Bayesian Model Search for Nonstationary Periodic Time Series

Beniamino Hadj-Amar, Bärbel Finkenstädt, Mark Fiecas, Francis Lévi and Robert Huckstepp *

October 24, 2018

Abstract

We propose a novel Bayesian methodology for analyzing nonstationary time series that exhibit oscillatory behaviour. We approximate the time series using a piecewise oscillatory model with unknown periodicities, where our goal is to estimate the change-points while simultaneously identifying the potentially changing periodicities in the data. Our proposed methodology is based on a trans-dimensional Markov chain Monte Carlo (MCMC) algorithm that simultaneously updates the change-points and the periodicities relevant to any segment between them. We show that the proposed methodology successfully identifies time changing oscillatory behaviour in two applications which are relevant to e-Health and sleep research, namely the occurrence of ultradian oscillations in human skin temperature during the time of night rest, and the detection of instances of sleep apnea in plethysmographic respiratory traces.

Keywords: Bayesian Spectral Analysis; Reversible-Jump MCMC; Ultradian Sleep Cycles; Sleep Apnea;

*Beniamino Hadj-Amar (E-mail: B.Hadj-Amar@warwick.ac.uk), Bärbel Finkenstädt Rand (E-mail: B.F.Finkenstadt@warwick.ac.uk), Department of Statistics, University of Warwick, Coventry CV4 7AL, UK. Francis Lévi (E-mail: F.Levi@warwick.ac.uk ), Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK. Robert Huckstepp (E-mail: R.Huckstepp@warwick.ac.uk), School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK. Mark Fiecas (E-mail: mfiecas@umn.edu), University of Minnesota, School of Public Health, Division of Biostatistics, Minneapolis, MN 55455, USA.
1 Introduction

Identifying the periodicities present in a cyclical phenomenon allows us to gain insight into the sources of variability that drive the phenomenon. For example, respiratory traces obtained from a plethysmograph used on rodents in experimental sleep apnea research exhibit many abrupt changes in their periodic components as the rat naturally changes their breathing pattern in the course of its sleep-wake activities (Han et al. 2002, Nakamura et al. 2003). Similarly, human temperature, as measured by a wearable sensing device over several days at relatively high temporal resolution, may be subject to a different periodic behaviour during the night when the individual transitions between ultradian sleep stages (Carskadon et al. 2005, Komarzynski et al. 2018). While the theory and methods for analyzing the periodicities of time series data whenever the time series is stationary is relatively well-developed, the task of modelling time series that show regime shifts in periodicity, amplitude and phase remains challenging, as the timing of changes and the relevant periodicities are usually unknown.

There have been many developments for modeling stationary oscillatory time series data. Rife & Boorstyn (1976) and Stoica et al. (1989) addressed the problem of estimating the frequencies, phases and amplitudes of sinusoidal signals under the assumption of a known number of sinusoids, where inference is based on maximum likelihood frequency estimators. These models, however, require very long time series and a large separation in the frequencies that drive the process, which will not always be the case in practice (Djuric 1996, Andrieu & Doucet 1999). Quinn (1989), Yau & Bresler (1993) and Zhang & Wong (1993) tackled the problem of model selection on the number of sinusoidal signals by employing the Akaike information criterion (Akaike 1974) and the minimum description length principle (Rissanen 1978). Djuric (1996) showed that these procedures tend to estimate a wrong number of components when the sample size is small and the signal-to-noise ratio is low.
Bayesian approaches to modelling stationary oscillatory signals were explored for the first time by Bretthorst (1988, 1990) with applications to nuclear magnetic resonance spectroscopy. Dou & Hodgson (1995, 1996) presented a Bayesian approach that uses a Gibbs sampler to identify multiple frequencies that drive the signal. Their method required the number of frequencies to be fixed in advance, and model selection was achieved by choosing the most probable model based on the estimation of the parameters for all possible models. Bayesian model selection for stationary oscillatory signals based on posterior model probabilities were also investigated by Djuric (1996). Andrieu & Doucet (1999) introduced a more efficient reversible-jump MCMC method (Green 1995) that jointly tackles model selection and parameter estimation for an unknown number of stationary sinusoidal signals and avoids the computationally expensive numerical optimization of Dou & Hodgson (1995, 1996) by sampling the frequencies one-at-time via Metropolis-Hastings (M-H) steps. To the best of our knowledge, currently there is no extension of this methodology to analyze nonstationary oscillatory signals.

A formal statistical modeling framework for a specific class of nonstationary time series data, called locally stationary time series, was developed by Dahlhaus et al. (1997). Extending this framework to a Bayesian setting, Rosen et al. (2009) proposed an approach to model the log of the time-varying spectral density using a mixture of smoothing splines. Rosen et al. (2012) improved on this by splitting the time series into an unknown but finite number of segments of variable lengths, thereby avoiding the need to pre-select partitions, and to estimate the time-varying spectral density using a fixed number of smoothing splines. For a given partition of the time series, the likelihood function is approximated via a product of local Whittle likelihoods (Whittle 1957). The methodology was developed using a Bayesian framework and is based on the assumption that, conditional on the position and number of partitions, the time series are piecewise stationary, and the underlying spectral density for each partition is smooth over frequencies. However, exploratory analyses of the time series in both of our case studies revealed
spectral densities with very sharp peaks, often at several nearby frequencies, thus invalidating the assumption that the spectral density is smooth over frequencies. In addition, the frequency location of these sharp peaks changed over time.

In this article we propose a novel Bayesian methodology for modelling oscillatory data that show regime shifts in periodicity, amplitude and phase. In contrast to previous work our approach does not require pre-specifying the number of regimes or the order of the model within a regime. We assume that, conditional on the position and number of change-points, the time series can be approximated by a piecewise changing sinusoidal regression model. The timing and number of changes are unknown, along with the number and values of relevant periodicities in each segment. We develop a reversible jump MCMC technique that jointly explores the parameter space of the change-points and sub-models for all segments.

The paper is organized as follows. Section 2 and 3 present the model, the prior specifications and the general structure of our Bayesian approach. Sections 4 and 5 provide a detailed explanation of our sampling scheme and a simulation study to demonstrate the performance of our approach. In Section 6 we illustrate the use of our methodology in two data-rich scenarios related to sleep, circadian rhythm and e-Health research, namely the identification of the spectral properties of experimental breathing traces arising in sleep apnea research, and the analysis of human temperature data measured over several days by a wearable sensor. We conclude and discuss our current findings in Section 7.

2 The Model

Consider a time series realization $y_1, \ldots, y_n$ whose periodic behaviour may change at $k$ unknown time-points $s_{(k)} = (s_1, \ldots, s_k)$ where $k$ is also unknown. Assume that in each sub-interval $I_j = [s_{j-1}, s_j)$ there are $m_j$ relevant frequencies $\omega_j = (\omega_{j,1}, \ldots, \omega_{j,m_j})$, for $j = 1, 2, \ldots, k + 1$. Setting
$s_0 = 1$ and $s_{k+1} = n$, we can write the following sinusoidal model (Andrieu & Doucet 1999)

$$y_t = \sum_{j=1}^{k+1} f(t, \beta_j, \omega_j) \mathbb{1}_{[t \in I_j]} + \epsilon_t,$$  \hspace{1cm} (1)

where

$$f(t, \beta_j, \omega_j) = \alpha_j + \mu_j t + \sum_{l=1}^{m_j} \left( \beta^{(1)}_{j,l} \cos(2\pi \omega_j \cdot t) + \beta^{(2)}_{j,l} \sin(2\pi \omega_j \cdot t) \right),$$ \hspace{1cm} (2)

$\beta_j = (\alpha_j, \mu_j, \beta^{(1)}_{j,1}, \ldots, \beta^{(1)}_{j,m_j}), \beta_{j,l} = (\beta^{(1)}_{j,l}, \beta^{(2)}_{j,l})$, $\mathbb{1}_{[t]}$ denotes the indicator function, and $\mu_j$ and $\alpha_j$ may, if needed, account for a linear trend within each segment. For simplicity we assume independent zero-mean Gaussian errors with regime-specific variances

$$\epsilon_t \sim N(0, \sigma^2_j), \quad \text{for } t \in I_j \text{ and } j = 1, \ldots, k+1,$$

noting that, in principle, the methodology can in principle be adjusted to the non-Gaussian case.

The dimension of the model is given by the number of change-points $k$ and the number of frequency components in each regime denoted by $m(k) = (m_1, \ldots, m_{k+1})'$. Furthermore, let $\beta(k) = (\beta^1, \ldots, \beta^k)'$, $\omega(k) = (\omega^1, \ldots, \omega^k)'$, $\sigma^2(k) = (\sigma^2_1, \ldots, \sigma^2_k)'$ and $\theta(k) = (\beta^k, \omega^k, \sigma^2(k))'$.

Using Equation (1), the likelihood of $(k, m(k), s(k), \theta(k))$ given the data $y = (y_1, \ldots, y_n)'$ is

$$\mathcal{L}(k, m(k), s(k), \theta(k), y) = \prod_{j=1}^{k+1} \mathcal{L}(m_j, \theta_j | y_j), \quad y_j = (y_i : t \in I_j),$$ \hspace{1cm} (3)

where

$$\mathcal{L}(m_j, \theta_j | y_j) = (2\pi \sigma^2_j)^{-n/2} \exp \left[ -\frac{1}{2\sigma_j^2} \sum_{i \in I_j} \left( y_i - x_i(\omega_j)' \beta_j \right)^2 \right].$$ \hspace{1cm} (4)
The vector of basis functions \( x_t(\omega_j) \) is defined as

\[
x_t(\omega_j) = (1, t, \cos(2\pi \omega_j t), \sin(2\pi \omega_j t), \ldots, \cos(2\pi \omega_{jm} t), \sin(2\pi \omega_{jm} t))',
\]

where \( \theta_j = (\beta'_j, \omega'_j, \sigma'^2_j)' \) is the vector of parameters and \( n_j \) the number of observations of the \( j^{\text{th}} \) segment.

3 Bayesian Inference

Given some pre-fixed maximal numbers of change-points, \( k_{\text{max}} \), and frequencies per regime, \( m_{\text{max}} \), inference is achieved by assuming that the true model is unknown but comes from a finite class of models where each model \( M_k \), with \( k \) change-points, is parameterized by the vector

\[
( m_{(k)}, s_{(k)}, \theta_{(k)} ) \in \Pi_k, \quad \Pi_k \in \Pi.
\]

Let \( S_k = \{ s_{(k)} \in [1, n]^k : 1 < s_{1} < \cdots < s_k < n \} \) and \( \Omega_{m_j} = (0, 0.5)^{m_j} \) denote, respectively, the sample space for the locations of change-points and the frequencies of the \( j^{\text{th}} \) segment. The overall parameter space can be written as a finite union of subspaces

\[
\Pi = \bigcup_{k=0}^{k_{\text{max}}} \{ k \} \times \Pi_k, \quad \text{and} \quad \Pi_k = S_k \times \bigcap_{j=1}^{k+1} \{ m_j \} \times \bigcup_{m_j=1}^{m_{\text{max}}} \{ \mathbb{R}^{2m_j+2} \times \Omega_{m_j} \times \mathbb{R}^2 \}.
\]

Bayesian inference on \( k, m_{(k)}, s_{(k)} \) and \( \theta_{(k)} \) is based on the following factorization of the joint posterior distribution

\[
\pi(k, m_{(k)}, s_{(k)}, \theta_{(k)} | y) = \pi(k | y) \pi(m_{(k)} | k, y) \pi(s_{(k)} | m_{(k)}, k, y) \pi(\theta_{(k)} | s_{(k)}, k, m_{(k)}, y),
\]

6
where we use $\pi(\cdot)$ as generic notation for probability density or mass function, whichever appropriate. Sampling from it poses a multiple model selection problem, namely of the number of change-points and number of frequencies in each regime, which can be addressed by constructing a reversible-jump MCMC algorithm Green (1995). The algorithm in its basic structure iterates between the following two moves:

1. **Segment model move:** Given a partition of the data at $k$ locations $s_{(k)}$, inference on the parameters $m_{(k)}$ and $\theta_{(k)}$ is based on the conditional posterior

   $$
   \pi(m_{(k)}, \theta_{(k)} | k, s_{(k)}, y) = \prod_{j=1}^{k+1} \pi(m_j, \theta_j | k, s_{(k)}, y).
   $$

   A reversible-jump MCMC algorithm is performed in parallel on each of the $k + 1$ segments, where at each iteration the number of sinusoids $m_j$, the linear coefficients $\beta_j$, the frequencies $\omega_j$ and the residual variances $\sigma_j^2$ are sampled independently in each segment, for $j = 1, \ldots, k+1$. Notice that at this stage the algorithm will explore subspaces of variable dimensionality regarding the number of frequencies per segment, while the change-point model remains fixed.

2. **Change-point move:** This step performs a reversible-jump MCMC algorithm for change-point model search where the number $k$ and locations of change-points $s_{(k)}$ are sampled, along with the linear coefficients, number of frequencies and their values as well as the residual variances for any segments affected by the move.

   Our prior specifications assume independent Poisson distributions for the number of breakpoints $k$ and frequencies in each segment $m_j$, conditioned on $k \leq k_{\text{max}}$ and $1 \leq m_j \leq m_{\text{max}}$, respectively. Given $k$, a prior distribution for the positions of the change-points $s_{(k)}$ can be chosen.
as in Green (1995)

\[
\pi(s_k | k) = \frac{(2k + 1)!}{(n-1)^{2k+1}} \prod_{j=0}^{k} (s_{j+1} - s_j)(s_{s_k+1} - <s_k < n), \quad s_0 = 1, s_{k+1} = n. \tag{5}
\]

Conditional on \(k\) and \(m_k\), we choose a uniform prior for the frequencies \(\omega_j, l = 1, \ldots, m_j, \quad j = 1, \ldots, k + 1\). Analogous to a Bayesian regression (Bishop 2006), a zero-mean isotropic Gaussian prior is assumed for the coefficients of the \(j^{th}\) segment, \(\beta_j \sim \mathcal{N}(0, \sigma^2_\beta I), \quad j = 1, \ldots, k + 1\), where \(\sigma^2_\beta\) is a pre-specified large value, and the prior on the residual variance \(\sigma^2_j\) of the \(j^{th}\) partition is \(\text{Inverse-Gamma}(\frac{\nu_0}{2}, \frac{\gamma_0}{2})\), where \(\eta_0\) and \(\nu_0\) are fixed at small values.

4 Sampling Scheme for Nonstationary Periodic Processes

Here we provide the sampling scheme associated with the nonstationary periodic processes that we wish to model. An outline of the overall procedure is as follows. Start with an initial configuration of number of change-points \(k\), along with their locations \(s_k\); this yields a partition of the data \(y = (y_1, \ldots, y_{k+1})\). Initialize the number of frequencies in each regime \(m_k\) and their values \(\omega_k\), along with the coefficients \(\beta_k\) and residual variances \(\sigma^2_k\). At each iteration of the algorithm a segment model and a change-point model move are estimated. A random choice with probabilities (6) based on the current number of parameters will determine whether to attempt a birth, death or a within-model move. In particular, let \(z\) denote the current number of parameters, i.e. change-points \(k\) in the change-point model or frequencies \(m_j\) in the \(j^{th}\) segment model; then, the dimension may increase by one (birth step) with probability \(b_z\), decrease by one (death step) with probability \(d_z\) or remain unchanged (within step) with probability \(\mu_z = 1 - b_z - d_z\), where

\[
b_z = c \min \left\{ 1, \frac{\pi(z + 1)}{\pi(z)} \right\}, \quad d_{z+1} = c \min \left\{ 1, \frac{\pi(z)}{\pi(z + 1)} \right\}, \quad \mu_z = 1 - b_z - d_z, \tag{6}
\]
for some constant $c \in \left[0, \frac{1}{2}\right]$, and $\pi(z)$ is the prior probability of the model including $z$. Reversibility of the Markov chain is guaranteed for move types that involve a change in dimensionality as $b_z \pi(z) = d_{z+1} \pi(z + 1)$. Here we chose $c = 0.4$ but other values are legitimate as long as $c$ is not larger than 0.5, to assure that the sum of the probabilities does not exceed 1 for some values of $z$. Naturally, $b_{k_{\text{max}}} = b_{m_{\text{max}}} = 0$ and $d_{k_0} = d_{m_1} = 1$. The pseudocode of the overall algorithm that describes an iteration of the sampler is given in Algorithm 1. We next describe in more detail the specific procedures needed to update the moves.

**Algorithm 1**

1. For each segment $j = 1, \ldots, k + 1$, perform a segment model move (Section 4.1)
   
   Draw $U \sim \text{Uniform}(0, 1)$
   
   if $U \leq b_{m_j} \rightarrow \text{birth-step}$
   else if $b_{m_j} \leq U \leq d_{m_j} \rightarrow \text{death-step}$
   else $\rightarrow \text{within-step}$

2. Perform a change-point model move (Section 4.2):
   
   Draw $U \sim \text{Uniform}(0, 1)$
   
   if $U \leq b_k \rightarrow \text{birth-step}$
   else if $b_k \leq U \leq d_k \rightarrow \text{death-step}$
   else $\rightarrow \text{within-step}$

4.1 Updating a Segment Model

Given the number of change-points $k$ and their locations $s_{(k)}$, a segment model move is performed independently and in parallel on each of the $k + 1$ partitions. Hence, throughout this subsection the subscript relating to the $j^{th}$ segment may be dropped and a segment of interest is denoted by
\[ y = (y_a, \ldots, y_b)', \] which contains \( n \) observations. Assume that the current number of frequencies is set at \( m \); then, an independent random choice is made between attempting a birth, death or within-model step, with probabilities given in (6). An outline of these moves is as follows (further details are provided in the Appendix).

**Within-Model Move:** Conditioned on the number of frequencies \( m \), we sample the vector of frequencies \( \omega \) following Andrieu & Doucet (1999), i.e. by sampling the frequencies one-at-time using a mixture of M-H steps, with target distribution

\[
\pi(\omega | \beta, \sigma^2, m, y) \propto \exp \left[ -\frac{1}{2\sigma^2} \sum_{t=a}^{b} \left( y_t - x_t(\omega)' \beta \right)^2 \right] \mathbb{1}_{\omega \in \Omega_m},
\]

(7)

In particular, the proposal distribution is a combination of a Normal random walk centred around the current frequency and a sample from values of the Discrete Fourier transform of \( y \). The corresponding vector of linear parameters \( \beta \) is then updated in a M-H step, from the target posterior conditional distribution

\[
\pi(\beta | \omega, \sigma^2, m, y) \propto \exp \left[ -\frac{1}{2\sigma^2} \sum_{t=a}^{b} \left( y_t - x_t(\omega)' \beta \right)^2 \right] - \frac{1}{2\sigma^2} \beta' \beta,
\]

(8)

where the proposed values are drawn from normal approximations to their posterior conditional distribution. Finally, the residual variance \( \sigma^2 \) is then updated in a Gibbs step from

\[
\sigma^2_{\omega, \beta} \sim \text{Inverse-Gamma} \left( \frac{n + \nu_0}{2}, \frac{\gamma_0 + \sum_{t=a}^{b} \left( y_t - x_t(\omega)' \beta \right)^2}{2} \right),
\]

(9)

**Between-Model Moves:** For this type of move, the number of frequencies is either proposed to increase by one (birth) or decrease by one (death). If a birth move is proposed, we have that \( m^p = m^c + 1 \), where current and proposed values are denoted by the superscripts \( c \) and
$p$, respectively. The proposed vector of frequencies is constructed by proposing an additional frequency to include in the current vector. Conditional on the frequencies, the corresponding vector of linear coefficients and the residual variance are sampled as in the within-model move. If a death move is proposed, we have that $m^p = m^c - 1$. Hence, one of the current frequencies is randomly chosen to be removed. The proposed corresponding vector of linear coefficients is drawn, along with the residual variance. For both moves, the updates are jointly accepted or rejected in a M-H step.

4.2 Updating the Change-Point Model

This part of the algorithm identifies the number and locations of change-points. Suppose the number of change-points is currently set to some value $k$, then according to the probabilities given in (6) a random decision is made between adding, removing or moving a change-point. The rules for updating these types of moves are described below and more details are given in the Appendix.

**Within-Model Move:** An existing change-point is proposed to be relocated with probability $\frac{1}{k}$, obtaining say $s^c_j$. The update for the selected change-point is proposed from a mixture of a Normal random walk centred on the current change-point $s^c_j$ and a sample from a uniform distribution on the interval $[s^c_{j-1} + \psi_s, s^c_{j+1} - \psi_s]$. Here, we introduced $\psi_s$ as a fixed minimum time between change-points avoiding change-points being too close to each other. Rosen et al. (2012) used a similar scheme, but on a discrete-scale. The number of frequencies and their values are kept fixed, and, conditional on the relocation, the linear coefficients for the segments affected by the relocation are sampled. These updates are jointly accepted or rejected in a M-H step and the residual variances are updated in a Gibbs step.
**Between-Model Moves:** For this type of move, the number of change-points may either increase (birth) or decrease (death) by one. If a birth move is proposed, we have that $k^p = k^c + 1$. The new proposed change-point is drawn uniformly on $f(s_{(k^c)}, \psi_s)$, the support of $s_{(k^c)}$ given the constraints imposed by $\psi_s$, i.e. $f(s_{(k^c)}, \psi_s) = [1 + \psi_s, s_1^c - \psi_s] \cup [s_1^c + \psi_s, s_2^c - \psi_s] \cup \ldots \cup [s_k^c + \psi_s, n - \psi_s]$. The latter involves splitting an existing segment. The number of frequencies and their values in the proposed segments are selected from the current states. Two residual variances for the new proposed segments are then constructed from the current single residual variance. Finally, two new vectors of linear parameters are sampled. If a death move is proposed, we have that $k^p = k^c - 1$. Hence, a candidate change-point to be removed is selected from the vector of existing change-points, with probability $\frac{1}{k^c}$. The latter involves merging two existing partitions. The number of frequencies and their values in the proposed segments are selected from the current states. A single residual variance is constructed from the current variances relative to the segments affected by the relocation. Finally, a new vector of linear coefficient is drawn. For both type of moves, these updates are jointly accepted or rejected.

## 5 Simulation Study

In this section, we carry out a simulation study to explore the performance of our method, which will be referred to as AutoNOM (Automatic Nonstationary Oscillatory Modelling). We simulate a time series consisting of $n = 900$ data points from model (1) with $k = 2$ change-points located at positions $s_{(2)} = (300, 650)$, and fixed number of frequencies per regime $m_{(2)} = (3, 1, 2)$ (See Supplementary Material, Table 5 for further details of the parametrization). Figure 1 (top panel) shows a realization from this model. The prior means $\lambda_\omega$ and $\lambda_s$, say, on the number of frequencies and change-points, respectively, were set to 2, to reflect a fair degree of prior information on their numbers. We will discuss later our finding that AutoNOM was relatively
insensitive to these prior specifications. The maximum number of change-points $k_{\text{max}}$ is set to 15, and the maximum number of frequencies per regime $m_{\text{max}}$ is set to 10. Furthermore, we fixed $\psi_s = 20$ and $\phi_\omega = 0.25$ (Appendix A.1.2) for the uniform distribution for sampling the frequencies. The full estimation algorithm is run for 20,000 updates, 5,000 of which are discarded as burn-in period. The estimation took 390 seconds with a (serial) program written in Julia 0.52 on a Intel® Core™ i7-4790S Processor 16 GB RAM. The results, summarized in Table 1 clearly show that a model with two change-points has the highest estimated posterior probability (left panel) and that AutoNOM correctly identifies the right number of significant frequencies in each regime (right panel).

Figure 1 (middle panel) shows the estimated posterior distribution for the location of the change-points, conditioned on three segments. The posterior means of the change-point locations are $\hat{E}(s_1 | k = 2, y) = 298.7$ and $\hat{E}(s_2 | k = 2, y) = 650.1$. Figure 1 (bottom panel) shows that the estimated posterior distributions are an excellent match to the true frequencies.

Table 1: Simulation study. (left panel) posterior probabilities for number of change-points; (right panel) posterior probabilities for number of frequencies in each regime, conditioned on $k = 2$.

| $k$ | $\hat{\pi}(k | y)$ | $m$ | $\hat{\pi}(m_1 | k = 2, y)$ | $\hat{\pi}(m_2 | k = 2, y)$ | $\hat{\pi}(m_3 | k = 2, y)$ |
|-----|------------------|-----|----------------|----------------|----------------|
| 0   | .00              | 1   | .00            | .99            | .00            |
| 1   | .02              | 2   | .00            | .01            | .98            |
| 2   | .97              | 3   | .98            | .00            | .02            |
| 3   | .01              | 4   | .02            | .00            | .00            |
| 4   | .00              | 5   | .00            | .00            | .00            |

5.1 Sensitivity Analysis

To investigate the influence of the prior means $\lambda_\omega$ and $\lambda_s$ we simulate 10 realizations from the same model and run our estimation algorithm for combinations of values for $\lambda_\omega$ and $\lambda_s$, ranging
Figure 1: Simulation study. (Top) Simulated time series. (Middle) Estimated posterior distribution for the location of the change-points, conditioned on $k = 2$. The dotted vertical lines represent true location of change-points. (Bottom) Estimated posterior distribution of the frequencies for each different segment, conditioned on $k = 2$, $m_1 = 3$, $m_2 = 1$ and $m_3 = 2$. The dotted vertical lines represent true values of the frequencies.
from 0.1 to 10.0. Table 2 shows the average posterior probability of choosing the correct model, i.e. $\hat{\pi}(k = 2, m_1 = 3, m_2 = 1, m_3 = 2 | y)$. We also investigated the average mean squared error between the true underlying signal and the estimated signal (see Table 6, Supplementary Material). Both analyses suggest that, for this example, the choice of the prior means $\lambda_\omega$ and $\lambda_s$ has hardly noticeable impact on the results. However, our experience is that small values for these hyper-parameters are preferable as they prevent the algorithm from overfitting.

Table 2: Sensitivity analysis simulation study. Average probability of choosing the correct model from 10 replications with varying $\lambda_\omega$ and $\lambda_s$.

| $\lambda_\omega$ | $\lambda_s = .1$ | $\lambda_s = .2$ | $\lambda_s = 0.5$ | $\lambda_s = 1.0$ | $\lambda_s = 2.0$ | $\lambda_s = 5.0$ | $\lambda_s = 10.0$ |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| .1               | .98              | .97              | .97              | .99              | .93              | .91              | 1.0              |
| .2               | .99              | .99              | .99              | .99              | .99              | .99              | .98              |
| .5               | .99              | 1.0              | .98              | .95              | .98              | 1.0              | .99              |
| 1.0              | .99              | .94              | .93              | .99              | .99              | .99              | .99              |
| 2.0              | .99              | .99              | .99              | .99              | .99              | .99              | .99              |
| 5.0              | .97              | .96              | .96              | .97              | .97              | .93              | .98              |
| 10.0             | .94              | .95              | .95              | .91              | .93              | .95              | .85              |

### 5.2 Detecting Spectral Peaks

We simulate a time series from the same simulation model as above with the only difference that the residual variances were set equal to one for all segments and thus are smaller than above. The performance of AutoNOM is compared with two existing methods, namely the Bayesian adaptive spectral estimation for nonstationary time series proposed by Rosen et al. (2012), referred to as AdaptSPEC, and the frequentist piecewise vector autoregressive method of Davis et al. (2006), referred to as AutoPARM. Specifically, we explore the performances of these methodologies in identifying the number and location of change-points, and the number and location of frequency peaks in each estimated segment. AdaptSPEC requires the user to specify in advance the number
of basis function $J$ used for smoothing the periodogram in the segments. We run AdaptSPEC for two different specifications, namely $J = 7$ and $J = 15$ basis functions. The model is fitted with a total of 15,000 iterations, 5,000 of which are discarded as burn-in, by using the R package provided by the authors. Posterior samples of peak frequencies are obtained by considering the modes of the spectrum per MCMC iteration. AutoPARM is performed with default tuning parameters.

The modal number of change-points for AdaptSPEC is 2 for both $J = 7$ and $J = 15$, with posterior probability $\hat{\pi}(k = 2 \mid y)$ of 76% and 88%, respectively; the modal number of change-points for AutoNOM is 2 and AutoPARM identifies 2 change-points as well. Conditioned on the modal number of change-points, Table 3 displays the estimated location of changes (left panel) and frequency peaks (right panel) for the different compared methods, where we report the standard deviation for the estimate obtained from the empirical distribution of the posterior samples. Similarly, we show in Figure 2 the estimated location of the frequency peaks and their 95% credible intervals, for each of the three identified segments; dotted vertical lines represent the true location of the frequency peaks. Results for AutoNOM are conditioned on the modal number of frequencies per regime.

| Segment 1 | Segment 2 | Segment 3 |
|-----------|-----------|-----------|
| AN        | AN        | AN        |
| AS J7     | AS J7     | AS J7     |
| AS J15    | AS J15    | AS J15    |
| AP        | AP        | AP        |

Figure 2: Simulation study with unitary residual variances. Estimated frequency peaks for AutoNOM (AN), AdaptSPEC (AS $J = 7, 15$) and AutoPARM (AP); 95% credible intervals (horizontal lines) are also reported for Bayesian methods. Dotted vertical lines are true locations of the frequency peaks.
Table 3: Simulation study with unitary residual variances. Estimated change-points locations (left panel) and frequency peaks (right panel) for AutoNOM (AN), AdaptSPEC (AS \(J = 7, 15\)) and AutoPARM (AP); standard errors are also reported for Bayesian methods.

|       | \(s_1\) | \(s_2\) | \(\omega_{1,1}\) | \(\omega_{1,2}\) | \(\omega_{1,3}\) | \(\omega_{2,1}\) | \(\omega_{3,1}\) | \(\omega_{3,2}\) |
|-------|-------|-------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| True  | 300   | 650   | .042            | .067            | .143            | .083            | .046            | .067            |
| AN    | 300.87| 650.46| .042 (0.97)     | .067 (0.31)     | .142            | .083 (0.0004)   | .045            | .067 (0.0003)   |
| AS J7 | 319.65| 628.51| .083 (15.36)    | -.              | -.              | .082            | .057            | -.              |
|       |       |       |                 | .0030           |                 |                 |                 |                 |
| AS J15| 298.5 | 647.01| .057 (1.21)     | .141 (0.14)     | .088            | .056            | -.              | -.              |
|       |       |       |                 | .0030 (0.0030)  |                 | .0050           | .0030           |                 |
| AP    | 299   | 648   | .060            | .140            | -.              | .080            | .055            | -.              |

It becomes clear that the detection of periodicities by AdaptSPEC is affected by the specification of the number of spline basis functions used for the smoothing, where increasing the number of basis function yields a better performance for AdaptSPEC. The example also shows that smoothing by splines may lead to peaks in the periodogram to be over-smoothed and neighbouring close peaks to be merged. AutoPARM seems to also suffer from the latter problem.

When we increase the residual variance to the high levels set originally, AdaptSPEC fails to detect any change-points for both \(J = 7\) and \(J = 15\), with posterior probability \(\hat{\pi}(k = 0|y)\) of 69% and 93%, respectively, while AutoPARM finds 7 change-points and thus severely over-estimates their number. Our conclusion from this comparison is that although AdaptSPEC and AutoPARM may be well suited for time series processes with smooth time-varying spectra with few or no peaks, both methods are severely challenged in detecting changes in spectra that exhibit pronounced peakedness, possibly at nearby frequencies, as can be expected to occur in reality for the type of time series that we wish to analyze.
6 Case Studies

The development of our methodology was motivated by the following two case studies where dense physiological signals were observed which exhibit unknown periodicities whose role may change over time in a more or less abrupt manner and where their detection is of relevance to the health and well-being of the subject.

6.1 Analysis of Human Skin Temperature

The development of information and communication technologies, in particular widespread internet access and availability of mobile phones and tablets, allows considering new developments in the health care system. To address the issue of personalized medical treatment according to the circadian timing system of the patient, referred to as chronotherapy (Levi & Schibler 2007), a novel and validated non-invasive mobile e-Health platform pioneered by the French project PiCADO (Komarzynski et al. 2018) is used to record and teletransmit skin surface temperature as well as physical activity data (Huang et al. 2018) from an upper chest e-sensor. Figure 3 (a) shows an example of 4 days of 5-minutes temperature recording for a healthy individual. The circadian rhythms in core and skin surface temperature are usually 8-12 hours out of phase, with respective maxima occurring near 16:00 at day time, and near 2:00 at night (Krauchi & Wirz-Justice 1994). The early night drop in core body temperature, which is critical for triggering the onset of sleep (Van Someren 2006), results from the vasodilatation of the skin vessels and associated rise in skin surface temperature (Kräuchi 2002). Under the assumption of stationarity Komarzynski et al. (2018) analyzed the skin temperature time series identifying both strong 12 hours (circahemidian) and 24 hours (circadian) rhythms. Here, we applied our methodology to the skin-temperature time series shown in Figure 3 (a) for 300,000 iterations, discarding the first 100,000 updates as burn-in. The maximum number of change-points $k_{\text{max}}$ was set to 10, whereas
the maximum number of frequencies per regime $m_{\text{max}}$ was set to 5. The estimated number of change-points had a mode at 7, with $\hat{\pi}(k = 7 | y) = 0.97$ and their estimated posterior distributions are shown in Figure 3 (c). Inspecting them alongside the physical activity data we can see that the change points mainly correspond to the start and end points of the prolonged rest periods at nights showing that skin temperature alternates between day activity and night rest including sleep. Figure 4 shows the estimated posterior distribution of the frequencies for the sleep segments (2, 4, 6, 8) along with the square root of the estimated power of the corresponding frequencies, where the power of each is frequency $\omega_{j,l}$ is summarized by the sum of squares of the corresponding linear coefficients, i.e. $I(\omega_{j,l}) = \beta_{j,l}^{(1)^2} + \beta_{j,l}^{(2)^2}$ (Shumway & Stoffer 2005). Figure 3 (b) shows the piecewise fitted signal, along with a 95% credible interval obtained from the 2.5 and 97.5 empirical percentiles of the posterior sample using Equation (17). Cycles of approximately 3 hours appear in segments 2, 4 and 6; cycles that range approximately 1-1.5 hour appear in segments 2, 4, 8 and cycles of around 2 hours appear in segments 4 and 6 while some longer periods identified in segments 4, 6, 8 indicate the presence of a trend.

Stages of sleep are characterized by ultradian oscillations between rapid eye movements (REM) and non-rapid eye movements (non-REM). The biological functionality that regulates the alternations between these two types of sleep is not yet much understood (Altevogt et al. 2006). However, several physiological changes that occur over night differ between REM and non-REM phases, such as heart rate, brain activity, muscle tone and body temperature (Berlad et al. 1993, Pace-Schott & Hobson 2002). The body cycles between REM and non-REM sleep stages with an average length that ranges approximately between 70 to 120 minutes, and there are usually four to six of these sleep cycles each night (Carskadon et al. 2005, Shneerson 2009). Our analysis was able to use skin temperature data alone to detect periods of sleep throughout the day and identify oscillatory behaviour during the night, whose frequencies are compatible with ultradian oscillations between REM and different non-REM sleep stages.
Figure 3: Analysis of skin temperature of a healthy subject. Panel (a) are the time series of skin temperature corresponding physical activity. Panel (b) is the estimated signal (solid line) along with its 95% credible interval; vertical lines are the estimated locations of the change-points. Panel(c) is the estimated posterior density histogram of the locations of the changes, conditioned on $k = 7$ change-points. Rectangles on the time axis of each plot correspond to periods from 20.00 to 8.00. The variation in skin temperature finds analogies with the rest-activity pattern that alternates between day activity and night rest.
Figure 4: Spectral properties for segments corresponding to night rest. Estimated posterior distribution of the frequencies along with square root of the estimated power of the corresponding frequencies. The results are conditional on the modal number of frequencies per segment.
6.2 Characterizing Instances of Sleep Apnea in Rodents

Sleep apnea is the temporary (≥ 2 breaths) interruption of breathing during sleep. Moderate or severe (≥ 15 events per hour) sleep apnea, occurs in about 50% of men and 25% of women over the age of 40 (Heinzer et al. 2015), with 91% of people with sleep apnea being undiagnosed (Tan et al. 2016). Sleep apnea is linked to many diseases. Patients with sleep apnea are at increased risk of: cardiovascular events, e.g., stroke and heart attack (Lanfranchi et al. 1999); cancer (Nieto et al. 2012); liver disease, e.g., non-alcoholic fatty liver disease (Sundaram et al. 2016); diabetes (Harsch et al. 2004); metabolic syndrome (Parish et al. 2007); and cognitive decline, e.g., early onset mild cognitive impairment (MCI) and Alzheimer disease (AD) (Osorio et al. 2016), and increased risk of dementia in the elderly (Lal et al. 2012). The motivation of this research is to provide a statistical methodology that can be applied to analyze large breathing data sets resulting from in vivo plethysmograph studies in rats to characterize the occurrence of sleep apnea under different experimental conditions. If this could be attained, a concrete aid to the understanding of the pathological implications of this status could be provided to clinicians and experimental biologists.

An unrestrained whole-body plethysmograph is used to produce a breathing trace from freely behaving rats for periods of up to 3 hours. Plethysmographs were made using an 2L air-tight box connected to a pressure transducer, with an air pump and outlet valve producing a flow rate of 2L/min. Airflow pressure signals were amplified using Neurolog system (Digitimer) connected to a 1401 interface and acquired on a computer using Spike2 software (CED).

Apneas are subclassified as post-sigh apneas, if the preceding breath was at least 25% above the average amplitude of prior breaths, or spontaneous apneas, if there was no manifestation of a previous sigh (Davis & ODonnell 2013). Airflow traces from the plethysmograph are shown in Figure 5 (left panels) and consist of three time series, which will be referred to as (a), (b) and (c). They correspond to different actions for this rat: (a) is characterised by an alternation of
sniffing and normal breathing; (b) presents the trace of a spontaneous apnea, followed by normal breathing; (c) shows normal breathing followed by a sigh, and a post-sigh apnea. Each time series contains 20,000 observations where the signal was sampled at 2000 Hz, namely at a rate of 2000 observations per second.

![Graphs of respiratory traces](image_url)

Figure 5: Plots of the respiratory traces of a rat (left panels) and corresponding estimated posterior power (right panels). Panel (a) is characterised by an alternation of sniffing and normal breathing. Panel (b) is a plot of the trace of a spontaneous apnea, followed by normal breathing. Panel (c) shows normal breathing followed by a sigh, and a post-sigh apnea. Dotted vertical lines correspond to the estimated locations of the change-points.

Our procedure allows us to set an upper bound, \( \phi_\omega \) (Appendix A.1.2) for the uniform interval where the new frequencies are sampled. As the periodogram ordinates for these data were approximately zero for all frequencies larger than 0.01, we decided accordingly to set \( \phi_\omega = 0.01 \).
The locations of the changes (vertical lines) are displayed in Figure 5 (left panels). The posterior power of the frequencies, for each time series, is shown in Figure 5 (right panels). These results are conditional on the modal number of change-points and the modal number of frequencies per segment.

For each data set, we summarise in Table 4 the spectral properties of each partition by displaying the periodicities corresponding to the first two largest values of the estimated power. When the rat is sniffing, (a), the air flow trace oscillates with a dominant period of approximately 0.2 seconds, namely 5 cycles per second. Normal breathing, (a) and (b), is characterised by lower frequencies and lower magnitude than sniffing, by oscillating with a dominant period of around 0.5 seconds, namely around 2 cycles every second. Apneas, (b) and (c), appear to be characterised by higher frequencies than normal breathing but with a lower power, with dominant periods of around 0.25 and 0.35 seconds. Notice that in the first partition of (c), the highest value of the power corresponds to the frequency responsible for a sigh before apnea.

Table 4: Spectral properties of respiratory traces of a rat. Periodicities (in seconds) corresponding to the first two largest values of the estimated power, for each time series (a), (b) and (c).

|        | (a) |        | (b) |        | (c) |
|--------|-----|--------|-----|--------|-----|
|        | Period | Power  | Period | Power  | Period | Power  |
| Segment 1 | .19 | .2272 | .20 | .0015 | .25 | 5.2540 |
|         | .15 | .0883 | .25 | .0010 | .21 | .0915  |
| Segment 2 | .20 | .2613 | .49 | .0117 | .27 | .0107  |
|         | .28 | .0686 | .34 | .0099 | .42 | .0043  |
| Segment 3 | .48 | .0692 | .48 | .0145 | .24 | .0365  |
|         | .32 | .0491 | .28 | .0122 | .37 | .0095  |
| Segment 4 | .58 | .0400 | .78 | .0122 | .36 | .0253  |
|         | .47 | .0251 | -  | -     | .24 | .0155  |
| Segment 5 | .19 | .3231 | -  | -     | -   | -      |
7 Conclusions and Outlook

We have developed a novel Bayesian methodology for analyzing nonstationary time series that exhibit oscillatory behaviour. Our approach is based on the assumption that, conditional on the position and number of change-points, the time series are approximately piecewise stationary. The timing and number of changes are unknown, along with the number and values of relevant periodicities in each regime. Bayesian inference is performed via a reversible jump MCMC technique that jointly explores the change-points and sub-models for all segments.

We illustrated the utility of our methodology in two case studies. First, we analyzed human skin temperature time series data obtained from a wearable device, which exhibited unknown periodicities that changed over time in an abrupt manner. Our proposed methodology identified interesting oscillations whose frequencies are consistent with ultradian oscillations between REM and non-REM sleep stages. Second, we characterized the occurrence of sleep apnea in large breathing data sets resulting from in vivo plethysmograph studies on rodents. Our spectral investigation was able to distinguish very sharp peaks, corresponding to different nearby frequencies, that are responsible for the different actions of the rodent.

Our approach proposed here is limited to a single-subject analysis and does not take into account the within-subject dependence across multiple measurements or the variability between subjects. Future works could involve extending the methodology to be able to analyze larger studies involving multiple subjects, taking account of the within-subject dependence and the inclusion of covariates, such as sex or age.
Acknowledgements

We wish to thank Dr. Paul Jenkins and Jack Jewson for their insightful and valuable comments. B. Hadj-Amar was supported by the Oxford-Warwick Statistics Programme (OxWaSP) and the Engineering and Physical Sciences Research Council (EPSRC) under Grant Number EP/L016710/1. B. Finkenstädt and F. Lévi were supported by the Medical Research Council (MRC), Grant reference: MR/M013170/1. F. Lévi was partly supported by the Conseil Régional d’Île de France, the Conseil Régional de Champagne-Ardenne, Mairie de Paris and the Banque Publique d’Investissement (BPI France) through the Fonds Unique Interministériel 12 (PicADo, Contract 11017951), and the Institut de Recherches en Santé Publique from France (CLOCK-DOM1, grant 2014-BDCR-EC). R. Huckstepp was supported by the MRC, Grant Reference: MC/PC/15070.

A Appendix - Details of the Sampling Scheme

A.1 Updating the Segment Model

A.1.1 Within-Model Move

**Sampling**: To obtain samples from the conditional posterior distribution $\pi(\omega|\beta, \sigma^2, m, y)$ (see Equation (7)), we draw the frequencies one-at-time using a mixture of M-H steps. In order to explore the parameter space efficiently, we design a mixture distribution $q(\omega_{lp}|\omega_{lc})$, so that

$$ q(\omega_{lp}|\omega_{lc}) = \delta_\omega q_1(\omega_{lp}|\omega_{lc}) + (1 - \delta_\omega) q_2(\omega_{lp}|\omega_{lc}), \quad l = 1, \ldots, m, \quad (10) $$

where $q_1$ is defined in Equation (11) below, $q_2$ is the density of a univariate Normal $\mathcal{N}(\omega_{lc}, \sigma^2_\omega)$, $\delta_\omega$ is a positive real number such that $0 \leq \delta_\omega \leq 1$, and $c$ and $p$ refer to current and proposed
values, respectively. According to Equation (10) we carry out with probability $\delta_\omega$ a M-H step with proposal distribution $q_1 (\omega^p_l | \omega^c_l)$,

$$q_1 (\omega^p_l | \omega^c_l) \propto \sum_{h=0}^{\tilde{n}-1} I_h \begin{bmatrix} h/n \leq \omega^p_l < (h+1)/n \end{bmatrix},$$

(11)

where $\tilde{n} = \lfloor n/2 \rfloor$ and $I_h$ is the value of the squared modulus of the Discrete Fourier transform of the observations $y$ evaluated at frequency $h/n$

$$I_h = \left| \sum_{j=a}^{b} y_j \exp \left( -i 2\pi \frac{h}{n} \right) \right|^2.$$

In this way frequencies are proposed from regions in parameter space with high posterior density, yielding a Markov chain which converges quickly to its invariant distribution. The proposal distribution $q_1 (\omega^p_l | \omega^c_l)$ is independent of the current state $\omega^c_l$. The acceptance probability for this move is

$$\alpha = \min \left\{ 1, \frac{\pi (\omega^p_l | \theta, \sigma^2, m, y)}{\pi (\omega^c_l | \theta, \sigma^2, m, y)} \times \frac{q_1 (\omega^c_l)}{q_1 (\omega^p_l)} \right\},$$

where $\omega^p = (\omega^c_1, \ldots, \omega^c_{l-1}, \omega^p_l, \omega^c_{l+1}, \ldots, \omega^c_m)'$. On the other hand, with probability $1 - \delta_\omega$, we perform a random walk M-H step with proposal distribution $q_2 (\omega^p_l | \omega^c_l)$, whose density is a univariate Normal distribution with mean $\omega^c_l$ and variance $\sigma^2_\omega$, i.e. $\omega^p_l | \omega^c_l \sim N(\omega^c_l, \sigma^2_\omega)$. This perturbation around the current value $\omega^c_l$ allows a local exploration of the conditional posterior distribution. The acceptance probability for this move is

$$\alpha = \min \left\{ 1, \frac{\pi (\omega^p_l | \theta, \sigma^2, m, y)}{\pi (\omega^c_l | \theta, \sigma^2, m, y)} \right\}.$$

Setting $\delta_\omega$ to a relative low value integrates a fairly high acceptance rate with a quick exploration.
of the parameter space. For our experiments, we followed Andrieu & Doucet (1999) and set \( \sigma^2 = (1/5n)^2 \) and \( \delta = 0.2 \).

**Sampling \( \beta \):** Given values of \( \omega \) and \( \sigma^2 \), the vector of linear parameters \( \beta \) can be sampled via a M-H step from the target posterior conditional distribution \( \pi(\beta | \omega, \sigma^2, m, y) \) (see Equation (8)). The proposed vector of coefficients \( \beta^p \) can be drawn from normal approximations to their posterior conditional distribution (Gilks et al. 1995, Rosen et al. 2012), e.g

\[
\beta^p \sim \mathcal{N}_{2m+2}(\hat{\beta}^p, \hat{\Sigma}^p),
\]

(12)

where

\[
\hat{\beta}^p = \arg \max_{\beta^p} \pi(\beta^p | \omega, \sigma^2, m, y),
\]

and

\[
\hat{\Sigma}^p = \left\{-\frac{\partial^2 \log \pi(\beta^p | \omega, \sigma^2, m, y)}{\partial \beta^p \partial \beta^p} \bigg|_{\beta^p = \hat{\beta}^p}\right\}^{-1}.
\]

The proposal for \( \beta^p \) is independent of the current values \( \beta^c \), and the acceptance probability for this move is

\[
\alpha = \min \left\{1, \frac{\pi(\beta^p | \omega, \sigma^2, m, y)}{\pi(\beta^c | \omega, \sigma^2, m, y)} \times \frac{q(\beta^c)}{q(\beta^p)} \right\},
\]

where \( q(\beta^c) \) and \( q(\beta^p) \) denote the Normal proposal densities \( \mathcal{N}_{2m+2}(\hat{\beta}^c, \hat{\Sigma}^c) \) and \( \mathcal{N}_{2m+2}(\hat{\beta}^p, \hat{\Sigma}^p) \), respectively.

**A.1.2 Between-Model Moves**

The number of frequencies on a segment is proposed to either increase or decrease by one. Let \( \theta^c = (\beta^c, \omega^c, \sigma^{2c}) \) and assume the Markov chain is currently at \( (m^c, \theta^c) \). We propose a move to \( (m^p, \theta^p) \) by drawing from a proposal density \( q(m^p, \theta^p | m^c, \theta^c) \) and accepting this
update with probability

\[
\alpha = \min \left\{ 1, \frac{\mathcal{L}(m^p, \theta^p \mid y)}{\mathcal{L}(m^c, \theta^c \mid y)} \times \frac{\pi(m^p) \pi(\theta^p \mid m^p)}{\pi(m^c) \pi(\theta^c \mid m^c)} \times \frac{q(m^c, \theta^c \mid m^p, \theta^p)}{q(m^p, \theta^p \mid m^c, \theta^c)} \right\}. \tag{13}
\]

The proposed state \((m^p, \theta^p)\) is drawn by first drawing \(m^p\), followed by \(\omega^p, \beta^p\) and \(\sigma^{2p}\). In fact, we can rewrite the proposal density as

\[
q(m^p, \theta^p \mid m^c, \theta^c) = q(m^p \mid m^c) \times q(\theta^p \mid m^p, m^c, \theta^c)
\]

\[
\times q(\beta^p \mid \omega^p, m^p, m^c, \theta^c)
\]

\[
\times q(\sigma^{2p} \mid \beta^p, \omega^p, m^p, m^c, \theta^c).
\]

**Birth move:** If a birth move is proposed, we have that \(m^p = m^c + 1\). The proposed frequency vector \(\omega^p\) is constructed as

\[
\omega^p = (\omega^c_1, \ldots, \omega^c_{m^c}, \omega^p_{m^p})',
\]

namely by keeping the current vector of frequencies and proposing an additional frequency \(\omega^p_{m^p}\).

Alternatively to Andrieu & Doucet (1999), we choose to sample \(\omega^p_{m^p}\) uniformly on the interval \((0, \phi_\omega)\), where \(0 < \phi_\omega < 0.5\). The value of \(\phi_\omega\) can be chosen to reflect prior information about the significant frequencies that drive the variation in the data, for example by choosing \(\phi_\omega\) in the low frequencies range \((0 < \phi_\omega < 0.1)\). Additionally, for computational and/or modelling reasons, we would like not to sample frequencies that are too close to each other. Hence, we choose to draw a candidate value \(\omega^p_{m^p}\) uniformly from the union of intervals of the form \([\omega^c_l + \psi_\omega, \omega^c_{l+1} - \psi_\omega]\), for \(l = 0, \ldots, m_c\) and denoting \(\omega^c_0 = 0\) and \(\omega^c_{m_c+1} = \phi_\omega\). Here, \(\psi_\omega\) is a fixed minimum distance between frequencies, which is chosen larger than \(\frac{1}{n}\); in fact, when the separation of two frequencies is less than the Nyquist step (Priestley 1981), i.e. \(|\omega_l - \omega_{l+1}| < \frac{1}{n}\), the two frequencies are
indistinguishable (Dou & Hodgson 1995). Moreover, we sort the proposed vector of frequencies \( \omega^p \) to ensure identifiability and perform practical estimation, as suggested in Andrieu & Doucet (1999). For proposed \( \omega^p \) and given \( \sigma^{2c} \), the proposed vector of linear coefficients \( \beta^p \) is sampled following the same procedure of Section A.1.1, namely we draw \( \beta^p \) from a normal approximation to their posterior conditional distribution \( \pi(\beta^p | \omega^p, \sigma^{2c}, m^p, y) \). Finally, the residual variance \( \sigma^{2p} \) is sampled directly from its posterior conditional distribution \( \pi(\sigma^{2p} | \omega^p, \beta^p, m^p, y) \) (see Equation (9)). The proposed state \((m^p, \theta^p)\) is accepted or reject in a Metropolis-Hastings step with probability

\[
\alpha = \min \left\{ 1, \frac{L(\theta^p, m^p | y)}{L(\theta^c, m^c | y)} \times \frac{\pi(m^p) \pi(\theta^p | m^p)}{\pi(m^c) \pi(\theta^c | m^c)} \times \frac{d_{m^p} \cdot (\frac{1}{m^p}) \cdot q(\beta^c) \cdot q(\sigma^{2c})}{b_{m^c} \cdot q(\omega^p_{m^p}) \cdot q(\beta^p) \cdot q(\sigma^{2p})} \right\},
\]

where the likelihood function is given in Equation (4), \( \pi(m) \) is the density of the Poisson distribution truncated at \( m_{\text{max}} \), \( b_{m^c} \) and \( d_{m^p} \) are defined in Equation (6), \( q(\omega^p_{m^p}) \) is the density of the uniform proposal for sampling the additional frequency, \( q(\beta^c) \) and \( q(\beta^p) \) are the multivariate Normal proposal densities \( N_{2m^c+2}(\hat{\beta}^c, \hat{\Sigma}^c) \) and \( N_{2m^p+2}(\hat{\beta}^p, \hat{\Sigma}^p) \), respectively; \( q(\sigma^2_p) \) and \( q(\sigma^2_c) \) are the Inverse-Gamma proposal densities defined in Equation (9).

**Death move:** If a death move is proposed, then \( m^p = m^c - 1 \). A vector of frequencies \( \omega^p \) is constructed by randomly selecting with probability \( \frac{1}{m^c} \) one of the current frequencies as the candidate frequency for removal. Given \( \omega^p \) and \( \sigma^{2c} \), a vector of linear coefficients \( \beta^p \) is sampled from a normal approximation to its posterior conditional distribution. Finally, conditioned on \( \omega^p \) and \( \beta^p \), the residual variance \( \sigma^{2p} \) is drawn its posterior Inverse-Gamma distribution. It is straightforward to see that the acceptance probability for the death step has the same form as the birth step, with the proper change of labelling of the variables, and the ratio terms inverted.
A.2 Updating the Change-Point Model

A.2.1 Within-Model Move

Let \( s^c_k = (s^c_1, \ldots, s^c_k)' \) be the current vector of change-points locations, \( m^c_k = (m^c_1, \ldots, m^c_{k+1})' \) be the current vector of number of frequencies, \( \omega^c_j = (\omega^c_1, \ldots, \omega^c_{k+1})' \) be the current vector of frequencies. Let \( \beta^c_k = (\beta^c_1, \ldots, \beta^c_{k+1})' \) and \( \sigma^2_k = (\sigma^2_1, \ldots, \sigma^2_{k+1})' \) be the current vectors of linear coefficients and residual variances, respectively.

Let us also define \( \theta^c_k = (\beta^c_k, \omega^c_k, \sigma^2_k)' \). Following Green (1995), a change-point, \( s^c_j \) say, is randomly selected with probability \( \frac{1}{k} \) from the existing set of change-points. In order to explore the parameter space in an efficient way and similar to above we construct a mixture distribution \( q(s^p_j | s^c_j) \), as

\[
q(s^p_j | s^c_j) = \delta_s q_1(s^p_j | s^c_j) + (1 - \delta_s) q_2(s^p_j | s^c_j)
\]

where \( q_1 \) is the density of a Uniform \( [s^c_{j-1} + \psi, s^c_{j+1} - \psi] \) and \( q_2 \) is the density of a univariate Normal \( N(s^c_j, \sigma^2_j) \) and \( \delta_s \) is a positive real number such that \( 0 \leq \delta_s \leq 1 \). We propose with probability \( \delta_s \) a candidate value \( s^p_j \) from the above uniform distribution where \( \psi \) is a fixed minimum time between change-points avoiding change-points being too close to each other. On the other hand, with probability \( (1 - \delta_s) \), \( s^p_j \) arises as a Normal random walk proposal centered at the current change-point \( s^c_j \). The proposed vector of change-points locations is denoted by

\[
s^p_k = (s^p_1, \ldots, s^p_{j-1}, s^p_j, s^p_{j+1}, \ldots, s^p_k)',
\]

and hence the proposed value \( s^p_j \) induces a new proposed data partition on \( [s^c_{j-1}, s^c_{j+1}] \) corresponding to \( [s^c_{j-1}, s^p_j) \) and \( (s^p_j, s^c_{j+1}) \). We denote the vectors of observations belonging to these two proposed segments as \( y^p_j \) and \( y^p_{j+1} \), which include \( n^p_j \) and \( n^p_{j+1} \) observations, respectively. Given \( s^p_k \), the proposed number of frequencies \( m^p_j, m^p_{j+1} \) are set equal to the current ones \( m^c_j \).
m_{j+1}^c$, so that $m_{j+1}^p = m_{j+1}^c$. Similarly, the proposed pair of frequency vectors $\omega_j^p, \omega_{j+1}^p$ is chosen equal to the current pair $\omega_j^c, \omega_{j+1}^c$ in the corresponding segments, i.e. $\omega_{(k)}^p = \omega_{(k)}^c$. The proposed vectors $\beta_j^p, \beta_{j+1}^p$ are sampled from normal approximations to their posterior conditional distributions $\pi(\beta_j^p | \omega_j^p, \sigma_{j+1}^c, m_j^p, y_j^p)$ and $\pi(\beta_{j+1}^p | \omega_{j+1}^p, \sigma_{j+1}^c, m_{j+1}^p, y_{j+1}^p)$ and are accepted in a M-H step with probability

$$\alpha = \min\left\{ 1, \frac{L(k, m_{(k)}^p, s_{(k)}^p, \theta_{(k)}^p | y)}{L(k, m_{(k)}^c, s_{(k)}^c, \theta_{(k)}^c | y)} \times \frac{\pi(s_{(k)}^p | k) \pi(\theta_{(k)}^p | m_{(k)}^p, k)}{\pi(s_{(k)}^c | k) \pi(\theta_{(k)}^c | m_{(k)}^c, k)} \times \prod_{h=j}^{j+1} q(\beta_{h}^p) \right\},$$

where the likelihood is specified in Equation (3) and $q(\beta_{h}^c)$ and $q(\beta_{h}^p)$ are multivariate Gaussian as in Equation (12). Note that the likelihood ratio and the prior ratio differ from one only for the two segments affected by the move of the change-points. Finally, the residual variances $\sigma_j^p, \sigma_{j+1}^p$ are drawn from their posterior conditional distributions $\pi(\sigma_j^p | \omega_j^p, \beta_j^p, m_j^p, y_j^p)$, $\pi(\sigma_{j+1}^p | \omega_{j+1}^p, \beta_{j+1}^p, m_{j+1}^p, y_{j+1}^p)$ in a Gibbs step.

### A.2.2 Between-Model Moves

Let $\xi_{(k')}^c = (s_{(k')}^c, m_{(k')}^c, \theta_{(k')}^c)'$ and assume the Markov chain is at $(k^c, \xi_{(k')}^c)$. We propose a move to $(k^p, \xi_{(k')}^p)$ by first drawing $k^p$, followed by sampling the change-point locations $s_{(k')}^p$. The latter involves either merging two segments (death) or splitting a segment (birth). The number of frequencies and their values in the proposed segments are selected from the current state. We draw $\beta_{(k')}^p$ and jointly update the entire state $(k^p, \xi_{(k')}^p)$. Hence, we propose a move to
by drawing from a proposal density of the form

\[ q(k^p, \xi_c^{(k^p)}, k^c, \xi_c^{(k^c)}) = q(k^p | k^c) \times q(\xi_c^{(k^p)} | k^p, k^c, \xi_c^{(k^c)}) \]

\[ \times q(m_c^{(k^p)}, \omega_c^{(k^p)} | s_c^{(k^p)}, k^p, k^c, \xi_c^{(k^c)}) \]

\[ \times q(\sigma_{c c}^{(k^p)} | m_c^{(k^p)}, \omega_c^{(k^p)}, s_c^{(k^p)}, k^p, k^c, \xi_c^{(k^c)}) \]

\[ \times q(\beta_{c s}^{(k^p)} | \sigma_{c c}^{(k^p)}, m_c^{(k^p)}, \omega_c^{(k^p)}, s_c^{(k^p)}, k^p, k^c, \xi_c^{(k^c)}) \].

**Birth move:** If a birth move is proposed, we have that \( k^p = k^c + 1 \). We draw a new change-point uniformly on \( f(s_{c c}^{(k^c)}, \psi_s) \), the support of \( s_{c c}^{(k^c)} \) given the constraints imposed by \( \psi_s \), i.e. \( f(s_{c c}^{(k^c)}, \psi_s) = [1 + \psi_s, s_1^c - \psi_s] \cup [s_1^c + \psi_s, s_2^c - \psi_s] \cup \ldots \cup [s_k^c + \psi_s, n - \psi_s] \). Hence, the new proposed location \( \tilde{s}_j \) is sampled from a Uniform \( \{ f(s_{c c}^{(k^c)}, \psi_s) \} \), where the proposal density is given by

\[ q(s_c^{(k^p)} | k^p, k^c, \xi_c^{(k^c)}) = \frac{1}{(n - 2\psi_s (k^c + 1) - 1)}. \tag{15} \]

As the proposed location \( \tilde{s}_j \) will lie within an existing interval \( (s_j^c, s_{j+1}^c) \) with probability one, we can define the proposed change-points location vector as

\[ s_c^{(k^p)} = (s_1^c, \ldots, s_j^c, \tilde{s}_j, s_{j+1}^c, \ldots, s_k^c)'. \]

The number of frequencies \( m_j^{(k^p)}, m_{j+1}^{(k^p)} \) corresponding to the two newly proposed segments \( [s_j^c, \tilde{s}_j] \) and \( [\tilde{s}_j, s_{j+1}^c] \) are set equal to the current number of frequencies on the whole segment \( (s_j^c, s_{j+1}^c) \). Therefore, we can construct the proposed vector of the number of frequencies \( m_c^{(k^p)} \) and the
proposed vector of frequencies \( \omega_{(k^p)} \) as

\[
m_{(k^p)} = (m_1^c, \ldots, m_{j-1}^c, m_j^c, m_{j+1}^c, \ldots, m_{k^c+1}^c),
\]

\[
\omega_{(k^p)} = (\omega_1^{c'}, \ldots, \omega_{j-1}^{c'}, \omega_j^{c'}, \omega_{j+1}^{c'}, \ldots, \omega_{k^c+1}^{c'}). 
\]

The proposed vector of residual variances \( \sigma_{(k^p)}^{2p} \) is

\[
\sigma_{(k^p)}^{2p} = (\sigma_{1^c}^{2c}, \ldots, \sigma_{j-1}^{2c}, \sigma_j^{2c}, \sigma_{j+1}^{2c}, \ldots, \sigma_{k^c+1}^{2c}),
\]

where the residual variances \( \sigma_{j}^{2p}, \sigma_{j+1}^{2p} \) for the split partition are constructed following Green (1995), namely as a perturbation of the current variance \( \sigma_j^{2c} \). Specifically, we draw \( u \sim \text{Uniform}(0, 1) \) and let \( \sigma_{j}^{2p}, \sigma_{j+1}^{2p} \) be deterministic transformations of \( \sigma_j^{2c} \), i.e

\[
\sigma_{j}^{2p} = \frac{u}{1-u} \sigma_j^{2c}, \quad \sigma_{j+1}^{2p} = \frac{1-u}{u} \sigma_j^{2c}. \tag{16}
\]

Finally, the proposed vector of linear coefficients \( \beta_{(k^p)}^p \) is

\[
\beta_{(k^p)}^p = (\beta_1^{c'}, \ldots, \beta_{j-1}^{c'}, \beta_j^{c'}, \beta_{j+1}^{c'}, \ldots, \beta_{k^c+1}^{c'}),
\]

where the vectors \( \beta_j^p, \beta_{j+1}^p \) are drawn from normal approximations of their posterior conditional distribution, as in Section 4.1. The proposed move to the state \( (k^p, \xi_{(k^p)}^p) \) is accepted with probability

\[
\alpha = \min \left\{ 1, \frac{\mathcal{L}(k^p, \xi_{(k^p)}^p | \mathbf{y})}{\mathcal{L}(k^c, \xi_{(k^c)}^c | \mathbf{y})} \times \frac{\pi(k^p) \pi(\xi_{(k^p)}^p | k^p)}{\pi(k^c) \pi(\xi_{(k^c)}^c | k^c)} \times \frac{d_{k^p} \cdot \frac{1}{k^p} \cdot \frac{1}{2} \cdot q(\beta_{k^c}^c)}{d_{k^c} \cdot q(s_{(k^p)}^p) \cdot \prod_{h=j+1}^{j+1} q(\beta_{k^c}^h) \times J_{\sigma^2}}, \right\}.
\]
where the likelihood function is provided in Equation (3), \( q(s^p_{(k)p}) \) is the uniform density defined in Equation (15); \( q(\beta^c_h), q(\beta^p_h) \) are the multivariate Normal proposal densities, and the Jacobian \( J_{\sigma^2} \) is

\[
J_{\sigma^2} = \frac{\partial (\sigma^2_{j}, \sigma^2_{j+1})}{\partial (\sigma^2_{j+c}, u)} = 2 \left( \frac{\sigma^2_{j} + \sigma^2_{j+1}}{\sqrt{\sigma^2_{j+c}}} \right)^2.
\]

The numerator of the proposal ratio is better understood by looking at the details of the death step, which are given below.

**Death Move:** If a death step is proposed, then \( k^p = k^c - 1 \). A candidate change-point \( s^c_j \) to be removed is sampled uniformly from the vector of existing change-points; that is, we propose to remove \( s^c_j \) with probability \( \frac{1}{k^c} \). Then, the proposed vector of change-points locations \( s^p_{(k)} \) is defined as

\[
s^p_{(k)} = (s^c_1, \ldots, s^c_{j-1}, s^c_{j+1}, \ldots, s^c_{k^c})'.
\]

The number \( m^p_j \) and the vector of relevant frequencies \( \omega^p_j \) of the newly merged segment \( [s^c_{j-1}, s^c_{j+1}] \) are selected by drawing an index at random from \( \{j, j+1\} \), obtaining say \( j^* \), and setting the proposed parameters equal to the current ones relative to the selected index. That is, we set \( m^p_j = m^c_j \) and \( \omega^p_j = \omega^c_j \). Hence, the proposed vectors of number of frequencies \( m^p_{(k)p} \) and their values \( \omega^p_{(k)p} \) are constructed as follows

\[
m^p_{(k)p} = (m^c_1, \ldots, m^c_{j-1}, m^c_j, m^c_{j+2}, \ldots, m^c_{k^c+1})',
\]

\[
\omega^p_{(k)p} = (\omega^c_1', \ldots, \omega^c_{j-1}', \omega^c_j, \omega^c_{j+2}', \ldots, \omega^c_{k^c+1}')'.
\]

The residual variance \( \sigma^2_{j} \) of the newly merged segment is obtained by inverting the transformation of Equation (16). Specifically, we construct \( \sigma^2_{j} = \sqrt{\sigma^2_{j+c} \sigma^2_{j+1}} \) and set the proposed vector of residual variances \( \sigma^2_{(k)p} \) as

\[
\sigma^2_{(k)p} = (\sigma^2_{1}, \ldots, \sigma^2_{j-1}, \sigma^2_{j}, \sigma^2_{j+2}, \ldots, \sigma^2_{k^c+1})'.
\]
The proposed vector of linear coefficients $\beta_{(k,p)}$ is

$$\beta_{(k,p)} = (\beta_1^{c'}, \ldots, \beta_{j-1}^{c'}, \beta_j^{p'}, \beta_{j+2}^{c'}, \ldots, \beta_{k'+1}^{c'})',$$

where the vector of coefficients $\beta_j^{p'}$ is drawn from Normal approximation to its posterior conditional distribution. The acceptance probability for the death step has the same form of the birth step, with the proper change of labelling of the variables, and the ratio terms inverted.

**References**

Akaike, H. (1974), ‘A new look at the statistical model identification’, *IEEE Transactions on Automatic Control* **19**(6), 716–723.

Altevogt, B. M., Colten, H. R. et al. (2006), *Sleep disorders and sleep deprivation: an unmet public health problem*, Washington (DC): National Academies Press.

Andrieu, C. & Doucet, A. (1999), ‘Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC’, *IEEE Transactions on Signal Processing* **47**(10), 2667–2676.

Berlad, I., Shlitzer, A., Ben-Haim, S. & Lavie, P. (1993), ‘Power spectrum analysis and heart rate variability in Stage 4 and REM sleep: evidence for state-specific changes in autonomic dominance’, *Journal of Sleep Research* **2**(2), 88–90.

Bishop, C. M. (2006), *Pattern Recognition and Machine Learning (Information Science and Statistics)*, New York: Springer-Verlag.
Bretthorst, G. L. (1988), *Bayesian spectrum analysis and parameter estimation*, Vol. 48, New York: Springer-Verlag.

Bretthorst, G. L. (1990), ‘Bayesian analysis. I. Parameter estimation using quadrature NMR models’, *Journal of Magnetic Resonance* **88**(3), 533–551.

Carskadon, M. A., Dement, W. C. et al. (2005), ‘Normal human sleep: an overview’, *Principles and Practice of Sleep Medicine* **4**, 13–23.

Dahlhaus, R. et al. (1997), ‘Fitting time series models to nonstationary processes’, *The Annals of Statistics* **25**(1), 1–37.

Davis, E. M. & ODonnell, C. P. (2013), ‘Rodent models of sleep apnea’, *Respiratory Physiology & Neurobiology* **188**(3), 355–361.

Davis, R. A., Lee, T. C. M. & Rodriguez-Yam, G. A. (2006), ‘Structural break estimation for nonstationary time series models’, *Journal of the American Statistical Association* **101**(473), 223–239.

Djuric, P. M. (1996), ‘A model selection rule for sinusoids in white Gaussian noise’, *IEEE Transactions on Signal Processing* **44**(7), 1744–1751.

Dou, L. & Hodgson, R. (1995), ‘Bayesian inference and Gibbs sampling in spectral analysis and parameter estimation. I’, *Inverse Problems* **11**(5), 1069.

Dou, L. & Hodgson, R. (1996), ‘Bayesian inference and Gibbs sampling in spectral analysis and parameter estimation: II’, *Inverse Problems* **12**(2), 121.

Gilks, W. R., Richardson, S. & Spiegelhalter, D. (1995), *Markov chain Monte Carlo in practice*, Boca Raton: CRC Press.
Green, P. J. (1995), ‘Reversible jump Markov chain Monte Carlo computation and Bayesian model determination’, *Biometrika* **82**(4), 711–732.

Han, F., Subramanian, S., Price, E. R., Nadeau, J. & Strohl, K. P. (2002), ‘Periodic breathing in the mouse’, *Journal of Applied Physiology* **92**(3), 1133–1140.

Harsch, I. A., Schahin, S. P., Brückner, K., Radespiel-Tröger, M., Fuchs, F. S., Hahn, E. G., Konturek, P. C., Lohmann, T. & Ficker, J. H. (2004), ‘The effect of continuous positive airway pressure treatment on insulin sensitivity in patients with obstructive sleep apnoea syndrome and type 2 diabetes’, *Respiration* **71**(3), 252–259.

Heinzer, R., Vat, S., Marques-Vidal, P., Marti-Soler, H., Andries, D., Tobback, N., Mooser, V., Preisig, M., Malhotra, A., Waeber, G. et al. (2015), ‘Prevalence of sleep-disordered breathing in the general population: the Hypnoolaus study’, *The Lancet Respiratory Medicine* **3**(4), 310–318.

Huang, Q., Cohen, D., Komarzynski, S., Li, X.-M., Innominato, P., Lévi, F. & Finkenstädt, B. (2018), ‘Hidden Markov models for monitoring circadian rhythmicity in telemetric activity data’, *Journal of The Royal Society Interface* **15**(139), 20170885.

Komarzynski, S., Huang, Q., Innominato, P. F., Maurice, M., Arbaud, A., Beau, J., Bouchahda, M., Ulusakarya, A., Beaumatin, N., Breda, G. et al. (2018), ‘Relevance of a mobile internet platform for capturing inter-and intrasubject variabilities in circadian coordination during daily routine: pilot study’, *Journal of Medical Internet Research* **20**(6), e204.

Kräuchi, K. (2002), ‘How is the circadian rhythm of core body temperature regulated?’, *Clinical Autonomic Research* **12**(3), 147–149.
Krauchi, K. & Wirz-Justice, A. (1994), ‘Circadian rhythm of heat production, heart rate, and skin and core temperature under unmasking conditions in men’, *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **267**(3), R819–R829.

Lal, C., Strange, C. & Bachman, D. (2012), ‘Neurocognitive impairment in obstructive sleep apnea’, *Chest* **141**(6), 1601–1610.

Lanfranchi, P. A., Braghiroli, A., Bosimini, E., Mazzuero, G., Colombo, R., Donner, C. F. & Giannuzzi, P. (1999), ‘Prognostic value of nocturnal Cheyne-Stokes respiration in chronic heart failure’, *Circulation* **99**(11), 1435–1440.

Levi, F. & Schibler, U. (2007), ‘Circadian rhythms: mechanisms and therapeutic implications’, *Annual Review of Pharmacology and Toxicology* **47**, 593–628.

Nakamura, A., Fukuda, Y. & Kuwaki, T. (2003), ‘Sleep apnea and effect of chemostimulation on breathing instability in mice’, *Journal of Applied Physiology* **94**(2), 525–532.

Nieto, F. J., Peppard, P. E., Young, T., Finn, L., Hla, K. M. & Farré, R. (2012), ‘Sleep-disordered breathing and cancer mortality: results from the Wisconsin sleep cohort study’, *American Journal of Respiratory and Critical Care Medicine* **186**(2), 190–194.

Osorio, R. S., Ducca, E. L., Wohlleber, M. E., Tanzi, E. B., Gumb, T., Twumasi, A., Tweardy, S., Lewis, C., Fischer, E., Koushyk, V. et al. (2016), ‘Orexin-A is associated with increases in cerebrospinal fluid phosphorylated-tau in cognitively normal elderly subjects’, *Sleep* **39**(6), 1253–1260.

Pace-Schott, E. F. & Hobson, J. A. (2002), ‘The neurobiology of sleep: genetics, cellular physiology and subcortical networks’, *Nature Reviews Neuroscience* **3**(8), 591.
Parish, J. M., Adam, T. & Facchiano, L. (2007), ‘Relationship of metabolic syndrome and obstructive sleep apnea’, *Journal of Clinical Sleep Medicine* 3(5), 467.

Priestley, M. B. (1981), *Spectral analysis and time series*, London: Academic Press.

Quinn, B. G. (1989), ‘Estimating the number of terms in a sinusoidal regression’, *Journal of Time Series Analysis* 10(1), 71–75.

Rife, D. C. & Boorstyn, R. R. (1976), ‘Multiple tone parameter estimation from discrete-time observations’, *Bell System Technical Journal* 55(9), 1389–1410.

Rissanen, J. (1978), ‘Modeling by shortest data description’, *Automatica* 14(5), 465–471.

Rosen, O., Stoffer, D. S. & Wood, S. (2009), ‘Local spectral analysis via a Bayesian mixture of smoothing splines’, *Journal of the American Statistical Association* 104(485), 249–262.

Rosen, O., Wood, S. & Stoffer, D. S. (2012), ‘AdaptSPEC: Adaptive spectral estimation for nonstationary time series’, *Journal of the American Statistical Association* 107(500), 1575–1589.

Shneerson, J. M. (2009), *Sleep medicine: a guide to sleep and its disorders*, Hoboken: John Wiley & Sons.

Shumway, R. H. & Stoffer, D. S. (2005), *Time Series Analysis and Its Applications (Springer Texts in Statistics)*, New York: Springer-Verlag.

Stoica, P., Moses, R. L., Friedlander, B. & Soderstrom, T. (1989), ‘Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements’, *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37(3), 378–392.
Sundaram, S. S., Halbower, A., Pan, Z., Robbins, K., Capocelli, K. E., Klawitter, J., Shearn, C. T. & Sokol, R. J. (2016), ‘Nocturnal hypoxia-induced oxidative stress promotes progression of pediatric non-alcoholic fatty liver disease’, *Journal of Hepatology* **65**(3), 560–569.

Tan, A., Cheung, Y. Y., Yin, J., Lim, W.-Y., Tan, L. W. & Lee, C.-H. (2016), ‘Prevalence of sleep-disordered breathing in a multiethnic Asian population in Singapore: A community-based study’, *Respirology* **21**(5), 943–950.

Van Someren, E. J. (2006), ‘Mechanisms and functions of coupling between sleep and temperature rhythms’, *Progress in Brain Research* **153**, 309–324.

Whittle, P. (1957), ‘Curve and periodogram smoothing’, *Journal of the Royal Statistical Society. Series B (Methodological)* **19**(1), 38–63.

Yau, S.-F. & Bresler, Y. (1993), ‘Maximum likelihood parameter estimation of superimposed signals by dynamic programming’, *IEEE Transactions on Signal Processing* **41**(2), 804–820.

Zhang, Q. & Wong, K. M. (1993), ‘Information theoretic criteria for the determination of the number of signals in spatially correlated noise’, *IEEE Transactions on Signal Processing* **41**(4), 1652–1663.
A SUPPLEMENTARY MATERIAL

A.1 Details of simulation study

Table 5: Illustrative example. Parameter values for simulation from model 1.

| Frequencies | Linear coefficients | Trends and variances |
|-------------|----------------------|----------------------|
| $\omega_{1,1}$ | 1/24 | $\beta_{1,1}$ | (2.0, 3.0) | $\mu_1$ | .010 |
| $\omega_{1,2}$ | 1/15 | $\beta_{1,2}$ | (4.0, 5.0) | $\mu_2$ | .000 |
| $\omega_{1,3}$ | 1/7 | $\beta_{1,3}$ | (1.0, 2.5) | $\mu_3$ | -.005 |
| $\omega_{2,1}$ | 1/12 | $\beta_{2,1}$ | (4.0, 3.0) | $\sigma^2_1$ | 4.0^2 |
| $\omega_{3,1}$ | 1/22 | $\beta_{3,1}$ | (2.5, 4.0) | $\sigma^2_2$ | 3.5^2 |
| $\omega_{3,2}$ | 1/15 | $\beta_{3,2}$ | (4.0, 2.0) | $\sigma^2_3$ | 2.8^2 |

The value of the intercept was set to zero for every segment.

Table 6: Sensitivity analysis illustrative example. Average MSE from 10 replications with varying $\lambda_\omega$ and $\lambda_s$

| | $\lambda_s = .1$ | $\lambda_s = .2$ | $\lambda_s = .5$ | $\lambda_s = 1.0$ | $\lambda_s = 2.0$ | $\lambda_s = 5.0$ | $\lambda_s = 10.0$ |
|----------------------|----------------|----------------|----------------|------------------|------------------|------------------|------------------|
| $\lambda_\omega = .1$ | .349 | .354 | .400 | .302 | .307 | .360 | .355 |
| $\lambda_\omega = .2$ | .435 | .360 | .347 | .392 | .369 | .387 | .386 |
| $\lambda_\omega = .5$ | .370 | .378 | .378 | .400 | .404 | .362 | .340 |
| $\lambda_\omega = 1.0$ | .416 | .361 | .447 | .382 | .340 | .396 | .387 |
| $\lambda_\omega = 2.0$ | .446 | .370 | .430 | .321 | .407 | .346 | .420 |
| $\lambda_\omega = 5.0$ | .470 | .404 | .346 | .400 | .329 | .350 | .393 |
| $\lambda_\omega = 10.0$ | .319 | .394 | .364 | .337 | .391 | .324 | .428 |

Table 6 displays the average mean squared error

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (\hat{f}_i - f_i)^2,$$
to assess the distance between the true underlying signal $f_t$ and the estimated signal $\hat{f}_t$. The latter is obtained by averaging across models of differing number of components, in contrast to model selection. Specifically, if we run our procedure for $S$ iterations, then the estimated signal $\hat{f}_t$ is defined as

$$\hat{f}_t = \frac{1}{S} \sum_{s=1}^{S} \sum_{j=1}^{k(s) + 1} f(t, \beta_j^{(s)}, \omega_j^{(s)}) \mathbb{1}_{[t \in I_j^{(s)}]}, \quad t = 1, \ldots, n,$$

where the superscript $(s)$ denotes the $s^{th}$ sample of the Markov chain.