Journal of the American Statistical Association

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Empirical Frequency Band Analysis of Nonstationary Time Series

Scott A. Bruce , Cheng Yong Tang , Martica H. Hall & Robert T. Krafty

To cite this article: Scott A. Bruce , Cheng Yong Tang , Martica H. Hall & Robert T. Krafty (2020) Empirical Frequency Band Analysis of Nonstationary Time Series, Journal of the American Statistical Association, 115:532, 1933-1945, DOI: 10.1080/01621459.2019.1671199

To link to this article: https://doi.org/10.1080/01621459.2019.1671199

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2020, VOL. 115, NO. 532, 1933–1945: Theory and Methods
https://doi.org/10.1080/01621459.2019.1671199

Taylor & Francis
Taylor & Francis Group

Empirical Frequency Band Analysis of Nonstationary Time Series

Scott A. Bruce[a], Cheng Yong Tang[b], Martica H. Hall[c], and Robert T. Krafty[d]

[a]Department of Statistics, George Mason University, Fairfax, VA; [b]Department of Statistical Science, Temple University, Philadelphia, PA; [c]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA; [d]Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA

ABSTRACT

The time-varying power spectrum of a time series process is a bivariate function that quantifies the magnitude of oscillations at different frequencies and times. To obtain low-dimensional, parsimonious measures from this functional parameter, applied researchers consider collapsed measures of power within local bands that partition the frequency space. Frequency bands commonly used in the scientific literature were historically derived, but they are not guaranteed to be optimal or justified for adequately summarizing information from a given time series process under current study. There is a dearth of methods for empirically constructing statistically optimal bands for a given signal. The goal of this article is to provide a standardized, unifying approach for deriving and analyzing customized frequency bands. A consistent, frequency-domain, iterative cumulative sum based scanning procedure is formulated to identify frequency bands that best preserve nonstationary information. A formal hypothesis testing procedure is also developed to test which, if any, frequency bands remain stationary. The proposed method is used to analyze heart rate variability of a patient during sleep and uncovers a refined partition of frequency bands that best summarize the time-varying power spectrum. Supplementary materials for this article are available online.

ARTICLE HISTORY
Received September 2018
Accepted September 2019

KEYWORDS
Frequency band estimation; Heart rate variability; Locally stationary; Spectrum analysis

1. Introduction

The frequency-domain properties of many biomedical time series signals have been found to contain valuable information. Frequency-domain properties of a nonstationary time series are characterized through its time-varying power spectrum, which is a bivariate function of time and frequency. Power at frequencies in close proximity to each other are not always distinguishable across time. Practitioners often use this characteristic to partition the frequency space into frequency bands and create collapsed measures of power within these bands. These frequency bands provide practical, low-dimensional summaries of the information contained in the continuous power spectrum. Some examples of biomedical signals where frequency bands have been developed and found to contain interpretable physiological information are heart rate variability (HRV) (Hall et al. 2004), electroencephalography (EEG) (Klimesch 1999), gait variability (Moore, MacDougall, and Ondo 2008), and center-of-pressure trajectory (Cabeza-Ruiz et al. 2011).

In the scientific literature, standard frequency bands used for analysis have largely been developed through manual inspection of time series data. This is accomplished by noting prominent oscillatory patterns in the data and forming frequency bands that largely account for these dominant waveforms. For example, frequency-domain analysis of HRV data began in the late 1960s (Billman 2011) and led to the development of three primary frequency bands used to summarize power spectra: very low frequency (VLF) ($\leq$ 0.04 Hz), low frequency (LF) (0.04–0.15 Hz), and high frequency (HF) (0.15–0.4 Hz) (Malik et al. 1996). Within these frequency bands, collapsed measures of power are used to summarize the frequency-domain properties of the data and provide a basis for comparison across studies. However, fixed bands that are not allowed to vary across subjects, covariates, experimental variables, or other settings may not adequately summarize the power spectrum under all settings. For example, Klimesch et al. (1998) showed that in the study of EEG signals, different frequency bands within the fixed alpha band (8–13 Hz) reflect quite different cognitive processes. This indicates that the *number* of frequency bands needed to adequately summarize the power spectrum may be influenced by experimental factors. Doppelmayr et al. (1998) also noted that the *endpoints* for the alpha frequency band may vary from subject to subject and proposes a data-adaptive definition for the alpha frequency band that allows for variation across subjects. Both of these examples illustrate the need for a standardized, quantitative approach to frequency band estimation that provides a data-driven determination of the number of frequency bands and their respective endpoints.

In the applied harmonic analysis and signal processing literature, time-frequency analysis of time-dependent signals has been a major focus over the last 50 years. Flandrin (1998) provided an introduction to many of the algorithms and corresponding theories developed, including the short-time Fourier transform, continuous wavelet transform, and Wigner–Ville distribution. These methods allow practitioners to estimate the

CONTACT Robert T. Krafty rkrafty@pitt.edu Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA 15261.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/r/JASA.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/JASA.

time-varying behavior of oscillatory components of time series data, which can be used to understand prominent frequency components of a signal at any given point in time. More specialized methods have also been developed to estimate the time-varying amplitudes and frequencies of the "elementary" oscillations of a signal (Huang et al. 1998; Wu and Huang 2009; Daubechies, Lu, and Wu 2011; Daubechies, Wang, and Wu 2016). While these methods provide quantitative approaches to time-frequency analysis, only a few of the proposals in this space address the need for quantitative frequency band estimation. Tiganj et al. (2012) developed an online algorithm for estimating a single prominent time-varying frequency band, but the simple local band-limited model is not appropriate for signals possibly comprised of multiple band-limited signals. More recently, Cohen, Pollak, and Eldar (2016) provided an algorithm for multiple frequency band estimation assuming the signal is composed of multiple, uncorrelated, cyclostationary bandpass processes. Cyclostationary processes are such that their mean and covariance functions are periodic with respect to time (Gardner, Napolitano, and Paura 2006) and arise in a wide array of applications (e.g., telemetry, radar, sonar, radio astronomy, etc.). However, cyclostationarity may not be a reasonable assumption in the study of biomedical signals where aperiodic nonstationary patterns may also exist.

Time-frequency analysis has also been studied in the statistical literature in the context of evolutionary power spectra (Priestley 1965) and locally stationary processes (Dahlhaus 1997). For example, wavelet-based methods have been successfully developed and applied specifically for the purpose of adaptive estimation of evolutionary power spectra (Neumann and von Sachs 1997). Such methods provide an appropriate level of smoothing across both time and frequency and can reveal important localized features of power spectra. However, such methods do not directly achieve the goal of quantitative frequency band estimation without further development. Rather than focusing on *frequency* localization, many of the proposals have largely been focused on *temporal* localization. For example, temporal change point detection in the power spectrum for nonstationary time series has been widely studied (Adak 1998; Ombao, von Sachs, and Guo 2005; Davis, Lee, and Rodriguez-Yam 2006; Kirch, Muhsal, and Ombao 2015). Recent proposals seek to identify temporal changes as well as the specific spectral characteristics responsible for those changes, such as the auto-spectra and cross-spectra (Preuss, Puchstein, and Dette 2015) and frequency bands of auto-spectra and cross-coherences (Schröder and Ombao 2019). For example, the FreSpeD algorithm of Schröder and Ombao (2019) detects temporal change points and allows for identified change points to be directly attributed to behavior within or between specific frequency bands. This is tremendously useful to practitioners studying biomedical signals who want to determine which frequency band (or bands) are responsible for specific temporal changes in the power spectrum. However, in this method, frequency bands are determined by the user's desired frequency resolution, resulting in many equally sized and spaced frequency bands. It is possible that a more parsimonious segmentation of the frequency space may exist and provide a more appropriate summarization of the power spectrum.

In this article, we introduce a framework for estimating both the number of frequency bands and their endpoints that best preserves the nonstationary dynamics of the power spectrum. First, we introduce a frequency-domain scan and hypothesis test (FRESH) procedure that can be used to detect and validate changes in nonstationary behavior across frequencies. An iterative algorithm is then provided, which uses this procedure to optimally partition the frequency space into bands containing similar nonstationary dynamics. We show that our testing procedure is consistent in detecting a change point in the frequency domain due to changes in the time-varying characteristics across frequency components. Finally, another test statistic is offered to determine which frequency bands, if any, remain "stationary" in the sense that the power spectrum is constant through time. We refer to the methodology presented in this article as empirical band analysis (EBA), since it offers a data-driven approach to frequency band estimation, which provides a parsimonious representation of the time-varying frequency-domain characteristics of a time series.

This article makes several novel contributions. First, the proposed framework offers a significant contribution to the scientific literature by providing a convenient and justifiable approach to customized frequency band detection, as an alternative to the standard practice of subjective determination through visual inspection. The framework is applicable to a wide array of practical settings where time-varying dynamics exist. Second, the statistical methodology introduced in this article also broadens the scope of change point detection problems that can be addressed by accounting for time-varying dynamics through time-domain localization of the FRESH statistic used in the FRESH procedure. This novel extension allows for detection of changes across frequencies in the *time-varying* behavior of the power spectrum and varies significantly from traditional approaches that aim to assess changes in a single dimension irrespective of other dimensions. The associated technical development itself is also novel as it incorporates both time-domain localization and frequency-domain signal extraction. Finally, the proposed detection algorithm properly accounts for dependence in power spectrum estimates among neighboring frequencies and can be adapted to the specific signal and settings under study for improved estimation.

The rest of the article is structured as follows. Section 2 introduces the locally stationary time series model and offers a definition of the frequency-banded power spectrum. Section 3 provides an overview of the local multitaper spectral estimator, which is used as an initial step in the FRESH procedure. It should be noted that an EBA can be conducted using any initial consistent time-varying spectral estimator, but this article focuses on the local multitaper estimator due to favorable empirical and theoretical properties. Section 4 details the components of the EBA analytical framework, including the FRESH statistic and its properties, an iterative algorithm for identifying multiple frequency bands, and another test statistic to determine if the power spectrum within a frequency band is constant in time. Section 5 contains simulation results to evaluate the empirical performance of the search algorithm in estimating the number of frequency bands and their endpoints. Section 6 provides an application of the EBA frequency band estimation and testing

framework on full night HRV data for an individual participating in a sleep study for older adults who are the primary caregiver for an ill spouse. Section 7 offers a discussion of the results and concluding remarks. Proofs of the theorems are provided in the supplementary materials of this article.

2. Model

2.1. Locally Stationary Time Series

We consider modeling a zero-mean locally stationary time series, X_t, $t = 1, \ldots, T$, through its time-varying Cramér representation (Priestley 1965),

$$X_t = \int_{-1/2}^{1/2} A(t/T, \omega) \exp(2\pi i t\omega) dZ(\omega), \tag{1}$$

where $A(u, \omega)$ is a complex-valued function of scaled time $u \in [0, 1]$ and frequency $\omega \in [-1/2, 1/2]$ that is Hermitian and periodic with period 1 as a function of frequency, and where $Z(\omega)$ is a zero-mean orthogonal increment process with unit variance. The time-varying power spectrum is then defined as $f(u, \omega) = |A(u, \omega)|^2$ and can be interpreted as the contribution to the variance of X_t at time uT from oscillations at frequency ω.

To assure tractible estimation and inference, we assume regularity conditions on the distribution of the orthogonal incremental process and on second-order temporal dynamics (Dahlhaus 1997; Guo et al. 2003). First, we assume mixing conditions where all cumulants of dZ exist and are bounded.

Assumption 1 (Mixing). For each $k = 1, 2, \ldots$, there exists a constant $C_k \in \mathbb{R}$ and a function $\Lambda_k : \mathbb{R}^{k-1} \to \mathbb{C}$ such that $\text{cum}\{dZ(\omega_1), \ldots, dZ(\omega_k)\} = \Delta\left(\sum_{j=1}^{k} \omega_j\right) \Lambda_k\left(\omega_1, \ldots, \omega_{k-1}\right) d\omega_1 \cdots d\omega_k$, where $\text{cum}\{\cdots\}$ denotes the cumulant; $\Lambda_1 = 0, \Lambda_2(\omega) = 1, |\Lambda_k(\omega_1, \ldots, \omega_{k-1})| \leq C_k$; and $\Delta(\omega) = \sum_{j=-\infty}^{j=\infty} \delta(\omega + j)$ is the period 2π extension of the Dirac delta function.

Second, we assume that the second-order structure of X_t evolves smoothly over time in the sense that the second derivative of the transfer function with respect to time exists and is bounded.

Assumption 2 (Temporal smoothness).

$$\max_{u \in [0,1],\, \omega \in [-1/2, 1/2]} \left| \frac{\partial^2 A(u, \omega)}{\partial u^2} \right|^2 < \infty.$$

2.2. Frequency-Banded Power Spectrum

Frequency band determination requires assumptions about the power spectrum that constitute a frequency-banded structure. In practice, these assumptions are often made implicitly and may not be stated explicitly. Here, we offer a formal definition for the frequency-banded power spectrum that characterizes sets of frequencies with similar time-varying dynamics.

To characterize the nonstationary behavior of the power spectrum, we consider the demeaned time-varying power spectrum,

$$g(u, \omega) = f(u, \omega) - \int_0^1 f(u, \omega) du, \tag{2}$$

which represents the nonstationary dynamics of the power spectrum beyond a simple level shift. We will assume $g(u, \omega)$ admits a partition of the frequency space such that

$$g(u, \omega) = \begin{cases} g_1(u) & \text{for } \omega \in (0, \omega_1) \\ g_2(u) & \text{for } \omega \in [\omega_1, \omega_2) \\ \vdots & \\ g_p(u) & \text{for } \omega \in [\omega_{p-1}, 0.5). \end{cases} \tag{3}$$

This structure implies the existence of frequency bands where the nonstationary dynamics of the underlying power spectrum among frequency components within each band are the same, aside from a level shift. In this case, a parsimonious segmentation of the frequency space exists, and the goal of our analysis is to estimate the number of frequency bands, p, and the associated partition of the frequency space, $\boldsymbol{\omega}_p = [\omega_1, \omega_2, \ldots, \omega_{p-1}]'$.

3. Time-Varying Spectrum Estimation

The proposed approach for estimating the number and location of partition points for the frequency bands applies an iterative, frequency-domain scan and hypothesis testing procedure to an initial estimate of the time-varying spectrum. As previously discussed, numerous methods exist for estimating a time-varying power spectrum. In this article, we consider the local multitaper estimator, as it provides asymptotically consistent estimates that optimally balance bias and variance in finite samples to offer favorable resolution across frequency (Percival and Walden 1993). As will be demonstrated in Section 4, this estimator also provides frequency-domain scan and hypothesis test statistics with favorable asymptotic characteristics. However, it should be noted that the iterative, frequency-domain scan procedure discussed in Section 4.2 can be applied to any consistent time-varying spectrum estimator. The hypothesis testing procedure discussed in Section 4.1 and the accompanying test of stationarity discussed in Section 4.3 are also applicable to other estimators, if the resulting distribution of the frequency-domain scan and hypothesis test statistics can be obtained.

3.1. Piecewise Stationary Approximation

We begin our estimation procedure by noting that locally stationary time series can be well approximated by piecewise stationary time series (Adak 1998). Accordingly, we can approximate the time-varying power spectrum as a piecewise sum of B stationary power spectra,

$$f(u, \omega) \approx \sum_{b=1}^{B} \mathcal{I}_b(u) f_b(\omega), \tag{4}$$

where the indicator function $\mathcal{I}_b(u) = 1$ if $u \in \left(\frac{b-1}{B}, \frac{b}{B}\right]$ and 0 otherwise. As a result, we are able to estimate the time-varying power spectrum locally using stationary power spectrum estimation procedures within time blocks. Here, we consider equally sized, nonoverlapping temporal blocks (Dahlhaus 1997). The proposed procedure could be extended in future work for local estimators with unequally sized or overlapping time blocks (Adak 1998; Ombao et al. 2001), although the distributional properties of the frequency-domain scan and hypothesis test statistic may become nontrivial.

3.2. Local Spectrum Estimation

For convenience, assume T is a multiple of B, so $T_B = T/B$, the number of observations in each time segment, is integer-valued. The periodogram estimator for the power spectrum of the bth segment,

$$Y_b(\omega) = \frac{1}{T_B}\left|\sum_{t=1}^{T}\mathcal{I}_b(t/T)X_t \exp(-2\pi i\omega t)\right|^2, \quad (5)$$

is an asymptotically unbiased but noisy estimator such that the variance approaches a nonzero constant as $T_B \to \infty$. Further, although it is asymptotically unbiased, in finite samples it can possess a nontrivial bias, a phenomenon commonly referred to as spectral leakage. One way to overcome the limitations of the periodogram estimator is to use the multitaper approach by applying orthogonal tapers to the data, estimating the power spectrum using each of the tapered datasets to generate multiple spectral estimates, then averaging the estimates to obtain the final estimator of the power spectrum (Thomson 1982). The tapering of the time series mitigates finite sample bias and the averaging across multiple estimates with orthogonal tapers provides a consistent estimator.

Let the kth single taper estimator for the power spectrum of the bth stationary segment, also known as the kth direct estimator, be defined as

$$\hat{f}_b^{(k)}(\omega) = \left|\sum_{t=1}^{T} v_b^k(t)\mathcal{I}_b(t/T)X_t \exp(-2\pi i\omega t)\right|^2 \text{ for } k = 1,\ldots,K, \quad (6)$$

where $\mathbf{v_b^k} = \left[v_b^k(1), v_b^k(2),\ldots,v_b^k(T)\right]'$ is the kth data taper for the bth time segment. Then let the multitaper estimator be the average of the K single taper estimates

$$\hat{f}_b^{(mt)}(\omega) = \frac{1}{K}\sum_{k=1}^{K}\hat{f}_b^{(k)}(\omega). \quad (7)$$

We propose the use of sinusoidal tapers of the form

$$v_b^k(t) = \sqrt{\frac{2}{T_B+1}}\sin\frac{\pi k[t-(b-1)T_B]}{T_B+1}, \text{ for } k = 1,\ldots,K, \quad (8)$$

which are orthogonal for $t \in [(b-1)T_B+1,\ldots,bT_B]$. Sinusoidal tapers are more computationally efficient than the Slepian tapers proposed in Thomson (1982), which require numerical eigenvalue decomposition to construct the tapers, and can achieve similar spectral concentration with significantly less local bias (Riedel and Sidorenko 1995). Another advantage is that the bandwidth, $2W$, which is the minimum separation in frequency between approximately uncorrelated spectral estimates, can be fully determined by setting the number of tapers, K, appropriately (Walden, McCoy, and Percival 1995). More specifically,

$$2W = \frac{K+1}{T_B+1}. \quad (9)$$

Finally, we can estimate the time-varying power spectrum by obtaining the multitaper spectral estimate on each local segment,

$$\hat{f}^{(mt)}(u,\omega) = \sum_{b=1}^{B}\mathcal{I}_b(u)\hat{f}_b^{(mt)}(\omega). \quad (10)$$

In practice, this estimator requires appropriate selection of two tuning parameters: B and K. The number of time blocks, B, balances frequency and temporal properties. It should be selected small enough so that there are enough time points per block, $T_B = T/B$, to ensure sufficient frequency resolution and that the central limit theorem holds for the local tapered periodograms. It should be selected large enough so that data within each block are approximately stationary. In many applications, scientific guidelines exist for the selection of B. For example, it is recommended to use time blocks 2–5 min in length for analyzing HRV time series (Malik et al. 1996). Asymptotic results presented in Section 4 indicate that the optimal rate for B is $T^{1/2}$. In the absence of scientific guidelines, we recommend selecting B as the factor of T closest to $T^{1/2}$. The number of tapers, K, controls the smoothness of local spectral estimates. General guidance for multitaper estimation is to select K such that the product of the number of Fourier frequencies and the bandwidth, which is a function of K, is equal to some predefined value, the popular default of which is 4 (Percival and Walden 1993). In our setting, this leads to selecting $K = 8(T+B)/T - 1$.

4. Empirical Band Analysis

4.1. Frequency Scan and Hypothesis Test (FRESH) Statistic

A natural way to identify partition points separating frequency bands with different nonstationary dynamics is to take the sum of squared differences between spectral estimators from neighboring frequencies, or frequency bands. We introduce a FRESH statistic that can be used to search the frequency space for potential partition points.

For each time segment, $b = 1,\ldots,B$, define the Fourier frequencies as $\omega_k = k/T_B$, $k = 1,\ldots,N_B$, where $N_B = \lfloor T_B/2 \rfloor - 1$. The FRESH statistic over the frequency band of width δ starting at ω_0 is defined as

$$Q(\omega_0,\delta) = \sum_{b=1}^{B}\left[\hat{g}(b/B,\omega_0+\delta) - \bar{g}_{[\omega_0,\omega_0+\delta)}(b/B,\cdot)\right]^2, \quad (11)$$

where $\hat{g}(b/B,\omega) = \hat{f}^{(mt)}(b/B,\omega) - B^{-1}\sum_{b=1}^{B}\hat{f}^{(mt)}(b/B,\omega)$ is the estimator of the demeaned time-varying power spectrum, and $\bar{g}_{[\omega_0,\omega_0+\delta)}(b/B,\cdot)$ is the average of the demeaned time-varying power spectrum estimates for $\omega \in [\omega_0,\omega_0+\delta)$. In

practice, we can consider fixing ω_0 and varying δ to detect changes in the frequency space. Intuitively, we expect $Q(\omega_0, \delta)$ to be "large" when a partition point of the frequency space exists at $\omega_0 + \delta$. To illustrate this point, we consider the behavior of the test statistic in the special case where only one partition point exists.

Theorem 1 (Stochastic boundedness). Let X_t, t=1,…,T, be a locally stationary time series satisfying Assumptions 1 and 2, and where its demeaned time-varying spectrum consists of two frequency bands such that

$$g(u,\omega) = \begin{cases} g_1(u) \text{ for } \omega \in [0, \omega^*) \\ g_2(u) \text{ for } \omega \in [\omega^*, 0.5] \end{cases}$$

for $u \in [0,1]$. Then

$$Q(\omega_0,\delta) = \begin{cases} O_p(B^{-1}) + O_p(B/T) + O_p(BK/T) + O_p(K^{-1/2}) \\ \quad \text{for } \omega_0 < \omega_0 + \delta < \omega^* \\ B\left(\frac{\omega^*-\omega_0}{\delta}\right)^2 \int \left[g_2(u) - g_1(u)\right]^2 du \\ \quad + O_p(B^{-1}) + O_p(B/T) + O_p(BK/T) \\ \quad + O_p(K^{-1/2}) \text{ for } \omega_0 < \omega^* \le \omega_0 + \delta. \end{cases}$$

The first asymptotic term in Theorem 1, $O_p(B^{-1})$, follows from the piecewise stationary approximation. The next two terms, $O_p(B/T)$ and $O_p(BK/T)$, ensure the total length of the time series, T, grows faster than the number of segments, B, and the number of observations in each segment, $T_B = T/B$, grows faster than the number of tapers used to construct the multitaper estimator, K. These terms ensure the asymptotic distribution for the stationary approximation to the local spectrum in each segment hold (Thomson 1982; Percival and Walden 1993). Finally, the number of tapers, K, needs to grow in order for the variability of the spectrum estimates to tend to zero.

Note that the deterministic component of $Q(\omega_0, \delta)$ is maximized when $\omega_0 + \delta = \omega^*$, given that ω_0 is set to be sufficiently small (see Figure 1). Considering this asymptotic behavior leads to the following potential estimator for the frequency partition point: $\hat{\omega}^* = \omega_0 + \arg\max_{\delta} Q(\omega_0, \delta)$ for ω_0 set appropriately.

Furthermore, we can construct a significance test under the null hypothesis that only one frequency band exists (i.e., no frequency partition points exist) to determine the significance of an estimated frequency partition point, $\hat{\omega}^*$, based on the FRESH statistic.

Theorem 2 (Null distribution). Let $H_0 : g(u,\omega) = g_0(u)$ for $\omega \in [\omega_0, \omega_0 + \delta]$ be true. Under Assumptions 1 and 2, if $T \to \infty, B \to \infty, K \to \infty, \frac{T}{B} \to \infty, \frac{T}{BK} \to \infty$, then

$$Q(\omega_0,\delta) = Z_{\omega_0,\delta,B}[1 + o_p(1)]$$

uniformly in ω_0 and δ where $Z_{\omega_0,\delta,B} \sim \sum_{b=1}^{B} \sigma^2_{\omega_0,\delta,b} \chi^2_{1,b}$, $\chi^2_{1,b}$ $b = 1,\ldots,B$ are iid χ^2_1 random variables, and $\sigma^2_{\omega_0,\delta,b}$ is the variance of $\hat{g}(b/B, \omega_0 + \delta) - \bar{g}_{[\omega_0,\omega_0+\delta)}(b/B, \cdot)$.

Theorem 2 shows that the FRESH statistic is approximately distributed as the sum of scaled central χ^2_1 random variables when only one frequency band exists. We then can approximate the coefficients, $\sigma^2_{\omega_0,\delta,b}$, and then approximate the CDF of the distribution using the Satterhwaite–Welch method (Welch 1938; Satterthwaite 1946) to obtain p-values for given test statistic values (see Section B in the supplementary materials for details). We can also determine the distribution of the FRESH statistic under the specific alternative that two frequency bands exist, and the starting frequency, ω_0, and ending frequency, $\omega_0 + \delta$, are not in the same frequency band (see Theorem 3).

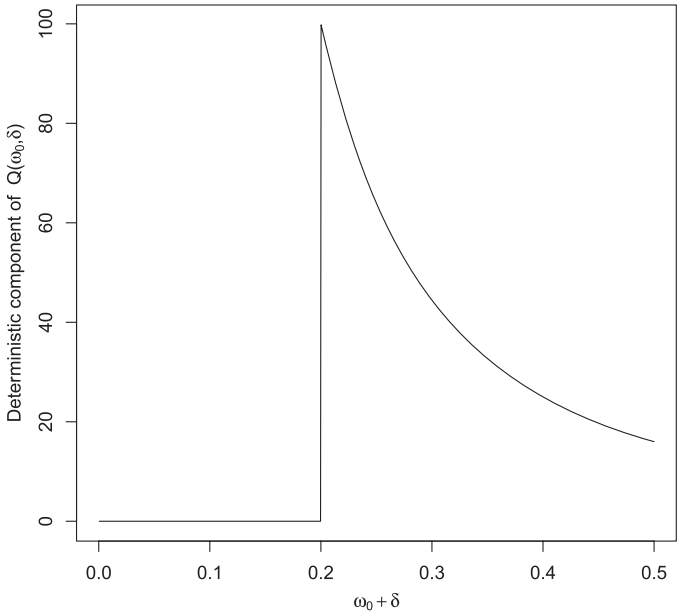


Figure 1. Deterministic component of $Q(0,\delta)$ for $\omega^* = 0.2, B = 100$, and $\int \left[g_2(u) - g_1(u)\right]^2 du = 1$.

Theorem 3 (Alternative distribution, two bands). Let the demeaned time-varying power spectrum consist of two frequency bands as in Theorem 1. Under Assumptions 1 and 2, if $T \to \infty, B \to \infty, K \to \infty, \frac{T}{B} \to \infty, \frac{T}{BK} \to \infty$, then

$$Q(\omega_0,\delta) = Z_{\omega_0,\delta,B}[1 + o_p(1)] \text{ for } \omega_0 < \omega^* \le \omega_0 + \delta,$$

uniformly in ω_0 and δ where $Z_{\omega_0,\delta,B} \sim \sum_{b=1}^{B} \sigma^2_{\omega_0,\delta,b} \chi^2_{1,b}$, $\chi^2_{1,b}$, $b = 1,\ldots,B$, are iid noncentral χ^2_1 random variables with noncentrality parameters

$$\lambda_{\omega^*,\omega_0,\delta,b} = \left(\frac{1}{\sigma^2_{\omega_0,\delta,b}}\right)\left(\frac{\omega^* - \omega_0}{\delta}\right)^2 \left[g_2(b/B) - g_1(b/B)\right]^2$$

and $\sigma^2_{\omega_0,\delta,b}$ is the variance of $\hat{g}(b/B, \omega_0 + \delta) - \bar{g}_{[\omega_0,\omega_0+\delta)}(b/B, \cdot)$.

Note here that the only difference in the distribution of the test statistic for the one frequency band and two frequency band settings is in the noncentrality parameter, λ, for the χ^2 random variables associated with each segment. When only a single frequency band exists, $\lambda = 0$ resulting in a central χ^2 distribution. However, when two frequency bands are present, the noncentrality parameters increase in proportion to the squared difference in the demeaned time-varying power spectrum between the two bands and the inverse of the variance of $\hat{g}(b/B, \omega_0 + \delta) - \bar{g}_{[\omega_0,\omega_0+\delta)}(b/B, \cdot)$.

Since the multitaper power spectrum estimator introduced in Section 3.2 is both consistent and asymptotically independent

across frequencies (Percival and Walden 1993), it can be shown that the demeaned time-varying power spectrum estimator also carries these qualities (see Section A of the supplementary materials) and accordingly the variances, $\sigma^2_{\omega_0,\delta,b}$, tend to zero. This means that the noncentrality parameters tend to infinity, resulting in divergence between the null and alternative distributions described above and an asymptotic power of 1 (see Corollary 1). Proof is available in Section A of the supplementary materials.

Corollary 1 (Power). Under the settings described in Theorem 1, the power to detect the true change point, ω^*, asymptotically tends to 1 such that, for all $q > 0$ and $\omega_0 < \omega^* \leq \omega_0 + \delta$, $P\left[Q(\omega_0, \delta) > q\right] \to 1$.

4.2. Iterative EBA Algorithm

The FRESH statistic and corresponding significance test introduced in Section 4.1 are shown to provide consistent estimation of a single frequency partition point over a particular range of frequencies. We extend this framework to detect multiple frequency partition points by introducing an iterative search algorithm that uses the FRESH statistic to efficiently explore the frequency space and identify all frequency partition points.

There are two significant challenges in extending this framework to allow for the estimation of possibly multiple frequency partition points. First, multiple frequencies need to be tested simultaneously to identify likely partition points, which means the error rate over the set of tests needs to be controlled appropriately. This can be addressed by considering simultaneous testing procedures for identifying likely partition points, such as the Hochberg step up procedure (Hochberg 1988), which controls the family-wise error rate (FWER). Additionally, the possibility of multiple partition points means FRESH statistics may exhibit multiple local maxima corresponding to likely partition points. Accordingly, each iteration of the algorithm is designed to identify the "nearest" frequency partition point to the starting frequency, reset the starting frequency to the newly identified partition point for the next iteration, and repeat this process until all partition points are identified.

The EBA algorithm begins by searching across all frequencies for possible partition points as in Section 4.1. Since multiple partition points may exist, all frequencies are tested for significance against the null distribution defined in Theorem 2, producing a set of FRESH statistics and p-values. The Hochberg step up procedure is then applied as follows to determine which frequencies will serve as candidate partition points. Suppose $p_{(1)}, p_{(2)}, \ldots, p_{(k)}$ are the set of ordered p-values (smallest to largest), and we desire to control FWER at level α. If $p_{(k)} < \alpha$, reject all null hypotheses corresponding to $p_{(1)}, p_{(2)}, \ldots, p_{(k)}$ and stop. Otherwise, "step up" to the next "more significant" p-value, $p_{(k-1)}$. If $p_{(k-1)} < \alpha/2$, reject all null hypotheses corresponding to $p_{(1)}, p_{(2)}, \ldots, p_{(k-1)}$ and stop. Otherwise, step up to the next p-value, $p_{(k-2)}$, and repeat this process. In general, step j determines if $p_{(j)} < \alpha/(k-j+1)$ and either rejects all null hypotheses for p-values $\leq p_{(j)}$ and stops or steps up to the next p-value. The frequencies corresponding to the rejected null hypotheses serve as the set of candidate partition points. If the set is empty, the algorithm stops. Otherwise, the smallest frequency in this set is taken as the first frequency partition point, since it is "nearest" to the starting frequency for the first iteration, $\omega_0 = 1/T_B$, and estimates the endpoint for the first frequency band covering the lowest frequencies.

New FRESH statistics are then computed such that the first frequency partition point serves as the new starting frequency and all larger frequencies are tested to identify additional partition points similarly. Since spectral estimates for frequencies very close to partition points may be correlated, these frequencies are excluded from consideration. This iterative process continues until no new partition points are identified, at which time the algorithm stops. Algorithm 1 provides pseudo code for implementing the iterative search procedure, which can estimate both the number of frequency bands and the partition points defining the frequency bands. A step-by-step illustration of the EBA algorithm using a simulated example dataset is also available in Section C of the supplementary materials to further demonstrate the operating characteristics of the algorithm.

Algorithm 1: EBAsearch

Data: Demeaned time-varying multitaper power spectrum estimates, $\hat{g}(b/B, \omega_k)$, for $b = 1, \ldots, B$ and $\omega_k = k/T_B$ for $k = 1, \ldots, N_B = \lfloor T_B/2 \rfloor - 1$, significance level for hypothesis testing, α, and number of tapers, K.

Result: Estimated number of partition points, $\widehat{p}$, and estimated partition points, $\widehat{\boldsymbol{\omega}}_{\widehat{p}} = \{\widehat{\omega}_1, \widehat{\omega}_2, \ldots, \widehat{\omega}_{\widehat{p}-1}\}$.

$\widehat{p} \leftarrow 1$, $\widehat{\boldsymbol{\omega}}_{\widehat{p}} \leftarrow \{\}$, $\epsilon \leftarrow \frac{K+1}{T_B+1}$, stop $\leftarrow 0$, $\omega_0 \leftarrow \omega_1$

while stop $\neq 1$ **do**

  Compute FRESH statistics, $Q(\omega_0, \delta_k)$, for $\delta_k = \omega_k - \omega_0$ and $\omega_k > \omega_0 + \epsilon$.

  Approximate p-values, p_k, to test $H_{0k} : g(u, \omega) = g(u)$ for $\omega \in [\omega_0, \omega_k]$ following Section B of the supplementary materials.

  Identify ω_k for which H_{0k} are rejected using Hochberg step up procedure controlling FWER at level α, $\mathbf{R}_\alpha = \{\omega_k : H_{0k} \text{ rejected}\}$.

  if $\mathbf{R}_\alpha = \{\}$ **then**

    stop $\leftarrow 1$

  else

    $\hat{\omega}^* \leftarrow \min \mathbf{R}_\alpha$

    $\widehat{\boldsymbol{\omega}}_{\widehat{p}} \leftarrow \widehat{\boldsymbol{\omega}}_{\widehat{p}} \cup \{\hat{\omega}^*\}$

    $\widehat{p} \leftarrow \widehat{p} + 1$

    $\omega_0 \leftarrow \hat{\omega}^*$

  end

end

return $\widehat{p}$, $\widehat{\boldsymbol{\omega}}_{\widehat{p}}$

4.3. Test for Local Stationarity Within Frequency Bands

Using the iterative EBA search algorithm, we are able to identify frequency bands that best preserve the nonstationary dynamics of the time series. In other words, among all frequencies within a given frequency band, the power spectrum should only vary across time. It is natural to then ask the question "which, if any, of the frequency bands are such that the power spectrum is

constant across both time and frequency?" Here, we introduce another test statistic to address this question. More specifically, for a given frequency band, $[\omega_0, \omega_0 + \delta)$, we want to test the hypothesis $H_0 : f(u,\omega) = f_0(u) = f_0$ for $\omega \in [\omega_0, \omega_0 + \delta)$ and $u \in [0,1]$. Note that this hypothesis can be characterized in terms of the demeaned time-varying power spectrum as $H_0 : g(u,\omega) = g_0(u) = 0$ for $\omega \in [\omega_0, \omega_0 + \delta)$ and $u \in [0,1]$. Assuming this to be true, a natural test statistic is the sum of squared values of $\bar{g}_{[\omega_0,\omega_0+\delta)}(b/B, \cdot)$ across time,

$$Q^S(\omega_0, \delta) = \sum_{b=1}^{B} \bar{g}^2_{[\omega_0,\omega_0+\delta)}(b/B, \cdot). \tag{12}$$

When the power spectrum for a given frequency band does not vary through time, then $\hat{g}(u,\omega)$ should be close to zero and thus $Q^S(\omega_0, \delta)$ should take on "small" values. Theorem 4 provides the distribution of this test statistic when the power spectrum is constant through time for a given frequency band.

Theorem 4 (Stationary null distribution). Let $H_0 : g(u,\omega) = g_0(u) = 0$ for $\omega \in [\omega_0, \omega_0 + \delta)$ be true. Under Assumptions 1 and 2, if $T \to \infty, B \to \infty, K \to \infty, \frac{T}{B} \to \infty, \frac{T}{BK} \to \infty$, then

$$Q^S(\omega_0, \delta) = Z^S_{\omega_0,\delta,B}[1 + o_p(1)]$$

uniformly in ω_0 and δ where $Z^S_{\omega_0,\delta,B} \sim \sum_{b=1}^{B} \sigma^{2S}_{\omega_0,\delta,b}\chi^2_{1,b}$, $\chi^2_{1,b}$ $b = 1, \ldots, B$ are iid central χ^2_1 random variables, and $\sigma^{2S}_{\omega_0,\delta,b}$ is the variance of $\bar{g}_{[\omega_0,\omega_0+\delta)}(b/B, \cdot)$.

We can approximate $\sigma^{2S}_{\omega_0,\delta,b}$ $b = 1, \ldots, B$ and obtain approximate p-values in the same manner as for the original FRESH statistic (see Section B of the supplementary materials for details).

5. Simulations

To evaluate the performance of the EBA search algorithm in finite samples, we consider three simulation settings,

$$f_1(u,\omega) = 1 \text{ for } \omega \in (0, 0.5), \tag{13}$$

$$f_2(u,\omega) = \begin{cases} 10 - 9u \text{ for } \omega \in (0, 0.15) \\ 1 \text{ for } \omega \in [0.15, 0.35) \\ 1 + 9u \text{ for } \omega \in [0.35, 0.5), \end{cases} \tag{14}$$

and

$$f_3(u,\omega) = \begin{cases} 5.5 + 4.5\sin(8\pi u - \pi/2) \text{ for } \omega \in (0, 0.15] \\ 5.5 + 4.5\cos(8\pi u) \text{ for } \omega \in (0.15, 0.35] \\ 5.5 + 4.5\cos(16\pi u) \text{ for } \omega \in (0.35, 0.5), \end{cases} \tag{15}$$

for $u \in [0,1]$. For the first setting, we want to demonstrate that the search algorithm does not produce an unreasonable number of false positives when no frequency partition points exist. For the second and third settings, we also want to demonstrate accurate estimation of the number of frequency bands and the corresponding frequency partition for both linear and nonlinear trends, respectively. Plots of individual realizations and multitaper spectrum estimates of these processes are available for illustration in Figures 2 and 3.

To evaluate accuracy in estimating the number of frequency bands, we directly compare the distribution of the estimated number of frequency bands, $\hat{p}$, for many replications to the true values, $p_1 = 1$ and $p_2 = p_3 = 3$. We use the Rand index (Rand 1971) to summarize the similarity between the estimated frequency partition, $\hat{\boldsymbol{\omega}}$, and the true partitions, $\boldsymbol{\omega}_1$ is the empty set and $\boldsymbol{\omega}_2 = \boldsymbol{\omega}_3 = [0.15, 0.35]'$. Let a be the number of pairs of Fourier frequencies in the same frequency band in $\hat{\boldsymbol{\omega}}$ and the same frequency band in $\boldsymbol{\omega}$ and b be the number of pairs of Fourier frequencies in different frequency bands in $\hat{\boldsymbol{\omega}}$ and different frequency bands in $\boldsymbol{\omega}$, then the Rand index, $R(\hat{\boldsymbol{\omega}}, \boldsymbol{\omega}) = (a+b)/\binom{N_B}{2}$, where $N_B = \lfloor T_B/2 \rfloor - 1$ is the number of Fourier frequencies. The Rand index can take on values from 0 to 1 with values close to 1 indicating good estimation of the true frequency partition. In what follows, we fix the FWER control level, $\alpha = 0.05$, and consider performance under different combinations of settings for the total length of the time series, T, the number of blocks assumed to be stationary, B, and the number of tapers used to estimate the power spectrum, K (see Table 1).

For the first simulation setting, Table 1 indicates the majority of replications across all settings correctly estimate only one frequency band. Since the search algorithm is designed to control FWER, the false positive rate remains under control even as the number of distinct frequencies tested increases. In some cases, FWER may be controlled more tightly than $\alpha = 0.05$ due to dependence among significance tests across frequencies (Efron 2007; Storey 2007; Causeur, Kloareg, and Friguet 2009). For the second and third settings, the performance in terms of the Rand index improves as TB^{-1} and B increase, so long as K is set appropriately to maintain a reasonable bandwidth (see Equation (9)). The improvement in performance can be seen in both the estimated number of partition points and the estimated partition through the distribution of Rand index values.

6. Analysis of Heart Rate Variability

Although one experiences reduced consciousness during sleep, the body is far from inactive. During a night of sleep, the body cycles between periods of non-rapid eye movement (NREM) sleep, which contains deep sleep, rapid eye movement (REM) sleep, in which dreaming typically occurs, and wakefulness. The physiological changes experienced by an individual during a night of sleep have been shown to be essential to the rejuvenating properties of sleep. One physiological measure commonly considered by sleep researchers and clinicians is HRV. HRV is a measure of the elapsed time between consecutive heart beats. The power spectrum of HRV provides indirect measures of autonomic nervous system activity and subsequently provides objective measures of stress and arousal (Hall et al. 2004). In this capacity, the time-varying spectrum of HRV has been used to better understand biological pathways connecting poor sleep to poor health.

We consider HRV during a night of sleep from a participant in the AgeWise Caregiver Study. The goal of the AgeWise Caregiver Study is to better understand sleep in individuals who serve as the primary caregiver for a spouse with a dementiating illness, which can guide interventions to improve caregiver

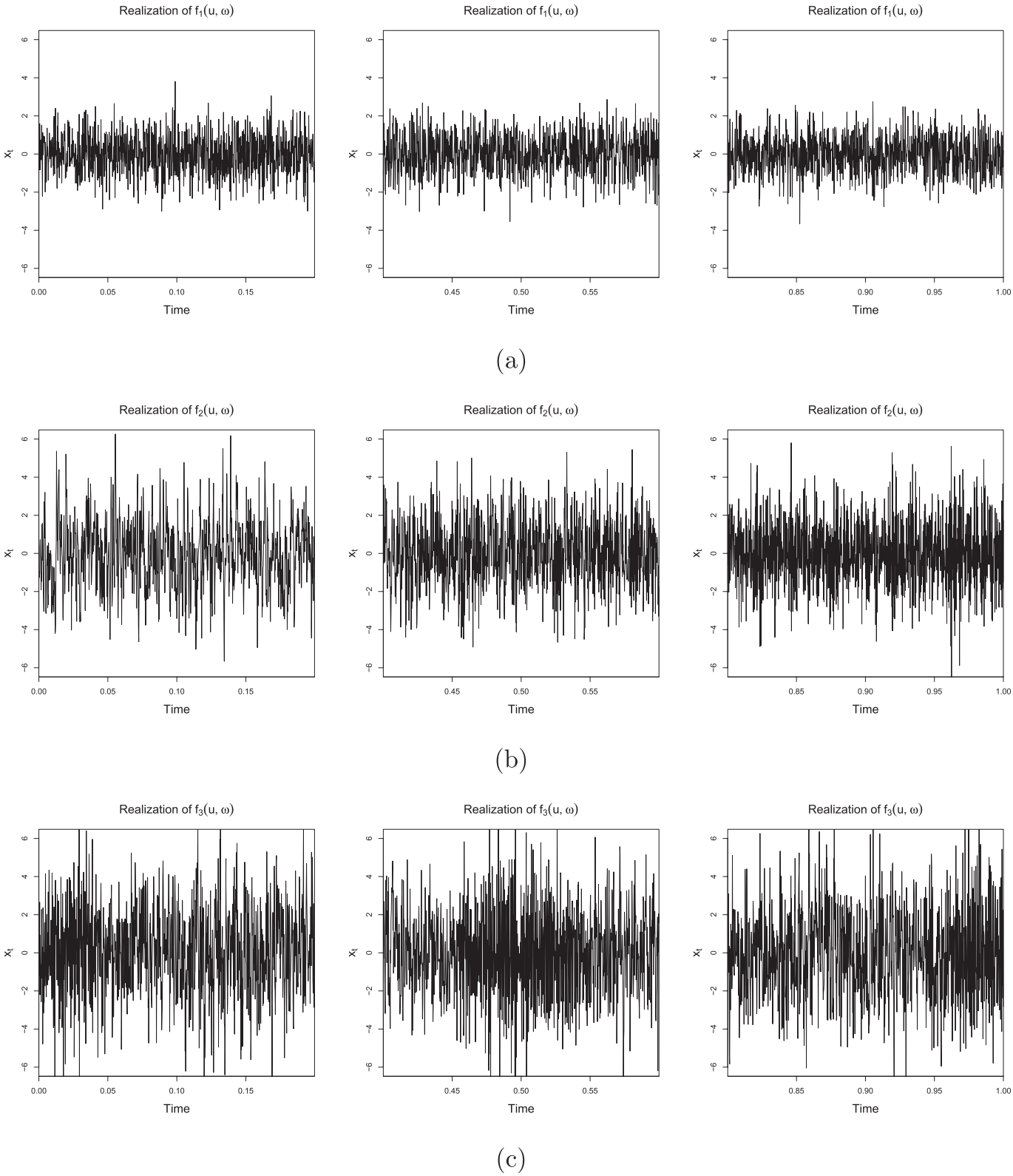


Figure 2. Each row displays the first, middle, and last 1000 observations (from left to right) for single realizations of $f_1(u,\omega)$ (a), $f_2(u,\omega)$ (b), and $f_3(u,\omega)$ (c) of total length $T = 5000$. Individual segments of the same realization are shown to illustrate time-varying behavior.

health through the treatment of poor sleep. The specific goal of the analysis considered here is to characterize the time-varying spectral HRV information for our participant, which can provide insight into his or her pattern of stress and arousal during sleep. The top panel of Figure 4 displays the detrended HRV time series for our participant. The participant was in bed for 8.7 hours, and the interpolated HRV series was sampled at 1 Hz, resulting in a time series of length $T = 31{,}420$. The bottom panel of Figure 4 displays the time series of sleep staging in bouts of REM, NREM and wakefulness for this participant, which

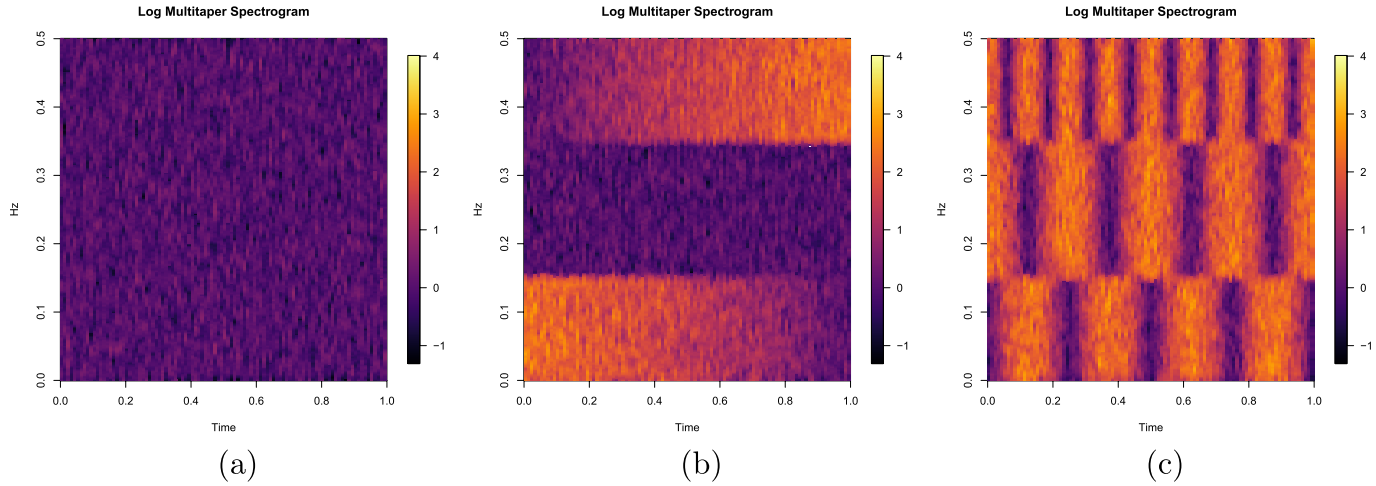


Figure 3. Multitaper estimates of the time-varying log power spectrum for single realizations of $f_1(u,\omega)$ (a), $f_2(u,\omega)$ (b), and $f_3(u,\omega)$ (c) for $T = 100{,}000$, $B = 100$, $K = 15$.

Table 1. Mean(SD) for the estimated number of frequency bands, $\hat{p}$, and Rand index values, $R(\hat{\boldsymbol{\omega}}, \boldsymbol{\omega})$, for $R = 1000$ replications.

Settings			White noise ($p = 1$)		Linear ($p = 3$)		Sinusoidal ($p = 3$)	
K	B	TB^{-1}	$\hat{p}_1$	$R(\hat{\boldsymbol{\omega}}_1, \boldsymbol{\omega}_1)$	$\hat{p}_2$	$R(\hat{\boldsymbol{\omega}}_2, \boldsymbol{\omega}_2)$	$\hat{p}_3$	$R(\hat{\boldsymbol{\omega}}_3, \boldsymbol{\omega}_3)$
10	10	500	1.011(0.104)	0.999(0.013)	2.062(0.310)	0.746(0.048)	2.900(0.622)	0.877(0.085)
		1000	1.005(0.071)	1.000(0.004)	2.028(0.312)	0.742(0.075)	3.001(0.501)	0.920(0.075)
	50	500	1.001(0.032)	1.000(0.003)	3.016(0.148)	0.988(0.022)	3.018(0.133)	0.982(0.007)
		1000	1.000(0.000)	1.000(0.000)	3.009(0.094)	0.994(0.004)	3.012(0.109)	0.991(0.004)
	100	500	1.001(0.032)	1.000(0.004)	3.008(0.089)	0.984(0.006)	3.003(0.055)	0.984(0.005)
		1000	1.000(0.000)	1.000(0.000)	3.003(0.055)	0.993(0.003)	3.001(0.032)	0.992(0.002)
15	10	500	1.005(0.071)	0.999(0.011)	2.652(0.739)	0.836(0.105)	3.197(0.533)	0.919(0.059)
		1000	1.007(0.083)	0.999(0.008)	2.476(0.663)	0.825(0.104)	3.179(0.446)	0.961(0.032)
	50	500	1.002(0.045)	1.000(0.008)	3.038(0.196)	0.969(0.010)	3.035(0.204)	0.975(0.009)
		1000	1.000(0.000)	1.000(0.000)	3.024(0.160)	0.985(0.005)	3.018(0.133)	0.988(0.005)
	100	500	1.000(0.000)	1.000(0.000)	3.016(0.133)	0.961(0.007)	3.016(0.126)	0.970(0.007)
		1000	1.001(0.032)	1.000(0.004)	3.006(0.077)	0.981(0.003)	3.007(0.083)	0.985(0.004)
20	10	500	1.011(0.104)	0.998(0.024)	3.222(0.515)	0.950(0.056)	3.228(0.526)	0.932(0.043)
		1000	1.007(0.083)	0.999(0.009)	3.244(0.532)	0.974(0.025)	3.208(0.503)	0.965(0.030)
	50	500	1.000(0.000)	1.000(0.000)	3.045(0.212)	0.948(0.012)	3.062(0.249)	0.957(0.013)
		1000	1.000(0.000)	1.000(0.000)	3.032(0.176)	0.974(0.006)	3.025(0.169)	0.980(0.006)
	100	500	1.001(0.032)	1.000(0.015)	3.035(0.184)	0.939(0.010)	3.024(0.153)	0.948(0.009)
		1000	1.001(0.032)	1.000(0.004)	3.015(0.137)	0.969(0.010)	3.015(0.122)	0.974(0.005)

was derived from the electroencephalogram (EEG) by a trained sleep technician using established guidelines (Iber et al. 2007). The sleep staging information will be used to closer investigate the results obtained from the EBA of the HRV time series.

To estimate the time-varying power spectrum of HRV, Malik et al. (1996) suggest dividing the full night HRV data into 2–5 min segments. The standard frequency bands used in the scientific literature to summarize the power spectrum of HRV are VLF ($\leq$0.04 Hz), LF (0.04–0.15 Hz), and HF (0.15–0.4 Hz). Figure 5(a) displays the multitaper power spectrum estimates for the full night HRV data for the chosen individual using 5 min intervals along with the traditional frequency bands. While these bands roughly categorize frequencies with different time-varying characteristics, it appears that the number of bands and their associated endpoints may not be very appropriate for summarizing the time-varying dynamics of the power spectrum of this particular dataset. For example, the time-varying characteristics of the higher frequency components within the HF band appear to exhibit different time-varying characteristics than the lower frequency components within the HF band. Additionally, the higher frequency components within the LF band and the lower frequency components within the HF band appear to have similar time-varying characteristics.

To see if a partition of the frequency space exists that can better preserve the nonstationary dynamics of the time series, we apply the iterative EBA search algorithm (see Figure 5(b)). Note that four frequency bands are estimated using the EBA algorithm, which supports the notion that more frequency bands are needed to summarize the time-varying power spectrum appropriately. The EBA algorithm splits the traditional HF band into two bands along the top of the ridge of high power that separates frequencies with different time-varying behavior. Overall, the EBA algorithm provides more homogeneous groupings of frequency components that exhibit similar time-varying dynamics compared to the traditional frequency bands for this particular dataset.

We can also use the test statistic from Section 4.3, $Q^S(\omega_0, \delta)$, to determine which, if any, frequency bands identified by the

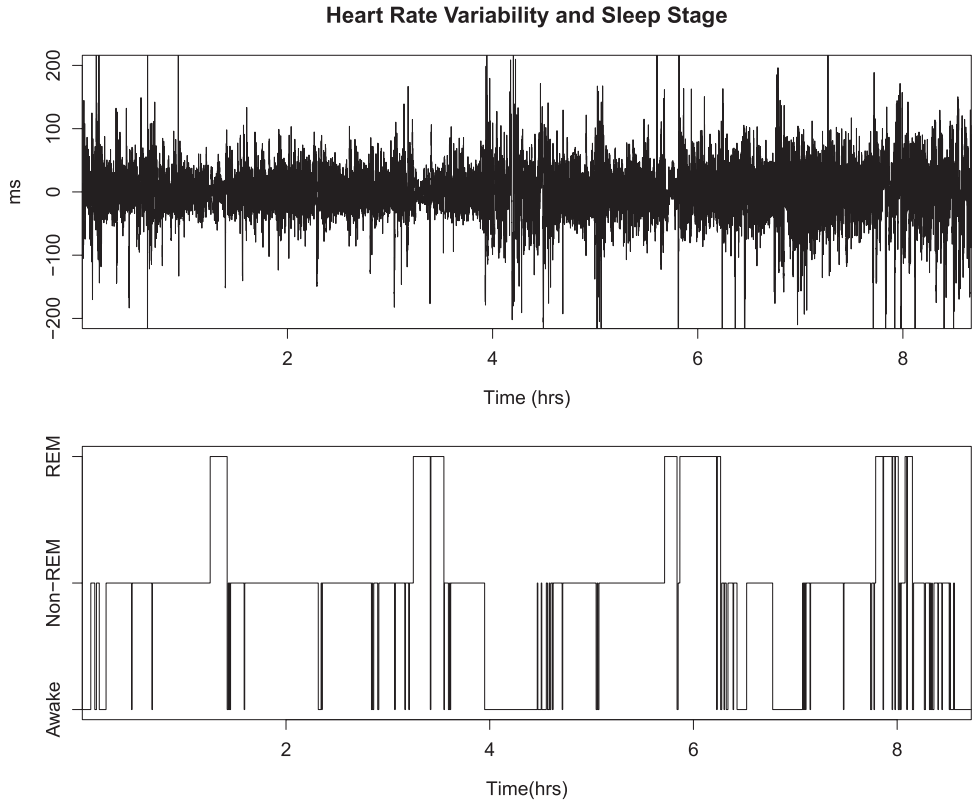
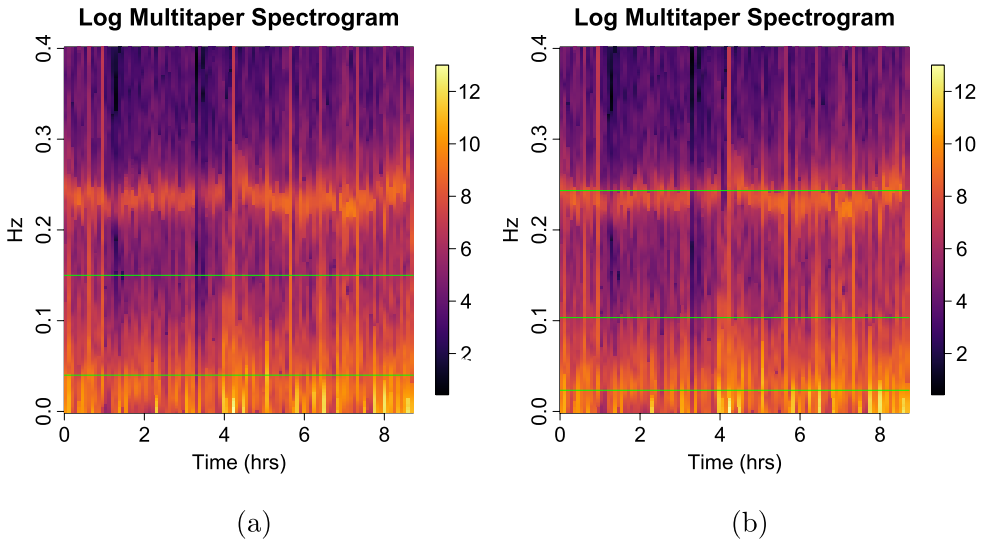


Figure 4. Detrended heart rate variability time series and corresponding sleep stage for selected participant.



(a) (b)

Figure 5. Log multitaper power spectrum estimates with traditional frequency bands (a) and estimated frequency bands using the search algorithm (b).

EBA algorithm exhibit a constant power spectrum through time. Using the full night HRV data, all of the frequency bands are found to exhibit time-varying dynamics as the p-values associated with $Q^S(\omega_0, \delta)$ are sufficiently small (see Table 2). Given that the body experiences varying levels of stress and arousal during NREM and REM sleep, we expect to see time-varying behavior in the frequency bands across the full night. However, we may see constant behavior in the frequency bands through time if we isolate portions of the full night data in which the individual is largely in either REM or NREM sleep. Figure 6 displays the estimated power spectra during the four clinically defined periods of NREM during the night. These periods occur between 0.3–1.2, 1.5–3.2, 4.7–5.6, and 7.3–7.8 hours after the participant went to bed, respectively.

Table 2. Approximate p-values for each frequency band to test $H_0 : g(u, \omega) = 0$ for the full night and during four periods of NREM throughout the night.

Band	Full night	NREM 1	NREM 2	NREM 3	NREM 4
(0.000, 0.023]	<0.0001	0.0060	0.0136	0.0046	0.0035
(0.023, 0.103]	<0.0001	0.0011	0.0096	0.0002	0.0008
(0.103, 0.243]	<0.0001	<0.0001	**0.8308**	0.0027	<0.0001
(0.243, 0.400]	<0.0001	<0.0001	**0.8503**	**0.1230**	<0.0001

NOTE: Bold indicates large p-values where we fail to reject the null hypothesis of stationarity.

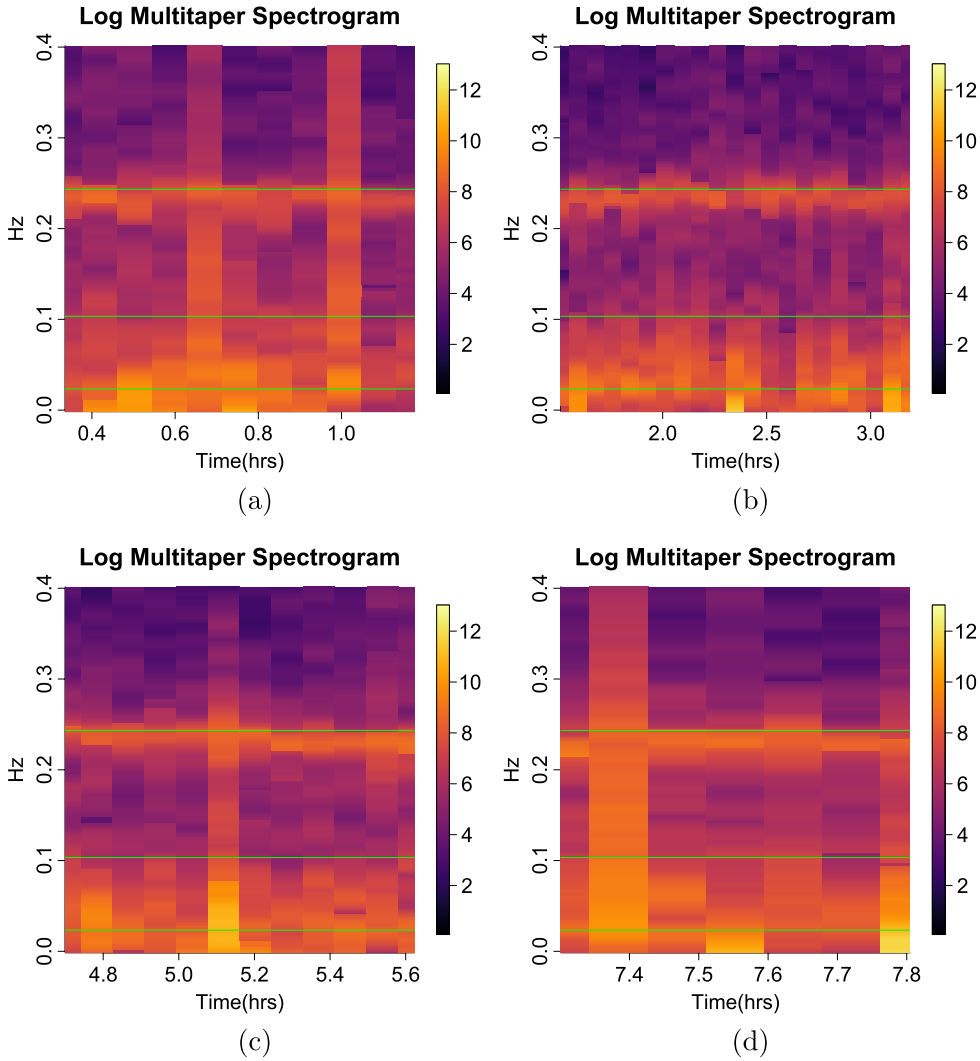


Figure 6. Sleep stage and power spectrum estimates with estimated frequency bands using EBA algorithm for the four periods of NREM throughout the night ((a)–(d), respectively).

During these portions of the night, the individual spends the majority of time in NREM sleep with brief interruptions of wakefulness. If we then calculate the demeaned time-varying power spectrum estimates and test statistics only using these portions of the night, we can get approximate p-values to determine if any frequency bands exhibit a constant power spectrum during each of these periods (see Table 2). We find that the periods of NREM associated with transition periods between wake and sleep (NREM 1 and NREM 4) exhibit time-varying behavior across all frequency bands. This finding is consistent with the scientific literature, as HRV power spectra have been shown to exhibit dynamic characteristics during transitional sleep periods (Trinder et al. 1992; Shinar et al. 2006). However, the two periods of NREM during the middle of the night (NREM 2 and NREM 3) exhibit time varying-behavior across some frequency bands, but not all. For NREM 2, the p-values for the two HF bands are sufficiently large, thus failing to reject the hypothesis that the power spectrum is constant during this time period. We also fail to reject the constant power spectrum hypothesis for the highest frequency band in NREM 3. High-frequency power is traditionally interpreted as an indirect measure of the parasympathetic nervous system, whose modulation is inversely related to stress and arousal (Malik et al. 1996). These results suggest that stress and arousal is stationary for this participant during NREM sleep within the middle of the night.

7. Discussion

The EBA framework introduced in this article offers a quantitative approach to identifying frequency bands that best preserve the nonstationary dynamics of the underlying time series. This framework allows for estimation of both the number of frequency bands and their corresponding endpoints through the use of a sensible FRESH statistic within in an iterative search algorithm. Another test statistic is also offered to determine which bands, if any, are stationary with respect to time. It would be interesting to consider other forms of the time-varying power spectrum estimator, such as unequally sized or overlapping segment piecewise stationary approximations, or smooth estimators, such as wavelet-based estimators, as potential enhancements to this framework. However, as mentioned previously, the distributional properties of the FRESH statistics in these settings may be nontrivial. Another interesting research direction would

be to extend the frequency domain partitioning framework to encompass both time and frequency domains. This could offer a unified solution that can identify time-localized frequency bands for characterizing time-varying power spectra.

We have focused on estimation of frequency bands for a single nonstationary time series, but this framework can be extended to address other interesting frequency-domain dimension reduction tasks. For example, covariate-dependent frequency band estimation for replicated time series could be employed to group frequencies with similar covariate-varying dynamics for parsimonious modeling of the association between the power spectrum and covariates. This extension would allow practitioners to estimate and use frequency bands that are most appropriate for a particular covariate under study. This framework could also be extended to the multivariate setting by estimating frequency bands which have a common spectral power among all or a particular subset of the time series using the spectral envelope. This extension could be very useful in the study of multichannel EEG and fMRI signals to analyze functional connectivity in the brain.

Supplementary Materials

Proofs: Proofs for theorems in this article.

Approximating the null distributions: Details on approximating the null distributions and corresponding p-values.

Illustrated example for algorithm: Step-by-step illustration of search algorithm for simulated dataset.

R code for algorithm implementation: R code containing all functions to run algorithm and example on simulated dataset. (.zip file)

Funding

This work is supported by National Institutes of Health grants R01GM113243 and P01AG020677 and by National Science Foundation grants IIS-1546087 and SES-1533956.

References

Adak, S. (1998), "Time-Dependent Spectral Analysis of Nonstationary Time Series," *Journal of the American Statistical Association*, 93, 1488–1501. [1934,1935,1936]

Billman, G. (2011), "Heart Rate Variability—A Historical Perspective," *Frontiers in Physiology*, 2, 86. [1933]

Cabeza-Ruiz, R., García-Massó, X., Centeno-Prada, R., Beas-Jiménez, J., Colado, J., and González, L.-M. (2011), "Time and Frequency Analysis of the Static Balance in Young Adults With Down Syndrome," *Gait & Posture*, 33, 23–28. [1933]

Causeur, D., Kloareg, M., and Friguet, C. (2009), "Control of the FWER in Multiple Testing Under Dependence," *Communications in Statistics—Theory and Methods*, 38, 2733–2747. [1939]

Cohen, D., Pollak, L., and Eldar, Y. C. (2016), "Carrier Frequency and Bandwidth Estimation of Cyclostationary Multiband Signals," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3716–3720. [1934]

Dahlhaus, R. (1997), "Fitting Time Series Models to Nonstationary Processes," *The Annals of Statistics*, 25, 1–37. [1934,1935,1936]

Daubechies, I., Lu, J., and Wu, H.-T. (2011), "Synchrosqueezed Wavelet Transforms: An Empirical Mode Decomposition-Like Tool," *Applied and Computational Harmonic Analysis*, 30, 243–261. [1934]

Daubechies, I., Wang, Y. G., and Wu, H.-T. (2016), "ConceFT: Concentration of Frequency and Time via a Multitapered Synchrosqueezed Transform," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 374, 20150193. [1934]

Davis, R. A., Lee, T. C. M., and Rodriguez-Yam, G. A. (2006), "Structural Break Estimation for Nonstationary Time Series Models," *Journal of the American Statistical Association*, 101, 223–239. [1934]

Doppelmayr, M., Klimesch, W., Pachinger, T., and Ripper, B. (1998), "Individual Differences in Brain Dynamics: Important Implications for the Calculation of Event-Related Band Power," *Biological Cybernetics*, 79, 49–57. [1933]

Efron, B. (2007), "Correlation and Large-Scale Simultaneous Significance Testing," *Journal of the American Statistical Association*, 102, 93–103. [1939]

Flandrin, P. (1998), *Time-Frequency/Time-Scale Analysis*, San Diego, CA: Academic Press. [1933]

Gardner, W. A., Napolitano, A., and Paura, L. (2006), "Cyclostationarity: Half a Century of Research," *Signal Processing*, 86, 639–697. [1934]

Guo, W., Dai, M., Ombao, H. C., and von Sachs, R. (2003), "Smoothing Spline ANOVA for Time-Dependent Spectral Analysis," *Journal of the American Statistical Association*, 98, 643–652. [1935]

Hall, M., Vasko, R., Buysse, D., Ombao, H., Chen, Q., Cashmere, J. D., Kupfer, D., and Thayer, J. F. (2004), "Acute Stress Affects Heart Rate Variability During Sleep," *Psychosomatic Medicine*, 66, 56–62. [1933,1939]

Hochberg, Y. (1988), "A Sharper Bonferroni Procedure for Multiple Tests of Significance," *Biometrika*, 75, 800–802. [1938]

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H. (1998), "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 454, 903–995. [1934]

Iber, C., Ancoli-Isreal, S., Chesson, A., and Quan, S. F. (2007), *The AASM Manual for the Scoring of Sleep and Associated Events*, Westchester, IL: American Academy of Sleep Medicine. [1941]

Kirch, C., Muhsal, B., and Ombao, H. (2015), "Detection of Changes in Multivariate Time Series With Application to EEG Data," *Journal of the American Statistical Association*, 110, 1197–1216. [1934]

Klimesch, W. (1999), "EEG Alpha and Theta Oscillations Reflect Cognitive and Memory Performance: A Review and Analysis," *Brain Research Reviews*, 29, 169–195. [1933]

Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., and Schwaiger, J. (1998), "Induced Alpha Band Power Changes in the Human EEG and Attention," *Neuroscience Letters*, 244, 73–76. [1933]

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., and Schwartz, P. J. (1996), "Heart Rate Variability. Standards of Measurement, Physiological Interpretation, and Clinical Use," *European Heart Journal*, 17, 354–381. [1933,1936,1941,1943]

Moore, S. T., MacDougall, H. G., and Ondo, W. G. (2008), "Ambulatory Monitoring of Freezing of Gait in Parkinson's Disease," *Journal of Neuroscience Methods*, 167, 340–348. [1933]

Neumann, M. H., and von Sachs, R. (1997), "Wavelet Thresholding in Anisotropic Function Classes and Application to Adaptive Estimation of Evolutionary Spectra," *The Annals of Statistics*, 25, 38–76. [1934]

Ombao, H. C., Raz, J. A., von Sachs, R., and Malow, B. A. (2001), "Automatic Statistical Analysis of Bivariate Nonstationary Time Series. In Memory of Jonathan A. Raz," *Journal of the American Statistical Association*, 96, 543–560. [1936]

Ombao, H., von Sachs, R., and Guo, W. (2005), "SLEX Analysis of Multivariate Nonstationary Time Series," *Journal of the American Statistical Association*, 100, 519–531. [1934]

Percival, D. B., and Walden, A. T. (1993), *Spectral Analysis for Physical Applications*, Cambridge: Cambridge University Press. [1935,1936,1937,1938]

Preuss, P., Puchstein, R., and Dette, H. (2015), "Detection of Multiple Structural Breaks in Multivariate Time Series," *Journal of the American Statistical Association*, 110, 654–668. [1934]

Priestley, M. B. (1965), "Evolutionary Spectra and Non-Stationary Processes," *Journal of the Royal Statistical Society*, Series B, 28, 228–240. [1934,1935]

Rand, W. M. (1971), "Objective Criteria for the Evaluation of Clustering Methods," *Journal of the American Statistical Association*, 66, 846–850. [1939]

Riedel, K. S., and Sidorenko, A. (1995), "Minimum Bias Multiple Taper Spectral Estimation," *IEEE Transactions on Signal Processing*, 43, 188–195. [1936]

Satterthwaite, F. E. (1946), "An Approximate Distribution of Estimates of Variance Components," *Biometrics Bulletin*, 2, 110–114. [1937]

Schröder, A. L., and Ombao, H. (2019), "FreSpeD: Frequency-Specific Change-Point Detection in Epileptic Seizure Multi-Channel EEG Data," *Journal of the American Statistical Association*, 114, 115–128. [1934]

Shinar, Z., Akselrod, S., Dagan, Y., and Baharav, A. (2006), "Autonomic Changes During Wake–Sleep Transition: A Heart Rate Variability Based Approach," *Autonomic Neuroscience*, 130, 17–27. [1943]

Storey, J. D. (2007), "The Optimal Discovery Procedure: A New Approach to Simultaneous Significance Testing," *Journal of the Royal Statistical Society*, Series B, 69, 347–368. [1939]

Thomson, D. J. (1982), "Spectrum Estimation and Harmonic Analysis," *Proceedings of the IEEE*, 70, 1055–1096. [1936,1937]

Tiganj, Z., Mboup, M., Chevallier, S., and Kalunga, E. (2012), "Online Frequency Band Estimation and Change-Point Detection," in *2012 1st International Conference on Systems and Computer Science (ICSCS)*, pp. 1–6. [1934]

Trinder, J., Whitworth, F., Kay, A., and Wilkin, P. (1992), "Respiratory Instability During Sleep Onset," *Journal of Applied Physiology*, 73, 2462–2469. [1943]

Walden, A. T., McCoy, E. J., and Percival, D. B. (1995), "The Effective Bandwidth of a Multitaper Spectral Estimator," *Biometrika*, 82, 201–214. [1936]

Welch, B. L. (1938), "The Significance of the Difference Between Two Means When the Population Variances Are Unequal," *Biometrika*, 29, 350–362. [1937]

Wu, Z., and Huang, N. E. (2009), "Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method," *Advances in Adaptive Data Analysis*, 1, 1–41. [1934]