldsymbol{\alpha}} := \mathbb{E}_{[\alpha|y]}\boldsymbol{\alpha}$ in (10) and (11). The resulting optimal point estimates $\hat{\delta}_\lambda$ now account for the scalar covariates $\boldsymbol{z}_i$, while the posterior predictive quantities (9) and $\widetilde{\mathcal{E}}_\lambda$ include the uncertainty due to the model parameters $\boldsymbol{\alpha}$.

4. Simulation study. We conduct two simulation studies: one that evaluates the BASOFR model against other SOFR models (Section 4.1) and one that assesses the decision analysis approach for window selection (Section 4.2), both using simulated datasets that resemble the NC data in our application study. Since the daily $PM_{2.5}$ trajectories are seasonal (see Section 5), we generate functional covariates $\{X_i\}_{i=1}^n$ with a seasonal pattern: each X_i follows a Gaussian process with mean function $\mu_i(t) = \sin(2\pi t/T + \phi_i)$ and covariance function $\mathrm{Cov}(X_i(t), X_i(t')) = \sigma_x^2 \exp\{-(t-t')^2/(2\ell^2)\}$. The period parameter T is fixed to induce an annual pattern ($T = 365/\text{maximal gestational length in days} = 365/295$), the offset $\phi_i \sim \mathrm{Unif}(0, 1)$ represents births at different times of year, and $\sigma_x = 0.7$ and $\ell = 0.01$ are chosen to visually resemble the $PM_{2.5}$ exposure curves. Each functional covariate is evaluated on a common and regular grid $(0, 0.01, 0.02, \ldots, 1)$. The online Supplementary Material includes results for smooth yet nonseasonal functional covariates, which is an easier setting for estimation and inference yet produces the same comparative results as those below (Gao and Kowal (2024)).

4.1. *BASOFR for point estimation and uncertainty quantification.* For a challenging estimation and inference scenario, we adopt the regression coefficient function in Figure 1, which presents both smooth and rapidly-changing features. The simulated datasets vary in the sample sizes and signal-to-noise ratios (SNR), with the SNR decreasing as n increases: $(n, \mathrm{SNR}) \in \{(50, 10), (100, 7), (500, 5), (10{,}000, 0.5)\}$. Using the aforementioned seasonal functional covariates, the response variables are simulated from (1) with $\mu = 0$ and σ determined based on the SNR, and the process is repeated to generate 50 datasets.

To compete with the BASOFR, we include BLISS (Grollemund et al. (2019)) and two Bayesian variations of the B-spline model (5)–(6). BLISS estimation follows the default recommendations to fit separate models with one to five levels and select the model with the lowest BIC. Due to the high computational cost (see Figure 3), we only include BLISS for $n \in \{50, 100\}$. Next, we modify (5)–(6) to include a global smoothness parameter $\lambda_k = \lambda$ and a diffuse inverse-Gamma prior on λ^2 (P-spline) and the local P-spline model from Figure 2 with independent and diffuse inverse-Gamma priors for each λ_k^2. Each model provides a point estimate of β via the posterior expectation and uncertainty quantification for β via 95% pointwise credible intervals. Unlike BLISS, the P-spline and local P-spline models *are* designed for high-dimensional data and are natural Bayesian competitors for SOFR.

Point estimation is evaluated using L_2-error (Figure 4) and uncertainty quantification is evaluated using mean credible interval widths and empirical coverage (Figure 5). Most notably, the proposed BASOFR model provides highly accurate point estimates and narrow interval estimates that achieve the nominal coverage, with the most substantial gains over competing methods occurring for larger sample sizes. By comparison, the P-spline and local P-spline intervals are far too conservative. Thus, neither global scale parameters nor local but independent scale parameters are sufficient for effective and adaptive inference: the *dependence* induced by the DHS prior in (6) is critical. Lastly, the narrow intervals provided by BLISS are far from achieving the nominal coverage and thus inadequate.

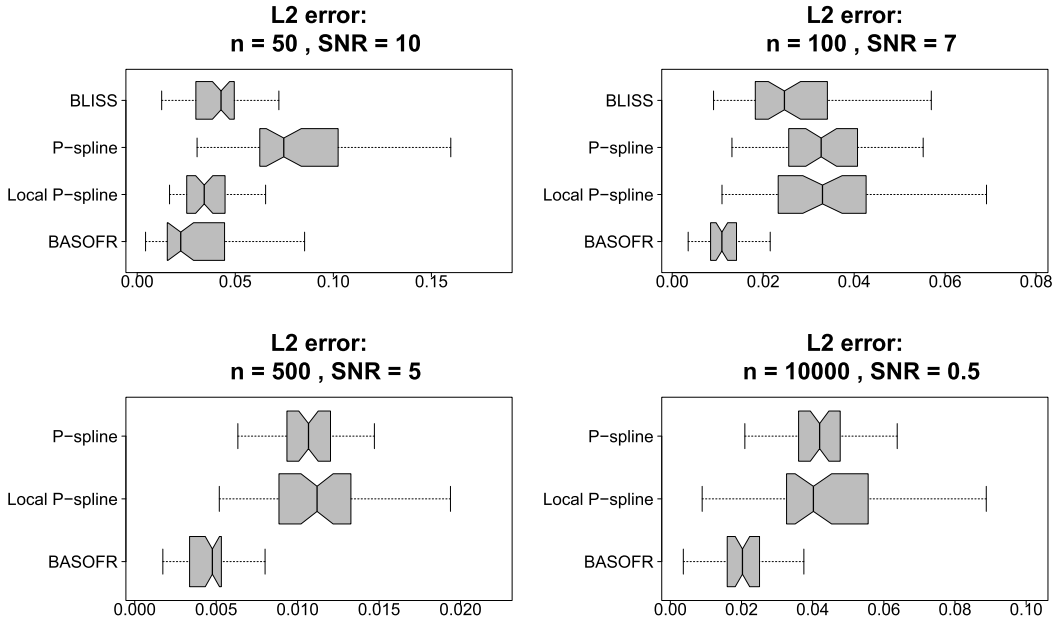

FIG. 4. *L_2-error for estimating the true regression function. BASOFR offers the most accurate point estimation with larger gains over competing methods as the sample size increases.*

4.2. *Decision analysis for selecting critical windows.* Next, we evaluate whether the proposed decision analysis approach (Section 3) is able to identify critical windows of susceptibility. We simulate 50 datasets from the SOFR model (1) with $n = 100{,}000$ observations with a low signal-to-noise ratio (SNR $= 0.5$) using a locally constant function for the true regression coefficient (see Figure 6) and the seasonal functional covariates $X_i(t)$. This design

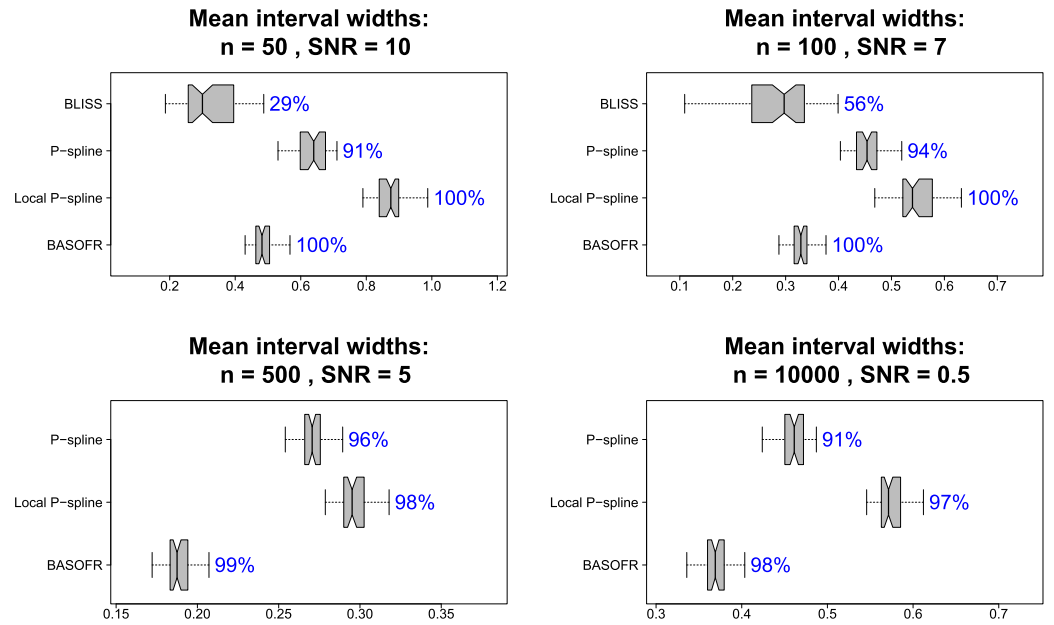

FIG. 5. *Mean credible interval widths (boxplots) and empirical pointwise coverage (blue annotations) for the 95% credible intervals computed under each model. BASOFR offers substantially more precise uncertainty quantification that maintains the nominal coverage.*

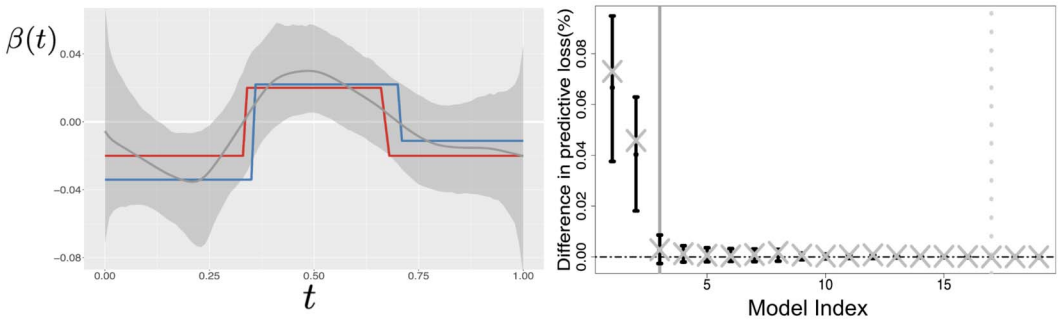

FIG. 6. Left: *True regression coefficient function* (*red*) *with the posterior mean* (*gray line*), 95% *credible interval* (*shade*), *and locally constant estimate* (*blue*) *for* β *under BASOFR for simulated data*. Right: *For each* λ *in the solution path of* (10), *expectations and* 80% *credible intervals for the percent difference in predictive mean squared error* $\widetilde{\mathcal{D}}_\lambda$ (*black lines*) *and the analogous empirical version* (*x-marks*) *based on* (11). *The vertical lines denote* $\lambda_{\min}$ (*dotted gray*) *and the simplest acceptable* λ (*solid gray*).

is constructed to mimic the output from the real data analysis (see Section 5, Figure 8) and especially the large sample size and low SNR. After fitting the BASOFR model, we apply the proposed decision analysis and select $\hat{\delta}_\lambda$ to be the simplest member of the acceptable family $\mathcal{A}_{0.1}$, that is, the locally constant approximation with the fewest changes in the local level.

First, we summarize the results on a single simulated dataset in Figure 6. Most notably, the posterior mean and credible intervals for β—as summaries of the BASOFR posterior—offer limited ability to describe the true regression function, both in terms of shape and effect direction. However, the decision analysis approach—which is based on the same BASOFR posterior—adequately recovers the truth. Figure 6 (right panel) also shows that other locally constant estimates are equally competitive, but the selected version is the simplest.

Next, we simulate 50 datasets from the same design and use the same two-stage procedure to fit the BASOFR and extract $\hat{\delta}_\lambda$. As a competing method, we select (positive or negative) windows based on whether the 95% posterior credible intervals for β exclude zero. This intuitive and popular strategy is based on the same BASOFR that is used in the decision analysis approach and thus differs only in the selection criteria.

To evaluate these approaches, we compute the true positive (TPR) and true negative (TNR) rates, defined here to be the correct detection of a truly positive (respectively, negative) window for the regression coefficient function β (Figure 7). Most notably, the proposed decision analysis approach is significantly better at selecting the critical windows of susceptibility, while selection based on credible intervals of β is far too conservative and thus underpowered. We emphasize that this important result applies for the BASOFR posterior credible intervals, which are substantially tighter (with the correct coverage) than competing interval estimates (Figure 5). Thus, alternative Bayesian SOFR models with less precise (wider) interval estimates would offer even less power to select these critical windows. We also compute the L_2-error of the point estimates from $\hat{\delta}_\lambda$ and $\hat{\beta}$, which confirms that the locally constant estimator does not sacrifice point estimation accuracy compared to the posterior mean.

The key takeaway from Figures 6–7 is that, despite using a smooth B-spline basis for β, the BASOFR posterior distribution *does* contain enough information to estimate a (true) locally constant regression function,but requires a careful decision analysis (Section 3)—rather than traditional posterior summaries (posteriors means and credible intervals)—to access it.

5. Prenatal $PM_{2.5}$ exposure and educational outcomes. Prenatal exposure to air pollution is related to a wide range of adverse birth, health, and behavioral outcomes in children. These outcomes include smaller fetal growth measurements (Leung et al. (2022)), low birth weight (Kloog et al. (2012), Srám et al. (2005)), infant negative affectivity (Rahman et al.

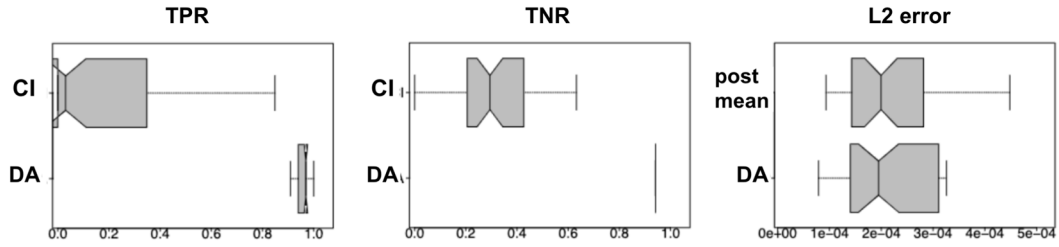

FIG. 7. *TPRs* (left) *and TNRs* (center) *show significantly better window selection for the proposed decision analysis* (*DA*) *approach, compared to credible interval-based selection* (*CI*)*, with no apparent loss in point estimation accuracy compared to the posterior mean* (right).

(2021)), and childhood asthma (Hazlehurst et al. (2021), Leon Hsu et al. (2015)), among many others. However, the link between air pollution exposure during pregnancy and later educational outcomes is less clear. Previous studies have demonstrated the adverse effects of prenatal air pollution exposure on neuropsychological development (Suades-González et al. (2015)), brain structure (Guxens et al. (2018)), memory function and attention (Chiu et al. (2016)), and autism diagnosis (Kalkbrenner et al. (2015)). These studies suggest a plausible link between prenatal air pollution exposure and adverse educational outcomes, but they do not directly consider educational outcomes or estimate the critical windows of susceptibility during pregnancy.

To address these limitations, we apply the BASOFR model and accompanying decision analysis to study the effects of prenatal exposure to $PM_{2.5}$ on educational outcomes. Specifically, we deploy the SOFR model (1) for standardized fourth EOG reading score $y_i \in \mathbb{R}$, $PM_{2.5}$ exposure during gestation $X_i : \mathcal{T}_i \to \mathbb{R}$, and other scalar covariates $\boldsymbol{z}_i \in \mathbb{R}^p$ (see Table 1) for a large cohort of mother-child pairs $i = 1, \ldots, n$ in NC.

The use of model (1) requires careful consideration of the domains $\mathcal{T}_i$ and the scalar covariates $\boldsymbol{z}_i$. First, each domain is subject-specific: $\mathcal{T}_i = [1, T_i]$, where T_i is the number of days in the gestational period for mother-child pair i. The gestation lengths range from 30 to 42 weeks, so the maximal domain is $\mathcal{T} = [1, T_{\max}]$, where $T_{\max} = 295$ days is the longest gestational period in the dataset. The B-spline basis is defined on this interval.

Next, the covariates $\boldsymbol{z}_i$ are given in Table 1 (with the exception of `Reading_Score` and `PrenatalPM2.5`). Each continuous covariate is centered and scaled, and the categorical variables are encoded using dummy variables. We modify (1) to include nonlinear additive effects for mother's age (`mAge`), length of gestation T_i (`Gestation`), and age-within-cohort (`Age_w_cohort`). The coefficients corresponding to both the linear and the nonlinear effects of the covariates $\boldsymbol{z}_i$ (except `Age_w_cohort`, see below) are assigned a hierarchical prior, $[\alpha_j \mid \sigma_j] \sim \mathcal{N}(0, \sigma_j^2)$, with $\sigma_j^{-2} \sim \text{Gamma}(0.01, 0.01)$ to encourages shrinkage and guard against the effects of multicollinearity among the correlated covariates $\boldsymbol{z}_i$ (see the online Supplementary Material; Gao and Kowal (2024)).

We motivate the additive terms as follows. Research has consistently shown that both teenage pregnancies and pregnancies among older mothers as well as deviations from an optimal gestational duration, whether shorter or longer, are linked with heightened health risks for the child (Chen et al. (2007), Goisis et al. (2017), Goldenberg et al. (2008), Olesen, Westergaard and Olsen (2003)). Since these health risks may also extend to the child's educational outcomes, we seek to account for both maternal age (`mAge`) and length of gestation T_i (`Gestation`) using more flexible additive models. To achieve this, we use piecewise continuous linear splines with knots carefully selected at maternal ages 18, 24, 29, and 34 and at weeks gestation 34, 37, 39, and 41. These knots were chosen to reflect key life stages for maternal age: adolescence (≤ 18), early adulthood (18–24), late twenties (24–29), early thirties (29–34), and advanced maternal age (≥ 34), as it is widely recognized that these life stages

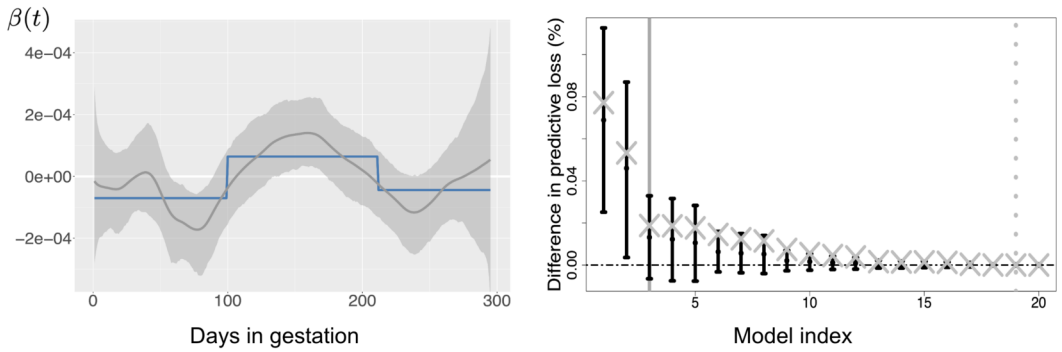

FIG. 8. Left: *Posterior mean* (*gray line*), 95% *credible interval* (*shade*), *and locally constant estimate* (*blue*) *for* β *under BASOFR for the NC data*. Right: *Expectations and* 80% *credible intervals for the percent difference in predictive mean squared error* $\widetilde{\mathcal{D}}_\lambda$ (*black lines*) *and the analogous empirical version* (*x-marks*) *based on* (11) *for each* λ *in the solution path of* (10). *The vertical lines denote* $\lambda_{\min}$ (*dotted gray*) *and the simplest acceptable* λ (*solid gray*).

can correspond to varied health and social circumstances (Gogoi (2014), Huang et al. (2023), Mirowsky and Ross (2002)). Similarly, for gestational length, the knots approximately delineate the transitions between early preterm ($\leq$ 34 weeks), late preterm (34–37 weeks), early term (37–39 weeks), full term (39–41 weeks), and post-term ($\geq$ 41 weeks) births, each stage associated with different health risks for the child (Boyle et al. (2012), Zhang and Kramer (2009)).

For age-within-cohort, we anticipate that older students may perform better on their standardized tests, but only up to a point: students who are more than one year older than their classmates may have repeated a grade or enrolled in kindergarten later for developmental reasons. Thus, we model `Age_w_cohort` as a nonlinear effect and, in particular, use the proposed adaptive B-spline model with dynamic shrinkage processes (5)–(7) for this term (as well as β in (1)). This specification encourages the nonlinear effect of `Age_w_cohort` to be smooth but can capture rapid changes, such as those expected around 52 weeks. Additional details and summary statistics for `Age_w_cohort` are in the online Supplementary Material (Gao and Kowal (2024)).

Posterior inference from the BASOFR model is based on 10,000 draws from the Gibbs sampler (after discarding a burn-in of 10,000). Traceplots show no lack of convergence, and effective sample sizes are sufficiently large.

First, we summarize our inference on the regression coefficient function β in Figure 8, which includes traditional posterior summaries of β (posterior means and 95% credible intervals) along with the proposed locally constant point estimate. We select $\hat{\delta}_\lambda$ to be the simplest member (i.e., the locally constant estimate with the fewest changes in the local level) of the acceptable family $\mathcal{A}_{0.1}$. Figure 8 (right panel) justifies this choice: the simplest member of the acceptable family indeed provides near-optimal prediction, compared to the other point estimators along the solution path of (10). Notably, the locally constant estimator substantially simplifies the shape of β and selects the critical windows of susceptibility. We refer the three locally constant regions in $\hat{\delta}_\lambda$ as R1, R2 and R3, which are similar but not identical to trimesters one, two, and three, respectively.

Despite the simplifications offered by the decision analysis, the interpretation of the regression coefficient function estimates requires some care. At first glance, Figure 8 suggests that $PM_{2.5}$ exposure is detrimental in R1 and R3 yet *favorable* in R2. Such a contradictory effect seems implausible. To investigate this outcome, we compute the estimated cumulative effect of exposure to $PM_{2.5}$, $\int_{\mathcal{T}_i} X_i(t)\hat{\beta}(t)\,dt$, for each mother-child pair $i = 1, \ldots, n$ (Figure 9), using both the posterior mean and the locally constant point estimator for β. The cumulative

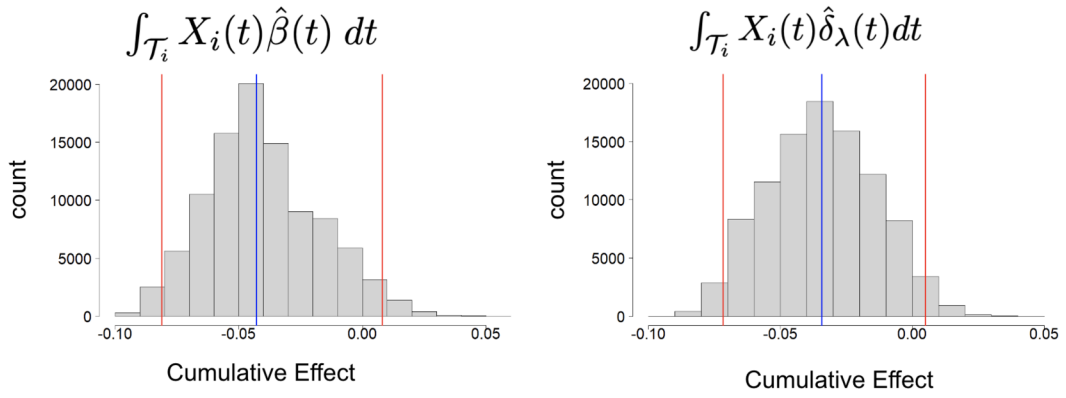

FIG. 9. *The estimated expected cumulative effects of* $PM_{2.5}$ *exposure during pregnancy across mother-child pairs* $i = 1, \ldots, n$ *using the posterior mean* $\hat{\beta}$ (left) *and* $\hat{\delta}_\lambda$ (right). *The sample median* (*blue*) *and* 95% *intervals* (*via sample quantiles*; *red*) *are annotated. The estimated cumulative effects agree and are significantly negative for nearly all mother-child pairs.*

effect of $PM_{2.5}$ exposure during gestation is *significantly negative* for nearly all mother-child pairs. Thus, as anticipated, exposure to $PM_{2.5}$ during gestation is negatively associated with fourth EOG reading scores (adjusting for z_i).

We further investigate *which* students were assigned positive or negative estimated cumulative effects $\int_{\mathcal{T}_i} X_i(t)\hat{\beta}(t)\,dt$ (Figure 10). Notably, the students for which this effect is (unexpectedly) positive were born almost exclusively in October–December. This result corresponds to a seasonal pattern in daily $PM_{2.5}$ exposure, which is further confirmed in Figure 10 (right panel): the birth month determines the average $PM_{2.5}$ exposure over each region R1, R2, and R3, with October–December corresponding uniquely to high exposures during R2 but low exposures during R1 and R3. Thus, the positive estimate in R2 is confounded by low exposures during R1 and R3 (Figure 8). We emphasize that seasonality is already included in the model via z_i: both birth month and age-within-cohort are included as nonlinear effects and capture overlapping yet mutually important notions of seasonality.

Our cumulative analysis—the estimated effects and windows selected (Figure 8), the overwhelmingly negative cumulative effects across mother-child pairs (Figure 9), and the seasonality patterns (Figure 10)—leads us to conclude that R1 and R3 represent the critical windows of susceptibility that are adversely associated with fourth EOG reading scores.

It is biologically plausible that $PM_{2.5}$ exposure has adverse effects on fetal development during early (R1) and late (R3) pregnancy. This may be attributed to the distinctive developmental milestones that occur in these stages. R1, roughly the first trimester, is a period of rapid cellular differentiation and organogenesis, with vital organs such as the heart, brain, and spinal cord beginning to form, while R3, roughly the third trimester, is characterized by rapid growth and maturation of organ systems, notably the central nervous system and the lungs (e.g., Sadler (2022)). Thus, R1 and R3 may be particularly susceptible to the adverse effects of harmful toxins. Additionally, our findings are consistent with other studies: Xiong et al. (2019) and Cheng et al. (2023) provide evidence of the unfavorable effects of air pollutant exposure during the first and third trimesters on birth defects. This research, coupled with studies suggesting that $PM_{2.5}$ exposure in the first and third trimesters are linked to lower intelligence quotient (IQ) as well as impaired behavioral and neuropsychological development in children (Chiu et al. (2016), Kalkbrenner et al. (2015), McGuinn et al. (2020), Sun et al. (2023)), underscores the significance of our results.

Lastly, we summarize the posterior inference for the scalar covariates z_i. Among the linear effects (Table 3), we find that lower fourth EOG reading scores are associated with

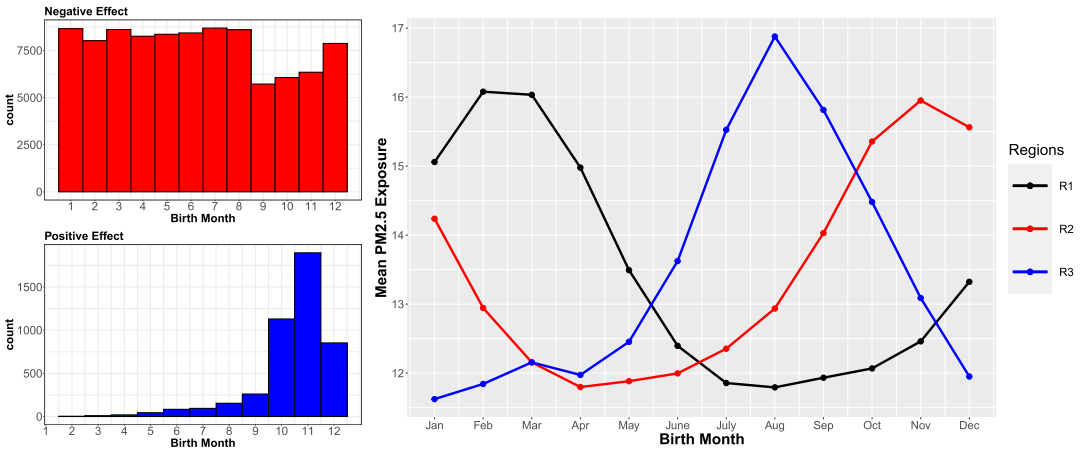

FIG. 10. *Distribution of the birth month for mother-child pairs with a negative* (top left) *or positive* (bottom left) *estimated expected cumulative effect of* $PM_{2.5}$ *exposure during pregnancy, along with the average* $PM_{2.5}$ *exposures over the windows identified by* $\hat{\delta}_\lambda$ (*see Figure* 9). *The positive effects represent a small fraction of mother-child pairs with birth months almost exclusively in October–December, which corresponds to high exposures during* $R2$ *but low exposures during* $R1$ *and* $R3$. *This seasonality is vital for interpreting* β.

lower mother's education level, presence of economic disadvantages, higher blood lead levels, smoking, and race/ethnicity and gender. The nonlinear effects are presented in Figure 11. Mother's age is positively associated with higher fourth EOG reading scores after age 24. The negative association prior to age 24 is perhaps explained by strong correlations between this younger age group and (lower) mother's education levels (see the online Supplementary Material; Gao and Kowal (2024)), which is already strongly associated with y_i. Gestational length is positively associated with fourth EOG reading scores until about 41 weeks, at which point the pregnancy is considered late term and accompanied by other health complications. Finally, age-within-cohort and birth month are highly correlated, and thus these effects must be interpreted jointly. The larger effects for birth months October–March are likely explained, in part, because those students are typically older within their cohort, which further explains why the age-within-cohort effect has only a small positive slope prior to week 52. However, for students at least one year older than their cohort, the age-within-cohort effect is significantly negative and includes rapid changes in the regression function, which justifies the choice of the adaptive B-spline model with dynamic shrinkage processes (5)–(7).

6. Conclusion. We developed a fully Bayesian modeling, computational, and decision analysis framework to study the effects of prenatal exposure to air pollution on educational outcomes. The proposed Bayesian adaptive scalar-on-function regression (BASOFR) model was designed to capture both smooth and abrupt changes in the association function, provide

TABLE 3

Posterior means and 95% *credible intervals for the scalar regression coefficients* $\boldsymbol{\alpha}$. *Intervals that exclude zero are annotated* (**)

Covariate	Regression coefficient estimates	Covariate	Regression coefficient estimates
noHS	−0.14 (−0.16, −0.13)**	Male	−0.13 (−0.14, −0.12)**
higherHS	0.28 (0.26, 0.29)**	EconDisadvantage	−0.27 (−0.29, −0.26)*
NH Black	−0.49 (−0.51, −0.48)**	Smoker	−0.07 (−0.09, −0.06)**
Hispanic	−0.06 (−0.08, −0.04)**	Blood_level	−0.028 (−0.033, 0.022)**

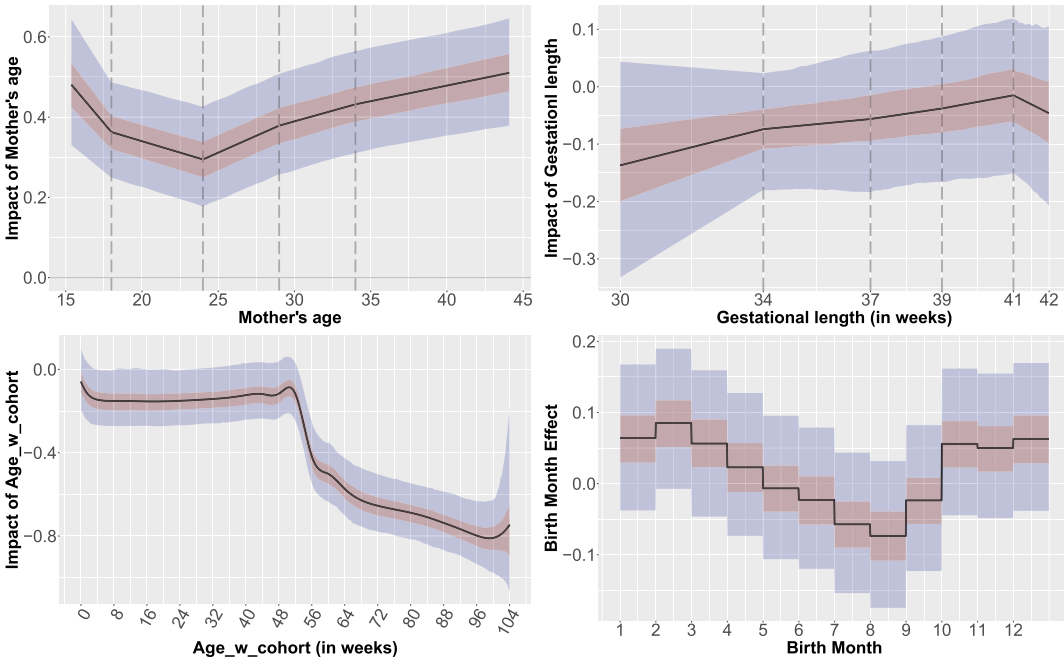

FIG. 11. *The posterior mean* (*line*) *and* 50% (*red*) *and* 95% (*blue*) *credible intervals for the nonlinear effects of mother's age at birth* (top left), *gestational length* (top right), *age-within-cohort* (bottom left), *and birth month* (bottom right) *on reading scores*.

full posterior uncertainty quantification, and maintain computational scalability in both the sample size and the number of observation points for each functional covariate. To complement the BASOFR model, we developed a decision analysis approach that produces locally constant point estimates of the regression coefficient function, which simultaneously: (i) extracts the critical windows of the functional domain, (ii) partially resolves fundamental interpretability issues for SOFR, and (iii) only adds minimal computational cost. Simulation studies demonstrated substantial advantages of the proposed approach for point estimation, uncertainty quantification, window selection, and computational scalability.

Using the proposed approach, we analyzed a large cohort ($n \approx 100{,}000$) of mother-child pairs in North Carolina to study the effects of $PM_{2.5}$ exposure during gestation on fourth end-of-grade reading scores. This analysis required careful consideration of the SOFR model output—including effect directions, cumulative effects, and seasonality—to identify two critical windows of susceptibility around trimesters one and three that correspond to adverse educational outcomes. Crucially, these results were enabled by our joint consideration of adaptive regression modeling, computational scalability, and interpretable posterior summaries via customized decision analysis. Our analysis included adjustments for important confounding variables and estimated nonlinear regression effects for mother's age, gestation length, birth month, and age-within-cohort.

We note that the estimated regression coefficient function (Figure 8), while complex, does not exhibit the same rapidly-changing features as in other examples (Figures 2, 6, and 11). However, the use of BASOFR remains justified: (i) the function shape was not known in advance, so the more flexible modeling capability is reassuring, and (ii) our simulation studies decisively showed that BASOFR delivers more accurate estimates and more precise and well-calibrated uncertainty quantification than competing methods. Regardless, the proposed Bayesian specification for adaptive regression coefficient functions was highly useful for inferring the nonlinear age-within-cohort effect, which exhibited both smooth and rapidly-

changing features (Figure 11). Thus, our adaptive Bayesian approach is useful not only in SOFR but also in function estimation and additive models more broadly.

The proposed functional regression model and accompanying decision analysis strategies offer promising extensions. First, these methods may be generalized for multiple functional predictors as well as functional response variables in function-on-function regression. Second, our decision analysis approach is broadly applicable for Bayesian SOFR and related distributed lag models, and thus is a useful addition to these models for more powerful window selection and interpretable model summaries. Furthermore, the decision analysis strategy may be altered to induce other structured point estimates, such as sparse or locally linear summaries, by varying the penalty term in (9). Lastly, our procedures may be applied to other datasets to estimate the critical windows of susceptibility for various exposures or interventions measured at high resolutions and paired with important outcomes of interest.

Acknowledgments. The authors would like to thank the review team for their constructive comments and Dr. Mercedes Bravo, Dr. Katherine B. Ensor, and Dr. Marie Lynn Miranda for their valuable insights.

Funding. Research was sponsored by the Army Research Office (W911NF-20-1-0184), the National Institute of Environmental Health Sciences of the National Institutes of Health (R01ES028819), and the National Science Foundation (SES-2214726). The content, views, and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office, the North Carolina Department of Health and Human Services, Division of Public Health, the National Institutes of Health, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation herein.

SUPPLEMENTARY MATERIAL

Supplement to "Bayesian adaptive and interpretable functional regression for exposure profiles" (DOI: 10.1214/23-AOAS1805SUPPA; .pdf). This supplement includes additional details about the computing, additional simulation results, and additional information and summary statistics for the North Carolina (NC) dataset.

R code (DOI: 10.1214/23-AOAS1805SUPPB; .zip). This section includes R code to reproduce the simulation analysis.

REFERENCES

BASHIR, A., CARVALHO, C. M., HAHN, P. R. and JONES, M. B. (2019). Post-processing posteriors over precision matrices to produce sparse graph estimates. *Bayesian Anal.* **14** 1075–1090.

BERGER, J. (1980). A robust generalized Bayes estimator and confidence region for a multivariate normal mean. *Ann. Statist.* 716–761.

BOSE, S., CHIU, Y.-H. M., HSU, H.-H. L., DI, Q., ROSA, M. J., LEE, A., KLOOG, I., WILSON, A., SCHWARTZ, J. et al. (2017). Prenatal nitrate exposure and childhood asthma. Influence of maternal prenatal stress and fetal sex. *Am. J. Respir. Crit. Care Med.* **196** 1396–1403.

BOYLE, E. M., POULSEN, G., FIELD, D. J., KURINCZUK, J. J., WOLKE, D., ALFIREVIC, Z. and QUIGLEY, M. A. (2012). Effects of gestational age at birth on health outcomes at 3 and 5 years of age: Population based cohort study. *BMJ* **344** e896. https://doi.org/10.1136/bmj.e896

BRAVO, M. A. and MIRANDA, M. L. (2021). Effects of accumulated environmental, social and host exposures on early childhood educational outcomes. *Environ. Res.* **198** 111241. https://doi.org/10.1016/j.envres.2021.111241

BROWN, P. J., VANNUCCI, M. and FEARN, T. (1998). Multivariate Bayesian variable selection and prediction. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 627–641.

CARDOT, H., FERRATY, F. and SARDA, P. (1999). Functional linear model. *Statist. Probab. Lett.* **45** 11–22.

CARVALHO, C. M., POLSON, N. G. and SCOTT, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika* **97** 465–480.

CHEN, X.-K., WEN, S. W., FLEMING, N., DEMISSIE, K., RHOADS, G. G. and WALKER, M. (2007). Teenage pregnancy and adverse birth outcomes: A large population based retrospective cohort study. *Int. J. Epidemiol.* **36** 368–373.

CHENG, Y., YIN, J., YANG, L., XU, M., LU, X., HUANG, W., DAI, G. and SUN, G. (2023). Ambient air pollutants in the first trimester of pregnancy and birth defects: An observational study. *BMJ Open* **13** e063712.

CHIU, Y.-H. M., HSU, H.-H. L., COULL, B. A., BELLINGER, D. C., KLOOG, I., SCHWARTZ, J., WRIGHT, R. O. and WRIGHT, R. J. (2016). Prenatal particulate air pollution and neurodevelopment in urban children: Examining sensitive windows and sex-specific associations. *Environ. Int.* **87** 56–65. https://doi.org/10.1016/j.envint.2015.11.010

DZIAK, J. J., COFFMAN, D. L., REIMHERR, M., PETROVICH, J., LI, R., SHIFFMAN, S. and SHIYKO, M. P. (2019). Scalar-on-function regression for predicting distal outcomes from intensively gathered longitudinal data: Interpretability for applied scientists. *Stat. Surv.* **13** 150.

FELDMAN, J. and KOWAL, D. R. (2022). Bayesian data synthesis and the utility-risk trade-off for mixed epidemiological data. *Ann. Appl. Stat.* **16** 2577–2602. https://doi.org/10.1214/22-AOAS1604

FIGUEIREDO, M. A. (2003). Adaptive sparseness for supervised learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **25** 1150–1159.

GAO, Y. and KOWAL, D. R. (2024). Supplement to "Bayesian adaptive and interpretable functional regression for exposure profiles." https://doi.org/10.1214/23-AOAS1805SUPPA, https://doi.org/10.1214/23-AOAS1805SUPPB

GOGOI, M. (2014). Association of maternal age and low socio-economic status of women on birth outcome. *Int. Res. J. Soc. Sci.* **3** 21–27.

GOISIS, A., REMES, H., BARCLAY, K., MARTIKAINEN, P. and MYRSKYLÄ, M. (2017). Advanced maternal age and the risk of low birth weight and preterm delivery: A within-family analysis using Finnish population registers. *Amer. J. Epidemiol.* **186** 1219–1226. https://doi.org/10.1093/aje/kwx177

GOLDENBERG, R. L., CULHANE, J. F., IAMS, J. D. and ROMERO, R. (2008). Epidemiology and causes of preterm birth. *Lancet* **371** 75–84.

GRIFFIN, J. E. and BROWN, P. J. (2005). Alternative prior distributions for variable selection with very many more variables than observations Technical Report Univ. Warwick, Centre for Research in Statistical Methodology.

GROLLEMUND, P.-M., ABRAHAM, C., BARAGATTI, M. and PUDLO, P. (2019). Bayesian functional linear regression with sparse step functions. *Bayesian Anal.* **14** 111–135.

GUXENS, M., LUBCZYŃSKA, M. J., MUETZEL, R. L., DALMAU-BUENO, A., JADDOE, V. W., HOEK, G., VAN DER LUGT, A., VERHULST, F. C., WHITE, T. et al. (2018). Air pollution exposure during fetal life, brain morphology, and cognitive function in school-age children. *Biol. Psychiatry* **84** 295–303.

HAHN, P. R. and CARVALHO, C. M. (2015). Decoupling shrinkage and selection in Bayesian linear models: A posterior summary perspective. *J. Amer. Statist. Assoc.* **110** 435–448.

HAZLEHURST, M. F., CARROLL, K. N., LOFTUS, C. T., SZPIRO, A. A., MOORE, P. E., KAUFMAN, J. D., KIRWA, K., LEWINN, K. Z., BUSH, N. R. et al. (2021). Maternal exposure to PM2. 5 during pregnancy and asthma risk in early childhood: Consideration of phases of fetal lung development. *Environ. Epidemiol.* **5**.

HUANG, C., JIANG, Q., SU, W., LV, F., ZENG, J., HUANG, P., LIU, W., LIN, M., LI, X. et al. (2023). Age-specific effects on adverse pregnancy outcomes vary by maternal characteristics: A population-based retrospective study in Xiamen, China. *BMC Public Health* **23** 326.

JAMES, G. M. (2002). Generalized linear models with functional predictors. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 411–432.

JAMES, G. M., WANG, J. and ZHU, J. (2009). Functional linear regression that's interpretable. *Ann. Statist.* **37** 2083–2108.

KALKBRENNER, A. E., WINDHAM, G. C., SERRE, M. L., AKITA, Y., WANG, X., HOFFMAN, K., THAYER, B. P. and DANIELS, J. L. (2015). Particulate matter exposure, prenatal and postnatal windows of susceptibility, and autism spectrum disorders. *Epidemiology* **26** 30–42.

KIM, S., SHEPHARD, N. and CHIB, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Rev. Econ. Stud.* **65** 361–393.

KLOOG, I., MELLY, S. J., RIDGWAY, W. L., COULL, B. A. and SCHWARTZ, J. (2012). Using new satellite based exposure methods to study the association between pregnancy PM2.5 exposure, premature birth and birth weight in Massachusetts. *Environ. Health* **11** 1–8.

KOWAL, D. R. (2021). Fast, optimal, and targeted predictions using parameterized decision analysis. *J. Amer. Statist. Assoc.* 1–12.

KOWAL, D. R. (2022a). Bayesian subset selection and variable importance for interpretable prediction and classification. *J. Mach. Learn. Res.* **23** 1–38.

KOWAL, D. R. (2022b). Subset selection for linear mixed models. *Biometrics* **00** 1–15.

KOWAL, D. R. and BOURGEOIS, D. C. (2020). Bayesian function-on-scalars regression for high-dimensional data. *J. Comput. Graph. Statist.* **29** 629–638.

KOWAL, D. R., BRAVO, M., LEONG, H., BUI, A., GRIFFIN, R. J., ENSOR, K. B. and MIRANDA, M. L. (2021). Bayesian variable selection for understanding mixtures in environmental exposures. *Stat. Med.* **40** 4850–4871. https://doi.org/10.1002/sim.9099

KOWAL, D. R., MATTESON, D. S. and RUPPERT, D. (2019). Dynamic shrinkage processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 781–804.

LEE, A., HSU, H.-H. L., CHIU, Y.-H. M., BOSE, S., ROSA, M. J., KLOOG, I., WILSON, A., SCHWARTZ, J., COHEN, S. et al. (2018). Prenatal fine particulate exposure and early childhood asthma: Effect of maternal stress and fetal sex. *J. Allergy Clin. Immunol.* **141** 1880–1886.

LEON HSU, H.-H., MATHILDA CHIU, Y.-H., COULL, B. A., KLOOG, I., SCHWARTZ, J., LEE, A., WRIGHT, R. O. and WRIGHT, R. J. (2015). Prenatal particulate air pollution and asthma onset in urban children. Identifying sensitive windows and sex differences. *Am. J. Respir. Crit. Care Med.* **192** 1052–1059.

LEUNG, M., WEISSKOPF, M. G., LADEN, F., COULL, B. A., MODEST, A. M., HACKER, M. R., WYLIE, B. J., WEI, Y., SCHWARTZ, J. et al. (2022). Exposure to PM 2.5 during pregnancy and fetal growth in eastern Massachusetts, USA. *Environ. Health Perspect.* **130** 017004.

MARX, B. D. and EILERS, P. H. C. (1999). Generalized linear regression on sampled signals and curves: A P-spline approach. *Technometrics* **41** 1–13.

MCGUINN, L. A., BELLINGER, D. C., COLICINO, E., COULL, B. A., JUST, A. C., KLOOG, I., OSORIO-VALENCIA, E., SCHNAAS, L., WRIGHT, R. J. et al. (2020). Prenatal PM2.5 exposure and behavioral development in children from Mexico City. *Neurotoxicology* **81** 109–115.

MIROWSKY, J. and ROSS, C. E. (2002). Depression, parenthood, and age at first birth. *Soc. Sci. Med.* **54** 1281–1298.

MORK, D. and WILSON, A. (2022). Treed distributed lag nonlinear models. *Biostatistics* **23** 754–771.

MORRIS, J. S. (2015). Functional regression. *Annu. Rev. Stat. Appl.* **2** 321–359.

MORRIS, J. S., BROWN, P. J., HERRICK, R. C., BAGGERLY, K. A. and COOMBES, K. R. (2008). Bayesian analysis of mass spectrometry proteomic data using wavelet-based functional mixed models. *Biometrics* **64** 479–489.

MORRIS, J. S. and CARROLL, R. J. (2006). Wavelet-based functional mixed models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 179–199. https://doi.org/10.1111/j.1467-9868.2006.00539.x

MÜLLER, H.-G. and STADTMÜLLER, U. (2005). Generalized functional linear models. *Ann. Statist.* **33** 774–805.

OLESEN, A. W., WESTERGAARD, J. G. and OLSEN, J. (2003). Perinatal and maternal complications related to postterm delivery: A national register-based study, 1978–1993. *Am. J. Obstet. Gynecol.* **189** 222–227. https://doi.org/10.1067/mob.2003.446

PICHENY, V., SERVIEN, R. and VILLA-VIALANEIX, N. (2019). Interpretable sparse SIR for functional data. *Stat. Comput.* **29** 255–267.

PUELZ, D., HAHN, P. R. and CARVALHO, C. M. (2017). Variable selection in seemingly unrelated regressions with random predictors. *Bayesian Anal.* **12** 969–989.

RAHMAN, F., COULL, B. A., CARROLL, K. N., WILSON, A., JUST, A. C., KLOOG, I., ZHANG, X., WRIGHT, R. J. and CHIU, Y.-H. M. (2021). Prenatal PM2. 5 exposure and infant temperament at age 6 months: Sensitive windows and sex-specific associations. *Environ. Res.* 112583.

RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis. Springer Series in Statistics*. Springer, New York, NY.

SADLER, T. W. (2022). *Langman's Medical Embryology*. Williams & Wilkins, Baltimore.

SCHWARTZ, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology* **11** 320–326. https://doi.org/10.1097/00001648-200005000-00016

SRÁM, R. J., BINKOVÁ, B., DEJMEK, J. and BOBAK, M. (2005). Ambient air pollution and pregnancy outcomes: A review of the literature. *Environ. Health Perspect.* **113** 375–382. https://doi.org/10.1289/ehp.6362

STRAWDERMAN, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Stat.* **42** 385–388.

SUADES-GONZÁLEZ, E., GASCON, M., GUXENS, M. and SUNYER, J. (2015). Air pollution and neuropsychological development: A review of the latest evidence. *Endocrinology* **156** 3473–3482. https://doi.org/10.1210/en.2015-1403

SUN, X., LIU, C., JI, H., LI, W., MIAO, M., YUAN, W., YUAN, Z., LIANG, H. and KAN, H. (2023). Prenatal exposure to ambient PM2. 5 and its chemical constituents and child intelligence quotient at 6 years of age. *Ecotoxicol. Environ. Saf.* **255** 114813.

TIBSHIRANI, R. J. and TAYLOR, J. (2011). The solution path of the generalized lasso. *Ann. Statist.* **39** 1335–1371.

WARREN, J., FUENTES, M., HERRING, A. and LANGLOIS, P. (2012). Spatial-temporal modeling of the association between air pollution exposure and preterm birth: Identifying critical windows of exposure. *Biometrics* **68** 1157–1167. https://doi.org/10.1111/j.1541-0420.2012.01774.x

WARREN, J. L., KONG, W., LUBEN, T. J. and CHANG, H. H. (2020). Critical window variable selection: Estimating the impact of air pollution on very preterm birth. *Biostatistics* **21** 790–806.

WILSON, A., CHIU, Y.-H. M., HSU, H.-H. L., WRIGHT, R. O., WRIGHT, R. J. and COULL, B. A. (2017). Bayesian distributed lag interaction models to identify perinatal windows of vulnerability in children's health. *Biostatistics* **18** 537–552.

XIONG, L., XU, Z., WANG, H., LIU, Z., XIE, D., WANG, A. and KONG, F. (2019). The association between ambient air pollution and birth defects in four cities in Hunan province, China, from 2014 to 2016. *Medicine* **98**.

ZHANG, X. and KRAMER, M. S. (2009). Variations in mortality and morbidity by gestational age among infants born at term. *J. Pediatr.* **154** 358–362.