The Annals of Applied Statistics
2024, Vol. 18, No. 1, 328–349
https://doi.org/10.1214/23-AOAS1791

ANOPOW FOR REPLICATED NONSTATIONARY TIME SERIES IN EXPERIMENTS

BY ZEDA LI[1,a], YU (RYAN) YUE[1,b] AND SCOTT A. BRUCE[2,c]

[1]*Baruch College, The City University of New York,* [a]*zeda.li@baruch.cuny.edu,* [b]*yu.yue@baruch.cuny.edu*
[2]*Department of Statistics, Texas A&M University,* [c]*sabruce@tamu.edu*

We propose a novel analysis of power (ANOPOW) model for analyzing replicated nonstationary time series commonly encountered in experimental studies. Based on a locally stationary ANOPOW Cramér spectral representation, the proposed model can be used to compare the second-order time-varying frequency patterns among different groups of time series and to estimate group effects as functions of both time and frequency. Formulated in a Bayesian framework, independent two-dimensional second-order random walk (RW2D) priors are assumed on each of the time-varying functional effects for flexible and adaptive smoothing. A piecewise stationary approximation of the nonstationary time series is used to obtain localized estimates of time-varying spectra. Posterior distributions of the time-varying functional group effects are then obtained via integrated nested Laplace approximations (INLA) at a low computational cost. The large-sample distribution of local periodograms can be appropriately utilized to improve estimation accuracy since INLA allows modeling of data with various types of distributions. The usefulness of the proposed model is illustrated through two real-data applications: analyses of seismic signals and pupil diameter time series in children with attention deficit hyperactivity disorder. Simulation studies, Supplementary Material (Li, Yue and Bruce (2024a)), and `R` code (Li, Yue and Bruce (2024b)) for this article are also available.

1. Introduction. The second-order frequency domain properties of time series, which can be quantified through the power spectrum, are essential to addressing scientific questions in a variety of fields. As a result, an increasing number of experiments collect time-series data across multiple subjects, which we refer to as replicated time series, in order to study the effects of experimental factors on second-order frequency patterns. Quite often, time series characteristics evolve over the course of the experiment, so it is essential to evaluate the effects of experimental factors in a time-dependent manner. The goal of this article is to introduce a flexible and computationally efficient method for analyzing replicated nonstationary time series commonly encountered in complex experiments.

1.1. *Existing spectral analysis approaches for replicated time series.* Spectral analysis of *replicated stationary* time series produced by designed experiments has been well studied in the time-series literature. Methods include parametric log-linear models (Diggle and al Wasel (1997)), functional semiparametric models (Iannaccone and Coles (2001); Krafty, Hall and Guo (2011)), wavelet-based models (Chau and von Sachs (2016)), tree-structured models (Freyermuth, Ombao and von Sachs (2010); Wang, Li and Bruce (2022)), and the smoothing spline analysis of power (ANOPOW) model (Stoffer et al. (2010)). Unfortunately, these methods are not capable of analyzing *replicated nonstationary* time series due to methodological and computational obstacles.

Received September 2022; revised May 2023.
Key words and phrases. Integrated nested Laplace approximations, locally stationary ANOPOW Cramér spectral representation, pupil diameter time series, replicated nonstationary time series, spectral analysis.

Few approaches in the literature are available for analyzing *replicated nonstationary* time series in the frequency domain. Qin, Guo and Litt (2009) considers a semiparametric time-frequency functional model for replicated locally stationary time series and uses a state-space representation for estimation. However, this model relies on the assumption of mutual independence among time series and cannot immediately accommodate experimental designs with dependent replications. In a recent study, Fiecas and Ombao (2016) propose a nonparametric approach that circumvents the parametric assumptions of Qin, Guo and Litt (2009) by smoothing across covariate and time nonparametrically. However, this model as well as the method of Qin, Guo and Litt (2009) are formulated for power spectra that evolve continuously and smoothly over covariate. This assumption becomes problematic when evaluating time series from experimental studies with multiple treatment groups that may not exhibit smooth changes across groups. Bruce et al. (2018) propose a nonparametric Bayesian method to quantify the association between time-varying spectra and covariate. Unfortunately, this approach can only accommodate a single covariate, which limits its use for studies involving multiple experimental factors, such as our motivating study. Assuming the error terms are Gaussian, Martinez et al. (2013) adopts the Bayesian wavelet-based semiparametric mixed-effects model developed by Morris and Carroll (2006) to characterize covariate-induced frequency patterns among time series. However, the assumption of Gaussianity may not hold for spectral analysis of time series, since local periodogram estimators of the power spectra are non-Gaussian. In addition, a common limitation of the Bayesian models of Bruce et al. (2018) and Martinez et al. (2013) is that they rely on sophisticated Markov chain Monte Carlo (MCMC) techniques and could become computationally infeasible when the study design is more complex and the data size is large. More recently, Yue et al. (2019) propose a generalized two-way functional ANOVA model that can handle two-dimensional functional effects and apply this model to a wide variety of applications. However, this model doesn't take advantage of the specific distributional characteristics of the local periodograms used to analyze the frequency domain behavior of nonstationary time series. The primary contribution of this work is the introduction of a flexible and computationally efficient analysis of power (ANOPOW) model for analyzing replicated nonstationary time series collected in experiments with fixed or mixed effects.

1.2. *Overview of the proposed method.* The proposed approach is based on a locally stationary ANOPOW representation for modeling replicated nonstationary time series that are expected to differ across categorical factors or groups. Under this model the time series replicates are assumed to have time-varying log-spectra, depending on the corresponding experimental factors, factor interactions, or possible random effects. Functional effects as functions of both time and frequency are modeled by a two-dimensional second-order random walk (RW2D) model in the framework of Gaussian Markov random fields (GMRF) (Rue and Held (2005)). Formulated in a Bayesian framework, the proposed approach approximates each nonstationary time series by a piecewise stationary time series to obtain localized estimates of the time-varying spectra and relies on integrated nested Laplace approximations (INLA) (Rue, Martino and Chopin (2009)) to provide accurate and fast approximations to posterior marginals of the experimental effects. INLA allows for modeling both Gaussian and non-Gaussian data, so the full large-sample distributional properties of the periodograms can be characterized by modeling the exponentially distributed local log periodograms directly. This approach shows improved estimation accuracy compared to alternatives relying on Gaussian approximations.

The proposed framework also provides inference via simultaneous credible intervals and zero-level contour avoiding functions proposed by Bolin and Lindgren (2015) for assessing the significance of experimental effects on time-varying dynamics of the power spectra. Computing these quantities typically requires integration of a high-dimensional joint posterior

distribution, which can be challenging in practice. Our approach circumvents this problem by treating the posterior approximations given by INLA as a mixture of Gaussian distributions. Consequently, the excursion method proposed by Bolin and Lindgren (2015) can be used to efficiently estimate simultaneous credible intervals and zero-level contour avoiding functions. This computational advantage is not offered by other Bayesian methods, such as those of Martinez et al. (2013) and Bruce et al. (2018).

The proposed ANOPOW model can be easily implemented using `R` (R Core Team (2023)). The INLA-based estimation procedure is implemented through the `R` package `INLA` which can be downloaded from www.r-inla.org. The excursion method for calculating the simultaneous credible intervals and the contour avoiding sets can be implemented by using the `R` package `excursions` (Bolin and Lindgren (2018)).

1.3. *Organization*. The article is organized as follows. Section 2 describes two motivating applications. Section 3 presents the locally stationary ANOPOW Cramér representation, the ANOPOW log-spectral model, and the priors on the functional effects. Section 4 discusses the INLA-based estimation procedure, including posterior inference and the construction of zero-level contour avoiding sets. Section 5 discusses extensions to incorporate random effects through a locally stationary mixed-effect ANOPOW Cramér representation. Section 6 presents the results of simulation studies. The proposed method is applied to the analyses of seismic signals and the pupil diameter time series in Sections 7 and 8, respectively. Section 9 concludes with some discussion and future work.

2. Motivating applications. We motivate our proposed method through two applications. The first application considers a collection of seismic signals under a simple one-factor model. The second application analyzes pupil diameter time series, obtained in a study of attention deficit hyperactivity disorder (ADHD) in children, which involves more complicated designs with two factors, nested structure, and random effects.

2.1. *Seismic signals*. Signals generated by earthquakes and explosions are often observed at seismic recording stations. Distinguishing between earthquakes and explosions is particularly important for monitoring a comprehensive explosive test-ban treaty. We analyze a collection of earthquake and explosion signals taken from Shumway and Stoffer (2017), which can be easily accessed through the `R` package `astsa` (Stoffer and Poison (2023)). The dataset includes eight earthquake and eight mining explosion time series signals of length $T = 2048$ recorded by stations in Scandinavia with a sampling rate of 40 Hz (see Figure 1). These signals have two phases, or arrivals, along a seismic recording station, that is, P phase ($t = 1, \ldots, 1024$) and S phase ($t = 1025, \ldots, 2048$). This dataset has been studied in the statistics literature (Fryzlewicz and Ombao (2009); Stoffer et al. (2010); Shumway and Stoffer (2017)). However, these methods largely focus on either identifying discriminatory features for classification purposes or exploring frequency-domain differences between the two groups, assuming time series are stationary within the P and S phases. In our analysis we study these nonstationary signals directly without separating them into P and S phases, which allows for investigating differences in the time-varying patterns of these signals.

2.2. *Pupil diameter time series in the ADHD study.* Another motivating example comes from a study of pupil diameter time series in children with attention deficit hyperactivity disorder (ADHD) (Rojas-Líbano et al. (2019)). ADHD is one of the most common childhood neuropsychiatric disorders and is characterized by inattention, impulsiveness, and hyperactivity. ADHD often persists into adulthood, and patients with untreated ADHD suffer from poor educational outcomes and familial relationships (Hamed, Kauer and Stevens (2015)).

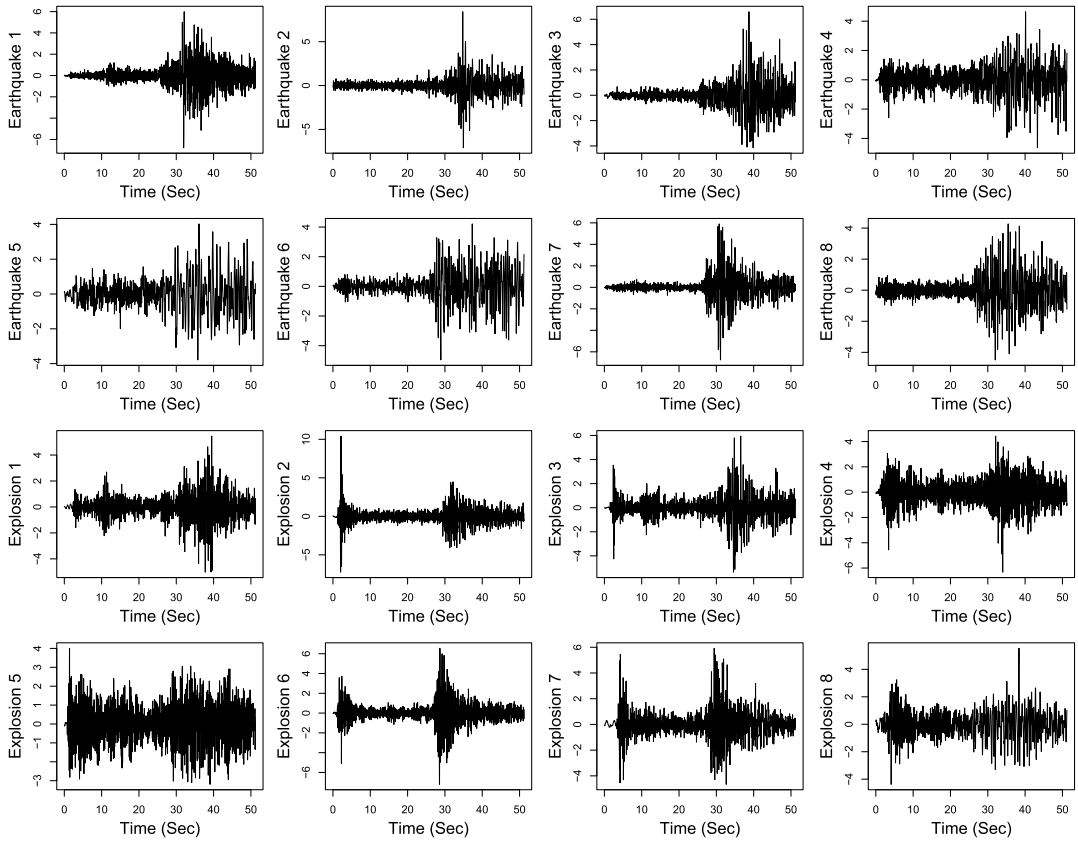

FIG. 1. *Seismic signals recorded by seismic recording stations in Scandinavia. Rows* 1–2: *Signals from earthquakes. Rows* 3–4: *Signals from explosions.*

The current diagnosis of ADHD is primarily based on observed behavior and reported symptoms, which is subjective and can be both inefficient and inaccurate (Wainstein et al. (2017)). Consequently, a reliable objective measure that characterizes the disorder is needed to ensure accurate diagnosis. Pupil diameter time series is a promising biomarker related to cognitive states, as it reflects neural responses modulated by brain state, which varies with arousal, attention, and behavior (Reimer et al. (2014)). Pupil dilation occurs in response to increases in arousal and mental effort, either triggered by external stimuli or spontaneously (Mathôt (2018)). Accordingly, pupil diameter time series collected during attentional experiments may offer indirect, objective measurements of attentiveness that can be used to better understand and diagnose ADHD.

A group of 50 children participated in the study. Twenty-eight of the children have previously been diagnosed with ADHD, and 22 children are healthy controls. Within the ADHD group, 11 of the 28 children completed the working memory task once without taking ADHD medication, and 17 of the 28 children completed the task twice, once with medication and once without medication. For the 17 ADHD children completing the working memory task twice, task order was randomized to avoid learning effects, and a one-day washout period was observed in between tasks to avoid crossover effects. We designate the groups accordingly as children with ADHD who complete the task without taking medicine (ADHD), children with ADHD who complete the task with taking medicine (mADHD), and children without ADHD (healthy control). Children performed multiple trials of the task during the experiment. For each trial three images (one- or two-dot arrays) and a distractor were sequentially presented, then a "probe" image was shown at the beginning of the fifth second. After the probe image

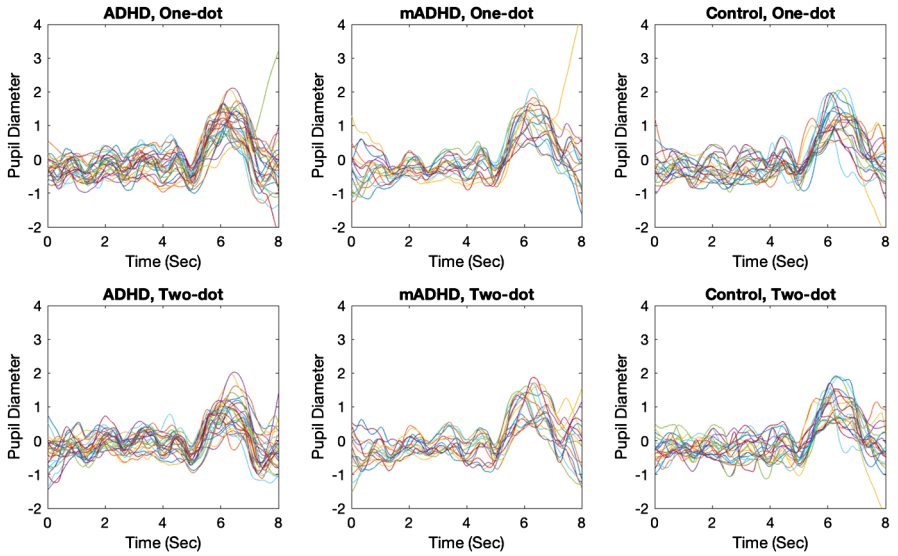

FIG. 2. *Downsampled, detrended, and filtered pupil diameter time series for three different groups. First column: Children with attention deficit hyperactivity disorder (ADHD) completing the task without taking medicine. Second column: ADHD children completing the task with taking medicine (mADHD). Third column: Children without ADHD (Control). Pupil diameter time series are obtained under two experimental cognitive loads: Images with one-dot (first row) and two-dot arrays (second row).*

was displayed, participants responded "yes" if the probe image had been presented in one of the trial's previous images or "no" if it had not been presented. Pupil diameter time series were collected for multiple trials. To increase the signal-to-noise ratio, a common practice is to average time series across multiple trials (Wainstein et al. (2017)). In our study we average trials for each subject by image type (one- or two-dot array) in order to study both temporal evolutions and the effect of varying cognitive loads. Consequently, two time series are obtained for each child within a certain group (see Figure 2). Pupil responses are dynamic during the task, so the data are inherently nonstationary, and the frequency domain properties provide physiological information about children's cognitive states.

The goal of our analysis is to understand how the time-varying frequency patterns of pupil diameter time series differ between children with ADHD and healthy controls. In addition, by taking into account the repeated measures design of the experiment, we are also able to investigate the effectiveness of medication in children with ADHD.

3. The model.

3.1. *Locally stationary ANOPOW Cramér representation.* Without loss of generality, we introduce the locally stationary ANOPOW Cramér representation in the setting of a two-factor factorial design; that is, each level of the factor α is crossed with each level of the factor β. Let $\{X_{jkrt}, t = 1, \ldots, T\}$ be a univariate locally stationary time series of length T obtained from the rth subject at the jth level of factor α and kth level of factor β in an experiment for $j = 1, \ldots, m_\alpha$, $k = 1, \ldots, m_\beta$, and $r = 1, \ldots, n_{jk}$. Formally, the locally stationary ANOPOW Cramér spectral representation of X_{jkrt} is

$$X_{jkrt} = \int_{-1/2}^{1/2} A_{jk}(t/T, \omega) \exp(2\pi i \omega t)\, dZ_{jkr}(\omega), \tag{1}$$

where Z_{jkr} are mutually independent identically distributed mean-zero orthogonal incremental processes such that $E\{dZ_{jkr}(\omega)\, d\overline{Z_{jkr}(\zeta)}\} = 1$ if $\omega = \zeta$ and zero otherwise, and $\overline{Z_{jkr}(\omega)}$

is the complex conjugate of $Z_{jkr}(\omega)$. The group-specific transfer function $A_{jk}(\nu,\omega)$ is a function of scaled time $\nu \in [0,1]$ and frequency $\omega \in \mathbb{R}$ and can be expressed as

$$A_{jk}(\nu,\omega) = A^{(\mu)}(\nu,\omega)A_j^{(\alpha)}(\nu,\omega)A_k^{(\beta)}(\nu,\omega)A_{jk}^{(\psi)}(\nu,\omega).$$

In this formulation $A^{(\mu)}$ is the transfer function of the grand mean effect, $A_j^{(\alpha)}$ is the transfer function of the main effect of α at the jth level, $A_k^{(\beta)}$ is the transfer function of the main effect of β at the kth level, and $A_{jk}^{(\psi)}$ is the transfer function of the corresponding interaction effect. In addition, $A^{(\mu)}(\nu,\omega)$, $A_j^{(\alpha)}(\nu,\omega)$, $A_k^{(\beta)}(\nu,\omega)$, and $A_{jk}^{(\psi)}(\nu,\omega)$ are complex-valued functions over $(\nu,\omega) \in [0,1] \times \mathbb{R}$ such that they are periodic and Hermitian as a function of frequency. Time-varying power spectra of the replicate-specific time series X_{jkrt}, $f_{jkr}(\nu,\omega)$, for all $r = 1,\ldots,n_{jk}$, and the group-level time-varying power spectrum at the jth level of α and the kth level of β, $f_{jk}(\nu,\omega)$, are given by

$$f_{jkr}(\nu,\omega) = f_{jk}(\nu,\omega) = \left|A^{(\mu)}(\nu,\omega)\right|^2 \left|A_j^{(\alpha)}(\nu,\omega)\right|^2 \left|A_k^{(\beta)}(\nu,\omega)\right|^2 \left|A_{jk}^{(\psi)}(\nu,\omega)\right|^2,$$

where $|\cdot|$ denotes the complex modulus.

To assure valid estimation and inference, we assume regularity conditions on the distribution of the orthogonal incremental processes and on the transfer functions (Brillinger (2001); Dahlhaus (1997); Guo et al. (2003)). First, we assume mixing conditions where all cumulants of dZ_{jkr} exist and are bounded for all orders.

ASSUMPTION 1. For each $s = 1,2,\ldots$, there exists a constant $C_s \in \mathbb{R}$ and a function $\Lambda_s : \mathbb{R}^{s-1} \to \mathbb{C}$ such that

$$\operatorname{cum}\{dZ_{jkr}(\omega_1),\ldots,dZ_{jkr}(\omega_s)\} = \Delta\left(\sum_{h=1}^{s}\omega_h\right)\Lambda_s(\omega_1,\ldots,\omega_{s-1})\,d\omega_1\cdots d\omega_s,$$

where $\operatorname{cum}\{\cdot\}$ denotes cumulant, $\Lambda_1 = 0$, $\Lambda_2(\omega) = 1$, $|\Lambda_s(\omega_1,\ldots,\omega_{s-1})| \le C_s$, and $\Delta(\omega) = \sum_{h=-\infty}^{\infty}\delta(\omega+h)$ is the period 2π extension of the Dirac delta function.

Second, we assume that the second-order structure of X_{jkrt} evolves smoothly over time and frequency.

ASSUMPTION 2. The transfer functions $A^{(\mu)}(\nu,\omega)$, $A_j^{(\alpha)}(\nu,\omega)$, $A_k^{(\beta)}(\nu,\omega)$, $A_{jk}^{(\psi)}(\nu,\omega)$ have continuous up to second-order partial derivatives with respect to ν and ω, and $\partial^d A_{jk}(\nu,\omega)/\partial\omega^d = \partial^d \overline{A_{jk}}(\nu,\omega)/\partial\omega^d$ for $d = 0,1$.

3.2. *Log-spectral ANOPOW model.* In many studies, such as our motivating study, scientific interest lies in the ratio of power at different frequencies. This is equivalent to looking at linear combinations of the log power spectrum. Thus, we consider a log-spectral ANOPOW model. In the stationary time series setting, similar log-spectra models have been considered in Diggle and al Wasel (1997), Krafty, Hall and Guo (2011), and Chau and von Sachs (2016). In particular, we define functions $\mu(\nu,\omega) = \log|A^{(\mu)}(\nu,\omega)|^2$, $\alpha_j(\nu,\omega) = \log|A_j^{(\alpha)}(\nu,\omega)|^2$, $\beta_k(\nu,\omega) = \log|A_k^{(\beta)}(\nu,\omega)|^2$, and $\psi_{jk}(\nu,\omega) = \log|A_{jk}^{(\psi)}(\nu,\omega)|^2$. The transfer function model induces a ANOPOW log-spectral model such that

$$\log f_{jkr}(\nu,\omega) = \mu(\nu,\omega) + \alpha_j(\nu,\omega) + \beta_k(\nu,\omega) + \psi_{jk}(\nu,\omega).$$

For identifiability we consider the following constraints:

$$(2)\qquad \alpha_{m_\alpha}(\nu,\omega) = 0, \qquad \beta_{m_\beta}(\nu,\omega) = 0, \qquad \psi_{m_\alpha k}(\nu,\omega) = 0, \qquad \psi_{jm_\beta}(\nu,\omega) = 0$$

for all $(\nu, \omega) \in [0, 1] \times \mathbb{R}$. These constraints make μ the mean response function at the m_αth level of α and the m_βth level of β, and the remaining main effects and interaction effects should be interpreted as deviations in the log power spectrum from the mean response function at different levels, that is, $j \neq m_\alpha$ and $k \neq m_\beta$.

3.3. *Priors on functional effects.* A two-dimensional second-order random walk (RW2D) prior is used to characterize the time-varying functional effects. Suppose the functional effects are observed at locations $\{(\nu_1, \omega_1), \ldots, (\nu_L, \omega_S)\}$, and let $N = L \times S$. Define N-variate vectors as $\boldsymbol{\mu} = [\mu(\nu_1, \omega_1), \ldots, \mu(\nu_L, \omega_S)]'$, $\boldsymbol{\alpha}_j = [\alpha_j(\nu_1, \omega_1), \ldots, \alpha_j(\nu_L, \omega_S)]'$, $\boldsymbol{\beta}_k = [\beta_k(\nu_1, \omega_1), \ldots, \beta_k(\nu_L, \omega_S)]'$, and $\boldsymbol{\psi}_{jk} = [\psi_{jk}(\nu_1, \omega_1), \ldots, \psi_{jk}(\nu_L, \omega_S)]'$. Let $\boldsymbol{g} = (g_1, \ldots, g_N)$ be a generic representation for any of these functional effects. Define $g_{\ell s}$ as the function observed at a particular location, (ν_ℓ, ω_s), within a two-dimensional grid of time-frequency locations for $\ell = 1, \ldots, L$ and $s = 1, \ldots, S$. Similarly, define $\boldsymbol{g}_{-\ell s}$ as the observed function for all locations other than (ν_ℓ, ω_s). The full conditional distributions of $g_{\ell s}$ for locations in the interior of the two-dimensional grid, that is, $\{(\ell, s) : \ell \notin (1, 2, L-1, L)$ and $s \notin (1, 2, S-1, S)\}$, are Gaussian with mean and precision

$$E(g_{\ell s}|\boldsymbol{g}_{-\ell s}) = \frac{1}{20}\left(8 \cdot \begin{smallmatrix} \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \bullet & \circ & \circ \\ \circ & \bullet & \triangle & \bullet & \circ \\ \circ & \circ & \bullet & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \end{smallmatrix} - 2 \cdot \begin{smallmatrix} \circ & \circ & \circ & \circ & \circ \\ \circ & \bullet & \circ & \bullet & \circ \\ \circ & \circ & \triangle & \circ & \circ \\ \circ & \bullet & \circ & \bullet & \circ \\ \circ & \circ & \circ & \circ & \circ \end{smallmatrix} - 1 \cdot \begin{smallmatrix} \circ & \circ & \bullet & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \bullet & \circ & \triangle & \circ & \bullet \\ \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \bullet & \circ & \circ \end{smallmatrix}\right), \qquad \text{Prec}(g_{\ell s}|\boldsymbol{g}_{-\ell s}) = 20\tau,$$

where τ is the smoothing parameter with a gamma prior. A smoothing parameter controls the roughness of the functional effect such that, as its inverse tends toward zero, the effect function tends toward a constant function of frequency and time with probability 1. The location of $g_{\ell s}$ is denoted by "$\triangle$" while the locations of neighbors that $g_{\ell s}$ depends on are denoted by "$\bullet$." The number in front of each grid denotes the weight given to the corresponding neighboring locations, and the expectation is a weighted sum over these locations. The RW2D prior is closely related to the thin-plate spline of Wahba (1990), and the particular choice of weights is derived from discretizing the penalty term in the thin-plate splines which penalizes roughness through the integrated squared second derivative of the functional effects; see Section 2 in Yue and Speckman (2010) for more details. Thus, our proposed method is in line with the existing spectral analysis literature that also smooths the power spectrum by penalizing the roughness through the integrated squared second derivative (Sørbye et al. (2009); Krafty and Collinge (2013); Rosen, Wood and Stoffer (2012); Li and Krafty (2019)).

With this form for the conditional expectation and precision, Yue and Speckman (2010) show that the resulting density of $\boldsymbol{g}$ has the form

$$\pi(\boldsymbol{g}) \propto |\boldsymbol{\Omega}_\tau|_+^{1/2} \exp\left(-\frac{1}{2}\boldsymbol{g}'\boldsymbol{\Omega}_\tau \boldsymbol{g}\right) \tag{3}$$

with $\boldsymbol{\Omega}_\tau = \tau \mathbf{R}$, where $\mathbf{R}$ is the sparse structure matrix reflecting the Markov property of the model and $|\boldsymbol{\Omega}_\tau|_+$ is the product of the nonzero eigenvalues of $\boldsymbol{\Omega}_\tau$. This model for $\boldsymbol{g}$ represents a GMRF that is characterized by specific conditional independence properties. In particular, for some $j, k = 1, \ldots, N$ and $j \neq k$, g_j and g_k are independent conditional on other variables $\boldsymbol{g}_{-jk}$, that is, $g_j \perp\!\!\!\perp g_k|\boldsymbol{g}_{-jk}$ if and only if $\boldsymbol{\Omega}_\tau[j, k] = 0$. The conditional distributions of nodes on the boundary are also provided in Section S3 of the Supplementary Material (Li, Yue and Bruce (2024a)).

Several remarks on the RW2D prior should be noted. First, the time-varying functional effects are assumed to be independent and have separate smoothing parameters to flexibly allow different levels of smoothness for different functional effects. Second, the thin-plate spline can be thought of as a two-dimensional analog of the cubic splines in one dimension. Sørbye et al. (2009) used the integrated Wiener process, which has a close resemblance to

the cubic splines, as a prior for the spectral analysis of a univariate stationary time series. Thus, our prior can be treated as an extension of the prior used in Sørbye et al. (2009) to replicated nonstationary time series. Third, it is possible to model the frequency-specific dependence patterns among multivariate time series by placing the RW2D prior on the squared coherences. However, we focus on analyzing *replicated univariate* nonstationary time series. As such, our RW2D prior considers smoothness across frequency and time in the functional effects rather than smoothness in the dependence among multivariate time series. Lastly, the RW2D prior is particularly well suited to capture slowly-varying dynamics in the time-varying functional effects. However, this prior is not the best option for characterizing time-varying functional effects with rapid or abrupt changes. Therefore, the regularity condition on smoothness of the functional effects (Assumption 2) is necessary. A possible topic for future research is the development of an ANOPOW model for replicated nonstationary time series with abrupt changes.

4. Estimation.

4.1. *Local log periodogram ANOPOW model.* Following Adak (1998), the proposed modeling approach uses a piecewise stationary process to approximate the locally stationary time series X_{jkrt}. For rescaled time $\nu \in [0, 1]$, a collection of partition points that divide the time series into L disjoint segments is defined by $\boldsymbol{\delta} = (\delta_0, \delta_1, \ldots, \delta_L)'$, where $\delta_0 = 0$ and $\delta_L = T$ such that X_{jkrt} is approximately stationary within segments $\{t : \delta_{\ell-1} < t \leq \delta_\ell\}$ for $\ell = 1, \ldots, L$. We define $\nu_\ell = (\delta_\ell + \delta_{\ell-1})/2$ as the midpoint of the ℓth segment and T_ℓ as the number of time points in the ℓth segment.

Given $\boldsymbol{\delta}$, define the local discrete Fourier transform of X_{jkrt} at frequency $\omega_{s\ell}$ within segment ℓ as $d^{(\ell)}_{jkrs} = T_\ell^{-1/2} \sum_{t=\delta_{\ell-1}+1}^{\delta_\ell} X_{jkrt} \exp(-2\pi i \omega_{s\ell} t)$ and the subsequent local periodogram as $I^{(\ell)}_{jkrs} = |d^{(\ell)}_{jkrs}|^2$, where $\omega_{s\ell} = s/T_\ell$, $s = 1, \ldots, S_\ell$ are the Fourier frequencies and $S_\ell = \lfloor (T_\ell - 1)/2 \rfloor$. In what follows, we use a simplified model and notation, based on a predetermined number of equally-sized time blocks, so that $S_\ell = S$, and thus $\omega_{s\ell} = \omega_s$ for all $\ell = 1, \ldots, L$. The following theorem summarizes the asymptotic properties of the local log periodogram under the locally stationary ANOPOW Cramér representation and establishes uniform convergence for the first two moments.

THEOREM 1. *Let X_{jkrt} have a locally stationary ANOPOW Cramér representation satisfying Assumptions 1 and 2. Define $\kappa^{(\ell)}_{jkrs} = \log I^{(\ell)}_{jkrs} - \log f_{jkr}(\nu_\ell, \omega_s)$. As $T \to \infty$, $\min\{T_\ell\} \to \infty$, and $\max\{T_\ell^2\} = O(T)$, we have*:

1. *$\kappa^{(\ell)}_{jkrs}$ are asymptotically independent for $s = 1, \ldots, S$ and are asymptotically distributed as $\log(\chi^2_2/2)$ for $s = 1, \ldots, S-1$ and $\log(\chi^2_1)$ for $s = 0, S$.*
2. *Let $\gamma \approx 0.577$ be the Euler–Mascheroni constant, and let $\gamma_s = \gamma$ for $s \neq 0, S$ and $\gamma_0 = \gamma_S = (\log 2 + \gamma)/\pi$. Define $\sigma_s^2 = \pi^2/6$ for $s = 1, \ldots, S-1$, and $\sigma_0^2 = \sigma_S^2 \approx 4.935$. Then $E(\kappa^{(\ell)}_{jkrs}) = -\gamma_s + O(1/\min\{T_\ell\}) + O(\max\{T_\ell\}/T)$, and $\mathrm{Var}(\kappa^{(\ell)}_{jkrs}) = \sigma_s^2 + O(1/\min\{T_\ell\}) + O(\max\{T_\ell\}/T)$.*

Proof for Theorem 1 is made available in Section S4 of the Supplementary Material (Li, Yue and Bruce (2024a)). The large-sample distributional properties of the local log periodogram, described in Theorem 1, lead to the following log periodogram ANOPOW model:

$$\log I^{(\ell)}_{jkrs} = \mu(\nu_\ell, \omega_s) + \alpha_j(\nu_\ell, \omega_s) + \beta_k(\nu_\ell, \omega_s) + \psi_{jk}(\nu_\ell, \omega_s) + \epsilon^{(\ell)}_{jkrs}, \tag{4}$$

where $\epsilon_{jkrs}^{(\ell)}$ are asymptotically independent $\log(\chi_2^2/2)$ variates for $s = 1, \ldots, S-1$, and $\log(\chi_1^2)$ variates for $s = 0, S$. Theorem 1 can be used to guide the determination for the number of time blocks to adequately characterize both the time- and frequency-varying aspects of the functional effects. This suggests that the number of time blocks should be on the order of the number of time points in each block, provided that T is sufficiently large.

The proposed modeling approach relies on INLA for estimation and inference, which can directly model responses characterized by the exponential family of distributions (Rue, Martino and Chopin (2009)), as in (4). Let the time-frequency grid $z_h = (\nu_\ell, \omega_s)$ have a one-to-one correspondence between h and (ℓ, s) such that $h = (\ell-1)S + s$ for $\ell = 1, \ldots, L$ and $s = 1, \ldots, S$. Define $\log \boldsymbol{I}_{jkr} = \{\log I_{jkrz_1}, \ldots, \log I_{jkrz_N}\}$ with $I_{jkrz_h} = \{I_{jkrs}^{(\ell)}\}$. Based on (4), we consider the following likelihood in our INLA procedure:

$$\log \boldsymbol{I}_{jkr} | \boldsymbol{\mu}_{jk} \sim \boldsymbol{\mu}_{jk} + \log(\chi_2^2/2), \quad \boldsymbol{\mu}_{jk} = \boldsymbol{\mu} + \boldsymbol{\alpha}_j + \boldsymbol{\beta}_k + \boldsymbol{\psi}_{jk}, \tag{5}$$

where $\boldsymbol{\mu} = \{\mu(z_1), \ldots, \mu(z_N)\}$, $\boldsymbol{\alpha}_j = \{\alpha_j(z_1), \ldots, \alpha_j(z_N)\}$, $\boldsymbol{\beta}_k = \{\beta_k(z_1), \ldots, \beta_k(z_N)\}$, and $\boldsymbol{\psi}_{jk} = \{\psi_{jk}(z_1), \ldots, \psi_{jk}(z_N)\}$.

It should be noted that from Theorem 1, a bias-corrected log periodogram, which is extensively used in the existing spectral analysis literature (Guo et al. (2003); Qin, Guo and Litt (2009); Krafty, Hall and Guo (2011); Rosen, Wood and Stoffer (2012)), can also be considered. Define the bias-corrected local log periodogram as $y_{jkrs}^{(\ell)} = \log I_{jkrs}^{(\ell)} + \gamma_s$. Then the following *approximate* log periodogram ANOPOW model can also be used:

$$y_{jkrs}^{(\ell)} = \mu(\nu_\ell, \omega_s) + \alpha_j(\nu_\ell, \omega_s) + \beta_k(\nu_\ell, \omega_s) + \psi_{jk}(\nu_\ell, \omega_s) + \epsilon_{jkrs}^{\prime(\ell)}, \tag{6}$$

where $\epsilon_{jkrs}^{\prime(\ell)}$ are asymptotically independent with zero mean and variance $\pi^2/6$. This formulation is typically adopted for computational efficiency and for facilitating estimation and theoretical development. Consequently, under this model penalized least-squared estimators (Qin, Guo and Litt (2009); Krafty, Hall and Guo (2011)) or estimators assuming Gaussian error terms (Martinez et al. (2013)) are typically used. The proposed framework can also accommodate the model in (6) with Gaussian error terms such that

$$\boldsymbol{y}_{jkr} | \boldsymbol{\mu}_{jk}, \theta_\epsilon \sim N(\boldsymbol{\mu}_{jk}, \boldsymbol{\Sigma}_{\theta_\epsilon}), \quad \boldsymbol{\mu}_{jk} = \boldsymbol{\mu} + \boldsymbol{\alpha}_j + \boldsymbol{\beta}_k + \boldsymbol{\psi}_{jk}, \tag{7}$$

where $\boldsymbol{y}_{jkr} = \{y_{jkrz_1}, \ldots, y_{jkrz_N}\}$ with $y_{jkrz_h} = \{y_{jkrs}^{(\ell)}\}$. The error variance $\boldsymbol{\Sigma}_{\theta_\epsilon} = \theta_\epsilon \mathbf{I}_{LS \times LS}$ is an $LS \times LS$ diagonal matrix with hyperparameter θ_ϵ on the diagonal. Inverse gamma priors are placed on θ_ϵ to complete the model specification. However, it is well known that estimators using the full distributional properties of the data are generally preferred when available (Krafty and Collinge (2013); Li and Krafty (2019)). Simulation studies presented in a later section and in the Supplementary Material (Li, Yue and Bruce (2024a)) support this notion, as the estimation procedure based on the true large-sample distributional properties of the local log periodograms in (4) produce smaller average squared errors, compared to that of the approximate Gaussian model in (6).

4.2. *Posterior inference with INLA.* Let $\boldsymbol{y}_{jkr}^*$ be a generic notation that could represent either $\boldsymbol{y}_{jkr}$ or $\log \boldsymbol{I}_{jkr}$. The joint posterior distribution of our model is given by

$$\begin{aligned}\pi(\boldsymbol{\mu}, \boldsymbol{\alpha}_j, \boldsymbol{\beta}_k, \boldsymbol{\psi}_{jk}, \boldsymbol{\tau} | \boldsymbol{y}^*) \propto & \prod_{jkr} \pi(\boldsymbol{y}_{jkr}^* | \boldsymbol{\mu}, \boldsymbol{\alpha}_j, \boldsymbol{\beta}_k, \boldsymbol{\psi}_{jk}) \\ & \times \pi(\boldsymbol{\mu} | \boldsymbol{\tau}_\mu) \prod_j \pi(\boldsymbol{\alpha}_j | \boldsymbol{\tau}_{\alpha_j}) \prod_k \pi(\boldsymbol{\beta}_k | \boldsymbol{\tau}_{\beta_j}) \prod_{jk} \pi(\boldsymbol{\psi}_{jk} | \boldsymbol{\tau}_{\psi_j}) \\ & \times \pi(\boldsymbol{\tau}_\mu) \pi(\boldsymbol{\tau}_\alpha) \pi(\boldsymbol{\tau}_\beta) \pi(\boldsymbol{\tau}_\psi),\end{aligned}$$

subject to the linear constraints (2), where $\boldsymbol{\tau}_\mu$, $\boldsymbol{\tau}_\alpha$, $\boldsymbol{\tau}_\beta$, $\boldsymbol{\tau}_\psi$ are sets of smoothing parameters of functional effects $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\psi}$, respectively, $\pi(\boldsymbol{y}^*_{jkr}|\cdot)$ is the likelihood function specified in (5) or (7), $\pi(\boldsymbol{\mu}|\cdot)$, $\pi(\boldsymbol{\alpha}_j|\cdot)$, $\pi(\boldsymbol{\beta}_k|\cdot)$ and $\pi(\boldsymbol{\psi}_{jk}|\cdot)$ are the RW2D priors taken on the effect functions, and the remaining $\pi(\cdot)$ denote the prior distributions on smoothing parameters.

Let $\boldsymbol{y}^* = \operatorname{vec}\{\boldsymbol{y}^*_{jkr}\}$ be vector of all response values with length n_y, $\boldsymbol{g} = (\boldsymbol{\mu}', \boldsymbol{\alpha}'_1, \ldots, \boldsymbol{\alpha}'_{m_A}, \boldsymbol{\beta}'_1, \ldots, \boldsymbol{\beta}'_{m_B}, \boldsymbol{\psi}'_{11}, \ldots, \boldsymbol{\psi}'_{m_A m_B})'$ be a vector of functional effect parameters with length n_g, and $\boldsymbol{\tau} = (\boldsymbol{\tau}'_\mu, \boldsymbol{\tau}_\alpha, \boldsymbol{\tau}'_\beta, \boldsymbol{\tau}'_\psi)'$ be a collection of smoothing parameters with length n_τ. Then the conditional posterior distribution of $\boldsymbol{g}$ is given by

$$\pi(\boldsymbol{g}|\boldsymbol{\tau}, \boldsymbol{y}^*) \propto \exp\left\{-\frac{1}{2}\boldsymbol{g}'\boldsymbol{\Omega}_\tau \boldsymbol{g} + \sum_{j=1}^{n_y} \log \pi(y^*_j|\boldsymbol{g}, \boldsymbol{\tau})\right\},$$

where $\boldsymbol{\Omega}_\tau$ is a block diagonal matrix that consists of precision matrices for each of the functional effects. We utilize INLA techniques in order to compute the posterior marginals of interest $\pi(g_j|\boldsymbol{y}^*) = \int \pi(g_j|\boldsymbol{\tau}, \boldsymbol{y}^*)\pi(\boldsymbol{\tau}|\boldsymbol{y}^*)\,d\boldsymbol{\tau}$, $j = 1, \ldots, n_g$. Interested readers can find more details on INLA in Rue, Martino and Chopin (2009) and Wang, Yue and Faraway (2018). Our INLA procedure is comprised of the following steps:

1. Propose an approximation $\tilde{\pi}(\boldsymbol{\tau}|\boldsymbol{y}^*)$ to the joint posterior of the smoothing parameters $\pi(\boldsymbol{\tau}|\boldsymbol{y}^*)$ by Laplace approximation such that

$$\tilde{\pi}(\boldsymbol{\tau}|\boldsymbol{y}^*) \propto \left.\frac{\pi(\boldsymbol{g}, \boldsymbol{\tau}, \boldsymbol{y}^*)}{\tilde{\pi}_G(\boldsymbol{g}|\boldsymbol{\tau}, \boldsymbol{y}^*)}\right|_{\boldsymbol{g}=\boldsymbol{g}^*(\boldsymbol{\tau})},$$

where $\tilde{\pi}_G(\boldsymbol{g}|\boldsymbol{\tau}, \boldsymbol{y}^*)$ is the Gaussian approximation to $\pi(\boldsymbol{g}|\boldsymbol{\tau}, \boldsymbol{y}^*)$ obtained by matching the modal configuration and the curvature at the mode $g^*(\boldsymbol{\tau})$ such that

$$\tilde{\pi}_G(\boldsymbol{g}|\boldsymbol{\tau}, \boldsymbol{y}^*) \propto \exp\left\{-\frac{1}{2}[\boldsymbol{g} - \boldsymbol{g}^*(\boldsymbol{\tau})]'[\boldsymbol{\Omega}_\tau + \operatorname{diag}(\boldsymbol{c})][\boldsymbol{g} - \boldsymbol{g}^*(\boldsymbol{\tau})]\right\},$$

where $\boldsymbol{c}$ is the second-order Taylor expansion of $\sum_{j=1}^{n_y} \log \pi(y^*_j|\boldsymbol{g}, \boldsymbol{\tau})$. The approximated marginals of each $\pi(\tau_i|\boldsymbol{y}^*)$ for $i = 1, \ldots, n_\tau$ can then be obtained by summing out the remaining variables $\boldsymbol{\tau}_{-i}$ from $\tilde{\pi}(\boldsymbol{\tau}|\boldsymbol{y}^*)$ using the density points that have already been evaluated in the grid exploration of $\tilde{\pi}(\boldsymbol{\tau}|\boldsymbol{y}^*)$ to construct an interpolation. The marginals are then computed via numerical integration (Martins et al. (2013)).

2. Propose a Laplace approximation $\tilde{\pi}(g_j|\boldsymbol{\tau}, \boldsymbol{y}^*)$ to the marginals of the conditional distribution of g_j, given the data and the smoothing parameters $\pi(g_j|\boldsymbol{\tau}, \boldsymbol{y}^*)$, for $j = 1, \ldots, n_g$,

$$\pi_{LA}(g_j|\boldsymbol{\tau}, \boldsymbol{y}^*) \propto \left.\frac{\pi(\boldsymbol{g}, \boldsymbol{\tau}, \boldsymbol{y}^*)}{\tilde{\pi}_{GG}(\boldsymbol{g}_{-j}|g_j, \boldsymbol{\tau}, \boldsymbol{y}^*)}\right|_{\boldsymbol{g}_{-j}=\boldsymbol{g}^*_{-j}(g_j,\boldsymbol{\tau})},$$

where $\tilde{\pi}_{GG}$ is the Gaussian approximation to the full conditional of $\boldsymbol{g}_{-j}$ and $\boldsymbol{g}^*_{-j}(g_i, \boldsymbol{\tau})$ is the modal configuration.

3. Numerically, integrate out $\boldsymbol{\tau}$ from $\pi(g_j|\boldsymbol{\tau}, \boldsymbol{y}^*)$ to get

$$\tilde{\pi}(g_j|\boldsymbol{y}^*) \approx \sum_k w_k \tilde{\pi}_{LA}(g_j|\boldsymbol{\tau}^{(k)}, \boldsymbol{y}^*), \tag{8}$$

where w_k are proportional to $\tilde{\pi}(\boldsymbol{\tau}^{(k)}|\boldsymbol{y}^*)$. The evaluation points $\boldsymbol{\tau}^{(k)}$ can be chosen in different ways, depending on the importance of computational efficiency in a given setting (Martins et al. (2013)).

4.3. *Testing for functional effects*. In traditional ANOVA models, it is often of interest to test whether or not effects are significant. For the proposed model, we could test the significance of the main effect such that $H_0 : \alpha_1(\nu, \omega) = \alpha_2(\nu, \omega) = \cdots = \alpha_{m_\alpha - 1}(\nu, \omega) = 0$. If this hypothesis is rejected, it is of subsequent interest to identify the collection of time points and frequencies for which at least one of the individual functions are different. From a Bayesian point of view, we want to identify the set of (ν, ω) values such that $\alpha_j(\nu, \omega) \neq 0$ with posterior probability $1 - \rho$. To find this set, we consider the *level-zero contour avoiding function*. For a given functional effect, $g(z)$ with $z = (\nu, \omega)$, we define the positive excursion set of $g(z)$ as $\Lambda^+(g) = \{z \in [0, 1] \times [0, 1/2]; g(z) > 0\}$, the negative excursion set of $g(z)$ as $\Lambda^-(g) = \{z \in [0, 1] \times [0, 1/2]; g(z) < 0\}$, and the pair of joint contour excursion sets with probability $1 - \rho$ as

$$\big(M_\rho^+(g), M_\rho^-(g)\big) = \underset{(D^+, D-)}{\arg\max} \big\{ \big| D^+ \cup D^- \big| : P\big(D^+ \subseteq \Lambda^+(g), D^- \subseteq \Lambda^-(g)\big) \geq 1 - \rho \big\},$$

where the sets (D^+, D^-) are open. Then the level-zero contour avoiding set is given by $E_\rho(g) = M_\rho^+(g) \cup M_\rho^-(g)$, which is the largest set such that, for z in that set, the effect function $g(z)$ is different from zero with probability $1 - \rho$. The level-zero contour avoiding function is then defined as $F(z) = \sup\{1 - \rho; z \in E_\rho\}$, which can be used to identify the collection of time-frequency locations exhibiting a nonzero functional effect with high probability. It should be noted that the level-zero contour avoiding function is based on simultaneous inference rather than pointwise inference for each time-frequency location. Also, this construction can be generalized to consider level-c contour avoiding sets with $c \neq 0$ to test for practical significance of functional effects above a particular nonzero threshold c.

Constructing zero-level contour avoiding functions requires integrating the joint posterior distribution $\pi(\boldsymbol{g}|\boldsymbol{y}^*) = \int \pi(\boldsymbol{g}|\boldsymbol{\tau})\pi(\boldsymbol{\tau}|\boldsymbol{y}^*)\, d\boldsymbol{\tau}$, which can be computationally challenging due to large sample sizes or complex experimental designs. However, the proposed method can efficiently estimate the zero-level contour avoiding function, using the excursion method introduced in Bolin and Lindgren (2015), since every posterior approximation given by INLA can be viewed as a mixture of Gaussians. More specifically, the level-zero contour avoiding function can be estimated as $F(z) = \sum_k w_k F_k(z)$, where $F_k(z)$ is the level-zero contour avoiding function calculated for the conditional posterior $\pi(\boldsymbol{g}|\boldsymbol{y}^*, \boldsymbol{\tau}^{(k)})$ for a fixed parameter configuration $\boldsymbol{\tau}^{(k)}$ with corresponding weights w_k, as in (8). If $\pi(\boldsymbol{g}|\boldsymbol{y}^*, \boldsymbol{\tau}^{(k)})$ is Gaussian, the computation of $F_k(z)$ only requires the ability to compute excursion probabilities of multivariate Gaussian distributions. This can be done efficiently using the sequential method of Bolin and Lindgren (2015). If $\pi(\boldsymbol{g}|\boldsymbol{y}^*, \boldsymbol{\tau}^{(k)})$ is non-Gaussian, we can use Gaussian approximations to $\pi(\boldsymbol{g}|\boldsymbol{y}^*, \boldsymbol{\tau}^{(k)})$, which are provided by the INLA procedure introduced in the previous section.

5. Incorporating mixed effects. In many applications, such as the motivating ADHD study, multiple time series are collected from each subject. Our proposed model can also incorporate mixed effects that account for possible dependence among replicated time series. Without loss of generality, consider a one-way mixed-effect ANOPOW model such that $\{X_{jqrt}, t = 1, \ldots, T\}$ is the rth univariate nonstationary time series of length T obtained from the qth subject at the jth level of factor α. We have the locally stationary mixed-effects ANOPOW Cramér representation

$$X_{jqrt} = \int_{-1/2}^{1/2} A_{jq}^*(\nu, \omega) \exp(2\pi i \omega t)\, dZ_{jqr}(\omega) \tag{9}$$

for $j = 1, \ldots, m_\alpha$, $q = 1, \ldots, Q$, $r = 1, \ldots, R$, where Z_{jqr} are mutually independent, identically distributed, mean-zero unit-variance orthogonal incremental processes that satisfy Assumption 1. We use superscript $*$ to emphasize that the transfer function in this representation

is different from that of (1) since $A^*_{jq}(\nu,\omega)$ is a random function rather than a deterministic function. In (9) $A^*_{jq}(\nu,\omega)$ has the following representation:

$$A^*_{jq}(\nu,\omega) = A^{(\mu)}(\nu,\omega)A^{(\alpha)}_j(\nu,\omega)A^{(\eta)}_{jq}(\nu,\omega),$$

where $A^{(\mu)}$ is the transfer function representing the grand mean effect, $A^{(\alpha)}_j$ is the transfer function representing the main effect at jth level, and $A^{(\eta)}_{jq}$ is the transfer function representing subject-specific random effects that accounts for the dependence and variability among replicates. $A^{(\mu)}(\nu,\omega)$ and $A^{(\alpha)}_j(\nu,\omega)$ are complex-valued deterministic functions over $(\nu,\omega)\in[0,1]\times\mathbb{R}$ such that they are periodic and Hermitian as a function of frequency, while $A^{(\eta)}_{jq}$ is a complex-valued random function over $(\nu,\omega)\in[0,1]\times\mathbb{R}$ such that it is periodic and Hermitian as a function of frequency. Additionally, $A^{(\eta)}_{jq}$ and $A^{(\eta)}_{jq'}$ are independent and identically distributed, given $q\neq q'$. We assume that $A^{(\mu)}(\nu,\omega)$, $A^{(\alpha)}_j(\nu,\omega)$, and $A^{(\eta)}_{jq}(\nu,\omega)$ satisfy Assumption 2 and place an additional condition on $A^{(\eta)}_{jq}(\nu,\omega)$ to ensure the subject-specific spectra are bounded away from zero.

ASSUMPTION 3. There exists an $C>0$ such that $\sup_{\nu,\omega}\Pr\{|A^{(\eta)}_{jq}(\nu,\omega)|^2<C\}=0$.

Under this assumption the time series X_{jqrt} exists with probability one such that X_{jqrt} has zero mean and power spectrum $|A^{(\mu)}(\nu,\omega)|^2|A^{(\alpha)}_j(\nu,\omega)|^2E\{|A^{(\eta)}_{jq}(\nu,\omega)|^2\}$ (Krafty, Hall and Guo (2011)). Conditional on A^*_{jq}, time series X_{jqrt} is mean zero and the replicate-specific power spectrum and population-level average power spectrum at the jth level can be expressed as

$$f_{jqr}(\nu,\omega) = \big|A^{(\mu)}(\nu,\omega)\big|^2\big|A^{(\alpha)}_j(\nu,\omega)\big|^2\big|A^{(\eta)}_{jq}(\nu,\omega)\big|^2,$$

$$f_j(\nu,\omega) = \big|A^{(\mu)}(\nu,\omega)\big|^2\big|A^{(\alpha)}_j(\nu,\omega)\big|^2,$$

respectively. Without loss of generality, we assume replicate-specific spectra are parameterized such that $E\{\log|A^{(\eta)}_{jq}(\nu,\omega)|^2\}=0$ for $\nu\in[0,1]$ and $\omega\in\mathbb{R}$ to facilitate inference on log spectra. Define $\mu(\nu,\omega)=\log|A^{(\mu)}(\nu,\omega)|^2$, $\alpha_j(\nu,\omega)=\log|A^{(\alpha)}_j(\nu,\omega)|^2$, and $\eta_{jq}(\nu,\omega)=\log|A^{(\eta)}_{jq}(\nu,\omega)|^2$, the mixed-effect log-spectral model is then given by

$$\log f_{jqr}(\nu,\omega)=\mu(\nu,\omega)+\alpha_j(\nu,\omega)+\eta_{jq}(\nu,\omega).$$

The first central moments of the log-spectra is $E\{\log f_{jqr}(\nu,\omega)\}=\mu(\nu,\omega)+\alpha_j(\nu,\omega)$, and the covariance function of the log-spectral random effect, which captures the within-subject variability, is $E\{\eta_{jq}(\nu_\ell,\omega_s)\eta_{jq}(\nu_{\ell'},\omega_{s'})\}$ for $\ell,\ell'=1,\ldots,L$ and $s,s'=1,\ldots,S$.

The power spectrum defined through the locally stationary ANOPOW Cramér representation in (1) is deterministic, and the asymptotic properties of the local periodograms are established in Theorem 1. Analogous to Theorem 1, Theorem 2, whose proof is in the Supplementary Material (Li, Yue and Bruce (2024a)), investigates the asymptotic properties of the local periodograms of collections of time series that can be modeled by the locally stationary mixed-effects ANOPOW Cramér representation. This allows the local log periodograms to be characterized by the local log periodogram ANOPOW models, namely, models (4) and (6).

THEOREM 2. *Let X_{jqrt} have a locally stationary mixed-effects ANOPOW Cramér representation, as in (9), satisfying Assumptions 1–3. Define $\kappa^{(\ell)}_{jqrs}=\log I^{(\ell)}_{jqrs}-\log f_{jqr}(\nu_\ell,\omega_s)$. As $T\to\infty$, $\min\{T_\ell\}\to\infty$, and $\max\{T_\ell^2\}=O(T)$, we have*:

1. $\kappa_{jqrs}^{(\ell)}$ *are asymptotically independent for* $s = 1, \ldots, S$ *and are asymptotically distributed as* $\log(\chi_2^2/2)$ *for* $s = 1, \ldots, S-1$ *and* $\log(\chi_1^2)$ *for* $s = 0, S$.

2. *Let* $\gamma \approx 0.577$ *be the Euler–Mascheroni constant such that* $\gamma_s = \gamma$ *for* $s \neq 0, S$ *and* $\gamma_0 = \gamma_S = (\log 2 + \gamma)/\pi$. *Define* $\sigma_s^2 = \pi^2/6$ *for* $s = 1, \ldots, S-1$, *and* $\sigma_0^2 = \sigma_S^2 \approx 4.935$. *Then* $E(\kappa_{jqrs}^{(\ell)}) = -\gamma_s + O(1/\min\{T_\ell\}) + O(\max\{T_\ell\}/T)$, *and* $\mathrm{Var}(\kappa_{jqrs}^{(\ell)}) = \sigma_s^2 + O(1/\min\{T_\ell\}) + O(\max\{T_\ell\}/T)$.

The empirical properties of this model will be further explored in simulation studies provided in the Supplementary Material (Li, Yue and Bruce (2024a)) and demonstrated in the analysis of pupil diameter time-series data in ADHD study mentioned in Section 2.

6. Simulation studies. In this section the empirical performance of the proposed methods are evaluated and compared with an alternative ordinary least squares (OLS) method. The alternative approach first obtains replicate-specific local log-spectral estimates by applying a smoothing procedure, such as the spline smoother of Wahba (1980), to replicate-specific local log periodograms, while ignoring the experimental design. Then by considering each frequency individually, the design structure is characterized by a design matrix, and ordinary least squares is used to estimate the frequency-specific functional effects locally within each time block. We denote this alternative approach as SmoothOLS. The proposed ANOPOW model that directly utilizes the true large-sample distribution of the local log periodograms (4) is denoted as ANOPOW-D, and the proposed ANOPOW model based on the approximation of the large-sample distribution of the local log periodograms (6) is denoted as ANOPOW-A.

We consider a two-way ANOPOW model in which each time series is generated from the frequency domain model used in Guo and Dai (2006). The replicate-specific transfer function is specified as

$$A_{jkr}\left(\frac{t}{T}, \frac{s}{T}\right) = A^{(\mu)}\left(\frac{t}{T}, \frac{s}{T}\right) A_j^{(\alpha)}\left(\frac{t}{T}, \frac{s}{T}\right) A_k^{(\beta)}\left(\frac{t}{T}, \frac{s}{T}\right) A_{jk}^{(\psi)}\left(\frac{t}{T}, \frac{s}{T}\right)$$

for $j = 1, \ldots, 3$ and $k = 1, \ldots, 3$, and $r = 1, \ldots, n_{jk}$, where

$$\begin{aligned}
&A^{(\mu)}(\nu, \omega) = \left[1.2\cos(\pi\omega)\right]^2 + 0.4\sin(2\pi\nu) + 0.7,\\
&A_1^{(\alpha)}(\nu, \omega) = 0.5A^{(\mu)}(\nu, \omega), \qquad A_2^{(\alpha)}(\nu, \omega) = 1.5A^{(\mu)}(\nu, \omega), \qquad A_3^{(\alpha)}(\nu, \omega) = 0,\\
&A_1^{(\beta)}(\nu, \omega) = \left[1.3\cos(2\pi\omega)\right]^2 + 0.4\cos(2\pi\nu) + 0.8,\\
&A_2^{(\beta)}(\nu, \omega) = 2A_1^{(\beta)}(\nu, \omega), \qquad A_3^{(\beta)}(\nu, \omega) = 0,\\
&A_{11}^{(\psi)}(\nu, \omega) = A_{12}^{(\psi)}(\nu, \omega) = \left[0.6\sin(2\pi\omega)\right]^2 + 0.4\sin(2\pi\nu) + 1,\\
&A_{21}^{(\psi)}(\nu, \omega) = A_{22}^{(\psi)}(\nu, \omega) = 2\left\{\left[0.6\sin(2\pi\omega)\right]^2 + 0.4\sin(2\pi\nu) + 1\right\},
\end{aligned}$$

and in addition, $A_{13}^{(\psi)} = A_{13}^{(\psi)} = A_{31}^{(\psi)} = A_{32}^{(\psi)} = A_{33}^{(\psi)} = 0$. Then, following Guo and Dai (2006), a finite sample of time series $\{X_{jkrt}, t = 1, \ldots, T\}$ are simulated as

$$X_{jkrt} = \sum_{s=1}^{T} A_{jkr}\left(\frac{t}{T}, \frac{s}{T}\right) \exp\left(\frac{2\pi st}{T} i\right) Z_{jkr}(s), \tag{10}$$

where $Z_{jkr}(s)$ are i.i.d. complex Gaussian variates with mean zero and variance $1/T$ and $Z_{jkr}(s) = Z_{jkr}^*(T-s)$ for $s/T \neq 0, 0.5$, and $Z_{jkr}(s)$ are i.i.d. real-valued Gaussian variates with zero mean and variance $1/T$ for $s/T = 0, 0.5$.

TABLE 1

Based on 100 *repetitions, mean of average squared errors (ASE* $\times 10^2$*) obtained through the proposed ANOPOW model based on the approximation of the large-sample distribution of the local log periodograms (ANOPOW-A), the proposed ANOPOW model that directly utilizes the true large-sample distribution of the local log periodograms (ANOPOW-D), and the alternative smoothed ordinary least squares method (SmoothOLS) are reported. The smallest mean ASE for each component within each setting is highlighted in bold*

n	Grid	Method	μ	α_1	α_2	β_1	β_2	ψ_{11}	ψ_{12}	ψ_{21}	ψ_{22}
5	(16×64)	ANOPOW-A	5.8	4.7	4.5	4.2	4.2	8.2	8.4	8.4	8.3
		ANOPOW-D	**4.9**	**4.1**	**4.0**	**3.4**	**3.4**	**6.3**	**6.7**	**6.6**	**6.6**
		SmoothOLS	27.8	14.5	14.5	14.2	14.1	27.2	27.1	27.1	27.0
	(32×32)	ANOPOW-A	4.5	4.4	**4.2**	9.5	9.6	5.9	5.9	5.7	5.8
		ANOPOW-D	**3.9**	**4.2**	**4.2**	**8.2**	**8.1**	**4.5**	**4.6**	**4.7**	**4.6**
		SmoothOLS	27.4	14.8	14.7	18.5	17.9	28.5	28.5	28.3	28.9
	(64×16)	ANOPOW-A	5.7	6.7	**6.3**	26.7	26.5	7.3	7.3	7.2	7.1
		ANOPOW-D	**5.5**	**6.4**	**6.3**	**24.0**	**24.0**	**6.5**	**6.6**	**6.5**	**6.4**
		SmoothOLS	33.1	21.1	20.7	40.3	40.1	36.3	36.1	35.9	36.6
10	(16×64)	ANOPOW-A	4.7	3.5	3.6	3.1	3.1	5.8	6.0	5.8	5.9
		ANOPOW-D	**4.2**	**3.4**	**3.4**	**2.5**	**2.6**	**4.7**	**5.0**	**4.7**	**4.8**
		SmoothOLS	24.7	7.9	8.0	7.7	7.5	13.8	14.0	14.0	13.9
	(32×32)	ANOPOW-A	3.7	**3.4**	**3.4**	7.1	7.1	3.7	3.6	3.5	3.6
		ANOPOW-D	**3.4**	3.5	3.5	**6.0**	**5.9**	**2.8**	**2.8**	**2.6**	**2.7**
		SmoothOLS	24.0	7.8	7.8	11.1	10.9	14.8	14.7	14.6	14.5
	(64×16)	ANOPOW-A	**5.2**	**5.2**	**5.3**	22.5	22.7	6.0	6.1	6.2	6.0
		ANOPOW-D	**5.2**	5.5	5.4	**19.0**	**19.2**	**5.2**	**5.5**	**5.6**	**5.4**
		SmoothOLS	29.3	12.8	12.9	31.6	31.5	18.9	18.7	18.9	18.6

We fix $T = 1024$ and consider two different sample sizes $n_{jk} = n = 5, 10$. In order to investigate the impact of different block sizes, the raw local periodograms under three different equally-spaced time-frequency grids $\{(\nu_\ell, \omega_s), \ell = 1, \ldots, L, s = 1, \ldots, S\}$, including $(L \times S) = (16 \times 64)$, (32×32), and (64×16) are obtained. Average squared error (ASE) is used to assess performance of these functional effect estimators. ASE is obtained by averaging the squared errors across the discrete time-frequency grid. For example, the ASE of $\hat{\alpha}_1(\nu, \omega)$ is computed as $\text{ASE} = [T(S+1)]^{-1} \sum_{t=1}^{T} \sum_{s=0}^{S} [\hat{\alpha}_1(t/T, \omega_s) - \alpha_1(t/T, \omega_s)]^2$. Simulation results are reported based on 100 repetitions. In addition to the mean ASEs reported in the Table 1, violin plots are used to visualize the distribution of the ASEs and are provided in Section S1.5 of the Supplementary Material (Li, Yue and Bruce (2024a)). Simulation studies on an additional time-domain model, an extension of the mixed-effect model in Krafty, Hall and Guo (2011), and simulations to evaluate the effectiveness of the zero-level contour avoiding function in identifying time-frequency locations where functional effects are statistically significant, are also presented in Section S1 of the Supplementary Material (Li, Yue and Bruce (2024a)).

The true functional effects and the corresponding ANOPOW-D estimates, based on a single realization of this process with time-frequency grid $(L \times S) = (16 \times 64)$ and $n = 10$, are provided in the Supplementary Material (Li, Yue and Bruce (2024a)), which demonstrates that the proposed approach can accurately capture the complex time-varying dynamics of the functional effects. The mean of ASEs for estimating the time-varying functional effects are presented in Table 1. Several important observations can be concluded from the results. First, estimation accuracy across all methods improves as n increases. Second, the proposed methods, namely, ANOPOW-A and ANOPOW-D, have smaller ASEs, compared to SmoothOLS, for estimating all time-varying functional effects in all settings. Third, ANOPOW-D has

smaller ASEs for estimating $\beta_k(\nu,\omega)$ and $\psi_{jk}(\nu,\omega)$ and has similar ASEs for estimating $\mu(\nu,\omega)$ and $\alpha_j(\nu,\omega)$, compared with ANOPOW-A. Note that $\beta_k(\nu,\omega)$ and $\psi_{jk}(\nu,\omega)$ exhibit more dynamic changes over frequencies than that of $\mu(\nu,\omega)$ and $\alpha_j(\nu,\omega)$. However, these functional effects have the same rate of change in their time-varying behavior. This indicates that ANOPOW-D has improved estimation accuracy for estimating more dynamic frequency patterns, compared to ANOPOW-A, which shows the advantages of utilizing the full distributional properties of local periodograms over the Gaussian approximation for estimating complex frequency dynamics. Lastly, for all methods considered, using the time-frequency grid $(L \times S) = (32 \times 32)$ produces the smallest ASEs for estimating $\mu(\nu,\omega)$, $\alpha_j(\nu,\omega)$, and $\psi_{jk}(\nu,\omega)$, while using the time-frequency grid $(L \times S) = (16 \times 64)$ has the smallest ASEs for estimating $\beta_k(\nu,\omega)$. This can also be attributed to the fact that $\beta_k(\nu,\omega)$ has more complex local frequency patterns, which requires higher resolution across frequencies to obtain accurate estimates.

7. Analysis of seismic signals. We use the proposed methodology to analyze the seismic signals described in Section 2.1. We consider the following one-way log-spectral ANOPOW model:

$$\log f_{jr}(\nu,\omega) = \mu(\nu,\omega) + \alpha_j(\nu,\omega), \quad j = 1,2, r = 1,\ldots,8,$$

where $\alpha_j(\nu,\omega)$ is the main effect function for each type of signal. We let $j = 1$ representing the mining explosion signals and $j = 2$ representing the earthquake signals. For identifability, let $\alpha_2(\nu,\omega) = 0$ such that $\alpha_1(\nu,\omega)$ represents the change in the power spectra for explosions vs. earthquakes and $\mu(\nu,\omega)$ represents the time-varying power spectra for earthquakes. The proposed INLA estimation procedure with time-frequency grid (32×64) is used to fit the model. However, the results appear to be insensitive to the choice of time-frequency grid, such as (64×32) (not shown). Goodness-of-fit analyses using graphical posterior predictive checks (Gabry et al. (2019)) and the cross-validated probability integral transform (Held, Schrödle and Rue (2010)) are also presented in Supplementary Material (Li, Yue and Bruce (2024a)) to provide evidence for model adequacy.

Figure 3 presents the posterior means for the group-level time-varying log spectra of earthquake and explosion signals, the main effect function $\alpha_1(\nu,\omega)$, and the corresponding 95% level-zero contour avoiding function. A positive value of $\alpha_1(\nu,\omega)$ implies that the spectral power of explosions is greater than that of earthquakes at (ν,ω) and vice versa. At the beginning of the P phase (zero to 10 seconds), the spectral power of explosions is significantly greater than that of earthquakes for nearly the entire frequency range. For the rest of the P phase (10–25 seconds), the spectral power of explosions is greater than that of earthquakes in the range of frequencies 5 Hz $\leq \omega \leq$ 20 Hz, which means explosions have relatively stronger high-frequency power. For the S phase, the two groups differ over the lower-to-middle frequency range. More specifically, the spectral power of earthquakes is greater than that of explosions at the lower frequencies (0 Hz $\leq \omega \leq$ 4 Hz) and is less than that of explosions at middle frequencies (4 Hz $\leq \omega \leq$ 15 Hz). These results are slightly different from the findings in Stoffer et al. (2010), which indicate the two groups differ over the entire frequency range. One possible explanation for the differences between our results and those of Stoffer et al. (2010) is that they considered only a portion of the S phase and treated this portion as a stationary time series, which may not reflect the underlying local frequency patterns.

8. Analysis of pupil diameter time series in children with ADHD. We use the proposed methodology to analyze the pupil diameter time series described in Section 2.2. Our analysis considers eight-second pupil diameter time series sampled at 1000 Hz during the visuospatial working memory task. Pupil time series were down-sampled to 62.5 Hz, detrended,

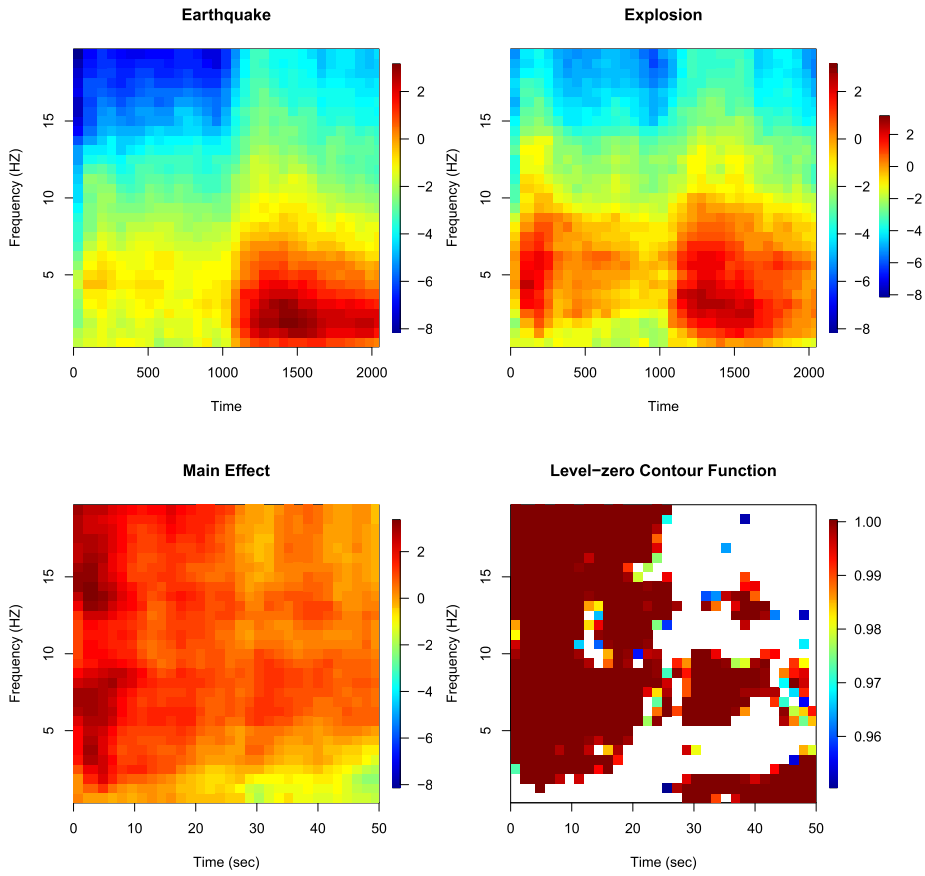

FIG. 3. *Row* 1: *Estimated time-varying log spectra of earthquake and explosion signals*. *Row* 2: *Estimated functional main effect and its* 95% *level-zero contour avoiding function*.

and filtered (see Figure 2). Interested readers can find more details on the experimental design, working memory task, and data processing in the data publication (Rojas-Líbano et al. (2019)). In our analysis missing values were estimated using linear spline interpolation, and trials with more than 50% of the data missing were excluded from the analysis. Based on the nature of this study, we consider the following two-way functional mixed-effect ANOPOW model:

$$\log f_{jkqr}(\nu,\omega) = \mu(\nu,\omega) + \alpha_j(\nu,\omega) + \beta_k(\nu,\omega) + \eta_{jkq}(\nu,\omega)$$

for $q = 1,\ldots,n_{jk}$, $j = 1,2,3$, $k = 1,2$, and $r = 1$. We denote α_j as the main functional effects of group assignments with $j = 1,2,3$ representing ADHD, mADHD, and healthy controls, respectively; β_k represents the main functional effects of cognitive load with $k = 1,2$ representing two-dot array and one-dot array, respectively; η_{jkq} represents the random effects whose covariance accounts for the variability in the subject-level log-spectra. It should be noted that $\eta_{1kq} = \eta_{2kq}$ for the 17 subjects who completed the tasks both on and off medication. Under this design $n_{11} = n_{12} = 28$, $n_{21} = n_{22} = 17$, and $n_{31} = n_{32} = 22$. For identifiability we set $\alpha_3(\nu,\omega) = 0$ and $\beta_2(\nu,\omega) = 0$.

The results of our analysis are presented in the following sections. In Section 8.1 we examine the estimated mean function and functional effects as time-frequency surfaces. In Section 8.2 we explore time-varying power collapsed within certain frequencies of the functional effects. Goodness-of-fit analyses using graphical posterior predictive checks (Gabry et al. (2019)) and the cross-validated probability integral transform (Held, Schrödle and Rue

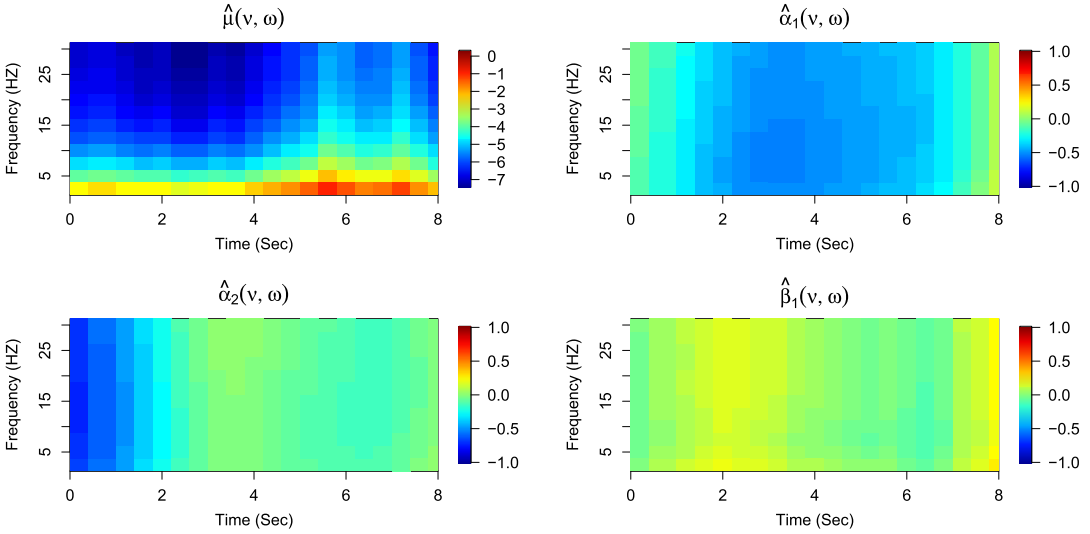

FIG. 4. *Estimated functional effects of pupil diameter time series where* $\hat{\mu}(v, \omega)$ *is the estimated functional grand mean,* $\hat{\alpha}_1(v, \omega)$ *is the estimated functional effect of ADHD (relative to Control),* $\hat{\alpha}_2(v, \omega)$ *is the estimated functional effect of mADHD (relative to Control), and* $\hat{\beta}_1(v, \omega)$ *is the estimated functional effects of cognitive load (two-dot relative to one-dot).*

(2010)) are also presented in Supplementary Material (Li, Yue and Bruce (2024a)) to provide evidence for model adequacy.

8.1. *Time-varying functional effects.* Figure 4 displays the estimates of functional effects of the mean, ADHD, mADHD, and cognitive load, respectively, while Figure 5 presents the corresponding 95% zero-level contour avoiding functions. Four important findings can be summarized from the time-varying functional effects; two findings are confirmatory and supported in the existing literature, and two are new, previously unexplored findings made possible by estimation of the time-varying functional effects and the proposed ANOPOW modeling framework. First, the mean function shows an increase in log power at all frequencies, especially from zero Hz to 20 Hz, after the "probe" image appears (five seconds after the

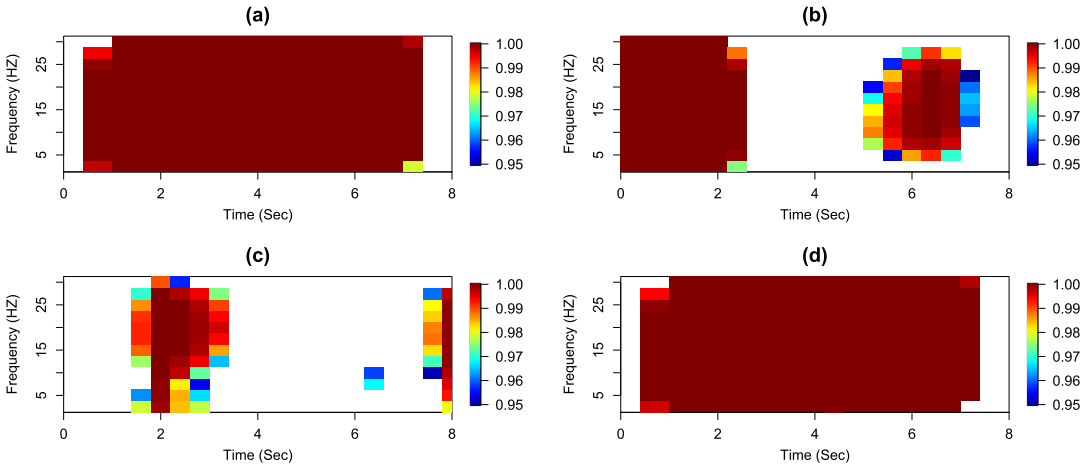

FIG. 5. *The 95% zero-level contour avoiding functions*: (*a*) *ADHD functional effects*, (*b*) *mADHD functional effects*, (*c*) *cognitive load functional effects, and* (*d*) *contrast between ADHD group and mADHD group. Red areas correspond to time-frequency points with values that are significantly different from zero.*

task starts). This finding is in line with studies in the physiological literature (Nakayama and Shimizu (2004)). An increase in power of the pupillary signal corresponds to an increase in mental workload, further indicating that retrieving information from memory requires more mental workload compared to memorization. Second, the zero-level contour avoiding functions indicate significant differences in power spectra across these groups. In particular, children with ADHD are associated with decreased power across all frequencies and times compared to healthy controls, which may provide a way of diagnosing ADHD using pupil diameter time series. This finding suggests that healthy controls exhibit greater variability in pupil response to external stimuli during visuospatial tasks.

Third, the pupil diameter time series dynamics for the mADHD group more closely resemble those of the control group, compared to the ADHD group. This can be seen more clearly in Figure 4, where $\hat{\alpha}_1(\nu, \omega)$ represents the estimated functional effect for ADHD relative to control and $\hat{\alpha}_2(\nu, \omega)$ represents the estimated functional effect for mADHD relative to control. Note that cognitive load is expected to be highest just before (three to five seconds) and just after (five to seven seconds) the "probe" image is displayed. During this time, $|\hat{\alpha}_1(\nu, \omega)| > |\hat{\alpha}_2(\nu, \omega)|$ across all frequencies, which indicates that the group-level power spectrum for the mADHD group more closely resembles that of the control group. This provides empirical evidence that medication modulates the dynamics of pupil diameter time series such that these dynamics more closely resemble healthy controls under cognitive load during visuospatial tasks in children with ADHD. While previous studies using this dataset (Wainstein et al. (2017)) have found similarities in pupil response between mADHD and controls after probe image onset, our findings reveal broader similarities between the two groups before and after probe image onset, which together represent period of high cognitive load during the task. Lastly, the functional effects of cognitive load are significantly greater than zero during the memorization portion of the working memory task (one to three seconds). We see that $\beta_1(\nu, \omega) > 0$ for all frequencies during this time period, which suggests that memorizing two-dot arrays requires higher cognitive workload, compared to one-dot arrays, since higher cognitive load is associated with increased power.

8.2. *Frequency band measures*. The proposed method provides an estimate of the group-level log spectra over the entire frequency range, allowing one to conduct inference on frequency-collapsed functionals. In particular, power within several frequency bands, including the low-frequency (LF) band (zero to four Hz) and the high-frequency (HF) band (four

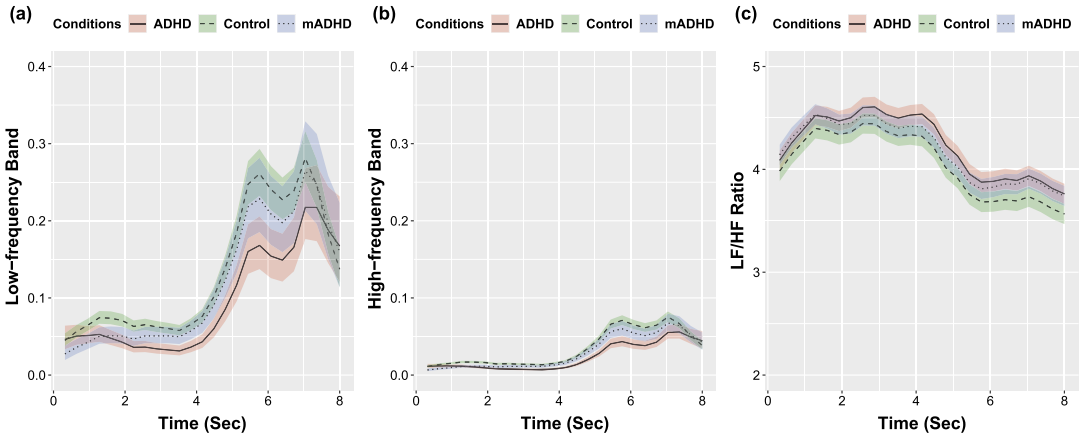

FIG. 6. *Left*: *The low-frequency* (*LF*) *band with* 95% *pointwise credible intervals*. *Middle*: *The high-frequency* (*HF*) *band with* 95% *pointwise credible intervals*. *Right*: *The time-varying LF/HF ratio*.

to eight Hz), provides important information about pupil diameter time series. Frequency-band collapsed measures are computed as integrals of the power spectra. The time-varying LF and HF bands are given by $f^{\mathrm{LF}}(\nu) = \int_0^4 f(\nu, \omega)\, d\omega$ and $f^{\mathrm{HF}}(\nu) = \int_4^8 f(\nu, \omega)\, d\omega$, respectively. Figure 6 presents the time-varying LF band $f^{\mathrm{LF}}(\nu)$ and the HF band $f^{\mathrm{HF}}(\nu)$ of ADHD, mADHD, and control groups along with the 95% simultaneous credible intervals. We see both LF and HF increase after the "probe" image delivery. Our results also show that children with ADHD are associated with decreased LF and HF power, while mADHD and control groups have relatively similar HF and LF. Figure 6 also shows the time-varying LF/HF ratio. Interestingly, the ADHD group has a higher LF/HF ratio than that of mADHD and control groups. Recent studies have shown lower LF/HF ratio of pupil diameter time series to be associated with higher cognitive load (Duchowski et al. (2020)). This is supported by our findings in which you can see a decrease in the LF/HF ratio across all groups during the period just before and after probe image onset (four to six seconds). However, our results show that the control group has significantly lower LF/HF ratios during this period, compared to ADHD. This may suggest that the inability to focus associated with ADHD may lead to lower cognitive load during visuospatial tasks for ADHD children, compared to healthy controls. To the best of our knowledge, this represents a new preliminary finding and possible direction of future scientific research that should be further investigated in future studies.

9. Discussion. This paper introduces a flexible and computationally efficient analysis of power model for analyzing replicated nonstationary time series collected in a designed experiment. The estimation is implemented through INLA, which allows for modeling the exponentially distributed local log periodograms directly, which provides more accurate estimates of functional effects than models using Gaussian approximations of local log periodograms. We conclude this article by discussing some limitations and future research directions. First, our analysis of replicated nonstationary time series assumes that all time series have the same sampling rate. However, time series, in practice, can be sampled at different rates. For example, one could be interested in the joint analysis of different signals, such as heart rate variability, electroencephalograms (EEG), and electromyograms. It would be interesting to develop methods that allow for spectral analysis of time series with different sampling rates. One possible way is incorporating the Bayesian approach of Zhang (2020). Second, the proposed method considers multiple univariate nonstationary time series. However, in many applications, such as EEG and functional magnetic resonance imaging, multivariate or high-dimensional time series are observed. It is of great interest to extend the proposed model to handle replicated multivariate time series. The methods for spectral analysis of multivariate time series, such as the one in Li et al. (2021), provide possibilities to extend our framework to multivariate or high-dimensional time series. Third, the proposed approach utilizes the Fourier transform of individual time series, which is subject to some limitations. For example, it cannot capture oscillatory information beyond the second moment, such as time-irreversibility and kurtosis, and inability to accommodate heavy-tail dependence and infinite variance. One solution to this problem is to adopt the so-called copula spectral density kernel that inherits the robustness properties of quantile regression and does not require moment assumptions (Kley et al. (2016); Li (2023)).

Funding. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM140476 and the National Science Foundation under Grant Number CDS&E-MSS-2152950. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation.

SUPPLEMENTARY MATERIAL

Supplement A: Supplementary material to "ANOPOW for replicated nonstationary time series in experiments" (DOI: 10.1214/23-AOAS1791SUPPA; .pdf). The pdf file contains additional simulation studies (Section S1), goodness-of-fit analyses for the two applications (Section S2), more technical details (Section S3), and the proofs of theorems (Section S4).

Supplement B: R code (DOI: 10.1214/23-AOAS1791SUPPB; .zip). The zip file contains R code to implement the proposed method.

REFERENCES

ADAK, S. (1998). Time-dependent spectral analysis of nonstationary time series. *J. Amer. Statist. Assoc.* **93** 1488–1501. MR1666643 https://doi.org/10.2307/2670062

BOLIN, D. and LINDGREN, F. (2015). Excursion and contour uncertainty regions for latent Gaussian models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 85–106. MR3299400 https://doi.org/10.1111/rssb.12055

BOLIN, D. and LINDGREN, F. (2018). Calculating probabilistic excursion sets and related quantities using excursions. *J. Stat. Softw.* **86** 1–20. https://doi.org/10.18637/jss.v086.i05

BRILLINGER, D. R. (2001). *Time Series: Data Analysis and Theory*. *Classics in Applied Mathematics* **36**. SIAM, Philadelphia, PA. Reprint of the 1981 edition. MR1853554 https://doi.org/10.1137/1.9780898719246

BRUCE, S. A., HALL, M. H., BUYSSE, D. J. and KRAFTY, R. T. (2018). Conditional adaptive Bayesian spectral analysis of nonstationary biomedical time series. *Biometrics* **74** 260–269. MR3777946 https://doi.org/10.1111/biom.12719

CHAU, J. and VON SACHS, R. (2016). Functional mixed effects wavelet estimation for spectra of replicated time series. *Electron. J. Stat.* **10** 2461–2510. MR3545466 https://doi.org/10.1214/16-EJS1181

DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916 https://doi.org/10.1214/aos/1034276620

DIGGLE, P. J. and AL WASEL, I. (1997). Spectral analysis of replicated biomedical time series. *J. R. Stat. Soc., Ser. C* **46** 31–71. With discussion and a reply by the authors. MR1452286 https://doi.org/10.1111/1467-9876.00047

DUCHOWSKI, A. T., KREJTZ, K., GEHRER, N. A., BAFNA, T. and BÆKGAARD, P. (2020). The low/high index of pupillary activity. In *Proceedings of the* 2020 *CHI Conference on Human Factors in Computing Systems* 1–12.

FIECAS, M. and OMBAO, H. (2016). Modeling the evolution of dynamic brain processes during an associative learning experiment. *J. Amer. Statist. Assoc.* **111** 1440–1453. MR3601700 https://doi.org/10.1080/01621459.2016.1165683

FREYERMUTH, J.-M., OMBAO, H. and VON SACHS, R. (2010). Tree-structured wavelet estimation in a mixed effects model for spectra of replicated time series. *J. Amer. Statist. Assoc.* **105** 634–646. MR2724848 https://doi.org/10.1198/jasa.2010.tm09132

FRYZLEWICZ, P. and OMBAO, H. (2009). Consistent classification of nonstationary time series using stochastic wavelet representations. *J. Amer. Statist. Assoc.* **104** 299–312. MR2504379 https://doi.org/10.1198/jasa.2009.0110

GABRY, J., SIMPSON, D., VEHTARI, A., BETANCOURT, M. and GELMAN, A. (2019). Visualization in Bayesian workflow. *J. Roy. Statist. Soc. Ser. A* **182** 389–402. MR3902665 https://doi.org/10.1111/rssa.12378

GUO, W. and DAI, M. (2006). Multivariate time-dependent spectral analysis using Cholesky decomposition. *Statist. Sinica* **16** 825–845. MR2281304

GUO, W., DAI, M., OMBAO, H. C. and VON SACHS, R. (2003). Smoothing spline ANOVA for time-dependent spectral analysis. *J. Amer. Statist. Assoc.* **98** 643–652. MR2011677 https://doi.org/10.1198/016214503000000549

HAMED, A. M., KAUER, A. J. and STEVENS, H. E. (2015). Why the diagnosis of attention deficit hyperactivity disorder matters. *Front. Psychiatry* **6** 168. https://doi.org/10.3389/fpsyt.2015.00168

HELD, L., SCHRÖDLE, B. and RUE, H. (2010). Posterior and cross-validatory predictive checks: A comparison of MCMC and INLA. In *Statistical Modelling and Regression Structures* 91–110. Physica-Verlag, Heidelberg. MR2664630 https://doi.org/10.1007/978-3-7908-2413-1_6

IANNACCONE, R. and COLES, S. (2001). Semiparametric models and inference for biomedical time series with extra-variation. *Biostatistics* **2** 261–276.

KLEY, T., VOLGUSHEV, S., DETTE, H. and HALLIN, M. (2016). Quantile spectral processes: Asymptotic analysis and inference. *Bernoulli* **22** 1770–1807. MR3474833 https://doi.org/10.3150/15-BEJ711

KRAFTY, R. T. and COLLINGE, W. O. (2013). Penalized multivariate Whittle likelihood for power spectrum estimation. *Biometrika* **100** 447–458. MR3068445 https://doi.org/10.1093/biomet/ass088

KRAFTY, R. T., HALL, M. and GUO, W. (2011). Functional mixed effects spectral analysis. *Biometrika* **98** 583–598. MR2836408 https://doi.org/10.1093/biomet/asr032

LI, Z. (2023). Robust conditional spectral analysis of replicated time series. *Stat. Interface* **16** 81–96. MR4465931

LI, Z., BRUCE, S. A., WUTZKE, C. J. and LONG, Y. (2021). Conditional adaptive Bayesian spectral analysis of replicated multivariate time series. *Stat. Med.* **40** 1989–2005. MR4229834 https://doi.org/10.1002/sim.8884

LI, Z. and KRAFTY, R. T. (2019). Adaptive Bayesian time-frequency analysis of multivariate time series. *J. Amer. Statist. Assoc.* **114** 453–465. MR3941268 https://doi.org/10.1080/01621459.2017.1415908

LI, Z., YUE, Y. (R.) and BRUCE, S. A. (2024a). Supplementary materials to "ANOPOW for replicated nonstationary time series in experiments." https://doi.org/10.1214/23-AOAS1791SUPPA.

LI, Z., YUE, Y. (R.) and BRUCE, S. A. (2024b). R code for "ANOPOW for replicated nonstationary time series in experiments." https://doi.org/10.1214/23-AOAS1791SUPPB.

MARTINEZ, J. G., BOHN, K. M., CARROLL, R. J. and MORRIS, J. S. (2013). A study of Mexican free-tailed bat chirp syllables: Bayesian functional mixed models for nonstationary acoustic time series. *J. Amer. Statist. Assoc.* **108** 514–526. MR3174638 https://doi.org/10.1080/01621459.2013.793118

MARTINS, T. G., SIMPSON, D., LINDGREN, F. and RUE, H. (2013). Bayesian computing with INLA: New features. *Comput. Statist. Data Anal.* **67** 68–83. MR3079584 https://doi.org/10.1016/j.csda.2013.04.014

MATHÔT, S. (2018). Pupillometry: Psychology, physiology, and function. *J. Cogn.* **1**.

MORRIS, J. S. and CARROLL, R. J. (2006). Wavelet-based functional mixed models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 179–199. MR2188981 https://doi.org/10.1111/j.1467-9868.2006.00539.x

NAKAYAMA, M. and SHIMIZU, Y. (2004). Frequency analysis of task evoked pupillary response and eye-movement. In *Proceedings of the* 2004 *Symposium on Eye Tracking Research and Applications. ETRA*'04 71–76. Assoc. Comput. Mach., New York, NY, USA. https://doi.org/10.1145/968363.968381

QIN, L., GUO, W. and LITT, B. (2009). A time-frequency functional model for locally stationary time series data. *J. Comput. Graph. Statist.* **18** 675–693. MR2572632 https://doi.org/10.1198/jcgs.2009.06109

R CORE TEAM (2023). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria.

REIMER, J., FROUDARAKIS, E., CADWELL, C. R., YATSENKO, D., DENFIELD, G. H. and TOLIAS, A. S. (2014). Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron* **84** 355–362.

ROJAS-LÍBANO, D., WAINSTEIN, G., CARRASCO, X., ABOITIZ, F. and CROSSLEY, N. (2019). A pupil size, eye-tracking and neuropsychological dataset from ADHD children during a cognitive task. *Sci. Data* **25**. https://doi.org/10.1038/s41597-019-0037-2

ROSEN, O., WOOD, S. and STOFFER, D. S. (2012). AdaptSPEC: Adaptive spectral estimation for nonstationary time series. *J. Amer. Statist. Assoc.* **107** 1575–1589. MR3036417 https://doi.org/10.1080/01621459.2012.716340

RUE, H. and HELD, L. (2005). *Gaussian Markov Random Fields*: *Theory and Applications. Monographs on Statistics and Applied Probability* **104**. CRC Press, Boca Raton, FL. MR2130347 https://doi.org/10.1201/9780203492024

RUE, H., MARTINO, S. and CHOPIN, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 319–392. MR2649602 https://doi.org/10.1111/j.1467-9868.2008.00700.x

SHUMWAY, R. H. and STOFFER, D. S. (2017). *Time Series Analysis and Its Applications*: *With R Examples*, 4th ed. *Springer Texts in Statistics*. Springer, Cham. MR3642322 https://doi.org/10.1007/978-3-319-52452-8

SØRBYE, S. H., HINDBERG, K., OLSEN, L. R. and RUE, H. (2009). Bayesian multiscale feature detection of log-spectral densities. *Comput. Statist. Data Anal.* **53** 3746–3754. MR2749919 https://doi.org/10.1016/j.csda.2009.03.020

STOFFER, D. and POISON, N. (2023). Astsa: Applied statistical time series analysis. R package version 1.15.

STOFFER, D. S., HAN, S., QIN, L. and GUO, W. (2010). Smoothing spline ANOPOW. *J. Statist. Plann. Inference* **140** 3789–3796. MR2674166 https://doi.org/10.1016/j.jspi.2010.04.043

WAHBA, G. (1980). Automatic smoothing of the log-periodogram. *J. Amer. Statist. Assoc.* **75** 122–132.

WAHBA, G. (1990). *Spline Models for Observational Data. CBMS-NSF Regional Conference Series in Applied Mathematics* **59**. SIAM, Philadelphia, PA. MR1045442 https://doi.org/10.1137/1.9781611970128

WAINSTEIN, G., ROJAS-LÍBANO, D., CROSSLEY, N. A., CARRASCO, X., ABOITIZ, F. and OSSANDÓN, T. (2017). Pupil size tracks attentional performance in attention-deficit/hyperactivity disorder. *Sci. Rep.* **8228**. https://doi.org/10.1038/s41597-019-0037-2

WANG, X., YUE, Y. R. and FARAWAY, J. J. (2018). *Bayesian Regression Modeling with INLA. Chapman & Hall/CRC Computer Science and Data Analysis Series*. CRC Press, Boca Raton, FL. MR3752633

WANG, Y., LI, Z. and BRUCE, S. A. (2022). Adaptive Bayesian sum of trees model for covariate-dependent spectral analysis. *Biometrics*. https://doi.org/10.1111/biom.13763

YUE, Y., BOLIN, D., RUE, H. and WANG, X.-F. (2019). Bayesian generalized two-way ANOVA modeling for functional data using INLA. *Statist. Sinica* **29** 741–767. MR3931386

YUE, Y. and SPECKMAN, P. L. (2010). Nonstationary spatial Gaussian Markov random fields. *J. Comput. Graph. Statist.* **19** 96–116. With supplementary data available online. MR2654402 https://doi.org/10.1198/jcgs.2009.08124

ZHANG, S. (2020). Nonparametric Bayesian inference for the spectral density based on irregularly spaced data. *Comput. Statist. Data Anal.* **151** 107019, 14. MR4107729 https://doi.org/10.1016/j.csda.2020.107019