Adaptive Bayesian sum of trees model for covariate-dependent spectral analysis

Yakun Wang1 | Zeda Li2 | Scott A. Bruce3

1Department of Statistics, George Mason University, Fairfax, Virginia, USA
2Paul H. Chook Department of Information System and Statistics, Baruch College, The City University of New York, New York, USA
3Department of Statistics, Texas A&M University, College Station, Texas, USA

Correspondence
Scott A. Bruce, Department of Statistics, Texas A, and M University, College Station, TX, USA.
Email: sabruce@tamu.edu

Funding information
National Institute of General Medical Sciences, Grant/Award Number: R01GM140476

Abstract
This paper introduces a flexible and adaptive nonparametric method for estimating the association between multiple covariates and power spectra of multiple time series. The proposed approach uses a Bayesian sum of trees model to capture complex dependencies and interactions between covariates and the power spectrum, which are often observed in studies of biomedical time series. Local power spectra corresponding to terminal nodes within trees are estimated nonparametrically using Bayesian penalized linear splines. The trees are considered to be random and fit using a Bayesian backfitting Markov chain Monte Carlo (MCMC) algorithm that sequentially considers tree modifications via reversible-jump MCMC techniques. For high-dimensional covariates, a sparsity-inducing Dirichlet hyperprior on tree splitting proportions is considered, which provides sparse estimation of covariate effects and efficient variable selection. By averaging over the posterior distribution of trees, the proposed method can recover both smooth and abrupt changes in the power spectrum across multiple covariates. Empirical performance is evaluated via simulations to demonstrate the proposed method’s ability to accurately recover complex relationships and interactions. The proposed methodology is used to study gait maturation in young children by evaluating age-related changes in power spectra of stride interval time series in the presence of other covariates.

KEYWORDS
Bayesian backfitting, gait variability, multiple time series, reversible jump Markov Chain Monte Carlo, Whittle likelihood

1 | INTRODUCTION

The frequency-domain properties of time series have often been found to contain valuable information. For example, frequency-domain analysis of biomedical time series, such as gait variability, heart rate variability (HRV), and electroencephalography (EEG), provides interpretable information about underlying physiological processes (Hausdorff et al., 1999; Hall et al., 2004; Klimesch, 1999). In many studies, biomedical time series are collected from multiple participants in conjunction with multiple covariates to explore connections between prominent oscillatory patterns in the time series and various clinical and behavioral outcomes. These relationships are often complex and highly interactive. As a result, a flexible, adaptive method that can estimate the association between power spectra and multiple covariates is needed to better understand the complex relationships between physiological processes and important measures of health and functioning.
A prime example and motivating application for this paper comes from a study of maturation in gait dynamics in young children (Hausdorff et al., 1999). Immature gait in very young children results in unsteady walking patterns and frequent falls (Shumway-Cook & Williams, 1995). While gait is relatively mature by age 3, neuromuscular control continues to develop well beyond this age (Preis et al., 1997). Accordingly, it is of interest to assess if gait dynamics continue to become more steady and regular beyond age 3, in conjunction with improving neuromuscular control. To assess gait variability and posture control in younger children, stride interval time series consisting of stride times during normal walking were observed in 50 children between the ages of 3 and 14 years (Hausdorff et al., 1999). For illustration, Figure 1 displays demeaned stride interval time series for three participants age 4, 7, and 11 years old. In addition to age, other covariates were also collected that may influence gait, such as gait speed and gender. In quantifying the association between age and the power spectra of stride interval time series, we seek to better understand the maturation of gait dynamics and variability associated with developing neuromuscular control with age in the presence of other related covariates.

In the time series literature, spectral analysis of multiple time series has received much attention in recent years. To quantify the association between a single covariate and power spectra, Fiecas and Ombao (2017) and Krafty et al. (2017) proposed methods that can capture a smooth covariate effect on the power spectrum. Bruce et al. (2018) proposed an adaptive Bayesian method that can capture both smooth and abrupt changes in power spectra across a covariate. Li et al. (2021) adapted the method of Bruce et al. (2018) for covariate-dependent spectral analysis of replicated multivariate time series. However, these methods are not readily extendable to incorporate multiple covariates, which hinders their applicability to many important studies. Existing methods that can account for multiple covariates are either parametric (Diggle & Al Wasel, 1997) or semi-parametric (Chau & von Sachs, 2016; Iannaccone & Coles, 2001; Qin et al., 2009; Stoffer et al., 2010). These approaches characterize covariate effects via design matrices within a linear modeling framework, and thus cannot immediately accommodate complex dependencies and interactions among covariates and power spectra. One exception is the approach proposed by Bertolacci et al. (2022), which introduces a mixture modeling approach with covariate-dependent mixture weights to account for complex covariate effects. However, a thin-plate Gaussian process prior is imposed on mixture weights, which is both smooth and stationary, and may not be appropriate for modeling abrupt spectral dynamics. Moreover, this method does not provide a means for variable selection when a large number of covariates are observed. The goal of this paper is to introduce a flexible method that can capture both smooth and abrupt changes in power spectra across multiple covariates, without loss of interpretation, and simultaneously provide a tool for variable selection.

To capture complex smooth, abrupt, and interaction effects of covariates on power spectra in a parsimonious manner, we propose a tree-based covariate partitioning framework. Tree-based models are not new and have become extremely popular in recent years (Breiman, 2001;
Chipman et al., 2010). For example, Chipman et al. (2010) proposed a Bayesian additive regression tree (BART) model which can flexibly model complex covariate effects and interactions and demonstrates outstanding predictive performance (Chipman et al., 2013). Accordingly, BART has been widely applied in many different scientific domains for various types of outcomes (Blattenberger & Fowles, 2017; van der Merwe, 2018; Waldmann, 2016) including smooth functional response variables (Starling et al., 2020).

In this paper, a sum of trees model for the covariate-dependent power spectrum is introduced to simultaneously partition multiple covariates parsimoniously. A penalized linear spline model is used for local spectrum estimation within terminal nodes of the trees. The framework is formulated in a fully Bayesian setting, where the trees are random and fit using an iterative Bayesian backfitting Markov chain Monte Carlo (MCMC) procedure and reversible-jump techniques (Green, 1995) to evaluate various tree modifications.

The proposed methodology expands the scope of covariate-dependent power spectra that can be accurately recovered in three meaningful ways. First, the flexible sum of trees model can recover complex nonlinear relationships and interaction effects without assuming a particular form of the relationship a priori. Second, by averaging over the posterior distribution of trees, the proposed method can recover both smooth and abrupt covariate effects on the power spectrum. While methods that assume completely smooth or completely abrupt changes will likely perform better when these assumptions are valid, the proposed method generally outperforms these methods in the presence of both smooth and abrupt covariate effects without prior knowledge of the nature of such effects. Third, the proposed method can automatically accommodate mixed-type covariates (nominal, ordinal, discrete, continuous) through the underlying tree structures, as well as high-dimensional covariates by placing a sparsity-inducing Dirichlet hyperprior on the splitting proportions of the regression tree prior (Linero, 2018) for sparse estimation of covariate effects and variable selection.

The rest of the paper is organized as follows. Section 2 provides a definition for the covariate-dependent power spectrum. The tree-based modeling framework for the power spectrum is introduced in Section 3, and Section 4 proposes the adaptive Bayesian sum of trees model and MCMC sampling scheme. Simulation results for various covariate effects (e.g., linear, nonlinear, smooth, abrupt, high-dimensional) and interaction effects are provided in Section 5. Section 6 presents the application to the study of gait maturation in young children. Conclusions and future directions of this work are covered in Section 7.

## 2 COVARIATE-DEPENDENT POWER SPECTRUM

We consider modeling a collection of stationary time series $X_{t\ell}$ of length $t = 1,...,T$ and $P$-dimensional covariates $\omega_\ell = (\omega_{p\ell},...,\omega_{pT})'$ for $p = 1,...,P$ mixed-type covariates and $\ell = 1,...,L$ independent subjects. To evaluate the effect of multiple covariates on the power spectrum, an extension of the Cramér representation (Cramér, 1942) is considered, which allows the power spectrum to vary across frequencies and covariates $X_{t\ell} = \int_{-1/2}^{1/2} A(\omega_\ell, \nu) \exp(2\pi i t \nu) dZ_\ell(\nu)$, where $A(\omega_\ell, \nu)$ is a Hermitian and periodic complex-valued function of frequency $\nu \in \mathbb{R}$ and covariates $\omega_\ell$ such that $A(\omega_\ell, \nu) = A(\omega_\ell, -\nu)$, $A(\omega_\ell, \nu) = A(\omega_\ell, \nu + 1)$, and $A(\omega_\ell, \nu)$ is square integrable over frequencies $[-1/2, 1/2]$. $Z_\ell$ are zero-mean mutually independent and identically distributed orthogonal processes with unit variance. Regularity conditions on the distribution of $Z_\ell$ must also be assumed to ensure subsequently introduced estimators are well-behaved; we assume cumulants of $dZ_\ell$ exist and are bounded for all orders (Brillinger, 2002).

The covariate-dependent power spectrum is then defined as $f(\omega, \nu) = |A(\omega, \nu)|^2$ and can be interpreted as the contribution to the variance at frequency $\nu$ given covariate values $\omega$. We assume that $A$, and subsequently the spectrum $f$, are continuous functions of frequency $\nu$, but can have a finite number of discontinuities as functions of covariates $\omega$. This flexibility allows for modeling abrupt changes over the covariate space and differs from models assuming entirely smooth changes across covariates (Fiecas & Ombao, 2017; Krafty et al., 2017).

## 3 TREE-BASED MODELING OF THE POWER SPECTRUM

### 3.1 Tree-based piecewise partition

Our method begins by assuming time series with similar covariate values have similar underlying power spectra and can be partitioned by a tree structure accordingly. It should be noted that methods have been proposed to approximate nonstationary time series through piecewise stationary time series, where time series are divided into approximately stationary intervals for time-dependent spectral analysis (Adak, 1998; Rosen et al., 2012). Our model is different from these approaches as it partitions the covariate space instead of time. More recently, Bruce et al. (2018) introduced a time- and covariate-based piecewise stationary approximation for time- and covariate-dependent spectral analysis using a two-dimensional grid.
However, directly extending a two-dimensional grid to higher dimensions to accommodate multiple covariates can easily lead to over-parameterization and undue computational complexity. In particular, a grid-based partition tends to produce a finer partition than is necessary, since partition points for each covariate do not depend on the other covariate. This results in less efficient information sharing across series with similar covariates and less accurate estimation. Tree-based approaches represent a more flexible and parsimonious alternative for partitioning multiple covariates. If the true partition does have a grid structure, a tree-based model can still well-approximate the partition and is thus preferable.

For the tree-based partition, each terminal node is defined through a collection of splitting rules corresponding to the tree structure and represents a region of the covariate space that shares the power spectrum. Given a tree $U$ with $B$ terminal nodes, the tree-based partition of the covariate-dependent Cramer spectral representation is given by $X_{\ell t} \approx f^{1/2}_{-1/2} \sum_{b=1}^{B} \delta(\omega_{\ell}; U, b) A_{b}(\nu) \exp(2\pi i t \nu) dZ_{\ell}(\nu)$, where $f_{b}(\nu) = |A_{b}(\nu)|^{2}$ is the power spectrum for the $b$th terminal node, and $\delta$ is a function that identifies terminal node membership for each observation based on covariates such that $\delta(\omega_{\ell}; U, b) = 1$ if the $t$th observation falls into the $b$th terminal node and $\delta(\omega_{\ell}; U, b) = 0$ otherwise.

### 3.2 Local power spectrum estimation

We now introduce an estimator for local power spectra within terminal nodes of the tree. Let $N = [T/2] - 1$ and $\nu_{k} = k/T$ for $k = 1, ..., N$ be the Fourier frequencies. The periodogram estimator of the power spectrum for the $t$th time series is $I_{\ell}(\nu_{k}) = \frac{1}{T} \sum_{t=1}^{T} X_{\ell t} \exp(-2\pi i t \nu_{k} T)^{2}$. The Whittle likelihood (Whittle, 1952), derived from the large sample distribution of the periodogram, can then be used to approximate the overall likelihood for the tree as a product of individual likelihoods, assuming $T$ is sufficiently large,

$$
L(I_{1}, ..., I_{L}|f_{1}, ..., f_{B}) \approx \prod_{\ell=1}^{L} \prod_{b=1}^{B} \left\{ (2\pi)^{-N/2} \sum_{k=1}^{N} \exp\{-\delta(\omega_{\ell}; U, b) \log f_{b}(\nu_{k}) \} \right. \\
+ \left. \exp(\log I_{\ell}(\nu_{k}) - \log f_{b}(\nu_{k})) \right\}^{2} 
$$

for $I_{\ell} = \{I_{\ell}(\nu_{1}), I_{\ell}(\nu_{2}), ..., I_{\ell}(\nu_{N})\}$. Log power spectra within each terminal node $\log f_{b}(\nu)$ are modeled using a Bayesian penalized linear spline model (Rosen et al., 2012)

$$
\log f_{b}(\nu) \approx \alpha_{b} + \sum_{s=1}^{S} \beta_{s}^{(b)} \cos(2\pi s \nu),
$$

where the functions $\cos(2\pi s \nu)$ are the Demmler–Reinsch basis functions for periodic even splines observed on an evenly spaced grid (i.e., the Fourier frequencies) (Schwarz & Krivobokova, 2016, Section 3). Only the first $S < N$ basis functions are used to provide a low-rank approximation to the full linear smoothing spline (Eubank, 1999). In order to achieve good computational efficiency without sacrificing estimation accuracy (Krafty et al., 2017), $S = 7$ basis functions are used for subsequent simulations and real-data analyses, which provide good empirical performance.

Gaussian priors are assumed such that $\alpha_{0} \sim N(0, \sigma_{a}^{2})$ where $\sigma_{a}^{2}$ is a constant value, and $\beta(\nu) = (\beta_{1}, ..., \beta_{S}) \sim N(0, \tau_{b}^{2} D_{S})$, where $D_{S} = \text{diag}\{\sqrt{2\pi} s\}^{-2}$. $\tau_{b}$ is a smoothing parameter that controls the roughness of the log spectrum. The scaling for the smoothing parameter, $\{\sqrt{2\pi} s\}^{-2}$, provides regularization of the integrated squared first derivative of the log power spectrum (Li & Krafty, 2019). A half-$t$ prior is placed on $\tau_{b}$ (Gelman, 2006) such that $p(\tau_{b}) \propto [1 + (\tau_{b}/\xi_{b})^{-2}]^{-1/2}$ for $\tau_{b} > 0$, where $\xi_{b}$ is scale and degree of freedom parameters respectively to complete the Bayesian model specification. A two-step MCMC sampling scheme for $\alpha_{b}$, $\beta(\nu)$, and $\tau_{b}$ following Rosen et al. (2012) is presented below.

(1) Let $Z_{b}$ be a $N \times S$ matrix of basis functions for $b$th terminal node such that $Z_{b} \sim \cos(2\pi s \nu_{k})$. Given $\tau_{b}$, basis functions $Z_{b}$, and periodogram ordinates $I_{\ell}$ for $\ell = 1, ..., L$, $\alpha_{0}$ and $\beta(\nu)$ are sampled jointly in a Metropolis–Hastings (M–H) step from

$$
p(\alpha_{b}, \beta(\nu)|\tau_{b}^{2}, I_{\ell}, Z_{b}) \propto \exp \left\{ -\sum_{\ell=1}^{L} \sum_{k=1}^{N} \left[ \alpha_{b} + Z_{b}^{\nu} \beta(\nu) \right] - \alpha_{b}^{2} - \frac{1}{\tau_{b}^{2}} \beta(\nu) D_{S}^{-1} \beta(\nu) \right\}. 
$$

(2) By representing the half-$t$ prior as a scale mixture of inverse gamma distributions (Wand et al., 2011), we efficiently obtain draws of $\tau_{b}$ from its full conditional distribution by sampling from

$$
(a_{b} | \tau_{b}^{2}) \sim IG\left(\frac{\xi_{b} + 1}{2}, \frac{\xi_{b} \tau_{b}^{2}}{2} + \frac{1}{A_{b}^{2}}\right), 
$$

where $A_{b}$ is the scale factor for the $b$th basis function.
and

\[(\tau^2_b | a_b, \beta^{(b)}) \sim IG(\xi_c + S + 1, \frac{\beta^{(b)} \rho^{(b)} + \xi_c}{a_b})\]  \hspace{1cm} (5)

where \(a_b\) is a latent variable, and \(\xi_c\) and \(A^2_c\) are fixed hyperparameters of the inverse gamma distribution.

### 4 ADAPTIVE BAYESIAN SUM OF TREES MODEL

#### 4.1 Sum of trees model

Poor mixing of single tree models has been noted in many applications such that the MCMC algorithm becomes stuck in subsets of the covariate space representing local optima and cannot efficiently traverse the entire parameter space (Wu et al., 2007). This can happen when single tree models grow very large in an effort to approximate more complex relationships, thus restricting possible modifications due to low-sample size and an abundance of other splits. We follow Chipman et al. (2010) in developing a Bayesian sum of trees formulation to relieve this problem by constructing many shallow trees as “weak learners” for estimation.

Let \(M\) be the number of trees. A sum of trees model for the log power spectrum is then constructed as \(\log f(\omega, \nu) \approx \sum_{j=1}^{M} \sum_{b=1}^{B_j} \delta(\omega; U_j, b) \log f_{b_j}(\nu)\), where \(U_j\) represents the \(j\)th tree that has \(B_j\) terminal nodes for \(j = 1, \ldots, M\). \(M\) should be sufficiently large to ensure adequate flexibility for recovering complex covariate effects. Model specification for local power spectra \(\log f_{b_j}(\nu)\) within each tree then follows directly from the specification for the single tree model introduced in Section 3.2.

Hyperparameters for priors introduced in Section 3.2 and tree structures should be selected to ensure a collection of “weak learners”. More specifically, estimates of \(\log f_{b_j}(\nu)\) should be shrunk toward zero by selecting hyperparameters \(\sigma^2_\nu\) and \(A_\tau\) sufficiently small, and tree depth should be kept small to produce shallow trees. Particular selections used in simulation studies and application are discussed in subsequent sections, and as recommended in Chipman et al. (2010), choices can also be made by cross-validation from a range of reasonable values.

#### 4.2 Prior specification

Let \(\Phi_j = \{\log f_{1j}(\nu), \ldots, \log f_{B_j}(\nu)\}\) be the collection of log power spectra across terminal nodes for the \(j\)th tree. To complete the Bayesian model specification, priors are imposed on \(U_j\) and \(\Phi_j\) in order to allow the trees to be random and fit from the data. Assuming independence across terminal node parameters and trees a priori, priors can be specified as \(p(U_1, \Phi_1, \ldots, (U_M, \Phi_M)) = \prod_j p(U_j, \Phi_j) = \prod_j p(\Phi_j | U_j) p(U_j)\), where \(p(\Phi_j | U_j) = \prod_b p(\log f_{b_j}(\nu) | U_j)\). The priors for \(\Phi_j | U_j\) then correspond to the priors of \(\alpha_b, \beta^{(b)}\) and \(\tau_b\) for the local power spectrum estimator introduced in Section 3.2. For the priors on the tree structure \(U_j\), three probabilities need to be considered.

1. The probability of a node to be split is defined as \(\Pr(\text{SPLIT}) = \gamma(1 + d)^{-\theta}\), where \(\gamma \in (0, 1)\), \(\theta \in [0, \infty)\) and \(d = 0, 1, \ldots\) is the depth of a given node. This prior is a regularization of the tree depth to encourage each tree to be shallow. We set \(\gamma = 0.95\) and \(\theta = 2\) following Chipman et al. (2010). Ročková and Saha (2019) proposed a minor modification \(\Pr(\text{SPLIT}) \propto \gamma^d\) for some \(0 \leq \gamma < 1/2\) to achieve the optimal posterior convergence rate, which can also be adopted within the proposed framework.

2. The probability of selecting the \(p\)th covariate for splitting is denoted as \(s_p\) for \(p = 1, \ldots, P\). The proposed model allows for two possible prior specifications: a uniform prior, \(s_p = P^{-1}\), such that all covariates have the same probability to be selected, and a sparsity-inducing Dirichlet prior \((s_1, \ldots, s_P) \sim D\left(\frac{c}{P}, \ldots, \frac{c}{P}\right)\) (Linero, 2018). For the Dirichlet prior, \(c\) determines the degree of sparsity and Linero (2018) offered multiple approaches for modeling this parameter. We set \(c = 1\) as suggested by Linero (2018) for computational convenience.

3. The probability of selecting a cutpoint for a given covariate is uniform. A uniform prior is also desirable as it is invariant for monotone transformations of the covariate (Chipman et al., 2010). For continuous, discrete, and ordinal covariates, cutpoints are selected from a fixed number of evenly spaced points over the range of possible values. For categorical covariates without an intrinsic ordering, a cutpoint represents a particular mapping of categories to the left and right child nodes created by the split. A categorical variable with \(q\) categories then has \(2^q - 2\) cutpoints that can be selected.

#### 4.3 Bayesian backfitting Markov chain Monte Carlo

It is important to note that each tree captures particular features of the covariate-dependent power spectrum and depends on the features captured by other trees. While this
5 | SIMULATION STUDIES

We consider three simulation settings representing abrupt and smoothly varying dynamics with complex covariate effects and interactions in order to demonstrate strong finite-sample estimation accuracy, as well as the ability to adapt to sparse covariate effects and conduct variable selection.

5.1 | Settings

(1) Abrupt+Smooth: let \( \omega = [\omega_1, \omega_2] \) where \( \omega_1, \omega_2 \overset{i.i.d.}{\sim} U(0,1) \). An AR(1) process for the \( \ell \)th time series is specified as \( x_{\ell t} = \phi_\ell x_{\ell t-1} + \varepsilon_{\ell t}, \varepsilon_{\ell t} \sim N(0,1) \), where \( \phi_\ell = -0.7 + 1.4\omega_2 \) when \( 0 \leq \omega_2 < 0.5 \) and \( \phi_\ell = 0.9 - 1.8\omega_2 \) when \( 0.5 \leq \omega_2 \leq 1 \).

(2) AR-Friedman: let \( \omega = [\omega_1, ..., \omega_5] \) where \( \omega_1, ..., \omega_5 \overset{i.i.d.}{\sim} U(0,1) \). An AR(1) process for the \( \ell \)th time series is specified as \( x_{\ell t} = \phi_\ell x_{\ell t-1} + \varepsilon_{\ell t}, \varepsilon_{\ell t} \sim N(0,1) \), where \( \phi_\ell = 0.5 \sin(\pi \omega_1 \omega_2) - (\omega_3 - 0.5)^2 + 0.35 \sin(\omega_4) - 0.5 - 1.5 \omega_5 \).

(3) Adjusted-AdaptSPEC-X: let \( \omega = [\omega_1, \omega_2] \) where \( \omega_1, \omega_2 \overset{i.i.d.}{\sim} U(0,1) \). Each covariate vector \( \omega \) is mapped to a latent variable \( z_\ell \in \{1, 2, 3, 4\} \). Figure 2(A) shows the mapping from \( \omega \) to \( z \). An AR(2) process is then specified as \( x_{\ell t} = \phi_{\ell 1} x_{\ell t-1} + \phi_{\ell 2} x_{\ell t-2} + \varepsilon_{\ell t}, \varepsilon_{\ell t} \sim N(0,1) \), where the coefficients for the four latent region are defined as follows. If \( z_\ell = 1, (\phi_{\ell 1,1}, \phi_{\ell 2,1}) = (1.5, -0.75); \) if \( z_\ell = 2, (\phi_{\ell 1,2}, \phi_{\ell 2,2}) = (-0.8, 0); \) if \( z_\ell = 3, (\phi_{\ell 1,3}, \phi_{\ell 2,3}) = (-1.5, -0.75); \) if \( z_\ell = 4, (\phi_{\ell 1,4}, \phi_{\ell 2,4}) = (0.2, 0) \).

The first setting represents an AR(1) process in which the coefficient varies smoothly across one covariate and abruptly across another. The second setting contains complex linear and nonlinear covariate effects and interactions (Friedman, 1991) adapted for time series data. The third setting represents an abruptly-changing process over two dimensions, similar to that of Bertolacci et al. (2022). Power spectra for covariate-dependent AR(1) processes from the first two settings can be represented as \( f(\omega_\ell, \nu) = [1 - 2\phi_\ell \cos(2\pi \nu) + \phi_\ell^2]^{-1}, \) where \( \omega_\ell \) represents the covariate vector for the \( \ell \)th time series and \( \phi_\ell \) depends on \( \omega_\ell \) as defined above. For the third setting, covariate-dependent AR(2) processes can similarly be represented as \( f(\omega_\ell, \nu) = [1 + 2(\phi_{\ell 1} + \phi_{\ell 2}) \cos(2\pi \nu) - 2\phi_{\ell 1} \phi_{\ell 2} \cos(4\pi \nu) + \phi_{\ell 1}^2 + \phi_{\ell 2}^2]^{-1} \).
for the prior on $\tau_b$ in Equations (4) and (5). We use both $M = 5$ trees and $M = 50$ trees for estimation, and different numbers ($L$) and lengths ($T$) of time series are considered. The MCMC procedure is run for a total of 10,000 iterations with the first 5,000 discarded as burn-in. In order to assess convergence, trace plots for summary measures of the mean-squared residuals, estimated log power spectrum, and tree structures for all settings are available in Web Appendix B. These diagnostics appear to show convergence after approximately 5,000 iterations across all settings.

Estimates of the covariate-dependent power spectrum for a single run are presented in Figure 3 to visually illustrate the ability of the proposed method to capture abrupt and smooth changes in power spectra. The smooth change in the conditional power spectrum over $\omega_2$ is captured by averaging over the posterior distribution of tree structures, which contains many splits across the range of possible values of $\omega_2$. On the other hand, splits on $\omega_1$ are concentrated around the true abrupt change at $\omega_1 = 0.5$, and the posterior mean estimator of the conditional power spectrum over $\omega_1$ appropriately reflects the abrupt change in the conditional power spectrum.

Given an estimate of the covariate-dependent log power spectrum, $\hat{f}(\omega, \nu)$, the MSE, $\text{MSE} = (NL)^{-1} \sum_{\ell=1}^{L} \sum_{k=1}^{N} \left[ \log \hat{f}(\omega_\ell, \nu_k) - \log f(\omega_\ell, \nu_k) \right]^2$, will be used to evaluate estimation accuracy. True covariate-dependent log power spectra, $\log f(\omega, \nu)$ for all settings can be fully determined by the representations and coefficients presented in Section 5.1. The mean and standard deviation of the mean squared error (MSE) is presented in Table 1 for 100 replications of all settings. For comparison, AdaptSPEC-X (Bertolacci et al., 2022) is also used to estimate the covariate-dependent power spectrum. It is important to note that AdaptSPEC-X allows for modeling of time- and covariate-varying power spectra with time-varying means, which is a more general setting than what is considered in this work. Accordingly, we implement a simplified version of the AdaptSPEC-X mixture model without the time-varying mean and power spectra components to enable more accurate comparisons. We applied $C = 50$ mixture components to ensure sufficient flexibility in estimating covariate effects on power spectra.

These results show that both methods see improved estimation accuracy as the number ($L$) and length ($T$) of time series increase. For the proposed method, using 50 trees provides slight improvements in estimation accuracy compared to using 5 trees when $L > 100$ and $T > 100$ since the additional trees provide more flexibility to capture complexities in the covariate-dependent power spectra. However, the proposed method using either 5 or 50 trees produces smaller MSEs than AdaptSPEC-X for comparable settings. This can be partially attributed to the presence of abrupt changes across one or more covariates.

**FIGURE 2** (A) Presents the mapping of covariate values $\omega_1$ and $\omega_2$ to latent variable values $z$ for the adjusted-AdaptSPEC-X simulation setting. D1 and D2 denote two simulated time series with similar covariate values that are mapped to different latent variable values corresponding to different power spectra. For these two realizations, (B) displays the true log power spectra (red lines), log periodogram ordinates (gray points), and estimated log power spectra using the proposed Bayesian sum of trees model (blue lines) and the AdaptSPEC-X model (green lines). This figure appears in color in the electronic version of this paper, and any mention of color refers to that version.
in all settings, which are better captured by the proposed tree-based approach. To illustrate this point, Figure 2(B) shows the estimated power spectra for two simulated time series with similar covariate values separated by an abrupt change from the Adjusted-AdaptSPEC-X setting. As noted in Bertolacci et al. (2022), the thin-plate Gaussian process prior on the mixture weights, while flexible, is both smooth and stationary. Accordingly, the AdaptSPEC-X power spectrum estimates for these two time series are similar due to smoothing across similar covariate values induced by the thin-plate Gaussian process prior. On the other hand, the proposed Bayesian sum-of-trees model is able to accurately distinguish the abrupt change in the power spectra. Moreover, averaging over the posterior distribution of trees enables the proposed method to recover smooth changes well (Figure 3). Taken together, this results in superior estimation accuracy across all settings seen in Table 1.

When the number of covariates is small, the proposed method can also be compared with a generalized additive model (GAM) approach (Wood, 2017). This is an important benchmark for comparison, since the GAM modeling approach is frequently used in practice, and its smooth additive structure is well-equipped to characterize smooth covariate effects. Additional simulations comparing the proposed method with a GAM approach in simulations are available in Web Appendix C. These results show that methods assuming completely smooth effects, like GAM, can outperform the proposed method when smoothness assumptions are valid and sufficient data are available. However, the proposed method generally outperforms these methods in the presence of both smooth and abrupt covariate effects without prior knowledge of the nature of such effects.

5.3 Results: computation time

Simulations were carried out on a Windows 10 machine with an 8-core Intel i7 3.6 GHz processor and 64 GB RAM using R version 4.0.3 (R Core Team, 2021). The R code for implementing the proposed model is provided as a zip file in Supporting Information and is described in Web Appendix F. Computationally-intensive aspects of the methodology are written in C++ using RcppArmadillo (Eddelbuettel & Sanderson, 2014) for more efficient computation and reduced run times. Replications were
TABLE 1  Mean and standard deviation of mean squared error (MSE) over 100 replications for three simulation settings with different lengths (T) and number (L) of time series.

| L   | T   | Abrupt+smooth | Proposed Bayesian sum of trees model (5 trees) | AR-Friedman | Adj. AdaptSPEC-X |
|-----|-----|---------------|-----------------------------------------------|-------------|------------------|
| 100 | 100 | 0.0406(0.0062)| 0.0386(0.0059)                              | 0.1475(0.0249)|                  |
| 100 | 250 | 0.0217(0.0031)| 0.0226(0.0030)                              | 0.0762(0.0102)|                  |
| 100 | 500 | 0.0138(0.0019)| 0.0153(0.0019)                              | 0.0528(0.0063)|                  |
| 200 | 100 | 0.0320(0.0052)| 0.0338(0.0042)                              | 0.1446(0.0202)|                  |
| 200 | 250 | 0.0173(0.0030)| 0.0197(0.0020)                              | 0.0719(0.0070)|                  |
| 200 | 500 | 0.0118(0.0023)| 0.0138(0.0015)                              | 0.0483(0.0050)|                  |
| 500 | 100 | 0.0238(0.0030)| 0.0276(0.0022)                              | 0.1348(0.0140)|                  |
| 500 | 250 | 0.0124(0.0015)| 0.0160(0.0013)                              | 0.0663(0.0077)|                  |
| 500 | 500 | 0.0085(0.0013)| 0.0119(0.0012)                              | 0.0452(0.0060)|                  |

| L   | T   | Proposed Bayesian sum of trees model (50 trees) | AdaptSPEC-X model |
|-----|-----|-----------------------------------------------|------------------|
| 100 | 100 | 0.0448(0.0070)                              | 0.0419(0.0048)   |
| 100 | 250 | 0.0195(0.0026)                              | 0.0202(0.0025)   |
| 100 | 500 | 0.0118(0.0015)                              | 0.0124(0.0013)   |
| 200 | 100 | 0.0309(0.0031)                              | 0.0312(0.0035)   |
| 200 | 250 | 0.0146(0.0017)                              | 0.0162(0.0016)   |
| 200 | 500 | 0.0093(0.0012)                              | 0.0104(0.0009)   |
| 500 | 100 | 0.0209(0.0022)                              | 0.0236(0.0020)   |
| 500 | 250 | 0.0103(0.0015)                              | 0.0120(0.0011)   |
| 500 | 500 | 0.0070(0.0013)                              | 0.0080(0.0005)   |
| 100 | 100 | 0.0576(0.0113)                              | 0.0529(0.0082)   |
| 100 | 250 | 0.0437(0.0108)                              | 0.0406(0.0073)   |
| 100 | 500 | 0.0412(0.0126)                              | 0.0377(0.0070)   |
| 200 | 100 | 0.0510(0.0087)                              | 0.0474(0.0058)   |
| 200 | 250 | 0.0398(0.0089)                              | 0.0372(0.0060)   |
| 200 | 500 | 0.0378(0.0093)                              | 0.0365(0.0061)   |
| 500 | 100 | 0.0466(0.0053)                              | 0.0398(0.0039)   |
| 500 | 250 | 0.0379(0.0051)                              | 0.0323(0.0049)   |
| 500 | 500 | 0.0370(0.0050)                              | 0.0319(0.0047)   |

Note: Results are presented for the proposed Bayesian sum of trees model with number of trees \( M = 5 \) (top), \( M = 50 \) (middle), and the AdaptSPEC-X model (bottom) with mixture components \( C = 50 \).
efficiency can be investigated by assessing the estimated posterior probability for model inclusion of each covariate, which is the proportion of posterior draws where the covariate appears in at least one split rule for at least one tree. Table 2 displays the MSE and posterior probabilities of model inclusion for the AR-Friedman setting using both the uniform and Dirichlet hyperpriors for tree splitting proportions introduced in Section 4.2 with $M = 50$ trees. Tables for all simulation settings for the two different hyperpriors using different numbers of trees ($M = 5, 50$) and visualizations of the posterior model inclusion probabilities are available in Web Appendix B.

From this table, some important trends should be noted. First, there is a slight improvement in estimation accuracy when using the Dirichlet hyperprior. This can be attributed to the proposed method’s ability to recover important variables with high posterior probability and the Dirichlet hyperprior’s superior ability to adapt to the sparse covariate vectors considered. This can be seen in Table 2 by comparing model inclusion posterior probabilities for important and noise variables. Important variables are included in models with high posterior probability ranging from 46.9% to 100% on average. However, the Dirichlet hyperprior shrinks posterior probabilities for noise variables closer to zero for all covariate vector sizes considered by 47%–62% on average compared to the uniform hyperprior. Second, both estimation accuracy and variable selection are negatively impacted by increasing covariate vector size $P$. This is expected since the additional noise covariates result in sparser covariate vectors, for which resolving important variables is relatively more difficult. Also, additional noise covariates inject additional randomness into the proposed model, which leads to implicit regularization (Mentch & Zhou, 2020). Compared with a random forest model for generic functional responses (Fu et al., 2021), the proposed method offers improved inclusion of important variables and exclusion of noise variables on average for the simulation settings considered herein, especially for larger covariate vector sizes. A table detailing variable selection performance for the generic functional random forest model for these simulation settings can be found in Web Appendix B.

### Table 2

| $P=100$ | Uniform | Dirichlet |
|---------|---------|-----------|
| MSE     | 0.0162(0.0019) | 0.0141(0.0016) |
| Noise   | 0.2681(0.3457) | 0.1287(0.3284) |
| $\omega_1$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_2$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_3$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_4$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_5$ | 1.0000(0.0000) | 1.0000(0.0000) |

| $P=200$ | Uniform | Dirichlet |
|---------|---------|-----------|
| MSE     | 0.0188(0.0030) | 0.0143(0.0017) |
| Noise   | 0.2098(0.3224) | 0.0787(0.2631) |
| $\omega_1$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_2$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_3$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_4$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_5$ | 0.9999(0.0002) | 0.9942(0.0612) |

| $P=1000$ | Uniform | Dirichlet |
|---------|---------|-----------|
| MSE     | 0.0285(0.0060) | 0.0221(0.0069) |
| Noise   | 0.1259(0.2908) | 0.0660(0.2319) |
| $\omega_1$ | 0.9584(0.1692) | 0.8859(0.3158) |
| $\omega_2$ | 0.9441(0.2013) | 0.8349(0.3713) |
| $\omega_3$ | 0.6782(0.4358) | 0.5005(0.4796) |
| $\omega_4$ | 1.0000(0.0000) | 1.0000(0.0000) |
| $\omega_5$ | 0.6312(0.4583) | 0.4690(0.4859) |

### 6 GAIT MATURATION ANALYSIS

We now present the analytical results for the motivating gait maturation study described in the introduction (Goldberger et al., 2000). The current analysis considers the effect of age on gait variability to better understand gait maturation in young children in the presence of other factors that may influence gait, such as gender and gait speed. The data contain stride interval time series from 50 healthy children with equal numbers of girls and boys between 3 and 14 years old. The time series consists of $T = 256$ stride times during normal walking after removing the first 60 s and last 5 s to avoid warm-up and ending effects (Figure 1). More details of data processing can be found in Hausdorff et al. (1999). The proposed Bayesian sum of trees model was used to estimate the covariate-dependent power spectrum of stride interval time series using 5 trees and 10,000 total iterations with the first 5,000 iterations discarded as burn-in. See Web Appendix B for convergence diagnostics and Web Appendix D for graphical posterior predictive checks for this application, which demonstrate stable estimation and model adequacy. Age, gait speed, and gender are all included in the model with 100% probability, and their effects are explored in what follows.

Partial dependence (PD; Friedman, 2001) is the most widely used method for evaluating covariate effects in
machine learning models. However, there is an issue with multicollinearity in this dataset, as age and gait speed are significantly correlated ($r = 0.653, p < 0.0001$), which can render PD unreliable due to extrapolation of the response at predictor values far outside the multivariate envelope of the data (Apley & Zhu, 2020). Therefore, we use accumulated local effects (ALE) (Apley & Zhu, 2020) to characterize covariate effects. ALE presents the effect in a small interval of the interested feature, which can mitigate issues with multicollinearity by localizing the estimated effect within the envelope of the data. Let $\omega = (\omega_j, \omega_k)$ where $\omega_j$ denotes the $j$th covariate and $\omega_k$ denotes all other covariates. The ALE for $\omega_j = x$ on the power spectrum at frequency $\nu$ is defined as

$$ f_{j,ALE}(x, \nu) = \int_{z_{j,0}}^{x} E_{\omega_j|\nu} \left[ \frac{\delta f(\omega, \nu)}{\delta \omega_j} \right] dz_j - \text{constant} $$

(6)

where $Z_j = \{z_{0,j}, ..., z_{H,j}\}$ is a collection of $H + 1$ partition points over the effective support of $\omega_j$. The constant is a value to vertically center the plot. Let $f(z_{h,j}, x_{j}; \nu)$ be the estimated power spectrum from a single posterior draw for $\omega_j = z_{h,j}, h = 1, ..., H$ and $\omega_k = x_{j}$ on frequency $\nu$, the uncentered ALE can then be estimated by

$$ g_{j,ALE}(x, \nu) = \sum_{h=1}^{H} \frac{1}{n_j(h)} \sum_{i: i \in N_j(h)} [f(z_{h,j}, x_{j}; \nu) - f(z_{h-1,j}, x_{j}; \nu)] $$

(7)

where $h_j(x)$ is the index for the interval of the partition $Z_j$ to which the value $x$ belongs, $n$ is the total number of observations and $n_j(h)$ is the number of observations in the $h$th segment of the partition for $\omega_j$ such that $\sum_{h=1}^{H} n_j(h) = n$. $N_j(h) = \{z_{h-1,j}, z_{h,j}\}$ represents the $h$th interval of the partition for $\omega_j$. Then, the estimated centered ALE is

$$ \hat{f}_{j,ALE}(x, \nu) = \hat{g}_{j,ALE}(x, \nu) - \frac{1}{n} \sum_{i=1}^{n} \hat{g}_{j,ALE}(x^{(i)}, \nu). $$

(8)

The estimated centered ALE above can be computed for each of the posterior draws from the RJMCMC sampler, and the posterior mean and 95% credible intervals can provide the desired inference on the covariate effects. By partitioning covariates into $H = 5$ intervals containing equal numbers of observations, Figure 4A,B shows the posterior mean of the ALE for age and gait speed on the power spectrum. Corresponding 95% credible intervals are available in Web Appendix B. Two findings can be concluded from these plots. First, power over all frequencies decreases as age increases. This indicates variability in stride times decreases with age, which is consistent with previous findings (Hausdorff et al., 1999).

Second, we can observe that power in low frequencies (LF $0.05-0.25$ stride$^{-1}$) decreases much more with age relative to higher frequencies (HF $0.25-0.5$ stride$^{-1}$), which is also expected as LF power corresponds to fluctuations over relatively longer timescales and is indicative of less mature neuromuscular control (Hausdorff et al., 1999). It should be noted that Hausdorff et al. (1999) discretized age as a categorical variable and then considered an ANOVA model to test for age effects. Our model, on the other hand, considers the age-dependent ALE as a continuous surface, which provides a more comprehensive assessment of the association between frequency patterns of gait variability time series and age without subjectively categorizing age into different bins.

Figure 4C,D present the posterior mean of the ALE of age and gait speed on the LF/HF ratio $\frac{\text{LF}}{\text{HF}}(\omega) = \int_{0.05}^{0.25} f(\omega, \nu) d\nu \int_{0.25}^{0.5} f(\omega, \nu) d\nu$ along with 95% pointwise credible intervals. This can be computed by replacing the power spectrum $f(\omega, \nu)$ and estimated power spectrum $\hat{f}(z_{h,j}, x_{j}; \nu)$ in Equations (6)–(8) with the $\frac{\text{LF}}{\text{HF}}(\omega)$ and its corresponding estimates $\frac{\text{LF}}{\text{HF}}(\omega) = \sum_{\nu_k \in (0.05,0.25)} \hat{f}(\omega, \nu_k) / \sum_{\nu_k \in (0.25,0.5)} \hat{f}(\omega, \nu_k)$, where $\frac{\text{LF}}{\text{HF}}(\omega)$ can be expressed as $\frac{\text{LF}}{\text{HF}}(z_{h,j}, x_{j})$ when calculating the ALE for the $j$th covariate. While LF/HF(\omega) decreases significantly with age, we see relatively larger decreases beyond 7 years of age. Previous analyses of gait variability time series have produced contradictory results. Hausdorff et al. (1999) showed that the LF/HF ratio presents a significant decrease in children 7–14 years of age. However, Preis et al. (1997) showed that gait maturation occurs more rapidly in children 3–7 years of age and changes slowly after 7 years. Our results are more consistent with analyses suggesting LF/HF ratio decreases beyond 7 years of age.

For gait speed, we observe significantly less LF ($< 0.1$ stride$^{-1}$) power (and lower LF/HF ratio) among faster walkers with speeds above 1.1 m/s. Gender appears to have a much less effect on power spectra (see Web Appendix B), with only a very small frequency range from 0.05 stride$^{-1}$ to 0.15 stride$^{-1}$ significantly different in power between males and females. There are no significant differences in LF/HF ratio between males and females with a posterior mean ALE of $-0.1680$ and 95% credible interval $(-0.39,0.06)$.

We provide some possible extensions of the proposed method which can be used to accomplish other goals of analyzing gait variability time series. First, other gait maturation studies have focused on classifying subjects into different groups based on the covariates and gait variability time series through support vector machines (Wu & Krishnan, 2009a, 2009b; Wu & Shi, 2011). Web
Appendix E describes how the proposed method can be adapted for estimating unknown covariates through an inverse regression framework for interpretable regression and classification. Second, if disease status is considered as a covariate, then this extension can potentially aide in diagnosing certain diseases, such as Parkinson’s disease, which is another goal of many gait studies (Daliri, 2012; Khorasani & Daliri, 2014). Third, visualization of covariate effects is a major advantage of tree-based models. In Web Appendix B, a visualization of the tree structures for the gait maturation analysis is presented.

7 | DISCUSSION

This paper describes a novel adaptive Bayesian covariate-dependent model for the power spectrum of multiple time series. By using a sum of trees model to characterize covariate effects, this model is flexible and can automatically recover complex nonlinear associations and interactions as well as provide efficient variable selection.

This paper is one of the first approaches to analyzing the power spectrum of multiple stationary time series with multiple covariates in a completely nonparametric
manner, but it is not without limitations. An important direction for future research is to extend the proposed Bayesian sum of trees model to accommodate additional data features commonly encountered in practice, such as time- and covariate-dependent time series (Bertolacci et al., 2022), extra spectral variability due to clustering effects (Krafty, 2016), and missingness in covariate vectors. Alternative partitioning frameworks, such Voronoi tessellations (Payne et al., 2020) and binary space partitioning trees (Fan et al., 2019), and soft-decision trees (Linero & Yang, 2018) that better adapt to smooth effects may also be considered for capturing covariate effects in an even more flexible and parsimonious manner.

ACKNOWLEDGMENTS
The authors thank the Editor, Associate Editor, and two Referees for their detailed and insightful comments that greatly improved the paper. This work was supported by the National Institutes of Health (NIH) under Award Number R01GM140476. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing.

DATA AVAILABILITY STATEMENT
The data (Hausdorff and Peng, 1999) that support the findings in this paper are openly available in Physionet (Goldberger et al., 2000) at https://doi.org/10.13026/C2H014.

ORCID
Zeda Li https://orcid.org/0000-0002-0305-6240
Scott A. Bruce https://orcid.org/0000-0002-7904-4211

REFERENCES
Adak, S. (1998) Time-dependent spectral analysis of nonstationary time series. *Journal of the American Statistical Association*, 93(444), 1488–1501.
Apley, D.W. & Zhu, J. (2020) Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B, Statistical Methodology*, 82(4). 1059–1086.
Bertolacci, M., Rosen, O., Cripps, E. & Cripps, S. (2022) AdaptSPEC-X: Covariate-dependent spectral modeling of multiple nonstationary time series. *Journal of Computational and Graphical Statistics*, 31, 436–454.
Blattenberger, G. & Fowles, R. (2017) Treed avalanche forecasting: mitigating avalanche danger utilizing Bayesian additive regression trees. *Journal of Forecasting*, 36(2), 165–180.
Breiman, L. (2001) Random forests. *Machine Learning*, 45(1), 5–32.
Brillinger, D.R. (2002) *Time series: data analysis and theory*. Philadelphia: SIAM.
Bruce, S.A., Hall, M.H., Buysse, D.J. & Krafty, R.T. (2018) Conditional adaptive Bayesian spectral analysis of nonstationary biomedical time series. *Biometrics*, 74(1), 260–269.
Chau, J. & von Sachs, R. (2016) Functional mixed effects wavelet estimation for spectra of replicated time series. *Electronic Journal of Statistics*, 10(2), 2461–2510.
Chipman, H., George, E.I., Gramacy, R.B. & McCulloch, R. (2013) Bayesian treed response surface models. *WIREs Data Mining and Knowledge Discovery*, 3(4), 298–305.
Chipman, H.A., George, E.I. & McCulloch, R.E. (2010) BART: Bayesian additive regression trees. *The Annals of Applied Statistics*, 4(1), 266–298.
Cramér, H. (1942) On harmonic analysis in certain function spaces. *Arkiv för Matematik, Astronomi och Fysik*, 28B, 1–7.
Daliri, M.R. (2012) Automatic diagnosis of neuro-degenerative diseases using gait dynamics. *Measurement*, 45(7), 1729–1734.
Diggle, P.J. & Al Wasel, I. (1997) Spectral analysis of replicated biomedical time series. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 46(1), 31–71.
Eddelbuettel, D. & Sanderson, C. (2014) RcppArmadillo: accelerating R with high-performance C++ linear algebra. *Computational Statistics and Data Analysis*, 71, 1054–1063.
Eubank, R. (1999) *Nonparametric regression and spline smoothing*. Boca Raton, FL: CRC Press.
Fan, X., Li, B. & Sisson, S. (2019) Binary space partitioning forest.
Fan, X., Li, B. & Sisson, S. (2019) Binary space partitioning forest.
Fiecas, M. & Ombao, H. (2017) Modeling the evolution of dynamic brain processes during an associative learning experiment. *Journal of the American Statistical Association*, 112(516), 1440–1453.
Gelman, A. (2006) Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, 1(3), 515–534.
Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.-K. & Stanley, H.E. (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23), 215–220.
Green, P.J. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4), 711–732.
Hall, M., Vasko, R., Buysse, D., Ombao, H., Chen, Q., Cashmere, J.D., Kupfer, D. & Thayer, J.F. (2004) Acute stress affects heart rate variability during sleep. *Psychosomatic Medicine*, 66(1), 56–62.
Hausdorff, J.M. & Peng, C.-K. (1999) Gait maturation database. Version 1.0.0, Physionet. https://doi.org/10.13026/C2H014.
Hausdorff, J.M., Zemany, L., Peng, C.-K. & Goldberger, A.L. (1999) Maturation of gait dynamics: Stride-to-stride variability and its temporal organization in children. *Journal of Applied Physiology*, 86(3), 1040–1047.
Iannaccone, R. & Coles, S. (2001) Semiparametric models and inference for biomedical time series with extra-variation. *Biostatistics*, 2(3), 261–276.
Khorasani, A. & Daliri, M.R. (2014) HMM for classification of Parkinson’s disease based on the raw gait data. *Journal of Medical Systems*, 38(12), 1–6.
Klimesch, W. (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29(2), 169–195.

Krafty, R.T. (2016) Discriminant analysis of time series in the presence of within-group spectral variability. *Journal of Time Series Analysis*, 37(4), 435–450.

Krafty, R.T., Hall, M. & Guo, W. (2011) Functional mixed effects spectral analysis. *Biometrika*, 98, 583–598.

Krafty, R.T., Rosen, O., Stoffer, D.S., Buyse, D.J. & Hall, M.H. (2017) Conditional spectral analysis of replicated multiple time series with application to nocturnal physiology. *Journal of the American Statistical Association*, 112(520), 1405–1216.

Li, Z., Bruce, S.A., Wutzke, C.J. & Long, Y. (2021) Conditional adaptive Bayesian spectral analysis of replicated multivariate time series. *Statistics in Medicine*, 40(8), 1989–2005.

Li, Z. & Krafty, R.T. (2019) Adaptive Bayesian time-frequency analysis of multivariate time series. *Journal of the American Statistical Association*, 114(525), 453–465.

Linero, A.R. (2018) Bayesian regression trees for high-dimensional prediction and variable selection. *Journal of the American Statistical Association*, 113(522), 626–636.

Linero, A.R. & Yang, Y. (2018) Bayesian regression tree ensembles that adapt to smoothness and sparsity. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(5), 1087–1110.

Mentch, L. & Zhou, S. (2020) Randomization as regularization: a degrees of freedom explanation for random forest success. *Journal of Machine Learning Research*, 21(171), 1–36.

Payne, R.D., Guha, N., Ding, Y. & Mallick, B.K. (2020) A conditional density estimation partition model using logistic Gaussian processes. *Biometrika*, 107(1), 173–190.

Preis, S., Klemms, A. & Müller, K. (1997) Gait analysis by measuring ground reaction forces in children: changes to an adaptive gait pattern between the ages of one and five years. *Developmental Medicine and Child Neurology*, 39(4), 228–233.

Qin, L., Guo, W. & Litt, B. (2009) A time-frequency functional model for locally stationary time series data. *Journal of Computational and Graphical Statistics*, 18(3), 675–693.

R Core Team (2021) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rosen, O., Wood, S. & Stoffer, D. (2012) AdaptSPEC: adaptive spectral estimation for nonstationary time series. *Journal of the American Statistical Association*, 107, 1575–1589.

Rokčová, V. & Saha, E. (2019) On theory for BART. In *Proceedings of Machine Learning Research*, vol. 89, *Proceedings of Machine Learning Research*, pp. 2839–2848. Naha, Okinawa, Japan.

Schwarz, K. & Krivobokova, T. (2016) A unified framework for spline estimators. *Biometrika*, 103(1), 121–131.

Shumway-Cook, A. & Williams, M.W. (1995) *Motor control: theory and practical applications*, Baltimore: Williams & Wilkins.

Starling, J.E., Murray, J.S., Carvalho, C.M., Bukowski, R.K., Scott, J.G., “other” et al. (2020) BART with targeted smoothing: an analysis of patient-specific stillbirth risk. *Annals of Applied Statistics*, 14(1), 28–50.

Stoffer, D.S., Han, S., Qin, L. & Guo, W. (2010) Smoothing spline ANOPOW. *Journal of Statistical Planning and Inference*, 140(12), 3789–3796.

van der Merwe, S. (2018) Time series analysis of the southern oscillation index using Bayesian additive regression trees. *ArXiv*. https://arxiv.org/abs/1804.01809

Waldmann, P. (2016) Genome-wide prediction using Bayesian additive regression trees. *Genetics Selection Evolution (Paris)*, 48(1), 42.

Wand, M.P., Ormerod, J.T., Padoan, S.A. & Frühwirth, R. (2011) Mean field variational Bayes for elaborate distributions. *Bayesian Analysis*, 6(4), 847–900.

Whittle, P. (1952) The simultaneous estimation of a time series harmonic components and covariance structure. *Trabajos de Estadística*, 3(1), 43–57.

Wood, S. (2017) *Generalized additive models: an introduction with R* (2 ed.). Chapman and Hall.

Wu, Y. & Krishnan, S. (2009a) Computer-aided analysis of gait rhythm fluctuations in amyotrophic lateral sclerosis. *Medical and Biological Engineering and Computing*, 47(11), 1165–1171.

Wu, Y. & Krishnan, S. (2009b) Statistical analysis of gait rhythm in patients with Parkinson's disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(2), 150–158.

Wu, Y. & Shi, L. (2011) Analysis of altered gait cycle duration in amyotrophic lateral sclerosis based on nonparametric probability density function estimation. *Medical Engineering and Physics*, 33(3), 347–355.

Wu, Y., Tjelmeland, H. & West, M. (2007) Bayesian CART: prior specification and posterior simulation. *Journal of Computational and Graphical Statistics*, 16(1), 44–66.

**SUPPORTING INFORMATION**

A zip file containing R code for implementing the proposed Bayesian sum of trees model and a pdf file containing Web Appendices A–F are available with this paper at the Biometrics Website on Wiley Online Library. Web Appendix A referenced in Section 4.4 displays the sampling scheme details of MCMC algorithm. Web Appendix B referenced in Sections 5 and 6 contains the additional results for simulations and gait maturation study. Web Appendix C referenced in Section 5.2 describes simulation comparisons with a generalized additive model. Web Appendix D referenced in Section 6 presents graphical posterior predictive plots for model checking. Web Appendix E referenced in Section 6 presents an extension of the proposed method for inverse regression to estimate unknown covariates. Web Appendix F referenced in Section 5.3 describes the provided R code.

**How to cite this article:** Wang, Y., Li, Z. & Bruce, S.A. (2023) Adaptive Bayesian sum of trees model for covariate-dependent spectral analysis. *Biometrics*, 79, 1826–1839. [https://doi.org/10.1111/biom.13763](https://doi.org/10.1111/biom.13763)