



Periodic Variable Stars Modulated by Time-varying Parameters

Giovanni Motta¹, Darlin Soto², and Márcio Catelan^{3,4,5} ¹Department of Statistics, Texas A&M University, 3143 TAMU, 155 Ireland Street, College Station, TX 77843-3143, USA²Department of Statistics, Faculty of Mathematics, Pontificia Universidad Católica de Chile, 7820436 Macul, Santiago, Chile; dmsoto1@uc.cl³Instituto de Astrofísica, Facultad de Física, Pontificia Universidad Católica de Chile, 7820436 Macul, Santiago, Chile⁴Centro de Astro-Ingeniería, Pontificia Universidad Católica de Chile, 7820436 Macul, Santiago, Chile⁵Millennium Institute of Astrophysics, Nuncio Monseñor Sotero Sanz 100, 7500000, Santiago, Chile

Received 2021 July 25; revised 2021 November 3; accepted 2021 November 4; published 2022 January 26

Abstract

Many astrophysical phenomena are time-varying, in the sense that their brightness changes over time. In the case of periodic stars, previous approaches assumed that changes in period, amplitude, and phase are well described by either parametric or piecewise-constant functions. With this paper, we introduce a new mathematical model for the description of the so-called *modulated* light curves, as found in periodic variable stars that exhibit *smoothly* time-varying parameters such as amplitude, frequency, and/or phase. Our model accounts for a smoothly time-varying trend and a harmonic sum with smoothly time-varying weights. In this sense, our approach is flexible because it avoids restrictive assumptions (parametric or piecewise-constant) about the functional form of the trend and amplitudes. We apply our methodology to the light curve of a pulsating RR Lyrae star characterized by the Blazhko effect. To estimate the time-varying parameters of our model, we develop a semi-parametric method for unequally spaced time series. The estimation of our time-varying curves translates into the estimation of time-invariant parameters that can be performed by ordinary least squares, with the following two advantages: modeling and forecasting can be implemented in a parametric fashion, and we are able to cope with missing observations. To detect serial correlation in the residuals of our fitted model, we derive the mathematical definition of the spectral density for unequally spaced time series. The proposed method is designed to estimate smoothly time-varying trends and amplitudes, as well as the spectral density function of the errors. We provide simulation results and applications to real data.

Key words: RR Lyrae variable stars – Blazhko effect – Time series analysis

1. Introduction

RR Lyrae stars are important astrophysical tools for the measurement of distances and studies of the astrophysical properties of old stellar populations. They are moderately bright, evolved low-mass stars, currently in the core helium-burning phase, also known as the horizontal branch. Their periods are typically in the range of between about 0.2 and 1.0 day, which together with their characteristic light-curve shapes, allow them to be relatively easily identified in time series photometric surveys. An overview of their properties can be found in the monographs by Smith (1995) and Catelan & Smith (2015).

In spite of their astrophysical importance, RR Lyrae stars are still not fully understood. Indeed, one of the longest-standing problems in stellar astrophysics is also one that specifically affects RR Lyrae stars: the so-called “Blazhko effect” (Blazhko 1907). It is a long-term modulation of an RR Lyrae’s light curve, over timescales ranging from a few to hundreds of days (for recent reviews, see Catelan & Smith 2015; Gillet et al. 2019). The Blazhko effect is particularly common among fundamental-mode (ab-type) pulsators (e.g., Plachy et al. 2019), but is also present, to a lesser extent, in first-overtone (c-type) RR Lyrae stars (e.g., Netzel et al. 2018).

Over the decades since it was first described, the Blazhko effect has persistently defied theoretical explanations as to its

cause (e.g., Gillet et al. 2019). Gradual strengthening and weakening of turbulent convection in the stellar envelope (Stothers 2006), a 9:2 resonance between the fundamental and ninth-overtone radial modes (Buchler & Kolláth 2011), and interaction between fundamental and first-overtone modes in the “either-or” region of the instability strip (Gillet 2013) are the most recent candidates, but no consensus has yet been reached as to the root cause of the Blazhko effect, due in large part to the difficulties involved in the nonlinear hydrodynamical modeling of the phenomenon.

In this article, we introduce a model for time series observations of variable stars having a smoothly time-varying trend and amplitudes. More precisely, we develop a semi-parametric method for unequally spaced time series measuring the brightness of a modulated variable star. Our approach is flexible because it avoids assumptions about the functional form of the trend and amplitudes. The estimation of our time-varying curves translates into the estimation of time-invariant parameters that can be performed by ordinary least squares, with the following two advantages: modeling and forecasting can be implemented in a parametric fashion, and we are able to cope with missing observations. We also study the spectral density of the residuals obtained from the fit of our novel model.

In order to detect serial correlation in the residuals, in this paper we derive the definition of the spectral density for unequally spaced time series. There are many reasons why astronomical time series are not sampled equidistantly, and the gaps can be either regular or random. From the Earth, stars cannot be observed during the day, which introduces regular

gaps in the time series. Also, for about half a year, most objects become unobservable, as they are up on the sky at the same time as the Sun, which introduces yearly gaps. There could be clouds or high winds, forcing the closure of telescopes, producing random gaps. There could be high-priority alerts overriding the observations, or the telescope could be available only on certain nights.

In some cases, observations are unevenly spaced due to missing values. Astronomical data sets often contain missing values, and this limitation is sometimes due to incomplete observations or varying survey depths. Even if telescopes are recording and storing information systematically, that is, at a regular cadence, there are a few things that can alter the regular sampling. For example, an astronomer might decide to increase the exposure time if there are clouds obscuring the target, to try to increase the signal-to-noise ratio. Conversely, if the observing conditions are excellent, the astronomer might decide to decrease the exposure times (and hence the cadence) to avoid saturating the detector, for example. Missing values are usually handled via imputation, that is, the gap generated by the missing value is “filled in” by an estimated value. If observations are missing because of the survey, imputation can be performed using statistical models. For example, Feigelson et al. (2018) apply ARIMA models to fill in missing data in astronomical time series.

However, in astrostatistics, missing value problems are sometimes inherently brought about by the manner in which physical processes are recorded. In particular, telescopes are not located in the center of the solar system. Since the speed of light is finite, this results in a time delay between the arrival times of signals at our position and at the center of the solar system. This is typically corrected for by referring the times of observations to either Heliocentric Julian Dates (HJD) or Barycentric Julian Dates (BJD), which refer to the center of the Sun or the entire solar system, respectively. Thus, even if telescopes record data strictly evenly according to the local time at the observatory (e.g., one observation performed every night at local midnight), this correction will slowly change between observations, modifying what was initially a regular grid to an irregular one. Also, this correction is different for every source on the sky, even though sources close to each other may have very similar corrections. Therefore, for some astronomical data sets where missing values may arise from the manner in which observations of a physical process are collected, or even the nature of the physical process itself (e.g., sudden, extreme dimming events that may occasionally render an object impossible to detect for a certain amount of time), the imputation method may not be applicable (see Chattopadhyay 2017). Our novel approach, which involves the classical periodogram, has the advantage of not relying on any imputation method.

We shall divide the present study into six main sections. In Section 2 we introduce our novel model and clarify analogies and differences as compared with previous approaches. In Section 3, the present status of important ingredients of amplitude and frequency modulations is critically discussed. In Section 4 we present the method we adopt to estimate the time-varying parameters. In Section 5 we present a new method to estimate the spectral density of unequally spaced times series, which is needed for the analysis of the residuals. Section 6 provides simulation results, whereas Section 7 illustrates the advantages of using our novel method by means of an

application to an RR Lyrae variable star. Finally, our main conclusions are summarized in Section 8.

Through the paper we use bold uppercase letters to denote matrices, and bold slanted to denote vectors. We denote by \mathbf{I}_m the identity matrix of size m , by $\mathbf{0}_n$ a column-vector of zeros of length n , by $\text{tr}\{\mathbf{A}\}$ the trace of \mathbf{A} , by \mathbf{A}^\top the transpose of \mathbf{A} , by $\|\mathbf{A}\|$ the Frobenius norm $\|\mathbf{A}\| = [\text{tr}\{\mathbf{A}^\top\mathbf{A}\}]^{1/2}$, and by \mathbf{A}^{-1} the inverse of the square matrix \mathbf{A} , that is, the square matrix \mathbf{A}^{-1} such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

2. A Novel Time-varying Modulation Model for Variable Stars

Light curves of variable stars are typically fitted using harmonic models with a linear (or constant) trend and time-invariant amplitudes (see Equations (1) and (5) in Richards et al. 2011). This type of model would be inappropriate when the underlying trend and amplitudes change over time in a more complex way. Eilers et al. (2008) proposed a model with one harmonic component ($K=1$) where the trend and amplitudes vary smoothly over time. In this paper, we extend the model by Eilers et al. (2008) to the case of $K \geq 1$ harmonic components, where the amplitudes associated with each harmonic component vary smoothly over time. We estimate our model by means of P-splines (Eilers & Marx 1996), which are a combination of B-splines and penalties. The estimation of the time-varying curves translates into the estimation of time-invariant parameters that can be performed by the least-squares method, with the following three advantages: it is computationally fast, forecasting can be implemented in a parametric fashion, and we can cope with missing observations.

Compared to local smoothers (such as kernel smoothers), the main advantage of regression spline in the context of time series is that the unknown parameters are time-invariant and thus they can be estimated globally rather than locally. As a consequence, forecasting only requires good estimates of the global unknown parameters. We can think of regression splines with B-splines as a semi-parametric model in the sense that it contains parametric as well as nonparametric components. The parametric component is given by a finite number of parameters, whereas the nonparametric component is given by the basis functions. Another advantage of parametric and semi-parametric models over nonparametric models is the computing speed, as many nonparametric models are computationally intensive. Finally, the use of B-splines in regression allows us to rewrite the estimation problem as a least-squares fit, avoiding the use of numerical methods—such as Newton–Raphson—which can be time consuming.

Let $\{Y_i \equiv Y_i, i = 1, \dots, N\}$ be a set of observations occurring at certain discrete times t_1, \dots, t_N . In the case of equally spaced observations, $t_i = t_0 + i\Delta$ where i is an integer, and $\Delta > 0$ is the constant data spacing. Then $|t_i - t_k| = \Delta|i - k|$, and typically $\Delta = 1$. Astronomical light curves are often observed unequally in time, that is, the data spacing of observation times is not constant.

We decompose the observed light curve into the sum of a deterministic signal μ and a random noise z . The deterministic part $\mu(t)$ consists of a trend $m(t)$ and a modulated periodic signal. The modulated periodic signal is a linear combination of K cosines and sines, with weights given by the *modulating*

functions $g(t)$:

$$\begin{aligned} Y_i &= \mu(t_i) + z_i, \\ i &= 1, \dots, N, \{z_i\} \sim \text{WN}(0, \sigma_z^2), \\ \mu(t_i) &= m(t_i) + \sum_{k=1}^K \{g_{1,k}(t_i) \cos(w_k t_i) \\ &\quad + g_{2,k}(t_i) \sin(w_k t_i)\}, \end{aligned} \quad (1)$$

or in matrix notation $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{z}$, where $\mathbf{Y} = (Y_1, \dots, Y_N)^\top$ is the vector of observations at time $\mathbf{t} = (t_1, \dots, t_N)^\top$, $\boldsymbol{\mu} = [\mu(t_1), \dots, \mu(t_N)]^\top$ is the expectation of \mathbf{Y} , $m(t_i)$ is the smooth time-varying trend at time t_i , the $g(t_i)$'s are smooth time-varying amplitudes of the cosine and sine waves at time t_i , respectively, $w_k = 2\pi f_k$ is the angular frequency, and f_k is the ordinary frequency. Since the errors are zero mean, the expectation of the observed brightness at time t_i is equal to the deterministic part of the signal at time t_i , that is, $E[Y_i] = \mu(t_i)$.

We refer to $m(\cdot)$ as the ‘‘trend,’’ that is, the (typically) aperiodic change in the mean of the light curve. By contrast, we call ‘‘amplitudes’’ the functions $g(\cdot)$'s that weigh the periodic variation (of this average brightness) of cosine and sine waves. In Appendix A we summarize the standard mathematical definitions of amplitude modulation and frequency modulation. Both trend and amplitudes are typically restricted to be sinusoidal, whereas in this paper our trend $m(t)$ and our amplitude functions $g(t)$'s are general smooth functions and not necessarily sinusoidal. In Section 3 we clarify the mathematical connection between the standard modulation models and our novel modulation model in Equation (1).

The error vector $\mathbf{z} = (z_1, \dots, z_N)^\top$ is a *white noise* (WN) process with mean zero and variance σ_z^2 . That is, each error z_i , $i = 1, \dots, N$, follows a zero-mean WN process with variance σ_z^2 :

$$\begin{aligned} E[z_i] &= 0 \quad \text{and} \quad E[z_i z_j] \\ &= \delta_{i=j} \sigma_z^2, \quad \text{for all } i, j = 1 \dots, N, \end{aligned}$$

where $\delta_{i=j} = 1$ if $i = j$ and zero otherwise.

Our model in Equation (1) is defined in discrete time, and it focuses on the time domain. Kelly et al. (2014) adopt the continuous-time autoregressive moving average (CARMA) models to estimate the variability features of a light curve in the frequency domain. More specifically, Kelly et al. (2014) use the power spectral density (PSD) of CARMA models to account for irregular sampling and measurement errors. A stationary CARMA(p, q) process has the PSD

$$P(f) = \sigma^2 \left| \sum_{j=0}^q \beta_j (i2\pi f)^j \right|^2 \left/ \left| \sum_{k=0}^p \alpha_k (i2\pi f)^k \right|^2 \right.$$

To illustrate the importance of fitting models with time-varying parameters, Kelly et al. (2014) simulated a light curve that switches from one CARMA process to another. More precisely, they constructed a nonstationary light curve by generating two CARMA processes of the same order ($p = 5, q = 3$), but with different parameters (see Kelly et al. 2014, Section 4.3):

$$\boldsymbol{\theta}(t) = \begin{cases} \boldsymbol{\theta}_1 & t_1 \leq t < t_0 \\ \boldsymbol{\theta}_2 & t_0 \leq t \leq t_N, \end{cases}$$

where $\boldsymbol{\theta}(t) = [\alpha_1(t), \dots, \alpha_p(t), \beta_1(t), \dots, \beta_q(t), \sigma^2(t)]^\top$. The vector $\boldsymbol{\theta}(t)$ is a step-wise function that is constant before and

after t_0 . The approach based on piecewise-constant parameters is receiving growing interest in various areas of astrophysics. Wong et al. (2015) adopt a Poisson model for the photon counts. They define $\lambda(t_j, w_i)$ as the expected count per unit time and per unit wavelength averaged over the bin centered at (t_j, w_i) , and detect change points π such that $\{\lambda(t_j, w_i) | t_j \leq \pi\} \neq \{\lambda(t_j, w_i) | t_j > \pi\}$. Wong et al. (2015) estimate the number of change points and their values. Xu et al. (2021) develop a method for modeling a time series of images, and assume that the arrival times of the photons follow a Poisson process. They assume that all image stacks between any two adjacent change points (in the time domain) share the same unknown piecewise-constant function. Xu et al. (2021) estimate the piecumber and the locations of all of the change points (in the time domain), as well as all of the unknown piecewise-constant functions between any pairs of the change points.

Instead of considering parameters that are piecewise-constant functions of time, in this paper we allow the parameters to be *smooth* functions of *rescaled time*, permitting the process to be *locally stationary*. The framework of local stationarity introduced by Dahlhaus (1997), where the parameter curves are defined in rescaled time $u = t/T$, provides a meaningful asymptotic theory. Locally stationary $Y(t)$ means that if the functions $m(u)$ and $g(u)$ in Equation (1) are ‘‘smooth’’ and T is large, $m\left(\frac{t}{T}\right) \approx m\left(\frac{r}{T}\right)$ and $g\left(\frac{t}{T}\right) \approx g\left(\frac{r}{T}\right)$ for values of r close to t , that is, *locally* around t . More precisely, we assume that the functions $m(x)$ and $g(x)$ are Lipschitz continuous, that is, there exist constants C_m and C_g such that

$$\begin{aligned} |m(z) - m(u)| &\leq C_m |z - u| \quad \text{and} \\ \times |g(z) - g(u)| &\leq C_g |z - u|, \quad \text{for all } u, z \in [0, 1]. \end{aligned} \quad (2)$$

To define $m(u)$ and $g(u)$ as functions of rescaled time let us consider, for each fixed $u \in [0, 1]$ and increasing T , the sequence $t = t_T = \lfloor uT \rfloor$, where $\lfloor x \rfloor$ is the largest integer not exceeding x . Then we obtain the following uniform bound: $\sup_t |t/T - u| < 1/T$. A time series is stationary if the moments of the underlying stochastic process, such as expectation and variance, are time-invariant. The idea behind local stationarity is to allow for time-varying parameters, in a way that *locally* the process behaves as stationary. Lipschitz continuity is a smoothness assumption that implies uniform continuity. The model in Equations (1)–(2) is a locally stationary process written in rescaled time in a way such that, as T grows, we observe more and more ‘‘observations’’ of the same type around u . That is, if $m(\cdot)$ and $g(\cdot)$ are smooth we have

$$\begin{aligned} |m\left(\frac{t}{T}\right) - m(u)| &\leq C_m \left| \frac{t}{T} - u \right| < \frac{C_m}{T} \rightarrow 0 \quad \text{and} \\ \times \left| g\left(\frac{t}{T}\right) - g(u) \right| &\leq C_g \left| \frac{t}{T} - u \right| < \frac{C_g}{T} \rightarrow 0 \quad \text{as } T \rightarrow \infty, \end{aligned}$$

for all $t = \lfloor uT \rfloor$. The locally stationary framework is important for handling, in a meaningful way, the asymptotic theory arising in statistics for processes with time-varying parameters. Suppose that we observe $X_t = \mu(t) + z_t$, with $z_t \sim \text{WN}(0, \sigma_z^2)$ for $t = 1, \dots, T$. Inference in this case means studying the properties of an estimator for the unknown function $\mu(t)$ on the grid $\{1, \dots, T\}$. Given that μ changes over time, it is obvious

that an asymptotic approach where $T \rightarrow \infty$ is not suitable for describing a statistical method, since future “observations” $\{\mu(t), t > T\}$ do not necessarily contain any information on $\mu(t)$ on $\{1, \dots, T\}$. To overcome these problems, Dahlhaus (1996) suggested to consider a triangular array of data. In analogy with nonparametric regression, it seems natural to set down the asymptotic theory in a way that we “observe” $\mu(t)$ on a finer grid (but on the same interval), i.e., that we observe the process $Y_T(t) = \mu\left(\frac{t}{T}\right) + z_t$, where Y_T is now a triangular array and μ is now rescaled to the interval $[0, 1]$. Working in rescaled time is often adopted also within the estimation framework of regression splines (see Zhou et al. 1998, among others).

Time series analysis of nonstationary sequences can be deterministic or stochastic. A popular example of stochastic nonstationarity is the well-known class of integrated processes, where the observed times series can be made stationary by differencing. Feigelson et al. (2018) apply autoregressive integrated moving average (ARIMA) models to light curves of several variable stars, discussing their effectiveness for different temporal characteristics. The process $\{X_t\}$ is an ARIMA(p, d, q) process if $Y_t = (1 - B)^d X_t$, obtained by applying the operator $1 - B$ repeatedly d times, is a stationary ARMA(p, q) process. The most popular example of an ARIMA(p, d, q) process is the “random walk” $X_t = X_{t-1} + Z_t$, which is an ARIMA with $p = q = 0$ and $d = 1$.

In the next section, we review the models proposed by Benkő et al. (2011) and Benkő (2018) for Blazhko light curves. Interestingly, our model in Equation (1) generalizes the models by Benkő et al. (2011) and Benkő (2018) in the sense that the modulating functions $g(\cdot)$ are not confined to the class of parametric (sinusoidal or nonsinusoidal) functions.

3. Modeling Blazhko Light Curves

The Blazhko effect is a periodic amplitude and phase variation in the light curves of RR Lyrae variable stars. In astronomy, the Blazhko effect is usually interpreted as a modulation phenomenon. Modulation is the process of transmitting a low-frequency signal into a high-frequency wave, called the carrier wave, by changing its amplitude, frequency, or phase angle through the modulating signal. In Appendix A we review the main mathematical definitions underlying the modulation phenomenon in astrophysics.

In this section, we review the models proposed by Benkő et al. (2011) and Benkő (2018), respectively, and we compare them with our novel model in Equation (1). To describe Blazhko light curves, Benkő et al. (2011) proposed to fit the following model:

$$\begin{aligned} \mu^*(t) &= a_0^A a_0 + a_0 g^A(t) \\ &+ \sum_{k=1}^K [a_0^A a_k + a_k g^A(t)] \sin[2\pi k f_0 t + \varphi_k + k g^F(t)], \end{aligned} \quad (3)$$

where a_k and f_0 denote amplitude and frequency, respectively, and

$$\begin{aligned} g^M(t) &= \sum_{j=1}^{\ell^M} a_j^M \sin(2\pi j f_m t + \varphi_j^M), \\ M &= A \text{ or } F. \end{aligned} \quad (4)$$

More recently, Benkő (2018) introduced a similar model:

$$\begin{aligned} \mu^*(t) &= m_0 + \sum_{r=1}^{\ell} b_r \sin(2\pi r f_m t + \varphi_r^b) \\ &+ \sum_{k=1}^K [a_k + g_k^A(t)] \sin[2\pi k f_0 t + \varphi_k + g_k^F(t)], \end{aligned} \quad (5)$$

where b_r and φ_r^b denote amplitude and frequency of the modulating signal, respectively, and

$$g_k^M(t) = \sum_{j=1}^{\ell_k^M} a_{kj}^M \sin(2\pi j f_m t + \varphi_{kj}^M), \quad M = A \text{ or } F. \quad (6)$$

The functions $g^M(t)$ and $g_k^M(t)$ in Equation (4) and (6) are the modulating functions with subscripts $M=A$ and $M=F$ denoting amplitude and frequency modulation, respectively. The main pulsation frequency is denoted by f_0 , whereas f_m is the modulating frequency. In this paper we improve the models in Equations (3)–(4) and (5)–(6) from two different viewpoints. From the modeling viewpoint, we relax the assumption of parametric amplitude and frequency modulations. Assuming parametric amplitude and frequency modulations results in a Fourier sum with time-invariant amplitudes and time-invariant frequencies, whereas our time-varying amplitudes and frequencies do not obey any particular form. From the estimation viewpoint, we do not rely on the nonlinear least-squares algorithms, such as the Levenberg–Marquardt algorithm, that are typically used to fit parametric nonlinear models. These methods require initial values close to the solution, which in some applications are difficult to find.

Both models proposed by Benkő et al. (2011) and Benkő (2018) and given by Equations (3) and (5), respectively, are a special case of our model defined by Equation (1). To see this, let us define

$$\begin{aligned} v(t) &= a_0^A a_0 + a_0 g^A(t), \\ w_{1,k}(t) &= [a_0^A a_k + a_k g^A(t)] \sin[\varphi_k + k g^F(t)], \\ &k = 1, \dots, K, \\ w_{2,k}(t) &= [a_0^A a_k + a_k g^A(t)] \cos[\varphi_k + k g^F(t)], \\ &k = 1, \dots, K, \end{aligned} \quad (7)$$

and

$$\begin{aligned} u(t) &= m_0 + \sum_{r=1}^{\ell} b_r \sin(2\pi r f_m t + \varphi_r^b), \\ h_{1,k}(t) &= [a_k + g_k^A(t)] \sin[\varphi_k + g_k^F(t)], \\ &k = 1, \dots, K, \\ h_{2,k}(t) &= [a_k + g_k^A(t)] \cos[\varphi_k + g_k^F(t)], \\ &k = 1, \dots, K. \end{aligned} \quad (8)$$

We now show how Equations (7) and (8) allow to us compare our model in Equation (1) with the models proposed by Benkő et al. (2011) and Benkő (2018), respectively. Comparing the models in Equations (1) and (3), the time-varying trend and amplitudes of the model in Equation (1) are expressed as

$$\begin{aligned} m(t) &= v(t), \\ g_{\ell,k}(t) &= w_{\ell,k}(t), \quad \ell = 1, 2, \quad k = 1, \dots, K. \end{aligned} \quad (9)$$

At the same time, comparing the model in Equation (1) with the model in Equation (5), the time-varying trend and amplitudes

of the model in Equation (1) are

$$\begin{aligned} m(t) &= u(t), \\ g_{\ell,k}(t) &= h_{\ell,k}(t), \quad \ell = 1, 2, \quad k = 1, \dots, K, \end{aligned} \quad (10)$$

the ordinary frequency being $f_k = kf_0$.

As we can see in Equations (9) and (10), the functions $m(t)$ and $g_{\ell,k}(t)$ incorporate the amplitude and frequency modulation functions $g^M(t)$ and $g_k^M(t)$ in Equations (4) and (6). In this sense, the limitation of our approach is that it does not aim at identifying the amplitude and frequency modulating functions $g^M(t)$ and $g_k^M(t)$ in Equations (4) and (6). That said, the benefit of our approach from the estimation viewpoint is twofold. An important advantage of our model in Equation (1) over the models in Equations (3) and (5) is that the modulating frequency f_m does not need to be estimated. In other words, in order to describe statistically a Blazhko light curve using our model in Equation (1), we only need to estimate f_0 . If the observed time series is indeed a Blazhko light curve, the modulating frequency f_m is included in the nonparametric trend $m(t)$ and amplitude $g_{\ell,k}$ of our model in Equation (1). Moreover, assuming that the frequencies are known, for our model in Equation (1) we only need to estimate the functions $m(\cdot)$ and $g_{\ell,k}(\cdot)$, whereas for the model in Equations (3) and (5) the estimated parameters are the amplitudes a_0^A , a_0 , m_0 , a_k^s 's, b_r^s 's, and a_{kj}^M 's, and the phases φ_r^s 's, φ_k^s 's, φ_j^M 's, and φ_{kj}^M 's.

4. Estimation

In Section 4.1 we define estimators of the unknown trend $m(\cdot)$, amplitudes $\{g_{\ell,k}(\cdot), \ell = 1, 2, k = 1, \dots, K\}$, and variance σ_z^2 of the model in Equation (1), and in Section 4.2 we explain how to select the tuning parameters associated with the B -splines and the penalization used in the estimation method. We denote by N the sample size, $T = t_N - t_1$ the time span, J the number of B -splines basis, d the degree of the B -splines, K the number of harmonics components, r the order of the penalty, and M the number of replications in Monte Carlo simulations.

We performed our calculations using the R Language for Statistical Computing (R Core Team 2021). Our codes combine existing functions (available as part of R packages) with our own development. The computations implemented in this paper are available as a GitHub public code repository.⁶

4.1. Penalized Least Squares

As mentioned in Section 2, we use B -splines to estimate the trend and amplitudes of the model given by Equation (1). The smooth trend function $m(t_i)$ is modeled as a linear combination of B -splines basis

$$m(t_i) = \sum_{j=1}^J \alpha_j B_j(t_i), \quad i = 1, \dots, N,$$

which can be written in matrix notation as

$$\mathbf{m} = \mathbf{B}\boldsymbol{\alpha},$$

where $\mathbf{m} = [m(t_1), \dots, m(t_N)]^\top$, $\mathbf{B} = [B_{ij}] = [B_j(t_i)]$ is the $N \times J$ basis matrix ($i = 1, \dots, N, j = 1, \dots, J$) and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_J)^\top$. The exact definition of B -splines is given in Appendix B.

The smooth amplitude functions, $g_{\ell,k}(t_i)$, $\ell = 1, 2$, are modeled in the same way:

$$\begin{aligned} g_{1,k}(t_i) &= \sum_{j=1}^J \beta_{k,j} B_j(t_i), \\ g_{2,k}(t_i) &= \sum_{j=1}^J \gamma_{k,j} B_j(t_i), \quad k = 1, \dots, K. \end{aligned}$$

In matrix notation

$$\mathbf{g}_{1,k} = \mathbf{B}\boldsymbol{\beta}_k \text{ and } \mathbf{g}_{2,k} = \mathbf{B}\boldsymbol{\gamma}_k, \quad k = 1, \dots, K,$$

where $\boldsymbol{\beta}_k = (\beta_{k,1}, \dots, \beta_{k,J})^\top$, $\boldsymbol{\gamma}_k = (\gamma_{k,1}, \dots, \gamma_{k,J})^\top$, and $\mathbf{g}_{\ell,k} = [g_{\ell,k}(t_1), \dots, g_{\ell,k}(t_N)]^\top$, $\ell = 1, 2, k = 1, \dots, K$. Thus, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}_k$, and $\boldsymbol{\gamma}_k$, $k = 1, \dots, K$, are vectors associated with the trend and amplitudes, respectively. We define the $N \times N$ matrices \mathbf{C}_k and \mathbf{S}_k as

$$\begin{aligned} \mathbf{C}_k &= \text{diag}\{\cos(w_k t_1), \dots, \cos(w_k t_N)\} \text{ and} \\ \mathbf{S}_k &= \text{diag}\{\sin(w_k t_1), \dots, \sin(w_k t_N)\}, \\ k &= 1, \dots, K. \end{aligned}$$

Thus, the model for the expected value of \mathbf{Y} , in matrix notation, can be expressed as

$$E[\mathbf{Y}] = \boldsymbol{\mu} = \mathbf{B}\boldsymbol{\theta},$$

where \mathbf{B} is the $N \times c$ design matrix given by

$$\mathbf{B} = [\mathbf{B}|\mathbf{C}_1\mathbf{B}|\dots|\mathbf{C}_K\mathbf{B}|\mathbf{S}_1\mathbf{B}|\dots|\mathbf{S}_K\mathbf{B}],$$

with $c = J(2K + 1)$, and

$$\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}_1^\top, \dots, \boldsymbol{\beta}_K^\top, \boldsymbol{\gamma}_1^\top, \dots, \boldsymbol{\gamma}_K^\top)^\top$$

is the vector of regression coefficients of length c .

The *ordinary least squares* (OLS) estimator of $\boldsymbol{\theta}$ is the vector $\hat{\boldsymbol{\theta}}_{\text{OLS}}$, which minimizes the sum of squares:

$$M_\theta = \|\mathbf{Y} - \mathbf{B}\boldsymbol{\theta}\|^2.$$

Equating to zero the partial derivatives with respect to each component of $\boldsymbol{\theta}$ and assuming (as we shall) that $\mathbf{B}^\top \mathbf{B}$ is nonsingular, the estimator of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}}_{\text{OLS}} = (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{Y}.$$

The OLS estimate also maximizes the likelihood of the observations when the errors z_1, \dots, z_N are independent and identically distributed and Gaussian.

The size of the basis determines the amount of smoothing of the fitted curves. The larger the value of J , the bumpier the fitting will be. To avoid overfitting, Eilers & Marx (1996) proposed a penalty on the (high-order) finite differences of the coefficients:

$$\begin{aligned} M_\theta^* &= \|\mathbf{Y} - \mathbf{B}\boldsymbol{\theta}\|^2 + \tau_1 \|\mathbf{D}_r \boldsymbol{\alpha}\|^2 \\ &\quad + \sum_{k=1}^K \{\tau_{2k} \|\mathbf{D}_r \boldsymbol{\beta}_k\|^2 + \tau_{2k+1} \|\mathbf{D}_r \boldsymbol{\gamma}_k\|^2\}, \end{aligned}$$

where $\{\tau_k, k = 1, \dots, 2K + 1\}$ are positive regularization parameters that control the smoothness of the curve, penalizing the coefficients that are far apart from one another. If $\tau_k = 0$, $k = 1, \dots, 2K + 1$, we have the standard normal equations of linear regression with a B -splines basis. The larger the value of τ_k , the closer the coefficient $\boldsymbol{\theta}$ is to zero. When $\tau_k \rightarrow \infty$ we obtain a polynomial fit. The matrix \mathbf{D}_r constructs r th-order

⁶ <https://github.com/DarlinSoto/Modulation-models>

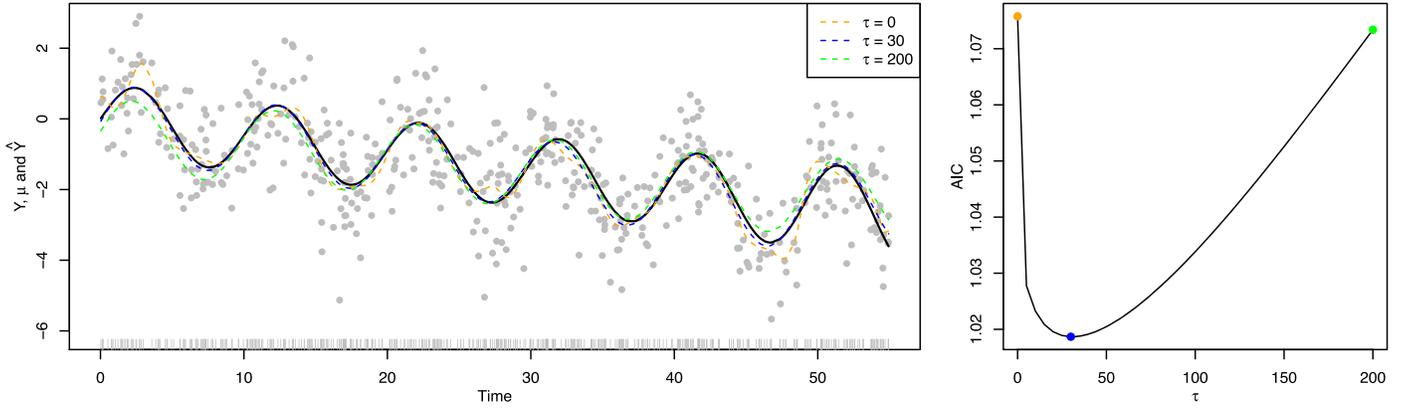


Figure 1. Automatic selection of the tunable parameters presented in Section 4.2. Left: data (gray dots) simulated according to the model defined by Equation (1), with $N = 500$, $\mu(t_i) = -0.05t_i - (-0.0002t_i + 0.0003t_i^2)\cos(0.2\pi t_i) + (1 - 0.0005t_i)\sin(0.2\pi t_i)$ (black curve), and where the errors follow a Gaussian distribution with zero mean and variance $\sigma_z^2 = 1$. Time is unequally spaced, obtained from a uniform distribution $U(0, 55)$ (gray ticks on the horizontal axis). We illustrate three estimates of Y corresponding to three different specifications of τ_j , with $j = 1, 2, 3$: $\tau_j = 0$ (orange curve), $\tau_j = 30$ (blue curve), and $\tau_j = 200$ (green curve). Right: values of the AIC in Equation (14), obtained from the simulated and estimated light curve, corresponding to forty-one equally spaced values of τ ranging from 0 to 200. The three points (orange, blue, and green) on the AIC curve correspond to the three fits presented in the left-hand plot of the figure.

differences of a vector η as

$$\mathbf{D}_r \eta = \Delta^r \eta.$$

The first difference of η , $\Delta^1 \eta$, is the vector with elements $\eta_{i+1} - \eta_i$. Repeated differencing applied to $\Delta \eta$ results in higher differences, such as $\Delta^2 \eta$ and $\Delta^3 \eta$.

The penalties can be represented as $\theta^\top \mathbf{P} \theta$ with the block-diagonal matrix $\mathbf{P} = \mathbf{T} \otimes \mathbf{D}_r^\top \mathbf{D}_r$, and $\mathbf{T} = \text{diag}\{\tau_1, \tau_2, \tau_3, \dots, \tau_{2K+1}\}$. Then, minimizing

$$M_\theta^* = \|\mathbf{Y} - \mathbf{B}\theta\|^2 + \theta^\top \mathbf{P} \theta$$

with respect to θ , the *penalized ordinary least squares estimator* (POLS) of θ is

$$\hat{\theta}_{\text{POLS}} = (\mathbf{B}^\top \mathbf{B} + \mathbf{P})^{-1} \mathbf{B}^\top \mathbf{Y}. \quad (11)$$

The prediction of Y at time t_i is given by

$$\hat{Y}_i = \hat{\mu}(t_i) = \mathbf{B}(t_i)^\top \hat{\theta}_{\text{POLS}}, \quad (12)$$

where $\mathbf{B}(t_i)$ is the i th row of \mathbf{B} , the residuals are $\hat{z} = \mathbf{Y} - \hat{\mathbf{Y}}$, with $\hat{\mathbf{Y}} = (\hat{Y}_1, \dots, \hat{Y}_N)^\top$, and the mean square error (MSE) is $\text{MSE} = N^{-1} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2$.

The estimators of the trend \mathbf{m} and amplitudes $\{\mathbf{g}_{\ell,k}, \ell = 1, 2, k = 1, \dots, K\}$, are

$$\hat{\mathbf{m}} = \mathbf{B}\hat{\alpha}, \quad \hat{\mathbf{g}}_{1,k} = \mathbf{B}\hat{\beta}_k, \quad \hat{\mathbf{g}}_{2,k} = \mathbf{B}\hat{\gamma}_k. \quad (13)$$

Another parameter of interest is the variance of the errors, σ_z^2 , which can be estimated by

$$\hat{\sigma}_z^2 = [N - \text{tr}(\hat{\mathbf{S}})]^{-1} \sum_{i=1}^N \{Y_i - \mathbf{B}(t_i)^\top \hat{\theta}_{\text{POLS}}\}^2$$

where $\hat{\mathbf{S}} = \mathbf{B}(\mathbf{B}^\top \mathbf{B} + \mathbf{P})^{-1} \mathbf{B}^\top$.

In addition to the point estimate, interval estimation for \hat{Y}_i is often of interest and is easy to construct. In Appendix C we derive parametric and nonparametric confidence intervals for \hat{Y}_i .

4.2. Automatic Selection of the Tunable Parameters

Before calculating the estimator in Equation (11), it is necessary to select the tuning parameters $\tau =$

$(\tau_1, \tau_2, \dots, \tau_{2K+1})^\top$. To choose the tuning parameters, we propose to use the Akaike information criterion (AIC).

The AIC penalizes the log-likelihood of a fitted model by considering the effective number of parameters. The definition of AIC given by Hastie et al. (2004) is

$$\text{AIC}(\tau) = \overline{\text{err}}(\tau) + 2 \frac{\text{df}}{N} \hat{\sigma}_0^2,$$

where $\overline{\text{err}}(\tau)$ corresponds to the mean square error in the case of Gaussian errors, df is the effective number of parameters, N is the number of observations used to fit the model, and $\hat{\sigma}_0^2$ is given by the variance of the residuals from the \hat{Y}_i that are computed when $\tau = \mathbf{0}_{2K+1}$.

The value for τ is chosen by minimizing the AIC, which is computed as

$$\text{AIC}(\tau) = \frac{1}{N} \sum_{i=1}^N \{Y_i - \mathbf{B}(t_i)^\top \hat{\theta}_{\text{POLS}}\}^2 + 2 \frac{\text{tr}(\hat{\mathbf{S}})}{N} \hat{\sigma}_0^2, \quad (14)$$

The AIC given by Equation (14) can also be used to select the number of B -splines J , the degree d of the B -spline, the order of penalty r , and the number of harmonic components K .

In Figure 1, we have generated $N = 500$ observations from the model described in Equation (1), with the Gaussian errors $\{z_i, 1 \leq i \leq 500\}$ being simulated using the R function `rnorm`. We consider the following artificial signal:

$$\begin{aligned} \mu(t_i) = & -0.05t_i - (-0.0002t_i + 0.0003t_i^2) \\ & \times \cos(0.2\pi t_i) + (1 - 0.0005t_i)\sin(0.2\pi t_i), \end{aligned}$$

with the errors following a Gaussian distribution with zero mean and variance $\sigma_z^2 = 1$. Time t is unequally spaced and was obtained from a uniform distribution $U(\theta_1, \theta_2)$ with $\theta_1 = 0$ and $\theta_2 = 55$ using the R function `runif`. In the first plot of Figure 1 the observations \mathbf{Y} are represented by the gray points, and the mean μ by the black curve. The orange, blue, and green curves illustrate three possible estimates for \mathbf{Y} obtained using the method described in the Section 4.1 with increasing smoothing parameters. The orange line is the fit obtained with $\tau_j = 0$, $j = 1, 2, 3$: the corresponding $\hat{\mathbf{Y}}$ matches the data well, but fits the true μ poorly because it is wiggly. The blue curve is

obtained using the smoothing parameters $\tau_j = 30$, $j = 1, 2, 3$, and the green curve is obtained using $\tau_j = 200$, $j = 1, 2, 3$. In the second plot of Figure 1, we observe that the optimal tuning parameters are $\tau_j = 30$, $j = 1, 2, 3$, and as the values of τ increase the obtained curve fits the observed data less closely.

5. Detecting Serial Correlation

A statistical model is an approximation to the true process that generates the observed data. After fitting the model given by Equation (1), it is necessary to check whether the residuals obtained from the fit behave like a white noise process. A significant departure from this assumption suggests the inadequacy of the assumed form of the model. Thus, it is important to assess whether the residuals follow a white noise process.

Detecting serial correlation becomes more challenging when the available observations are unequally spaced in time. If the observations are unequally spaced, so are the errors. In order to study the spectral density of the residuals obtained from the fitted model, in this section we derive the mathematical definition of the spectrum for unequally spaced time series.

Before presenting our approach, we briefly review the results given by Deeming (1975) about the relationships between the periodogram, the spectral density (PSD), and the autocorrelation function for *continuous* time series. Then, we extend the results given by Deeming (1975) to the case of *discrete* time series.

Let $\{\varepsilon_i\}$ be a *continuous*, zero-mean stationary times series with spectral density

$$P_\varepsilon(\lambda) = \int_{-\infty}^{\infty} r_\varepsilon(h) \exp(i\lambda h) dh, \quad -\infty < \lambda < \infty,$$

and autocovariance function

$$r_\varepsilon(h) = \int_{-\infty}^{\infty} P_\varepsilon(\lambda) \exp(-i\lambda h) d\lambda, \quad h \in \mathbb{R}.$$

Consider a time series $\varepsilon_1, \dots, \varepsilon_N$ with spectrum $P_\varepsilon(\cdot)$ and autocovariance function $r_\varepsilon(\cdot)$, and assume that the observations $\varepsilon_1, \dots, \varepsilon_N$ are obtained at *unequally spaced* times t_1, \dots, t_N , respectively. The periodogram of $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)^\top$ at frequency λ is defined as

$$I_\varepsilon(\lambda) = \sum_{k=1}^N \sum_{j=1}^N \varepsilon_k \varepsilon_j \exp(i\lambda[t_k - t_j]), \quad \lambda = 2\pi f. \quad (15)$$

Deeming (1975) proved that the expectation of the periodogram of $\{\varepsilon_i\}$ in Equation (15) is

$$E[I_\varepsilon(\lambda)] = P_\varepsilon(\lambda) \star W_\varepsilon(\lambda), \quad (16)$$

where $W_\varepsilon(\lambda)$ is the power spectral window given by

$$W_\varepsilon(\lambda) = \sum_{j=1}^N \sum_{k=1}^N \exp(i\lambda[t_k - t_j]),$$

and $P_\varepsilon(\lambda) \star W_\varepsilon(\lambda)$ is the continuous convolution of $P_\varepsilon(\lambda)$ with $W_\varepsilon(\lambda)$ defined as

$$P_\varepsilon(\lambda) \star W_\varepsilon(\lambda) = \int_{-\infty}^{\infty} P_\varepsilon(\omega) W_\varepsilon(\lambda - \omega) d\omega.$$

The following lemma states that it is possible to extend the result in Equation (16) to the case of a *discrete* zero-mean stationary times series that is generated according to *equally spaced* times but observed at *unequally spaced* times. The lemma applies to unequally spaced time points t_i with index i

belonging to a subset \mathcal{I} of the set $\mathbb{N} = \{1, 2, \dots\}$ of positive integers.

Lemma 1. *Let $\{\varepsilon_i, t_i = t_0 + i\Delta, \Delta > 0, i \in \mathcal{I} \subseteq \mathbb{N}\}$ be a zero-mean, stationary, discrete time series with spectral density*

$$P_\varepsilon(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \exp(i\lambda h \Delta) r_\varepsilon(h), \quad -\infty \leq \lambda \leq \infty, \quad (17)$$

with autocovariance function defined as $r_\varepsilon(h) = E[\varepsilon_k \varepsilon_j]$, with $k = j + |h|$, $h \in \mathbb{Z}$, that can be expressed in term of the spectral density in Equation (17) as

$$r_\varepsilon(h) = \frac{2\pi}{N_{\mathcal{I}}} \sum_{j=1}^{N_{\mathcal{I}}} \exp(-i\lambda_j h \Delta) P_\varepsilon(\lambda_j), \quad h \in \mathbb{Z}, \quad (18)$$

where $\lambda_j = 2\pi f_j$, with $f_j = j/(N_{\mathcal{I}} \Delta)$ and $N_{\mathcal{I}} = \max\{I\}$. Then, the expectation of the periodogram in Equation (15) obtained from $\{\varepsilon_i, i \in \mathcal{I}\}$ is

$$E[I_\varepsilon(\lambda)] = \frac{2\pi}{N_{\mathcal{I}}} P_\varepsilon(\lambda) \star W_\varepsilon(\lambda), \quad (19)$$

with power spectral window given by

$$W_\varepsilon(\lambda) = \sum_{k \in \mathcal{I}} \sum_{j \in \mathcal{I}} \exp(i\lambda[t_k - t_j]) \quad (20)$$

and $P_\varepsilon(\lambda) \star W_\varepsilon(\lambda)$ is the discrete convolution of $P_\varepsilon(\lambda)$ with $W_\varepsilon(\lambda)$ defined as

$$P_\varepsilon(\lambda) \star W_\varepsilon(\lambda) = \sum_{j=1}^{N_{\mathcal{I}}} P_\varepsilon(\omega_j) \star W_\varepsilon(\lambda - \omega_j),$$

$$\omega_j = 2\pi f_j, \quad f_j = \frac{j}{N_{\mathcal{I}} \Delta}.$$

Our result in Equation (19) differs from the result by Deeming (1975) in Equation (16). Deeming (1975) proved that the expectation of both discrete and continuous Fourier transforms of a *continuous stochastic process* $f(t)$; (in the sense of Equations (31) and (32) in Deeming 1975) is equal to the *continuous convolution* of the spectral density of $f(t)$ with a spectral window (see Equations (33) and (36) in Deeming 1975). In Lemma 1, instead, we prove that the expectation of the discrete Fourier transform of the *discrete stochastic process* ε_t is equal to the *discrete convolution* of the spectral density of ε_t with a spectral window (up to the constant $2\pi/N_{\mathcal{I}}$).

When the time series is generated according to an equally spaced stochastic process and the observations are equally spaced, the periodogram is an (asymptotically) unbiased estimator of the spectral density (see Priestley 1981, page 418). However, when the observations are unequally spaced it does not make sense to estimate the spectral density in the same way. This is due to the *power spectral window* $W_\varepsilon(\lambda)$ in Equations (19)–(20). Nevertheless, as we show in the following proposition, it is possible to disentangle the spectral density $P_\varepsilon(\lambda)$ from the spectral window $W_\varepsilon(\lambda)$.

Proposition 1. *Let $\mathcal{F}\{g_i\}[k]$ denote the Discrete Fourier Transform of the sequence of m numbers g_1, \dots, g_m into*

another sequence h_1, \dots, h_m , that is,

$$\begin{aligned} h_k &= \mathcal{F}\{g_j\}[k] = \sum_{j=1}^m g_j \exp(-ikj2\pi/m), \\ k &= 1, \dots, m. \end{aligned} \quad (21)$$

Accordingly, define $\mathcal{F}^{-1}\{h_k\}[j]$ as the Inverse Discrete Fourier Transform of the sequence h_1, \dots, h_m into another sequence g_1, \dots, g_m , that is,

$$\begin{aligned} g_j &= \mathcal{F}^{-1}\{h_k\}[j] = \frac{1}{m} \sum_{k=1}^m h_k \exp(ikj2\pi/m), \\ j &= 1, \dots, m. \end{aligned} \quad (22)$$

Assume that $\{\varepsilon_i, i \in \mathcal{I}\}$ satisfy the same conditions as in Lemma 1. Then we can write the spectral density $P_\varepsilon(\lambda_j)$ in Equation (17) at frequency $\lambda_j = 2\pi f_j$, with $f_j = j/(N_{\mathcal{I}}\Delta)$, as

$$\begin{aligned} P_\varepsilon(\lambda_j) &= \frac{N_{\mathcal{I}}}{2\pi} \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{E[I_\varepsilon(\lambda_j)]\}[k]}{\mathcal{F}\{W_\varepsilon(\lambda_j)\}[k]} \right\} [j], \\ j &= 1, \dots, N_{\mathcal{I}}. \end{aligned} \quad (23)$$

The proofs of Lemma 1 and Proposition 1 are given in Appendix D. Equation (23) suggests that in order to estimate $P_\varepsilon(\lambda)$, we need the value of Δ and $W_\varepsilon(\lambda)$, as well as the estimate of $E[I_\varepsilon(\lambda)]$. Notice that, in general, the time series denoted in this section as $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)^\top$ is possibly autocorrelated, whereas the errors $\mathbf{z} = (z_1, \dots, z_N)^\top$ of our model in Equation (1) are assumed to be serially uncorrelated. The following algorithm is proposed to establish whether the unequally spaced residuals obtained when fitting our model in Equation (1) are uncorrelated.

1. Obtain the residuals $\hat{\mathbf{z}} = \mathbf{Y} - \hat{\mathbf{Y}}$.
2. For each $j = 1, \dots, N_{\mathcal{I}}$, define $\hat{I}_{\hat{\mathbf{z}}}(\lambda_j)$ as the periodogram in Equation (15) computed upon the residuals $\hat{\mathbf{z}}$ obtained in step (1), with $\lambda_j = \frac{2\pi j}{N_{\mathcal{I}}\Delta}$.
3. For each $j = 1, \dots, N_{\mathcal{I}}$, define $W_{\hat{\mathbf{z}}}(\lambda_j)$ as the power spectral window in Equation (20) computed upon the residuals $\hat{\mathbf{z}}$ obtained in step (1), with $\lambda_j = \frac{2\pi j}{N_{\mathcal{I}}\Delta}$.
4. Smooth the periodogram obtained in step (2) over frequencies, and denote the smoothed periodogram by $\tilde{I}_{\hat{\mathbf{z}}}(\lambda_j)$.
5. Calculate the Discrete Fourier Transform in Equation (21) of the power spectral window and the periodogram obtained in steps (3) and (4), respectively.
6. For each frequency $\lambda_j = \frac{2\pi j}{N_{\mathcal{I}}\Delta}$, define the estimated spectral density of the errors \mathbf{z} as

$$\begin{aligned} \hat{P}_{\hat{\mathbf{z}}}(\lambda_j) &= \frac{N_{\mathcal{I}}}{2\pi} \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{\tilde{I}_{\hat{\mathbf{z}}}(\lambda_j)\}[k]}{\mathcal{F}\{W_{\hat{\mathbf{z}}}(\lambda_j)\}[k]} \right\} [j], \\ j &= 1, \dots, N_{\mathcal{I}}, \end{aligned} \quad (24)$$

where the inverse Fourier transform \mathcal{F}^{-1} is given by Equation (22).

7. If the estimated spectral density obtained in step (6) does not vary significantly over frequencies, conclude that the errors are uncorrelated over time.

6. Simulation Results

In this section we provide Monte Carlo simulations to illustrate the performance of the estimators $\hat{\mu}(t)$, $\hat{m}(t)$, $\hat{g}_{\ell,k}(t)$, $\ell = 1, 2, k = 1, \dots, K$, defined by Equations (12) and (13), and the estimator of the spectral density in Equation (23).

In Section 6.1 we simulate unequally spaced observations from the model in Equation (1), under two scenarios. In the first scenario both the trend and amplitudes are sinusoidal, whereas in the second scenario the trend and amplitudes are polynomial. In Section 6.2 we simulate a Blazhko light curve and fit the model in Equation (1). Finally in Section 6.3 we evaluate the performance of the estimator of the spectral density defined in Equation (23) of a discrete unequally spaced time series.

6.1. Simulating Our Novel Time-varying Model

In this section we generate the data according to the model described by Equation (1) with $K=2$, $N=500$, and time t is unequally spaced, obtained from a uniform distribution $U(\theta_1, \theta_2)$ with $\theta_1=0$ and $\theta_2=1$. In order to illustrate the flexibility of our novel method, we consider two different scenarios for the trend and amplitudes. In the first scenario, we simulate a sinusoidal trend and amplitudes as $m(t) = \sin(2\pi t)$, $g_{1,1}(t) = \cos(9\pi t)$, $g_{2,1}(t) = \sin(6\pi t)$, $g_{1,2}(t) = \cos(4\pi t)$, $g_{2,2}(t) = \sin(7\pi t)$, with frequencies $w_1 = 40\pi$ and $w_2 = 100\pi$. In the second scenario, we simulate a (global) polynomial trend and amplitudes as $m(t) = 0.2t - 5t^2 + 5.5t^3$, $g_{1,1}(t) = 4t^3 - 5t^2$, $g_{2,1}(t) = -0.5 - 0.5t + 2.5t^2 - 0.5t^3$, $g_{1,2}(t) = -t + t^2 + 1.3t^3$, $g_{2,2}(t) = 0.5 + 2t^2 - 3t^3$, with frequencies $w_1 = 30\pi$ and $w_2 = 40\pi$. In both scenarios, we assume that the error terms $\{z_i, i = 1, \dots, N\}$ follow a Gaussian distribution with zero mean and variance $\sigma_z^2 = 2$.

In both scenarios, we simulate $M=200$ realizations of the model in Equation (1). For each $j=1, \dots, M$, we compute the estimate $\hat{\theta}^{(j)}$ defined by Equation (11). In the first scenario, we select the smoothing parameter $\tau = (50, 1, 2, 10, 1)^\top$, a total number of B -splines $J=33$ of order $d=3$, and an order penalty $r=2$. In the second scenario, we choose the smoothing parameter $\tau = (3, 3, 3, 3, 3)^\top$, a total number of B -splines $J=6$ of order $d=3$, and an order penalty $r=4$. Figure 2 shows our estimates of μ , m , $\{g_{\ell,k}, \ell = 1, 2, k = 1, 2\}$, and their 95% confidence intervals. Figure 2 shows that our model in Equation (1) fits well the simulated data in both the sinusoidal and polynomial scenarios. That is, the trend and amplitudes are well fitted in both scenarios. The 95% confidence intervals are constructed in a nonparametric fashion using quantiles; see Appendix C.1 for more details.

6.2. Simulating a Blazhko RR Lyrae Light Curve Characterized by Amplitude Modulation

We simulate a Blazhko RR Lyrae light curve with amplitude modulation according to Benkő et al. (2011) as

$$\begin{aligned} Y_i &= \mu(t_i) + z_i, \quad i = 1, \dots, N = 1000, \\ \mu(t_i) &= \left[1 + \frac{U_m(t_i)}{U_c} \right] c(t), \\ c(t) &= a_0 + \sum_{k=1}^4 a_k \sin(2\pi k f_0 t + \varphi_k), \\ U_m(t) &= a_m \sin(2\pi f_m t + \varphi_m), \end{aligned} \quad (25)$$

where $c(t)$ is the carrier wave with four harmonic components, $U_m(t)$ is the modulating signal, $U_c = a_m/h$ is the amplitude of the unmodulated light curve, and $\{z_i, i = 1, \dots, N\}$ are the error terms.

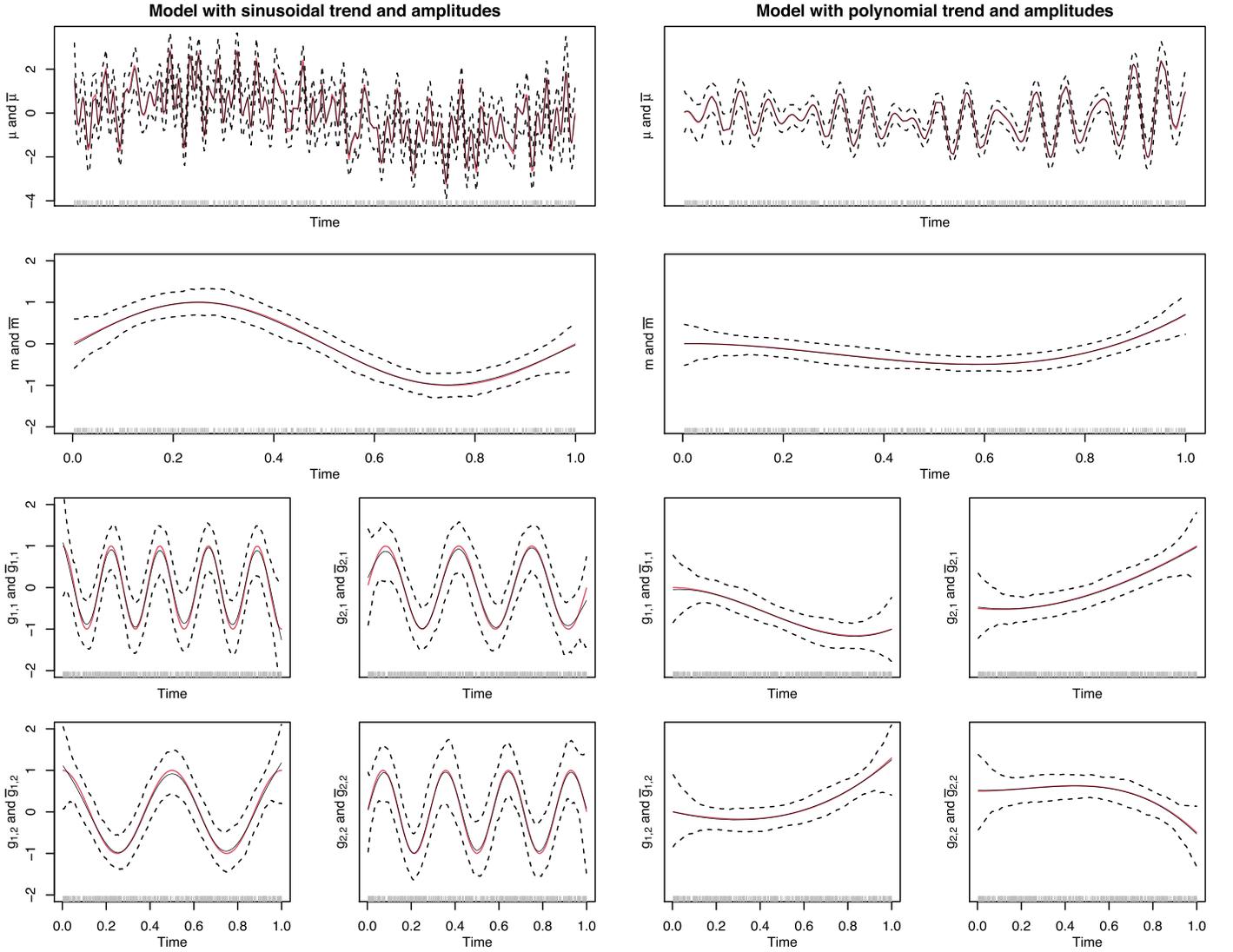


Figure 2. Simulation scenarios of Section 6.1: data generated from the model in Equation (1) with *sinusoidal* and *polynomial* time-varying trends and amplitudes. Time t is unequally spaced, obtained from the uniform distribution $U(0, 1)$. The first column shows the fit of the model in Equation (1) with a sinusoidal trend and amplitudes, whereas the second column shows the fit of the model in Equation (1) with a polynomial trend and amplitudes. From the $M = 200$ realizations of our estimators we compute, for each fixed t , three averages and confidence intervals. The first row shows the true $\mu(t)$ (red solid line), together with the average $\bar{\mu}(t) = \frac{1}{M} \sum_{j=1}^M \hat{\mu}^{(j)}(t)$ of the estimates $\hat{\mu}^{(j)}(t)$ (black solid line). The second row shows the true trend $m(t)$ together with the average $\bar{m}(t) = \frac{1}{M} \sum_{j=1}^M \hat{m}^{(j)}(t)$ of the estimates $\hat{m}^{(j)}(t)$. The third and fourth rows show the true amplitudes $g_{\ell,k}(t)$, $\ell = 1, 2$, $k = 1, 2$, together with the average $\bar{g}_{\ell,k}(t) = \frac{1}{M} \sum_{j=1}^M \hat{g}_{\ell,k}^{(j)}(t)$ of the estimates $\hat{g}_{\ell,k}^{(j)}(t)$. The nonparametric quantiles (black dashed lines) are the confidence intervals corresponding to the 2.5th and 97.5th order statistics, respectively; see Appendix C.1.

The values of the parameters used in Equation (25) and the time design are obtained from Benkő et al. (2011). In particular, $a_m = 0.1$ mag, $h = 1.2$, $a_0 = 0.01$ mag, $f_m = 0.05$ days $^{-1}$, $\varphi_m = 270^\circ$, and the values $\{a_k, \varphi_k, 1 \leq k \leq 4\}$ are presented in Table 1. We convert the Blazhko phase φ_m and the main phases $\{\varphi_k, 1 \leq k \leq 4\}$ in Equation (25) from degrees to radians using the R function `NISTdegToRadian` (available in the R package `NISTunits2016` by Gama 2016). The original time design $\{t_j, j = 1, \dots, 28,799\}$ in Benkő et al. (2011) is equally spaced. However, variable stars are often observed at irregular intervals. For this reason, in our simulation exercise we sample a subset of the original time points and use this subset to evaluate the performance of our method. We obtain the time design $\{t_i, i = 1, \dots, 1000\}$ in Equation (25) by sampling the original, equally spaced time design $\{t_j, j = 1, \dots, 28,799\}$. We end up with $N = 1000$ unequally

spaced observations ranging from $t = 0.03819$ days to $t = 69.37847$ days. The error terms are generated independently from a Gaussian distribution with zero mean and variance $\sigma_z^2 = 0.005$.

If we consider of our novel model in Equation (1), with time-varying trend and amplitudes specified as

$$\begin{aligned}
 m(t_i) &= a_0 [1 + U_m(t_i)/U_c], \\
 g_{1,k}(t_i) &= a_k \cos(\varphi_k) [1 + U_m(t_i)/U_c], \\
 k &= 1, \dots, 4, \\
 g_{2,k}(t_i) &= a_k \sin(\varphi_k) [1 + U_m(t_i)/U_c], \\
 k &= 1, \dots, 4,
 \end{aligned} \tag{26}$$

we can rewrite the model in Equation (25) as a special case of our model given by Equation (1). The main advantage of fitting the model in Equation (1) instead of the model in Equation (25)

Table 1
Parameters of Simulated RR Lyrae Star

k	kf_0 (days ⁻¹)	a_k (mag)	φ_k (°)
1	2	0.401	5.490
2	4	0.171	144.040
3	6	0.133	285.250
4	8	0.097	81.290

Note. Parameters (frequencies, amplitudes, and phases) obtained from Benkő et al. (2011), as explained in Section 6.2.

is that one does not need to estimate the parameters a_m , h , a_0 , φ_m , a_k , φ_k , $k = 1, \dots, 4, f_m$. Moreover, we do not need to adopt any specific functional form for $m(\cdot)$ and $g(\cdot)$, such as those given by Equation (26), because they are well approximated by B -splines.

We fit the model in Equation (1) with $K=4$ to the data generated according to the model in Equation (25). We assume that the frequencies f_k , $k = 1, \dots, 4$, of each harmonic component are known; see Table 1. Also, we use a total of $J = 18$ B -splines of degree $d=3$, an order penalty $r=1$, and the smoothing parameters $\tau = (5, 1, 0.1, 0.1, 0.1, 0.1, 1, 0.1, 4)^T$.

We fit the model in Equation (1) to the simulated data obtained from the model in Equation (25), and present the results in Figure 3. The first row shows the simulation of the amplitude modulated RR Lyrae light curve given by Equation (25); (gray points), together with the true and fitted curve (solid red and solid black lines, respectively). The second row shows the residuals, and the third and fourth rows show the true trend and amplitudes (red lines) given by Equation (26) and their fits (black lines). We observe from Figure 3 that the model in Equation (1) fits well the simulated data. That is, the trend and amplitudes are well fitted, and the residuals satisfy the assumption of zero mean and constant variance. The 95% confidence intervals are constructed in a parametric fashion; see Appendix C.2 for more details.

6.3. Estimating the Spectral Density of Unequally Spaced Time Series

In this section we estimate the spectral density of unequally spaced time series by means of our novel estimator in Equation (23). To this end, we simulate unequally spaced observations generated from the following AR(2) process:

$$\begin{aligned} \varepsilon_i &= \phi_1 \varepsilon_{i-1} + \phi_2 \varepsilon_{i-2} + z_i \\ z_i &\sim \mathcal{N}(0, \sigma_z^2) \\ t_i &= t_0 + i\Delta, \end{aligned} \quad (27)$$

where $i = 1, \dots, N$, with $N = 500$ equally spaced observations, starting time $t_0 = 0.67$, and $\Delta = 0.33$. In order to simulate a realistic AR(2) process, we use the coefficients of the sunspot numbers in Example 3.2.9 of Brockwell & Davis (2016), where $\phi_1 = 1.318$, $\phi_2 = -0.634$, and $\sigma_z^2 = 289.2$. These ϕ coefficients ensure the existence of a causal solution

$$\varepsilon_i = \sum_{j=0}^{\infty} \psi_j z_{i-j} \quad (28)$$

of Equation (27). The time series in Equation (28) is *causal* in the sense that ε depends upon current and past (rather than

future) values of the error term z . We simulate $M = 500$ times the AR(2) model given by Equation (27) obtaining the observations $\varepsilon_1^{(m)}, \dots, \varepsilon_N^{(m)}$, $m = 1, \dots, M$. Then, in order to obtain unequally spaced observations we use the following three steps.

1. We divide time into 50 blocks, where each block has 10 observations, in a way to preserve the original time series structure.
2. In order to preserve the autocorrelation between the observations, we select randomly 30 blocks and collect the time points corresponding to these blocks, obtaining a new set of time points $\{t_i^*, i = 1, \dots, n\}$, with $n = 300$ observations. In contrast to the simulation schemes of Sections 6.1 and 6.2 where time was sampled randomly, here data sets with uniformly sampled subsets are produced. While the former sampling is close to the data distribution of large ground-based surveys, the latter is the typical sampling of photometric space telescopes that are dedicated to high-cadence time series observations, such as Kepler (Koch et al. 2010).
3. Finally, we collect the observations $\varepsilon_i^{(m)}$ corresponding to the new set of time points $\{t_i^*, i = 1, \dots, n\}$ and rename them as $e_i^{(m)} = \varepsilon_i^{(m)}$, with $e_i^{(m)}$ being observed at time t_i^* , $i = 1, \dots, n$.

Thus, we obtain the unequally spaced observations $e_1^{(m)}, \dots, e_n^{(m)}$, which represent a subset of the equally spaced time series $\varepsilon_1^{(m)}, \dots, \varepsilon_N^{(m)}$. For each $m = 1, \dots, M$, and each fixed frequency $\lambda_j = 2\pi f_j$, $f = j/(N\Delta)$, $j = 1, \dots, N$, we compute the periodogram of $e_1^{(m)}, \dots, e_n^{(m)}$ as

$$I_e^{(m)}(\lambda_j) = \sum_{k=1}^n \sum_{d=1}^n e_k^{(m)} e_d^{(m)} \exp(i\lambda_j [t_k^* - t_d^*]),$$

the average of the periodograms $I_e^{(m)}(\lambda_j)$ as

$$\bar{I}_e(\lambda_j) = \frac{1}{M} \sum_{m=1}^M I_e^{(m)}(\lambda_j),$$

and the power spectral window of $e_1^{(m)}, \dots, e_n^{(m)}$ as

$$W_e(\lambda_j) = \sum_{d=1}^n \sum_{k=1}^n \exp(i\lambda_j [t_k^* - t_d^*]).$$

For each frequency λ_j , $j = 1, \dots, N$, replacing $E[I_e(\lambda_j)]$ with $\bar{I}_e(\lambda_j)$ and substituting $W_e(\lambda_j)$ in Equation (23), the estimated spectral density of the unequally spaced time series $e_1^{(m)}, \dots, e_n^{(m)}$ is given by

$$\hat{P}_e(\lambda_j) = \frac{N}{2\pi} \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{\bar{I}_e(\lambda_j)\}[k]}{\mathcal{F}\{W_e(\lambda_j)\}[k]} \right\} [j]. \quad (29)$$

A smooth version of the estimated spectral density in Equation (29) is

$$\tilde{P}_e(\lambda_j) = \frac{1}{N_\lambda} \sum_{i=1}^{N_\lambda} K_h(\lambda_j - \lambda_i) \hat{P}_e(\lambda_j). \quad (30)$$

The rescaled kernel function is defined as $K_h(x) = \frac{1}{h} K(x/h)$, where K is a second-order kernel and h is the bandwidth. For this application, we used the Gaussian kernel $K(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2)$ and a bandwidth $h = 0.3$.

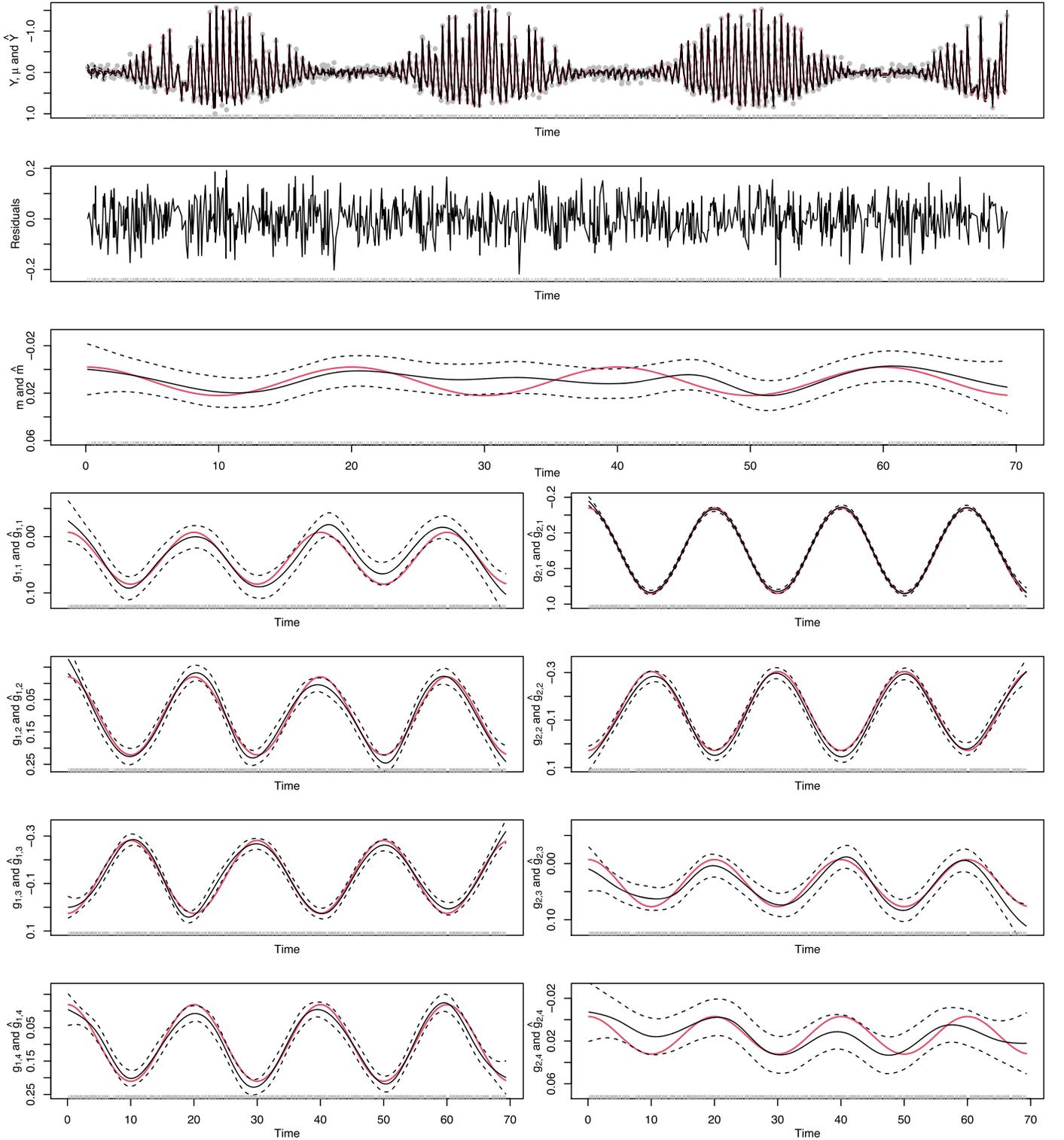


Figure 3. Simulated Blazhko RR Lyrae light curve characterized by amplitude modulation; see Section 6.2. The first row shows the light-curve data of an RR Lyrae star simulated according to the model given by Equation (25) (gray points), the curve $\mu(t)$ in Equation (25) (solid red line), the prediction \hat{V}_i (solid black line), and the parametric 95% confidence intervals (dashed black lines) obtained according to Appendix C.2. The second row shows the residuals. The third row shows the time-varying trend $m(t_i)$; (red solid lines) simulated according to Equation (26), together with the estimated trend $\hat{m}(t_i)$ (solid black lines) obtained according to Equation (13). The last four rows show the time-varying amplitudes $\{g_{\ell,k}(t_i)\}$, $\ell = 1, 2, k = 1, \dots, 4$ (red solid lines) simulated according to Equation (26), together with their estimates $\hat{g}_{\ell,k}(t_i)$ (solid black lines) given by Equation (13). The parametric 95% confidence intervals (dashed black lines) are obtained according to Appendix C.2.

Figure 4 compares the underlying spectral density $P_\varepsilon(\lambda_j)$ of the equally spaced time series $\{\varepsilon_i\}$, with the estimated spectral densities $\hat{P}_e(\lambda_j)$ and $\tilde{P}_e(\lambda_j)$ of the unequally spaced time series

$\{e_i\}$. The underlying spectral density of the equally spaced time series $\{\varepsilon_i\}$, generated by the AR(2) process in Equation (27), is given by

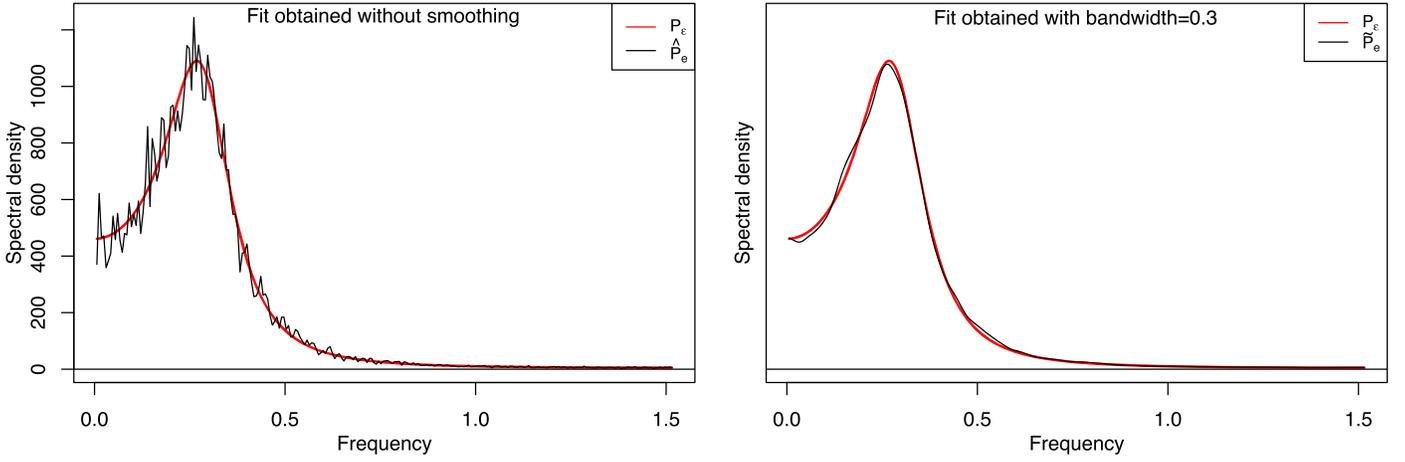


Figure 4. Estimated spectral density of the unequally spaced time series sampled by blocks in Section 6.3. Left: comparison between the true spectral density in Equation (31) (red line) and the estimated spectral density $\hat{P}_\varepsilon(\lambda_j)$ in Equation (29) (black line) for $j = 1, \dots, N/2$. Right: comparison between the true spectral density in Equation (31) (red line) and the smooth estimated spectral density $\tilde{P}_\varepsilon(\lambda_j)$ in Equation (30) (black line) for $j = 1, \dots, N/2$. The true spectral density corresponds to the equally spaced time series ε_j , which follows the AR(2) process given by Equation (27), whereas the estimated spectral density is computed from the unequally spaced observations $e_1^{(m)}, \dots, e_n^{(m)}$, $m = 1, \dots, M$. The unequally spaced time series $e_1^{(m)}, \dots, e_n^{(m)}$ is obtained as a subset of the time series $\varepsilon_1^{(m)}, \dots, \varepsilon_N^{(m)}$. In this example $N = 500$, $n = 300$, and $M = 500$.

$$P_\varepsilon(\lambda_j) = \frac{\sigma_\varepsilon^2}{2\pi} [1 + \phi_1^2 + \phi_2^2 + 2\phi_2 + 2(\phi_1\phi_2 - \phi_1)\cos(\lambda_j\Delta) - 4\phi_2\cos^2(\lambda_j\Delta)]^{-1}. \quad (31)$$

The estimated spectral density $\hat{P}_\varepsilon(\lambda_j)$ of the unequally time series $\{e_j\}$ is given in Equation (29), and its smooth version $\tilde{P}_\varepsilon(\lambda_j)$ in Equation (30). Figure 4 shows that the estimated spectral density of the unequally time series, $\tilde{P}_\varepsilon(\lambda_j)$, fits very well the true spectral density $P_\varepsilon(\lambda_j)$.

7. Application to Real Data

In this section, we fit our model in Equation (1) and the model proposed by Benkő (2018) in Equation (5) to the same light curve: the V783 Cyg, KIC 5559631. This time series has 61,351 unequally spaced observations and is available online from the Konkoly Observatory of the Hungarian Academy of Sciences webpage.⁷ We choose this particular light curve for two reasons. First, the Blazhko effect of the V783 Cyg time series is known to be characterized by a sinusoidal amplitude and frequency modulation (Benkő et al. 2014). The light curve of V783 Cyg can be described by $K = 15$ significant harmonics with a sinusoidal amplitude and frequency modulations (Benkő et al. 2014), which makes V783 Cyg an ideal target for comparing the fits obtained with the models in Equations (1) and (5). Second, these two modulations are well captured and fitted by our novel model in Equation (1).

In order to reduce the computational time and to satisfy the condition $t_i = t_0 + i\Delta$ with $\Delta > 0$ and $i \in \mathcal{I} \subseteq \mathbb{N}$ (which is required by Proposition 1), we analyze a ≈ 78 days segment of this light curve from $t = 827.44$ days to $t = 904.9$ days. For this segment, the time origin and the time spacing take the values $t_0 = 827.42$ days and $\Delta = 0.0204345$ days, respectively, with a total of $N = 2101$ unequally spaced observations.

When fitting the models in Equations (1) and (5), the main pulsation and modulation frequencies are not estimated: they take the values $f_0 = 1.611084$ days⁻¹ and $f_m = 0.036058$ days⁻¹ (see

Benkő et al. 2014), respectively. In addition to the $K = 15$ significant harmonics fitted by Benkő et al. (2014), we found, after pre-whitening and fitting our model in Equation (1), four significant frequencies taking the values $f'_{11} = 18.3254$ days⁻¹, $f'_{12} = 19.9365$ days⁻¹, $f'_{13} = 21.5476$ days⁻¹, and $f'_{14} = 23.1587$ days⁻¹. The values we obtain for $\{f'_j, 11 \leq j \leq 14\}$ demonstrate that these frequencies are not harmonics of the form kf_0 , which might suggest that these four are independent frequencies. Interestingly, however, we find that the latter belong to a set of fourteen “reflection frequencies” of the form $\{f'_j = 2f_N - (30 - j)f_0, 1 \leq j \leq 14\}$, where $f_N = 24.46$ days⁻¹ is the Nyquist frequency. Among these fourteen frequencies, only the last six $\{f'_j, 9 \leq j \leq 14\}$ exhibit significant peaks in the Lomb–Scargle periodogram (computed according to Lomb 1976). However, to avoid overfitting, we only consider the four frequencies $\{f'_j, 11 \leq j \leq 14\}$ corresponding to last four peaks of the estimated power spectrum (see the last row of Figure 5, bottom right panel). In summary, the only truly independent frequencies are f_0, f_m and f_N , the other frequencies $\{f_k = kf_0, k = 1, \dots, 15\}$ and $\{f'_j = 2f_N - (30 - j)f_0, 1 \leq j \leq 14\}$ being linear combinations (or harmonics) of those.

The frequencies f'_j do not depend on the Blazhko frequency f_m , as the information regarding the Blazhko effect is captured by the time-varying trend $m(\cdot)$ and amplitudes $\{g_{\ell,k}(\cdot), \ell = 1, 2, k = 1, \dots, K\}$; see Equations (9)–(10) and Figure 6. We use a different notation (f' rather than f) to avoid confusion, since in this case $f_N < 16f_0 = 25.77$ days⁻¹, and the four frequencies we are considering take values < 24 days⁻¹.

After fitting the models in Equations (1) and (5), we compute their residuals and estimate their spectral densities using Equation (24). To estimate the spectral densities according to the procedure in Section 5, we adopt the Gaussian kernel $K(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2)$ with a bandwidth $h = 7.2$. We fitted both models with a PC having a 2.7 GHz 12 core Intel Xeon E5 processor and 64 GB of 1866 MHz DDR3 memory. Fitting our novel model in Equation (1) required 17 minutes and 13 s, whereas fitting the model by Benkő (2018) in Equation (5) required 12 minutes and 28 s.

⁷ https://konkoly.hu/KIK/data_en.html

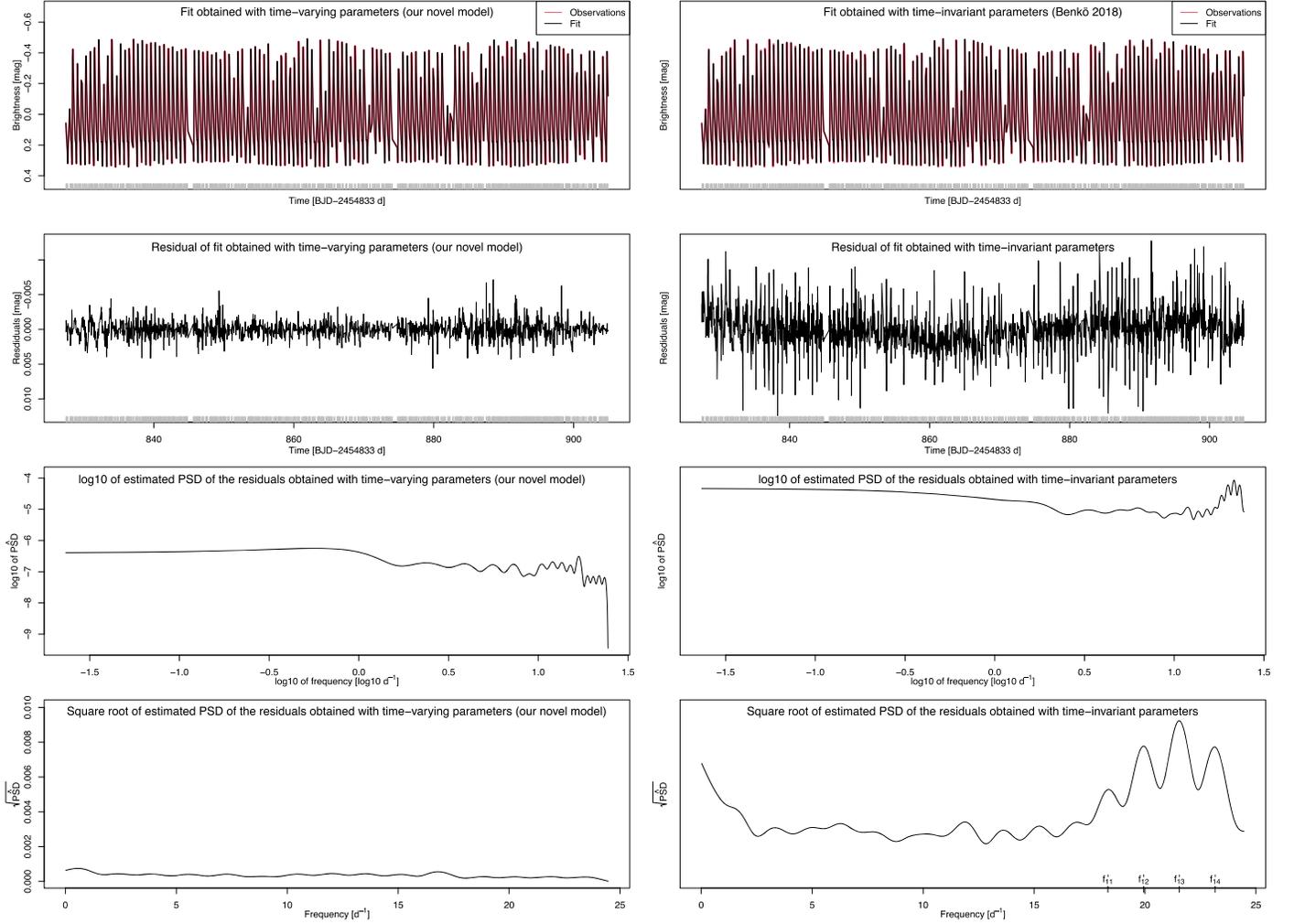


Figure 5. Comparison of fitted models to the light curve of V783 Cyg in Section 7.1. The first column corresponds to the fit of the model in Equation (1), whereas the second column corresponds to the fit of the model in Equation (5). From top to bottom: the first row shows the Brightness mag (red solid lines) together with the fits (black solid lines). The second row shows the residuals resulting from the fits. The last two rows show the spectral density of the residual obtained with Equation (24) under different transformations (log-10 scale and square root).

The description provided so far applies to both fits of models in Equations (1) and (5). We now provide, separately, computational details about the estimation of these two models. Then in Sections 7.1 and 7.2 we compare and interpret the fits.

To fit our novel model in Equation (1), we apply the methodology described in Section 4. When fitting our model in Equation (1) we consider two sets of harmonic components. The first set is given by the harmonic components with frequencies $\{f_k = kf_0, k = 1, \dots, 15\}$ provided by Benkő et al. (2014), weighted by our amplitudes $\{g_{\ell,k}(t_i), \ell = 1, 2, k = 1, \dots, 15\}$. For the second set, the harmonic components are characterized by the four amplitudes $\{g'_{\ell,j}(t_i), \ell = 1, 2, j = 11, \dots, 14\}$ weighting the corresponding four frequencies $\{f'_j, j = 11, \dots, 14\}$. That is, we fit the following extended version:

$$\begin{aligned} \mu(t_i) = & m(t_i) + \sum_{k=1}^{15} \{g_{1,k}(t_i) \cos(\omega_k t_i) \\ & + g_{2,k}(t_i) \sin(\omega_k t_i)\} \\ & + \sum_{j=11}^{14} \{g'_{1,j}(t_i) \cos(\omega'_j t_i) + g'_{2,j}(t_i) \sin(\omega'_j t_i)\} \end{aligned}$$

of the model in Equation (1), with $\omega_k = 2\pi f_k$ and $\omega'_j = 2\pi f'_j$. The resulting fitted model involves a total of $39J$ parameters. Before fitting our model, we select the smoothing parameters τ

and the number of B -splines J . These parameters are selected by the AIC criterion described in Section 4.2. To simplify the selection of the smoothing parameters τ , we consider the case $\tau_2 = \dots = \tau_{11}, \tau_{12} = \dots = \tau_{21}, \tau_{22} = \dots = \tau_{31}$, and $\tau_{32} = \dots = \tau_{39}$. We pick the smoothing parameter τ_1 over the grid $\{0, 0.1, 10\}$, the parameters $\{\tau_k, k = 2, \dots, 39\}$ over the grid $\{0, 0.1, 10, 100\}$, and the total number J of B -splines (of degree $d = 3$) over the grid $\{8, 13, 23, 33\}$. We apply the AIC formula in Equation (14). The lowest AIC value occurs for $J = 33$ B -splines, $\tau_1 = 0, \tau_k = 0.1, k = 2, \dots, 21, \tau_k = 0, k = 22, \dots, 31$, and $\tau_k = 10, k = 32, \dots, 39$.

To fit the model in Equation (5), we implement the Levenberg–Marquardt algorithm using the R function `nls.lm` (available in the R package `minpack.lm` by Elzhov et al. 2016), with $K = 15$ and $\ell = \ell_k^A = \ell_k^F = 1, k = 1, \dots, K$, for a total of 93 parameters.

7.1. Comparing the Accuracy of the Fits

The MSE corresponding to the fit of our model in Equation (1) is 0.000001, whereas the MSE of the model in Equation (5) is 0.000008. That is, the MSE of the model in

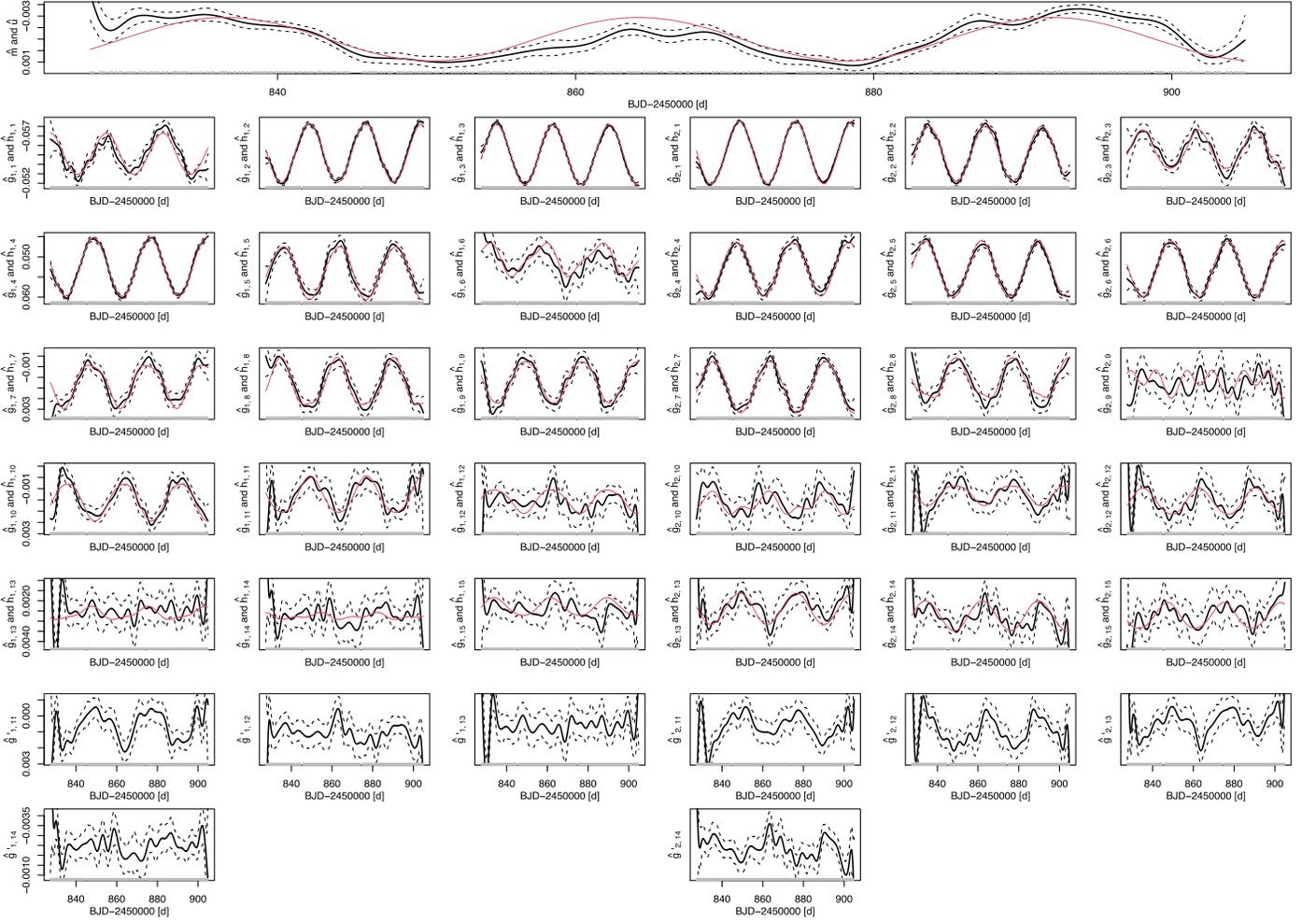


Figure 6. Comparing the estimated time-varying trend and amplitudes fitted to the light curve of V783 Cyg studied in Section 7.2. Red solid lines: estimates of $u(t)$ and $\{h_{\ell,k}(t), \ell = 1, 2, k = 1, \dots, 15\}$ defined in Equation (8) obtained with the model in Equation (5). Black solid lines: estimates of $m(t)$, $\{g_{\ell,k}(t), \ell = 1, 2, k = 1, \dots, 15\}$, and $\{g'_{\ell,j}(t), \ell = 1, 2, j = 11, \dots, 14\}$ obtained with our novel model in Equation (1). Black dashed lines: 95% confidence intervals for $m(t)$, $g_{\ell,k}(t)$, and $g'_{\ell,j}(t)$ obtained according to Appendix C.2.

Equation (1) is approximately 12.5% smaller than the MSE of the model in Equation (5). Fitting the model in Equation (1) involves 1287 parameters, whereas the number of parameters estimated with the model in Equation (5) is 93. The larger number of parameters needed to fit the model in Equation (1) is due to the semi-parametric form of the trend and amplitudes, which does not impose any particular shape to the underlying functions we estimate.

Figure 5 compares the fits of the models in Equations (1) and (5). The first row shows the fitted curves, the second row shows the residuals, and the third and fourth rows show the estimated spectral density of the residuals. Although the fitted curves (first row) look very similar, the residuals are significantly different. Indeed, the residuals obtained with the model in Equation (1) are compatible with the assumption of stationary and uncorrelated errors. By contrast, the residuals obtained with the model in Equation (5) exhibit a time-dependent trend. Moreover, the estimated spectral densities in the last two rows of Figure 5 show that the model in Equation (1) delivers residuals with a flat estimated spectral density, mimicking the behavior of the spectral density of white noise errors, whereas the model in Equation (5) shows that some harmonic components should be added to the model (see the peaks

between the frequencies 17 days^{-1} and 24 days^{-1} in the last row).

7.2. Comparing the Estimated Time-varying Parameters

In Section 3 we have showed that the model in Equation (5) is a special case of our novel model in Equation (1). To establish whether the fitted model in Equation (1) matches (or differs from) the fitted model in Equation (5), we now compare the estimates of $m(t)$ and $g_{\ell,k}(t)$ obtained by fitting the model in Equation (1) with the estimates of $u(t)$ and $h_{\ell,k}(t)$ defined in Equation (8) obtained by fitting the model in Equation (5).

Figure 6 shows the estimates $\hat{m}(t)$, $\{\hat{g}_{\ell,k}(t), \ell = 1, 2, k = 1, \dots, 15\}$, and $\{\hat{g}'_{\ell,j}(t), \ell = 1, 2, j = 11, \dots, 14\}$ (black lines), together with the estimates $\hat{u}(t)$ and $\{\hat{h}_{\ell,k}(t), \ell = 1, 2, k = 1, \dots, 15\}$ (red lines). The estimated trend $\hat{m}(t)$ is similar to the sinusoidal $\hat{u}(t)$. Similarly, the first eight estimated harmonic components $\{\hat{g}_{\ell,k}(t), \ell = 1, 2, k = 1, \dots, 8\}$ and $\{\hat{h}_{\ell,k}(t), \ell = 1, 2, k = 1, \dots, 8\}$ are very close to each other. The next seven estimated harmonic components $\{\hat{g}_{\ell,k}(t), \ell = 1, 2, k = 9, \dots, 15\}$ and $\{\hat{h}_{\ell,k}(t), \ell = 1, 2, k = 9, \dots, 15\}$ are still similar but in some cases are slightly different.

Nevertheless, these small differences do not have a significant impact on the fitted curves, because the last harmonic components have less of a contribution to the fit than the first ones. For the four estimated harmonic components associated with the frequencies $\{f'_j, j = 11, \dots, 14\}$, which were fitted only for the model in Equation (1)—and were not fitted for the model in Equation (5)—we observe that the four corresponding time-varying amplitudes $\{\hat{g}'_{\ell,j}(t), \ell = 1, 2, j = 11, \dots, 14\}$ are allowed to have either a sinusoidal or a nonsinusoidal form. This finding is in accordance with the form of Amplitude Modulation and Frequency Modulation of Blazhko stars described by Benkő (2018). Finally, in Figure 6, we see that most of the confidence intervals of $\hat{g}_{\ell,k}(t)$ contain $\hat{h}_{\ell,k}(t)$. Therefore we conclude the following. Although the modulation frequency f_m is not a parameter of our model in Equation (1), we are able to describe, through the estimated time-varying trend $\hat{m}(t)$ and amplitudes $\hat{g}_{\ell,k}(t)$, the Blazhko effect resulting from the amplitude and frequency modulation considered by the model in Equation (5).

8. Summary

In this article, we introduced a model for time series observations of variable stars that are modulated by smoothly time-varying mean magnitudes, amplitudes, and phases. Previous approaches assume that the underlying parameters are either time-invariant or piecewise-constant functions. From the modeling viewpoint, our approach is more flexible because it avoids assumptions about the functional form of the aforementioned time-dependent quantities. From the computational viewpoint, estimating our time-varying curves translates into the estimation of time-invariant parameters that can be performed by ordinary least squares.

An important challenge when dealing with astronomical time series is that observations are unequally spaced in time. In some cases, observations are unevenly spaced due to missing values. Missing values are sometimes handled via imputation, that is, the gap generated by the missing value is “filled in” by an estimated value. Our novel approach, which involves the classical periodogram, has the advantage of not relying on any imputation method.

We study the performance of our approach under several simulation scenarios. Finally, we apply our method to V783 Cyg (KIC 5559631), a well-known RR Lyrae star presenting the Blazhko effect. In this case, the effect is characterized by a sinusoidal amplitude and frequency modulation. When comparing the time-varying fit obtained with our novel model with the time-invariant fit obtained with the model proposed by Benkő (2018), we found that both amplitude and frequency modulations are well captured and fitted by our novel model, and also that our time-varying method outperforms the time-invariant fit. Indeed the estimation error obtained with our fit is significantly smaller than the error obtained with the time-invariant fit. In addition, the residuals obtained with our novel method are compatible with the assumption of stationary and uncorrelated errors, whereas the residuals obtained with the time-invariant model by Benkő (2018) exhibit a time-dependent trend and some significant spectral peaks.

In the future, we plan to extend our methodology in four important directions. First, we plan to apply our novel method

to the study of a larger sample of Blazhko RR Lyrae stars. Second, our approach can be extended to the analysis of other classes of variable stars presenting long-term changes in their light-curve shapes. Third, our fitting method does not require the period(s), amplitude(s), and phase(s) of the Blazhko effect to be determined, as we obtain instead the empirical functions $m(\cdot)$ and $g_{\ell,k}(\cdot)$. We are currently investigating what kind of (or how much more) information can be obtained from these empirical functions, as compared to conventional approaches. Finally, we aim to study Blazhko light curves characterized by more than one Blazhko frequency—V783 Cyg, which was addressed in some detail in this paper, is a special case, because this star does not show any additional Blazhko frequencies (Benkő et al. 2014).

The authors would like to thank two anonymous reviewers for helpful comments which led to significant improvements of the paper. D.S. was funded by the National Agency for Research and Development (ANID), Doctorado Nacional grant No. 2017-21171100. Support for G.M. and M.C. has been provided by ANID’s Millennium Science Initiative through grant No. ICN12 120009, awarded to the Millennium Institute of Astrophysics (MAS). M.C. acknowledges additional support by Proyecto Basal AFB-170002 and FONDECYT grant #1171273. We thank all of the participants of the Astronomical Data Science Workshop organized by Texas A&M University on 2020 February 17–18, as well as the participants of the IISA 2021 conference organized by the University of Illinois Chicago on 2021 May 20–23. We thank the Astrostatistics group at the Center for Astrophysics of Harvard University for constructive criticism of this manuscript. Special thanks go to József Benkő for sharing the parameters we used to simulate the Blazhko star of Section 6.2, and to Gergely Hajdu for useful discussions.

Software: The Language R for Statistical Computing (R Core Team 2021), the R package NISTunits (Gama 2016), and the R package minpack.lm (Elzhov et al. 2016).

Appendix A Modulation

The Blazhko effect is a periodic amplitude and phase variation in the light curves of RR Lyrae variable stars. In astronomy, the Blazhko effect is usually interpreted as a modulation phenomenon. Modulation is the process of transmitting a low-frequency signal into a high-frequency wave, called the *carrier wave*, by changing its amplitude, frequency, and/or phase angle through the *modulating signal*. The function of the carrier wave is to carry the message or modulating signal from the transmitter to the receiver. The superposition of the signal and the carrier wave results in the so-called *modulated signal*.

In this Appendix we review two types of modulation, as given in Benkő et al. (2011): amplitude modulation and frequency modulation. This will be helpful for a comparison between our model (Equation (1)) and the models proposed by Benkő et al. (2011) and Benkő (2018), in the case of RR Lyrae stars presenting the Blazhko effect.

A.1. Amplitude Modulation

Amplitude modulation (AM) changes the amplitude of the carrier signal. Let the carrier wave $c(t)$ be a sinusoidal signal of the form

$$c(t) = U_c \sin(2\pi f_c t + \phi_c),$$

where the constant parameters U_c , f_c , and ϕ_c are the amplitude, frequency, and phase of the carrier wave, respectively.

Let $U_m(t)$ represent a waveform that is the message to be transmitted, or the *modulating* signal. The transmitter uses the information signal $U_m(t)$ to vary the amplitude of the carrier U_c to produce the *amplitude modulated* signal U_{AM} :

$$\begin{aligned} U_{AM}(t) &= [U_c + U_m(t)] \sin(2\pi f_c t + \phi_c) \\ &= [U_c + U_m(t)] \frac{c(t)}{U_c} = \left[1 + \frac{U_m}{U_c} \right] c(t). \end{aligned} \quad (\text{A1})$$

In the simplest case, when the modulating signal is sinusoidal, that is,

$$U_m(t) = U_m^A \sin(2\pi f_m t + \phi_m^A), \quad (\text{A2})$$

the amplitude modulated signal in Equation (A1) is

$$\begin{aligned} U_{AM}(t) &= [U_c + U_m^A \sin(2\pi f_m t + \phi_m^A)] \\ &\quad \times \sin(2\pi f_c t + \phi_c). \end{aligned} \quad (\text{A3})$$

Clearly, a more complex example of amplitude modulation arises when $K \geq 1$, where K denotes the number of harmonic components. Suppose the carrier wave $c(t)$ is a linear combination of sine harmonics:

$$c(t) = a_0 + \sum_{k=1}^K a_k \sin(2\pi k f_0 t + \phi_k),$$

and the modulating signal is sinusoidal and given again by Equation (A2). Following the same idea as in Equation (A1), the amplitude modulated signal in Equation (A3) is

$$\begin{aligned} U_{AM}(t) &= \left[1 + \frac{U_m(t)}{U_c} \right] c(t) \\ &= \left[1 + \frac{U_m^A \sin(2\pi f_m t + \phi_m^A)}{U_c} \right] \\ &\quad \times \left[a_0 + \sum_{k=1}^K a_k \sin(2\pi k f_0 t + \phi_k) \right]. \end{aligned} \quad (\text{A4})$$

If we call $h = U_m^A/U_c$ and use the basic trigonometrical identities $\sin(a)\sin(b) = \frac{1}{2}[\cos(a-b) - \cos(a+b)]$ and $\sin(a) = \cos\left(a - \frac{\pi}{2}\right)$, Equation (A4) can be written as

$$\begin{aligned} U_{AM}(t) &= a_0 + \sum_{k=1}^K a_k \sin(2\pi k f_0 t + \phi_k) \\ &\quad + a_0 h \sin(2\pi f_m t + \phi_m^A) \\ &\quad + \sum_{k=1}^K \frac{a_k h}{2} \sin(2\pi(kf_0 - f_m)t + (\phi_k - \phi_m) + \pi/2) \\ &\quad - \sum_{k=1}^K \frac{a_k h}{2} \sin(2\pi(kf_0 + f_m)t + (\phi_k + \phi_m) + \pi/2). \end{aligned} \quad (\text{A5})$$

This example shows that when the time-varying amplitude $U_m(t)$ in Equation (A2) takes a sinusoidal form, the amplitude modulated model with *time-varying* amplitude in Equation (A4) can be written as a model with *time-invariant* parameters as in Equation (A5). This implies that, when frequencies and

phases are known, the parameters $\{a_k, 0 \leq k \leq K\}$ in Equation (A5) can be estimated by ordinary least squares.

A.2. Amplitude and Frequency Modulation

Frequency modulation (FM) changes the frequency of the carrier signal. We assume the sinusoidal carrier wave to be

$$c(t) = U_c \sin(\Theta(t)),$$

where $\Theta(t) = 2\pi f_c t + \phi_c$ is the angular part of the function. Suppose that the modulating signal is $U_m(t)$. Then the modulated angular part is given by

$$\Theta(t) = 2\pi f_c t + 2\pi k_{FM} \int_0^t U_m^F(\tau) d\tau + \phi_c,$$

where k_{FM} is the frequency deviation, and the *frequency modulated* signal is expressed as

$$U_{FM}(t) = U_c \sin\left(2\pi f_c t + 2\pi k_{FM} \int_0^t U_m^F(\tau) d\tau + \phi_c\right). \quad (\text{A6})$$

In the simplest case, when the modulating signal is represented by a sinusoidal wave with amplitude U_m^F and frequency f_m , the integral of such a signal is

$$\int_0^t U_m(\tau) d\tau = \frac{U_m^F}{2\pi f_m} \sin(2\pi f_m t + \phi_m),$$

and the frequency modulated signal in Equation (A6) is

$$U_{FM}(t) = U_c \sin\left(2\pi f_c t + \frac{k_{FM} U_m^F}{f_m} \sin(2\pi f_m t + \phi_m) + \phi_c\right). \quad (\text{A7})$$

In practice, modulated signals can be a mixture of amplitude and frequency modulations, which can be used to describe Blazhko RR Lyrae stars (Benkő et al. 2011). We review the simplest case when both AM and FM are sinusoidal. Combining the amplitude modulated signal in Equation (A3) and the frequency modulated signal in Equation (A7), the *amplitude and frequency modulated* signal is thus

$$\begin{aligned} U_{Comb}(t) &= [U_c + U_m^A \sin(2\pi f_m t + \phi_m)] \\ &\quad \times \sin\left(2\pi f_c t + \frac{k_{FM} U_m^F}{f_m} \sin(2\pi f_m t + \phi_m) + \phi_c\right). \end{aligned}$$

A.3. Blazhko Modulation

Amplitude and frequency modulations have been observed in Blazhko RR Lyrae stars (e.g., Benkő et al. 2010; Chadid et al. 2010; Poretti et al. 2010; Sódor et al. 2012). Assuming that the observed data sets are precise and long enough, Benkő et al. (2011) proposed an amplitude and frequency modulation model for Blazhko RR Lyrae stars given by

$$m_{Comb}^*(t) = \frac{m_{AM}^*(t)}{c^*(t)} m_{FM}^*(t), \quad (\text{A8})$$

where $c^*(t)$ is the carrier wave, and the functions $m_{AM}^*(t)$ and $m_{FM}^*(t)$ are the nonsinusoidal amplitude and frequency

modulations, given respectively by

$$m_{\text{AM}}^*(t) = \left[a_0^A + \sum_{p=1}^q a_p^A \sin(2\pi p f_m t + \varphi_p^A) \right] c^*(t), \quad (\text{A9})$$

$$m_{\text{FM}}^*(t) = a_0 + \sum_{k=1}^K a_k \sin[2\pi k f_0 t + k a_0^F + k \sum_{p=1}^q a_p^F \sin(2\pi p f_m t + \varphi_p^F) + \varphi_k]. \quad (\text{A10})$$

Here, the modulating signal used in the amplitude and frequency modulation is an arbitrary periodic signal represented by a Fourier sum with a constant frequency f_m . Superscripts A and F denote the amplitude modulation and frequency modulation parameters, respectively, and f_0 and f_m are the main pulsation and modulation frequencies, respectively.

Substituting Equations (A9) and (A10) in Equation (A8), we finally have

$$m_{\text{Comb}}^*(t) = \left[a_0^A + \sum_{p=1}^q a_p^A \sin(2\pi p f_m t + \varphi_p^A) \right] \times \left\{ a_0 + \sum_{k=1}^K a_k \sin[2\pi k f_0 t + k a_0^F + k \sum_{p=1}^q a_p^F \sin(2\pi p f_m t + \varphi_p^F) + \varphi_k] \right\}.$$

Appendix B B-splines

In this Appendix we define B -splines and give some details about the estimation method that we used in this manuscript. For more details we refer the reader to the book by de Boor (1978).

A B -spline curve $f(t)$ of degree d is defined as

$$f(t) = \sum_{j=1}^J P_j B_{j,d}(t), \quad (\text{B1})$$

where P_j are the control points and $B_{j,d}(t)$ are the B -spline basis functions. Let t_{\min} and t_{\max} be, respectively, the lower and upper bounds of the domain of interest. In order to build the B -spline basis of degree d , we first divide the domain into n intervals, with n being a positive integer, obtaining the $n+1$ knots $\xi_0, \xi_1, \dots, \xi_{n+1}$. Each knot satisfies $\xi_j < \xi_{j+1}$, for all j . Second, we define $2d$ additional knots $\xi_0, \xi_1, \dots, \xi_{d-1}, \xi_{n+d+1}, \dots, \xi_{n+2d-1}, \xi_{n+2d}$. Then, the j th B -spline basis, $B_{j,d}(t)$, can be defined recursively as

$$B_{j,d}(t) = \frac{t - \xi_{j-1}}{\xi_{j+d-1} - \xi_{j-1}} B_{j,d-1}(t) + \frac{\xi_{j+d} - t}{\xi_{j+d} - \xi_j} B_{j+1,d-1}(t), \quad j = 1, \dots, J, \quad (\text{B2})$$

with

$$B_{j,0}(t) = \begin{cases} 1 & t \in [\xi_{j-1}, \xi_j), \\ 0 & \text{otherwise} \end{cases} \quad (\text{B3})$$

being used to initialize the recursion. Thus, to build the B -spline curve given by Equation (B1), we need $n+2d+1$ knots, and the total number of B -splines basis functions is $J = n+d$.

To illustrate how to construct a B -spline basis, consider the case of degree $d=2$ and assume that the domain $[t_{\min}, t_{\max}]$ has been divided into $n=3$ intervals, obtaining the knots ξ_2, \dots, ξ_5 . In this instance, the $2d=4$ additional knots are defined as ξ_0, ξ_1, ξ_6 , and ξ_7 . Using Equation (B2), we obtain

$$\begin{aligned} B_{1,2}(t) &= \frac{t - \xi_0}{\xi_2 - \xi_0} B_{1,1}(t) + \frac{\xi_3 - t}{\xi_3 - \xi_1} B_{2,1}(t), \\ B_{2,2}(t) &= \frac{t - \xi_1}{\xi_3 - \xi_1} B_{2,1}(t) + \frac{\xi_4 - t}{\xi_4 - \xi_2} B_{3,1}(t), \\ B_{3,2}(t) &= \frac{t - \xi_2}{\xi_4 - \xi_2} B_{3,1}(t) + \frac{\xi_5 - t}{\xi_5 - \xi_3} B_{4,1}(t), \\ B_{4,2}(t) &= \frac{t - \xi_3}{\xi_5 - \xi_3} B_{4,1}(t) + \frac{\xi_6 - t}{\xi_6 - \xi_4} B_{5,1}(t), \\ B_{5,2}(t) &= \frac{t - \xi_4}{\xi_6 - \xi_4} B_{5,1}(t) + \frac{\xi_7 - t}{\xi_7 - \xi_5} B_{6,1}(t), \end{aligned}$$

where

$$\begin{aligned} B_{1,1}(t) &= \frac{t - \xi_0}{\xi_1 - \xi_0} B_{1,0}(t) + \frac{\xi_2 - t}{\xi_2 - \xi_1} B_{2,0}(t), \\ B_{2,1}(t) &= \frac{t - \xi_1}{\xi_2 - \xi_1} B_{2,0}(t) + \frac{\xi_3 - t}{\xi_3 - \xi_2} B_{3,0}(t), \\ B_{3,1}(t) &= \frac{t - \xi_2}{\xi_3 - \xi_2} B_{3,0}(t) + \frac{\xi_4 - t}{\xi_4 - \xi_3} B_{4,0}(t), \\ B_{4,1}(t) &= \frac{t - \xi_3}{\xi_4 - \xi_3} B_{4,0}(t) + \frac{\xi_5 - t}{\xi_5 - \xi_4} B_{5,0}(t), \\ B_{5,1}(t) &= \frac{t - \xi_4}{\xi_5 - \xi_4} B_{5,0}(t) + \frac{\xi_6 - t}{\xi_6 - \xi_5} B_{6,0}(t), \\ B_{6,1}(t) &= \frac{t - \xi_5}{\xi_6 - \xi_5} B_{6,0}(t) + \frac{\xi_7 - t}{\xi_7 - \xi_6} B_{7,0}(t), \end{aligned}$$

and the coefficients $\{B_{j,0}(t), j=1, \dots, 7\}$ are defined in Equation (B3).

Suppose we have N observations $\{t_1, \dots, t_N\}$, that might be either equally or unequally spaced, with $t_i \in [t_{\min}, t_{\max}]$ for all $i=1, \dots, N$. The B -splines basis matrix evaluated at time $\{t_1, \dots, t_N\}$, denoted by \mathbf{B} , is the $N \times J$ matrix with entries $\{B_{j,d}(t_i), i=1, \dots, N, j=1, \dots, J\}$, in a way that each row contains a B -spline basis. The j th B -spline basis function satisfies the following properties:

$$\begin{cases} B_{j,d}(t) > 0 & \xi_{j-1} < t < \xi_{j+d}, \\ B_{j,d}(t) = 0 & \xi_0 \leq t \leq \xi_{j-1} \text{ and } \xi_{j+d} \leq t \leq \xi_{n+2d}, \\ \sum_{j=1}^J B_{j,d}(t) = 1 & \xi_d < t < \xi_{n+d}, \\ \left. \frac{\partial^k B_{j,d}(t)}{\partial t^k} \right|_{t=\xi_\ell} & 0 \leq k \leq d-1 \text{ are continuous.} \end{cases}$$

For ease of notation we use, throughout our manuscript, $B_j(t)$ instead of $B_{j,d}(t)$. Let us now consider the example of estimating the mean function $\mu(t)$ of model $Y_i = \mu(t_i) + \varepsilon_i$ using B -splines. Let $\mathbf{Y} = (Y_1, \dots, Y_6)'$ be the available $N=6$ responses observed, respectively, at time $\{t_1, \dots, t_6\}$, with $t_{\min} = t_1$ and $t_{\max} = t_6$. Then assume that $\mu(t) = \sum_{j=1}^J P_j B_j(t)$, for all $t \in [t_1, t_6]$. We use here B -splines basis functions of degree $d=2$; in order to construct them, we

divide the domain $[t_1, t_6]$ into $n = 3$ intervals. Hence, the total number of knots ξ_0, \dots, ξ_7 is $n + 2d + 1 = 8$, and the total number of B -splines basis functions is $J = n + d = 5$. The 6×5 design matrix \mathbf{B} has entries $B_{ij} = B_j(t_i)$, with $i = 1, \dots, 6$ and $j = 1, \dots, 5$, which permits estimating the coefficients $\{P_j, j = 1, \dots, 5\}$ by ordinary least squares. Indeed, if $\mathbf{Y} = (Y_1, \dots, Y_6)^\top$ denotes the response vector and $\boldsymbol{\theta} = (P_1, \dots, P_5)^\top$ the parameter vector, we can rewrite the model as $\mathbf{Y} = \mathbf{B}\boldsymbol{\theta} + \mathbf{z}$, where $\mathbf{z} = (z_1, \dots, z_6)^\top$ is the error vector. The estimated parameters are defined as $\hat{\boldsymbol{\theta}} = (\hat{P}_1, \dots, \hat{P}_5)^\top = (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{Y}$, and the estimated mean as $\hat{\mu}(t) = \sum_{j=1}^5 \hat{P}_j B_j(t)$, for all $t \in [t_1, t_6]$.

Appendix C Confidence Intervals

C.1. Nonparametric Quantiles

We use the quantiles 0.025 and 0.975 to construct the confidence intervals in our simulations of Section 6.1. For t fixed, confidence intervals for $\mu(t)$, $m(t)$, and $g_{\ell,k}(t)$, $\ell = 1, 2$, $k = 1, \dots, K$, are calculated according to the following three steps:

1. We estimate the coefficients of interest α , β_k , and γ_k , $k = 1, \dots, K$, following the procedure described in Section 4.1, and we define the $J \times M$ matrices $\mathbf{A}_M = [\hat{\alpha}^{(1)}, \dots, \hat{\alpha}^{(M)}]$, $\mathbf{B}_{kM} = [\hat{\beta}_k^{(1)}, \dots, \hat{\beta}_k^{(M)}]$, and $\mathbf{G}_{kM} = [\hat{\gamma}_k^{(1)}, \dots, \hat{\gamma}_k^{(M)}]$, where $\hat{\alpha}^{(j)}$, $\hat{\beta}_k^{(j)}$, and $\hat{\gamma}_k^{(j)}$ correspond to the estimators of α , β_k , and γ_k , given by Equation (11) in the j th Monte Carlo simulation, $j = 1, \dots, M$.
2. We define the $M \times 1$ vectors $\hat{\mathbf{m}}^{(M)}(t)$, $\hat{\mathbf{g}}_{\ell,k}^{(M)}(t)$, $\ell = 1, 2$, $k = 1, \dots, K$, and $\hat{\mu}^{(M)}(t)$ as

$$\begin{aligned} \hat{\mathbf{m}}^{(M)}(t) &= \mathbf{B}(t)^\top \mathbf{A}_M \\ &= [\hat{m}^{(1)}(t), \dots, \hat{m}^{(M)}(t)], \end{aligned}$$

$$\begin{aligned} \hat{\mathbf{g}}_{1,k}^{(M)}(t) &= \mathbf{B}(t)^\top \mathbf{B}_{kM} \\ &= [\hat{g}_{1,k}^{(1)}(t), \dots, \hat{g}_{1,k}^{(M)}(t)], \quad k = 1, \dots, K \end{aligned}$$

$$\hat{\mathbf{g}}_{2,k}^{(M)}(t) = \mathbf{B}(t)^\top \mathbf{G}_{kM} = [\hat{g}_{2,k}^{(1)}(t), \dots, \hat{g}_{2,k}^{(M)}(t)], \quad k = 1, \dots, K$$

$$\begin{aligned} \hat{\mu}^{(M)}(t) &= \hat{\mathbf{m}}(t) + \sum_{k=1}^K \{\hat{\mathbf{g}}_{1,k}(t) \cos(w_k t) \\ &\quad + \hat{\mathbf{g}}_{2,k}(t) \sin(w_k t)\} = [\hat{\mu}^{(1)}(t), \dots, \hat{\mu}^{(M)}(t)], \end{aligned}$$

where $\hat{\mu}^{(j)}(t)$, $\hat{m}^{(j)}(t)$, and $\hat{g}_{\ell,k}^{(j)}(t/t_N)$, $\ell = 1, 2$,

$k = 1, \dots, K$ correspond to the estimators of $\mu(t)$, $m(t)$, and $g_{\ell,k}(t/t_N)$ given by Equations (12) and (13) in the j th Monte Carlo simulation, and $\mathbf{B}(t)$ corresponds to the vector formed by the B -splines evaluated at time t .

3. We calculate the empirical quantiles of order 0.025 and 0.975 of the $M \times 1$ vectors $\hat{\mathbf{m}}^{(M)}(t)$, $\hat{\mathbf{g}}_{\ell,k}^{(M)}(t)$, $\ell = 1, 2$, $k = 1, \dots, K$, and $\hat{\mu}^{(M)}(t)$.

C.2. Parametric Quantiles

We use the parametric quantiles to construct the confidence intervals for our simulation in Section 6.2 and our application in Section 7. Assuming that the error terms $\{z_i, i = 1, \dots, N\}$ follow a Gaussian distribution with zero mean and variance σ_z^2 , the $(1 - \alpha) \times 100\%$ prediction interval for $E[\hat{Y}_i] = \mathcal{B}^\top(t_i)E[\hat{\boldsymbol{\theta}}]$, with $i = 1, \dots, N$, is

$$\mathcal{B}^\top(t_i) \hat{\boldsymbol{\theta}} \pm z(1 - \alpha/2) \sqrt{\mathcal{B}^\top(t_i) \text{Var}[\hat{\boldsymbol{\theta}}] \mathcal{B}(t_i)},$$

where $z(1 - \alpha/2)$ denotes the $(1 - \alpha/2)$ quantile of the standard Gaussian distribution, and

$$\text{Var}[\hat{\boldsymbol{\theta}}] = \sigma_z^2 (\mathcal{B}^\top \mathcal{B} + \mathbf{P})^{-1} \mathcal{B}^\top \mathcal{B} (\mathcal{B}^\top \mathcal{B} + \mathbf{P})^{-1}.$$

The $(1 - \alpha) \times 100\%$ confidence interval for the trend $E[\hat{m}(t_i)] = \mathcal{B}(t_i)^\top \mathbf{Q}_m E[\hat{\boldsymbol{\theta}}]$ is

$$\begin{aligned} &\mathcal{B}(t_i)^\top \mathbf{Q}_m \hat{\boldsymbol{\theta}} \pm z(1 - \alpha/2) \\ &\times \sqrt{\mathcal{B}(t_i)^\top \mathbf{Q}_m \text{Var}[\hat{\boldsymbol{\theta}}] \mathbf{Q}_m^\top \mathcal{B}(t_i)}, \end{aligned}$$

and the $(1 - \alpha) \times 100\%$ confidence intervals for the amplitudes $E[\hat{g}_{\ell,k}(t_i)] = \mathcal{B}(t_i)^\top \mathbf{Q}_g(\ell, k) E[\hat{\boldsymbol{\theta}}]$, $\ell = 1, 2$, $k = 1, \dots, K$, are

$$\begin{aligned} &\mathcal{B}(t_i)^\top \mathbf{Q}_g(\ell, k) \hat{\boldsymbol{\theta}} \pm z(1 - \alpha/2) \\ &\times \sqrt{\mathcal{B}(t_i)^\top \mathbf{Q}_g(\ell, k) \text{Var}[\hat{\boldsymbol{\theta}}] \mathbf{Q}_g(\ell, k)^\top \mathcal{B}(t_i)}, \end{aligned}$$

where $\mathcal{B}(t_i)$ is the i th row of the matrix \mathcal{B} , and $\mathcal{B} = [\mathbf{B}|\mathbf{B}|\dots|\mathbf{B}]$ is a matrix of dimension $N \times c$. The $c \times c$ matrices $\mathbf{Q}_m, \{\mathbf{Q}_g(\ell, k), \ell = 1, 2, k = 1, \dots, K\}$ satisfy $\mathbf{Q}_m \boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top)^\top$ and

$$\begin{aligned} \mathbf{Q}_g(1, 1) \boldsymbol{\theta} &= (\mathbf{0}_J^\top, \beta_1^\top, \mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top)^\top, & \mathbf{Q}_g(2, 1) \boldsymbol{\theta} &= (\mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top, \gamma_1^\top, \mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top)^\top, \\ \mathbf{Q}_g(1, 2) \boldsymbol{\theta} &= (\mathbf{0}_J^\top, \mathbf{0}_J^\top, \beta_2^\top, \mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top)^\top, & \mathbf{Q}_g(2, 2) \boldsymbol{\theta} &= (\mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top, \mathbf{0}_J^\top, \gamma_2^\top, \mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top)^\top, \\ & & & \vdots \\ \mathbf{Q}_g(1, K) \boldsymbol{\theta} &= (\mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top, \beta_K^\top, \mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top)^\top, & \mathbf{Q}_g(2, K) \boldsymbol{\theta} &= (\mathbf{0}_J^\top, \dots, \mathbf{0}_J^\top, \gamma_K^\top)^\top. \end{aligned}$$

Appendix D Proofs

ORCID iDs

In this Appendix we prove the results in Lemma 1 and Proposition 1 (see Section 5).

Márcio Catelan  <https://orcid.org/0000-0001-6003-8877>

D.1. Proof of Lemma 1

Notice that the expectation of the periodogram in Equation (15) of the observations $\{\varepsilon_i, i \in \mathcal{I}\}$ is

$$E[I_\varepsilon(\lambda)] = \sum_{k \in \mathcal{I}} \sum_{j \in \mathcal{I}} E[\varepsilon_k \varepsilon_j] \exp(i\lambda[t_k - t_j]).$$

If we replace the expectation $E[\varepsilon_k \varepsilon_j]$ with the right-hand side of Equation (18), we obtain

$$\begin{aligned} E[I_\varepsilon(\lambda)] &= \frac{2\pi}{N_{\mathcal{I}}} \sum_{j=1}^{N_{\mathcal{I}}} P_\varepsilon(\omega_j) \sum_{k \in \mathcal{I}} \sum_{j \in \mathcal{I}} \exp(i[\lambda - \omega_j] \\ &\times [t_k - t_j]) = \frac{2\pi}{N_{\mathcal{I}}} \sum_{j=1}^{N_{\mathcal{I}}} P_\varepsilon(\omega_j) W_\varepsilon(\lambda - \omega_j) \\ &= \frac{2\pi}{N_{\mathcal{I}}} P_\varepsilon(\lambda) * W_\varepsilon(\lambda). \end{aligned} \quad (\text{D1})$$

D.2. Proof of Proposition 1

Let $\mathcal{F}\{g_j\}[k]$ denote the discrete Fourier transform (DFT) of the sequence of m numbers g_1, \dots, g_m into another sequence h_1, \dots, h_m , that is,

$$\begin{aligned} h_k &= \mathcal{F}\{g_j\}[k] = \sum_{j=1}^m g_j \exp(-ikj2\pi/m), \\ k &= 1, \dots, m, \end{aligned}$$

and $\mathcal{F}^{-1}\{h_k\}[j]$ denote the inverse DFT of the sequence h_1, \dots, h_m into another sequence g_1, \dots, g_m , that is,

$$\begin{aligned} g_j &= \mathcal{F}^{-1}\{h_k\}[j] = \frac{1}{m} \sum_{k=1}^m h_k \exp(ikj2\pi/m), \\ j &= 1, \dots, m. \end{aligned}$$

Let $\mathcal{F}\{g_j\}[k]$ and $\mathcal{F}\{\ell_j\}[k]$ be, respectively, the DFTs of the sequences $\{g_j\}$ and $\{\ell_j\}$ into the sequences $\{h_k\}$ and $\{m_k\}$. Then, the convolution theorem states that

$$\mathcal{F}\{g_j * \ell_j\}[k] = \mathcal{F}\{g_j\}[k] \mathcal{F}\{\ell_j\}[k]. \quad (\text{D2})$$

Applying the convolution theorem in Equation (D2) to Equation (D1), we obtain

$$\mathcal{F}\{P_\varepsilon(\lambda_j)\}[k] = \frac{N_{\mathcal{I}}}{2\pi} \frac{\mathcal{F}\{E[I_\varepsilon(\lambda_j)]\}[k]}{\mathcal{F}\{W_\varepsilon(\lambda_j)\}[k]}.$$

The inverse DFT of the last equation gives the result in Equation (23).

References

- Benkő, J. M. 2018, *MNRAS*, **473**, 412
- Benkő, J. M., Kolenberg, K., Szabó, R., et al. 2010, *MNRAS*, **409**, 1585
- Benkő, J. M., Plachy, E., Szabó, R., Molnár, L., & Kolláth, Z. 2014, *ApJS*, **213**, 31
- Benkő, J. M., Szabó, R., & Paparó, M. 2011, *MNRAS*, **417**, 974
- Blažko, S. 1907, *AN*, **175**, 325
- Brockwell, P., & Davis, R. 2016, *Introduction to Time Series and Forecasting* (2nd ed.; Heidelberg: Springer),
- Buchler, J. R., & Kolláth, Z. 2011, *ApJ*, **731**, 24
- Catelan, M., & Smith, H. A. 2015, *Pulsating Stars* (New York et al: Wiley)
- Chadid, M., Benkő, J. M., Szabó, R., et al. 2010, *A&A*, **510**, A39
- Chattopadhyay, A. K. 2017, *Incomplete Data in Astrostatistics* (New York: Wiley)
- Dahlhaus, R. 1996, *Stochastic Processes and their Applications*, **62**, 139
- Dahlhaus, R. 1997, *The Annals of Statistics*, **25**, 1
- de Boor, C. 1978, *A Practical Guide to Splines* (Berlin: Springer)
- Deeming, T. J. 1975, *Ap&SS*, **36**, 137
- Eilers, P. H. C., Gampe, J., Marx, B. D., & Rau, R. 2008, *Statistics in Medicine*, **27**, 3430
- Eilers, P. H. C., & Marx, B. D. 1996, *StaSc*, **11**, 89
- Elzhov, T. V., Mullen, K. M., Spiess, A.-N., & Bolker, B. 2016, *minpack.lm: R Interface to the Levenberg–Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds*, v.1.2-1, <https://CRAN.R-project.org/package=minpack.lm>
- Feigelson, E. D., Babu, G. J., & Caceres, G. A. 2018, *FrP*, **6**, 80
- Gama, J. 2016, *NISTunits: Fundamental Physical Constants and Unit Conversions from NIST*, v.1.0.1., <https://CRAN.R-project.org/package=NISTunits>
- Gillet, D. 2013, *A&A*, **554**, A46
- Gillet, D., Mauclaire, B., Lemoult, T., et al. 2019, *A&A*, **623**, A109
- Sódor, Á., Hajdu, G., Jurcsik, J., et al. 2012, *MNRAS*, **427**, 1517
- Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. 2004, *Math. Intell.*, **27**, 83
- Kelly, B. C., Becker, A. C., Sobolewska, M., Siemiginowska, A., & Uttley, P. 2014, *ApJ*, **788**, 33
- Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, *ApJL*, **713**, L79
- Lomb, N. R. 1976, *Ap&SS*, **39**, 447
- Netzel, H., Smolec, R., Soszyński, I., & Udalski, A. 2018, *MNRAS*, **480**, 1229
- Plachy, E., Molnár, L., Bódi, A., et al. 2019, *ApJS*, **244**, 32
- Poretti, E., Paparó, M., Deleuil, M., et al. 2010, *A&A*, **520**, A108
- Priestley, M. B. 1981, *Spectral Analysis and Time Series* (New York: Academic)
- R Core Team 2021, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, v.4.1.2, <https://www.R-project.org>
- Richards, J. W., Starr, D. L., Butler, N. R., et al. 2011, *ApJ*, **733**, 10
- Smith, H. A. 1995, *RR Lyrae Stars*, 27 (Cambridge: Cambridge Univ. Press)
- Stothers, R. B. 2006, *ApJ*, **652**, 643
- Wong, R. K. W., Kashyap, V. L., Lee, T. C. M., & van Dyk, D. A. 2015, *arXiv:1508.07083*
- Xu, C., Günther, H. M., Kashyap, V. L., Lee, T. C. M., & Zezas, A. 2021, *AJ*, **161**, 184
- Zhou, S., Shen, X., & Wolfe, D. A. 1998, *AnSta*, **26**, 1760