Stat

The ISI's Journal for the Rapid
Dissemination of Statistics Research

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

Improved nonparametric inference for multiple correlated periodic sequences

Ying Sun[a,*], Jeffrey D. Hart[b] and Marc G. Genton[c]

Received 14 July 2013; Accepted 21 July 2013

This paper proposes a cross-validation method for estimating the period as well as the values of multiple correlated periodic sequences when data are observed at evenly spaced time points. The period of interest is estimated conditional on the other correlated sequences. An alternative method for period estimation based on Akaike's information criterion is also discussed. The improvement of the period estimation performance is investigated both theoretically and by simulation. We apply the multivariate cross-validation method to the temperature data obtained from multiple ice cores, investigating the periodicity of the El Niño effect. Our methodology is also illustrated by estimating patients' cardiac cycle from different physiological signals, including arterial blood pressure, electrocardiography, and fingertip plethysmograph. Copyright © 2013 John Wiley & Sons, Ltd.

Keywords: cross-validation; model selection; multiple sequences; nonparametric estimation; period

1 Introduction

A number of nonparametric methods for period estimation in univariate time series have been proposed recently, including Hall et al. (2000), Hall & Yin (2003), Hall & Li (2006), Genton & Hall (2007), and Hall (2008). In these papers, to estimate the period of a periodic function, an appropriate random spacing of time points is needed in order to ensure consistency; in other words, unevenly spaced observations are needed. Considering that equally spaced data are prevalent in time-series analysis, Sun et al. (2012) proposed a cross-validation (CV) based nonparametric method for evaluating periodicity when time points are evenly spaced and a periodic sequence is observed from the model

$$Y_t = \mu_t + \varepsilon_t, \quad t = 1, \ldots, n,$$

where $Y_1, \ldots, Y_n$ are the observations; $\mu_1, \ldots, \mu_n$ are unknown periodic constants; and the errors $\varepsilon_1, \ldots, \varepsilon_n$ are independent and identically distributed (i.i.d.), and mean-zero random variables.

To make statistical inference on periodic sequences, the traditional approach is to estimate the values of the periodic sequence via parametric models, for example, trigonometric regression, where the estimated periodic component is a linear combination of a finite number of unknown regressors. These unknown parameters can be estimated by frequency domain methods based on the periodogram, such as in Walker (1971), Rice & Rosenblatt (1988), and Quinn & Thomson (1991). Moreover, Li (2012) proposed methods for detecting periodicity by using the quantile periodogram when the noise is asymmetric. Another direction of statistical inference for the periodic behavior in the frequency domain is periodicity testing (Canova, 1996). Typically, one says that a sequence exhibits periodicity if its

[a]Department of Statistics, University of Chicago, Chicago, IL 60637, USA
[b]Department of Statistics, Texas A&M University, College Station, TX 77843, USA
[c]CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia
*Email: sunwards@galton.uchicago.edu

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

The ISI's Journal for the Rapid
Dissemination of Statistics Research

spectral density has a peak or a large mass at a certain frequency. Then a test is developed to determine whether the cycle is significant by comparing the filtered variance with respect to the frequency to the variance of the original series. As commented by Sun et al. (2012), although these frequency domain methods are suitable for identifying sinusoidal functions that are highly correlated with the data, they are not appropriate for the problem of estimating the smallest period of a periodic function. The detailed discussion on the advantage of the CV method over the Quinn and Thomson periodogram-based method can be found in Sun et al. (2012). Furthermore, when little is known about the structure of such periodic sequences, it is important to have nonparametric methods for period estimation.

In many real-world problems, multivariate time series are available. For instance, several different variables are simultaneously recorded from a system under study, such as atmospheric temperature, pressure, and humidity in meteorology, or heart rate, blood pressure, and respiration in physiology. A multivariate time series could also be recorded from one variable but in spatially extended systems, such as in studies of meteorology, where temperature recordings are obtained from probes or satellites at different spatial locations. In this paper, we use a similar framework as in Sun et al. (2012) to develop nonparametric methods for period estimation when multiple correlated periodic sequences are observed at evenly spaced time points. The basic idea is to borrow information from other correlated sequences to improve the period estimation of interest.

One motivation for our methodology is the study of periodicity of El Niño effects, which are defined as sustained increases of at least $0.5\,^{\circ}$C in average sea surface temperatures over the east-central tropical Pacific Ocean. The El Niño phenomenon dramatically affects the weather in many parts of the world. It is therefore important to better understand its appearance. Various climate models and statistical models attempt to predict El Niño as a part of interannual climate variability. Sun et al. (2012) have applied their method to a series of sea surface temperatures to estimate the El Niño period. The El Niño Southern Oscillation (ENSO) is in the tropical Pacific Ocean area. In this paper, we investigate whether the ENSO can be detected in the North Atlantic region by analyzing temperature data from different ice cores in Greenland.

Let d be the number of sequences. We consider the following model:

$$\mathbf{Y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \ldots, n, \tag{1}$$

where $\mathbf{Y}_t = (Y_{1,t}, \ldots, Y_{d,t})^{\mathsf{T}} = (X_t, \mathbf{Z}_t^{\mathsf{T}})^{\mathsf{T}}$ is the observed vector with $X_t = Y_{1,t}$ and $\mathbf{Z}_t = (Y_{2,t}, \ldots, Y_{d,t})^{\mathsf{T}}$, $\boldsymbol{\mu}_t = (\mu_{1,t}, \ldots, \mu_{d,t})^{\mathsf{T}}$ is an unknown constant vector, and the error vectors $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{d,t})^{\mathsf{T}}$, $t = 1, \ldots, n$, are i.i.d. with mean zero, and covariance matrix $\boldsymbol{\Sigma}$ that can be partitioned into

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \boldsymbol{\sigma}_{12}^{\mathsf{T}} \\ \boldsymbol{\sigma}_{12} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

with $\sigma_1^2 = \mathrm{Var}(X_t)$, $\boldsymbol{\Sigma}_{22} = \mathrm{Var}(\mathbf{Z}_t)$ and $\boldsymbol{\sigma}_{12} = \mathrm{Cov}(X_t, \mathbf{Z}_t)$. It is assumed that $\mu_{s,i} = \mu_{s,i+mp_s}$ for $i = 1, \ldots, p_s, m = 1, 2, \ldots,$ $s = 1, \ldots, d$, and integers $p_1, \ldots, p_d$, each of which is at least 2. Of interest is estimating p_1, the period of the first sequence $X_1, \ldots, X_n$.

Our method of estimating the period of interest is also based on CV. The CV estimator proposed by Sun et al. (2012) for one periodic sequence evaluates a candidate period q by first "stacking," at the same time point, all data that are separated in time by a multiple of q and then computing a "leave-out-one-cycle" version of the variance for each of the q stacks of data. In other words, if we have k complete cycles where $k = n/q$, the stacked means of the other $k-1$ cycles are used to predict the ones left out, and then the averaged prediction errors are computed. For multiple correlated periodic sequences, suppose we are interested in estimating p_1 in model (1). Instead of the stacked means, we propose to use the conditional means, that is, conditional on other correlated sequences, to predict the left-out cycle in CV. The prediction errors tend to be smallest when the period candidate q_1 equals p_1. As CV is typically used

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

as a model selection tool, we also consider model selection criteria for multiple periodic sequences other than CV in estimating a period.

Sun et al. (2012) have described the asymptotic behavior of the one sequence CV method. They showed that when p_1 is sufficiently large, the period estimator $\hat{p}_1$ is virtually consistent, in the sense that $\lim_{n\to\infty} P\left(\hat{p}_1 = p_1\right)$ increases to 1 as p_1 increases. In this paper, we show that the multivariate CV has the same asymptotic properties. However, for finite samples, our simulation studies indicate that stronger correlation between sequences, better period estimation of other correlated sequences, and use of more correlated sequences all lead to substantial improvements in estimating the period of interest.

The rest of this paper is organized as follows. Section 2 describes the CV method of estimating the period and sequence values for multiple periodic sequences. Section 3 discusses asymptotic properties of the method. Simulations motivated by real data applications are reported in Section 4. Two applications to ice core data and physiological signals data are presented in Section 5. Concluding remarks are provided in Section 6, and a derivation of maximum likelihood estimates (MLEs) for bivariate sequence means is in the Appendix.

2 Methodology

Suppose we observe multiple sequences at equally spaced time points from model (1), where for each $s = 1, \ldots, d$, the sequence $\mu_{s,1}, \ldots, \mu_{s,n}$ is periodic with (smallest) period p_s and the $\boldsymbol{\varepsilon}_t$s are i.i.d. random vectors with zero means and covariance matrix $\boldsymbol{\Sigma}$. We propose a methodology for estimating p_1, the period of interest, in Section 2.1 and an alternative based on Akaike's information criterion (Akaike, 1973, AIC) in Section 2.2. Besides period estimation, we also discuss the estimation of sequence values for multiple periodic sequences.

2.1. Multivariate cross-validation method for period estimation

For one periodic sequence, Sun et al. (2012) proposed the CV method for period estimation. Let q be a candidate integer period with $2 \leqslant q \leqslant M_n$, where M_n is of smaller order than $\sqrt{n}$. For each $i = 1, \ldots, q$, they constructed an estimator of μ_i by stacking all data that are separated in time by a multiple of q. So at time points i, the data are $X_i, X_{i+q}, \ldots, X_{i+qk_{q,i}}$, where $k_{q,i}$ is the largest integer such that $i + qk_{q,i} \leqslant n$. For each relevant q, i, and j, define $X_{qij} = X_{i+(j-1)q}$. Let

$$\mathrm{CV}(q) = \frac{1}{n}\sum_{i=1}^{q}\sum_{j=1}^{k_{q,i}}\left(X_{qij} - \bar{X}_{qi}^{j}\right)^2, \tag{2}$$

where $\bar{X}_{qi}^{j}$ is the average of $X_{qi\ell}$, $\ell = 1, \ldots, k_{q,i}$, excluding X_{qij}. A period estimator $\hat{p}$ is defined to be the minimizer of $\mathrm{CV}(q)$ for $2 \leqslant q \leqslant M_n$.

Now, suppose we observe multiple sequences at equally spaced time points from model (1). Let q_s be a candidate integer period of the s-th sequence. For the first sequence, let $q = q_1$ for simplicity, and then $\bar{X}_{qi}^{j}$ is the average of $X_{qi\ell}$ excluding X_{qij}, $i = 1, \ldots, q, \ell = 1, \ldots, k_{qi}$ and $i + qk_{q,i} \leqslant n$. Define the averages for the sequences for $s = 2, \ldots, d$ in a similar way. For $i = 1, \ldots, q_s$, let $\bar{Y}_{s,q_s i}$ be the average of $Y_{s,q_s i\ell}$, $\ell = 1, \ldots, k_{q_s,i}$ and $i + q_s k_{q_s,i} \leqslant n$. Then define the residual for each of the sequences to be $\hat{\varepsilon}_{s,q_s i\ell} = Y_{s,q_s i\ell} - \bar{Y}_{s,q_s i}$. Now, by defining the residual separately for each sequence, we have a residual vector $\hat{\boldsymbol{\varepsilon}}_t$ at each time point t, $t = 1, \ldots, n$. Let

$$\mathrm{CV}_d(q) = \frac{1}{n}\sum_{i=1}^{q}\sum_{j=1}^{k_{q,i}}\left[X_{qij} - \left\{\bar{X}_{qi}^{j} + \boldsymbol{\sigma}_{12}^{\mathrm{T}}\boldsymbol{\Sigma}_{22}^{-1}\hat{\boldsymbol{\varepsilon}}_{qij}\right\}\right]^2, \tag{3}$$

Stat

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

The ISI's Journal for the Rapid Dissemination of Statistics Research

where $\hat{\boldsymbol{\varepsilon}}_{qij} = \hat{\boldsymbol{\varepsilon}}_{i+(j-1)q}$. Initially, we assume $\boldsymbol{\sigma}_{12}$ and $\boldsymbol{\Sigma}_{22}$ to be known and define a period estimator $\hat{p}_1$ to be the minimizer of $\mathrm{CV}_d(q)$ given $\hat{p}_2, \ldots, \hat{p}_d$, each of which is the minimizer of $\mathrm{CV}(q)$ in (2). In fact, we can choose $q_1, \ldots, q_d$ to minimize $\mathrm{CV}_d(q)$ in general.

The case of two sequences is a special case of the CV criterion in (3). If we observe two sequences $\{X_t : t = 1, \ldots, n\}$ and $\{Z_t : t = 1, \ldots, n\}$, criterion (3) can be simplified as

$$\mathrm{CV}_2(q) = \frac{1}{n} \sum_{i=1}^{q} \sum_{j=1}^{k_{q,i}} \left[X_{qij} - \left\{ \bar{X}^j_{qi} + \frac{\sigma_1}{\sigma_2} \rho \, \hat{\varepsilon}_{qij} \right\} \right]^2, \tag{4}$$

where $\sigma_1^2 = \mathrm{Var}(X_t), \sigma_2^2 = \mathrm{Var}(Z_t), \rho = \mathrm{Cov}(X_t, Z_t), \hat{\varepsilon}_{qij} = \hat{\varepsilon}_{i+(j-1)q}, i = 1, \ldots, q, j = 1, \ldots, k_{q,i}, r = 1, \ldots, q_2$, and $\ell = 1, \ldots, k_{q_2,r}$. Then we propose to estimate p_1 in the following way:

- S1. Apply the CV method (2) for one periodic sequence in Sun et al. (2012) to each sequence and obtain the period estimates $\hat{p}_1^{(0)}$ and $\hat{p}_2$ for p_1 and p_2, respectively.
- S2. Estimate the sequence values for each periodic sequence using stacked means corresponding to the estimated periods $\hat{p}_1^{(0)}$ and $\hat{p}_2$. We show in the Appendix that these estimators of the sequence values are $\sqrt{n}$-consistent.
- S3. Subtract the estimated sequence values from the observations, compute the sample covariance matrix from residual vectors, and obtain $\hat{\sigma}_1, \hat{\sigma}_2$ and $\hat{\rho}$.
- S4. Construct $\bar{X}^j_{qi} + \frac{\hat{\sigma}_1}{\hat{\sigma}_2} \hat{\rho} \, \hat{\varepsilon}_{qij}$ to predict X_{qij} and compute the averaged squared prediction errors in (4).
- S5. Choose q to minimize (4) and obtain the period estimate $\hat{p}_1^{(1)}$ of the first sequence.
- S6. Repeat steps 2–5 and choose the period estimate $\hat{p}_1 = \hat{p}_1^{(2)}$.

When $\boldsymbol{\varepsilon}_t$ has an elliptical distribution, the best linear predictor of X_t given Z_t is $\mu_{1,t} + \frac{\sigma_1}{\sigma_2} \rho \varepsilon_{2,t}$. This motivates the predictor $\bar{X}^j_{qi} + \frac{\hat{\sigma}_1}{\hat{\sigma}_2} \hat{\rho} \hat{\varepsilon}_{qij}$ of X_{qij}. When $\rho = 0$, Equation (4) reduces to Equation (2) for the one-sequence case, and the second sequence is not used. It is also worth pointing out that the multivariate CV method is computationally simple because the main computation of the prediction errors for each period candidate only involves calculating the stacked means.

2.2. Model selection criteria for multiple periods estimation

Cross-validation is typically used as a model selection tool. Sun et al. (2012) discussed period estimation for one sequence from a model selection point of view, as in fact each candidate period q corresponds to a model for the sequence consisting of the q parameters $\mu_1, \ldots, \mu_q$. Similarly, for multiple periodic sequences, if we assume that the error vectors in model (1) are i.i.d. multivariate normal, then AIC has the form

$$\mathrm{AIC}(q_1, \ldots, q_d) = -2 \log L \left(\hat{\boldsymbol{\mu}}_t, \hat{\boldsymbol{\Sigma}} | \mathbf{Y}_t, t = 1, \ldots, n \right) + 2 \sum_{s=1}^{d} q_s, \tag{5}$$

where $\hat{\boldsymbol{\mu}}_t$ and $\hat{\boldsymbol{\Sigma}}$ are the MLEs of $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}$, respectively, when the periods are assumed to be $q_1, \ldots, q_d$. Minimizing $\mathrm{AIC}(q_1, \ldots, q_d)$ provides estimators of $p_1, \ldots, p_d$. For the one-sequence case, it is straightforward to verify that the MLEs of $\mu_1, \ldots, \mu_q$ are the stacked means $\bar{Y}_{q1}, \ldots, \bar{Y}_{qq}$. For multiple correlated sequences, instead of simple stacked means for each sequence, the MLEs depend on the covariance matrix $\boldsymbol{\Sigma}$. The proof for the bivariate case is in the Appendix. We can see that only when $\rho = 0$ are the MLEs the usual stacked means. For $\rho \neq 0$, the MLEs depend on the value of ρ and the observations from both of the sequences.

The ISI's Journal for the Rapid Dissemination of Statistics Research

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

3 Theoretical properties of the bivariate CV method

Here, we provide some theoretical evidence for the effectiveness of our bivariate CV method. First of all, we can argue that, asymptotically, the proposed predictor of X_t is less noisy than that based on the one-series method of Sun et al. (2012). Let $\tilde{X}_q^{(t)}$ be the leave-out-one-cycle predictor of X_t using only data from the first sequence. The prediction error in this case is

$$X_t - \tilde{X}_q^{(t)} = \varepsilon_{1,t} + \left(\mu_{1,t} - \tilde{X}_q^{(t)}\right),$$

and in the two-sequence case, it is

$$X_t - \left(\tilde{X}_q^{(t)} + \frac{\hat{\sigma}_1}{\hat{\sigma}_2}\hat{\rho}\hat{\varepsilon}_{2,t}\right) = \varepsilon_{1,t} - \frac{\sigma_1}{\sigma_2}\rho\varepsilon_{2,t} + \left(\mu_{1,t} - \tilde{X}_q^{(t)}\right) + O_p\left(n^{-1/2}\right),$$

where for the latter error we have used a consistency result to be discussed near the end of this section.

Now, for $\delta_t = \varepsilon_{1,t} - (\sigma_1/\sigma_2)\rho\varepsilon_{2,t}$,

$$\mathrm{Var}(\delta_t) = \sigma_1^2 + \frac{\sigma_1^2}{\sigma_2^2}\rho^2\sigma_2^2 - 2\frac{\sigma_1}{\sigma_2}\rho\sigma_1\sigma_2\rho = \sigma_1^2\left(1-\rho^2\right) \leqslant \sigma_1^2.$$

Therefore, by using the second sequence, we can better predict X_t when $\rho \neq 0$, which in turn should lead to a less noisy CV criterion.

The first-stage CV criterion that uses information from both sequences is

$$\begin{aligned}
\mathrm{CV}_2(q) = \frac{1}{n}\sum_{t=1}^{n}\left(X_t - \tilde{X}_q^{(t)} - \frac{\hat{\sigma}_1}{\hat{\sigma}_2}\hat{\rho}\hat{\varepsilon}_{2,t}\right)^2 &= \mathrm{CV}_1(q) - \frac{2}{n}\frac{\hat{\sigma}_1}{\hat{\sigma}_2}\hat{\rho}\sum_{t=1}^{n}\left(X_t - \tilde{X}_q^{(t)}\right)\hat{\varepsilon}_{2,t} + \frac{1}{n}\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}\hat{\rho}^2\sum_{t=1}^{n}\hat{\varepsilon}_{2,t}^2 \\
&= \mathrm{CV}_1(q) - \frac{2}{n}\frac{\sigma_1}{\sigma_2}\rho\sum_{t=1}^{n}\left(X_t - \tilde{X}_q^{(t)}\right)\varepsilon_{2,t} + \frac{1}{n}\frac{\sigma_1^2}{\sigma_2^2}\rho^2\sum_{t=1}^{n}\varepsilon_{2,t}^2 + \xi_n \\
&= \mathrm{CV}_1(q) - \frac{2}{n}\frac{\sigma_1}{\sigma_2}\rho\sum_{t=1}^{n}\left(\mu_{1,t} - \tilde{X}_q^{(t)}\right)\varepsilon_{2,t} + \xi_n - \frac{2}{n}\frac{\sigma_1}{\sigma_2}\rho\sum_{t=1}^{n}\varepsilon_{1,t}\varepsilon_{2,t} + \frac{1}{n}\frac{\sigma_1^2}{\sigma_2^2}\rho^2\sum_{t=1}^{n}\varepsilon_{2,t}^2.
\end{aligned}$$

Now, the last two summands on the right-hand side of the last equation are free of q and hence irrelevant. Because $\tilde{X}_q^{(t)}$ and $\varepsilon_{2,t}$ are independent, we may use techniques as in Sun et al. (2012) to show that the effect of $\frac{2}{n}\frac{\sigma_1}{\sigma_2}\rho\sum_{t=1}^{n}\left(\mu_{1,t} - \tilde{X}_q^{(t)}\right)\varepsilon_{2,t}$ is asymptotically negligible. The term ξ_n is $O_p(n^{-1/2})$, and hence negligible, if we can show that the initial estimators of the two sequences of means are $\sqrt{n}$-consistent uniformly in t. This is carried out in the Appendix. So, the bottom line is that, under the conditions of Sun et al. (2012), the two-series CV method is asymptotically equivalent to the one sequence method.

Although the two methods are asymptotically equivalent, the simulation results of the next section provide ample evidence that the multiple series approach can improve upon the single series approach in moderate-sized samples.

4 Simulations

4.1. Bivariate case

To study whether another correlated periodic sequence improves the period estimation, we performed simulations with an intermediate value of the period, that is, $p_1 = 43$ months, which is the estimate of the El Niño period obtained by Sun et al. (2012). As it is well known that sea surface temperatures are negatively correlated with sea level pressures,

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

The ISI's Journal for the Rapid
Dissemination of Statistics Research

we create another series with $p_2 = 59$, let $\sigma_1 = \sigma_2 = 1$ in model (1), and consider three cases, $\rho = 0, -0.6, -0.8$. For all cases, the set over which the objective function $\mathrm{CV}_2(q)$ was searched was taken to be $\{12, 13, \ldots, 96\}$ months, or 1 to 8 years, which is the possible range of El Niño period, and the algorithm S1–S5 proposed in Section 2.1 was used to estimate p_1. The number of replications of each setting is 1000.

The left panel of Figure 1 shows how the probabilities of choosing $p_1 = 43$ increase as the sample size n increases for each $\rho = 0, -0.6, -0.8$ when $p_2 = 59$. It is clear that the convergence is faster for larger correlation. We can also see that for finite samples, the second sequence does improve the estimation of p_1, and more improvements could be obtained with stronger correlation. Then we consider different values of p_2 when $\rho = -0.8$. The probabilities of choosing $p_1 = 43$ are shown in the right panel of Figure 1 for $p_2 = 20, 59, 80$. For a fixed sample size n, the probability of choosing $p_1 = 43$ is higher for smaller values of p_2 because of more available cycles of the second sequence. Thus, for a fixed value of ρ, how much the second series helps for estimating p_1 depends on how well p_2 is estimated as well.

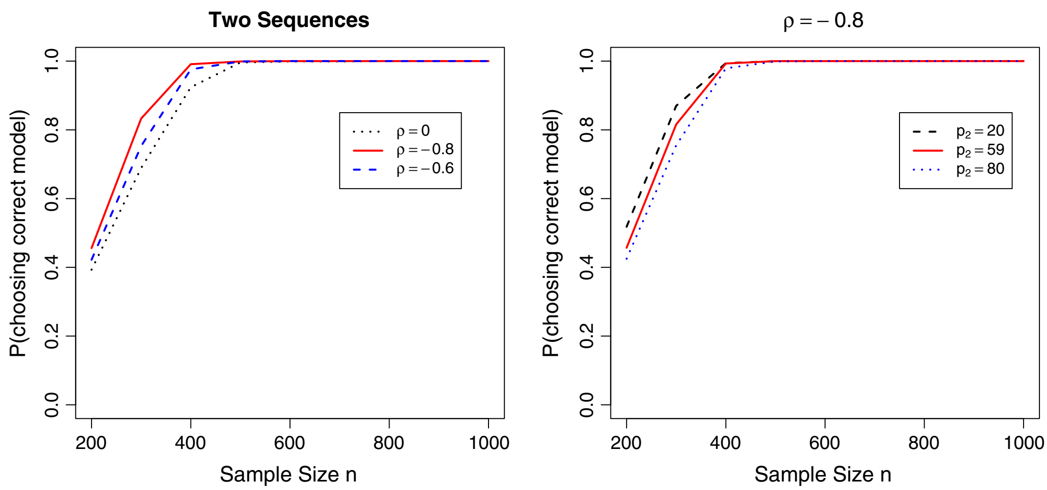

Figure 1. Left panel: $P(\hat{p}_1 = p_1)$ for the bivariate CV method at different correlation levels. Right panel: $P(\hat{p}_1 = p_1)$ for the bivariate CV method at different values of p_2.

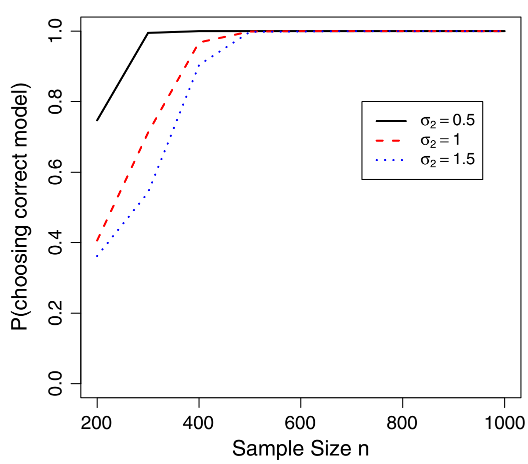

Figure 2. $P(\hat{p}_1 = p_1)$ for the bivariate CV method at different error levels of the second sequence. The values of σ_1 and ρ are 1 and -0.8, respectively, in each case.

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

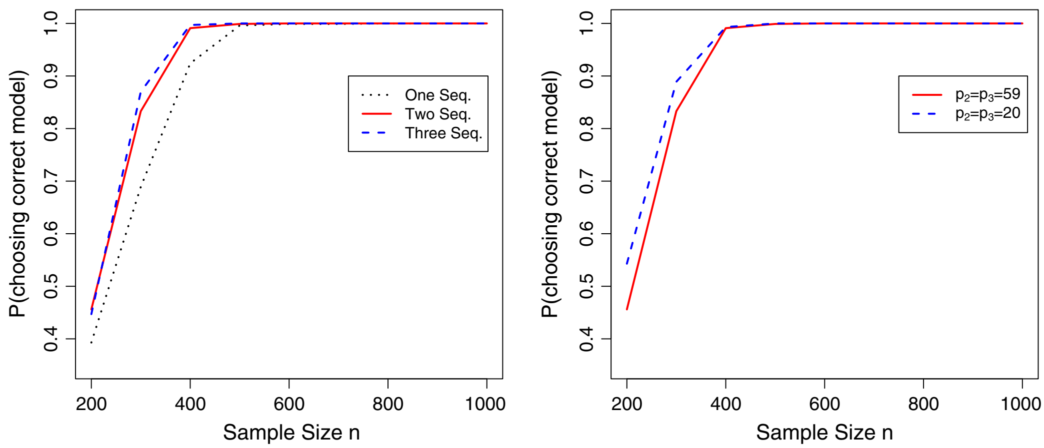

Figure 3. Left panel: $P\left(\hat{p}_1 = p_1\right)$ for the univariate, bivariate, and trivariate CV methods. Right panel: $P(\hat{p}_1 = p_1)$ for the trivariate CV method when $p_2 = p_3 = 59$ and $p_2 = p_3 = 20$.

Now, suppose $p_1 = p_2 = 43, \rho = -0.8$, and $\sigma_1 = 1$. We consider different error levels of the second sequence. Figure 2 shows the probabilities of choosing $p_1 = 43$ for the different error levels $\sigma_2 = 0.5, 1, 1.5$. The convergence is slower for larger error levels, which is further evidence that how well p_2 is estimated plays a role in the estimation of p_1.

4.2. Trivariate case

Suppose we observe three periodic sequences from model (1). Let $\sigma_1 = \sigma_2 = \sigma_3 = 1$, $\rho_{12} = -0.8$, which is the same as the correlation in the bivariate simulation, $\rho_{13} = 0.6$ and $\rho_{23} = -0.5$. Initially, we let $p_1 = 43$, $p_2 = p_3 = 59$. The probabilities of choosing $p_1 = 43$ are shown in the left panel of Figure 3. With three sequences, the convergence is faster than either one or two sequences, which shows that one more correlated sequence further improves the estimation of p_1, although the improvement from one to two sequences is apparently larger than improvements from two to three. We also consider a situation with $p_2 = p_3 = 20$. The right panel of Figure 3 shows that the convergence is faster for $p_2 = p_3 = 20$ than for $p_2 = p_3 = 59$. This is for the same reason as in the two sequences case, that is, more cycles are available for a smaller period when n is fixed.

4.3. AIC for bivariate case

For the bivariate case, the AIC in (5) becomes

$$\text{AIC}(q_1, q_2) = \frac{1}{1-\rho^2}\left\{\frac{1}{\sigma_1^2}\sum_{t=1}^{n}(X_t - \hat{\mu}_{1,t})^2 + \frac{1}{\sigma_2^2}\sum_{t=1}^{n}(Z_t - \hat{\mu}_{2,t})^2 - \frac{2\rho}{\sigma_1\sigma_2}\sum_{t=1}^{n}(X_t - \hat{\mu}_{1,t})(Z_t - \hat{\mu}_{2,t})\right\} + 2(q_1 + q_2),$$

where q_1 and q_2 are the period candidates for p_1 and p_2, and we choose the period estimates as the minimizer of $\text{AIC}(q_1, q_2)$. Again, let $p_1 = 43, p_2 = 59, \sigma_1 = \sigma_2 = 1$, and $\rho = -0.8$. The set for q_1 and q_2 over which the $\text{AIC}(q_1, q_2)$ was searched was taken to be $\{12, 13, \ldots, 96\}$, and the number of replications of each setting is 1000. In the simulation study, we assume that the parameters σ_1, σ_2, and ρ are known, and the mean parameters were estimated by the closed form MLEs in the Appendix, which are a function of the observations, ρ, and the ratio of σ_1

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

Stat
The ISI's Journal for the Rapid Dissemination of Statistics Research

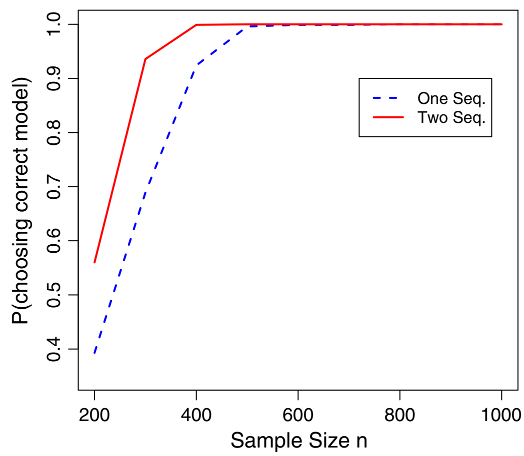

Figure 4. $P(\hat{p}_1 = p_1)$ for the bivariate case by the AIC method compared with the CV method for one sequence.

and σ_2. Then the criterion AIC(q_1, q_2) was minimized over the set $\{12, 13, \dots, 96\} \times \{12, 13, \dots, 96\}$. The probabilities of choosing $p_1 = 43$ are shown in Figure 4. Similarly, it is clear that the convergence is faster for the two-sequence case.

5 Applications

5.1. Ice core data

An ice core is a core sample removed from an ice sheet. It contains ice formed over a range of years, which enables the reconstruction of local temperature records. The Greenland Ice Sheet Project (GISP) was a decade-long project to drill ice cores in Greenland and has provided important insights into the climate variations and demonstrated the significance of the climatic record stored in ice sheets. It started in 1971 at Dye 3, where a 372-m deep core was recovered. After this, intermediate depth cores at various locations were drilled on the ice sheet. The first was a 398-m core at Milcent and another was a 405-m core at the Crete station in 1974. More information about GISP is provided by the National Oceanic and Atmospheric Administration (http://www.ncdc.noaa.gov/paleo/icecore/greenland/gisp/gisp.html).

A problem of great interest is to search for a signal of the El Niño effect from ice core series, as the ENSO is not expected in the North Atlantic region at all (North et al., 2012). However, North et al. (2012) also showed that the ENSO can be detected when several millennia of data are accumulated by analyzing one time series of 3600 yearly data from the Dye 3 station. In fact, such a long time series of ice core data is difficult to obtain; thus, instead of only using one sequence, we estimate the period of the ice core data from Milcent (Y_1) by borrowing information from the other two sequences from Dye 3 (Y_2) and Crete (Y_3) stations, respectively.

The annual data available for all three stations are from the years 1176–1872, 697 years in total. The three sequences are positively correlated, and the sample correlations are $\rho_{12} = 0.09$, $\rho_{13} = 0.28$, and $\rho_{23} = 0.20$. The univariate CV method proposed by Sun et al. (2012) yields a period estimate of $\hat{p}_1 = 14$ years for Milcent (see the left panel of Figure 5 for the CV curve), $\hat{p}_2 = 3$ years for Dye 3, and $\hat{p}_3 = 9$ years for Crete. However, the estimate $\hat{p}_1 = 14$ years has no clear interpretation in meteorology (North et al., 2012). Now, we add the other two sequences, Dye 3 and Crete, to estimate p_1. As seen in the CV curve in the right panel of Figure 5, the multivariate CV method gives

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

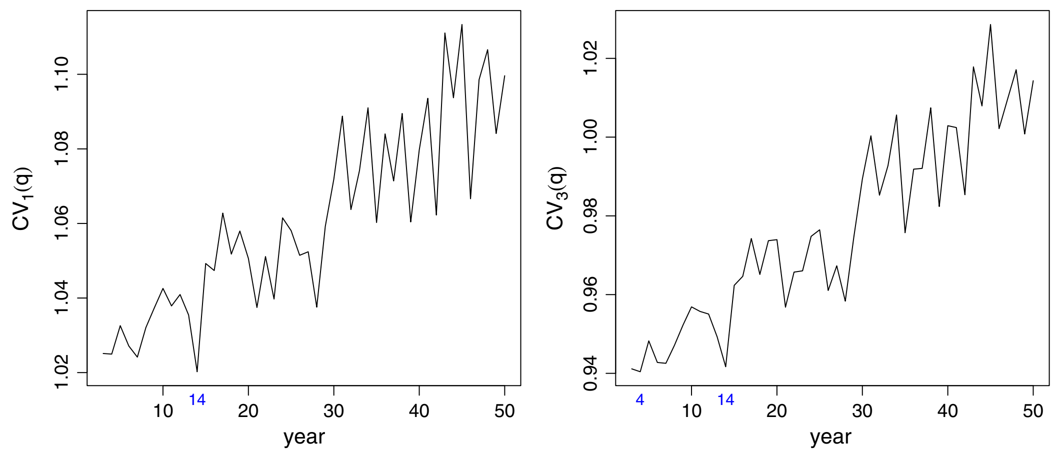

Figure 5. Left panel: the plot of the CV curve for estimating the period of the Milcent sequence. Right panel: the plot of the CV curve for estimating the period of the SST sequence when using Milcent, Dye 3, and Crete sequences.

$\hat{p}_1 = 4$ years, which is within the range of plausible El Niño periods; see Torrence & Webster (1999) and references therein. Even though the criterion $\text{CV}_3(4) = 0.9409$ is just slightly smaller than $\text{CV}_3(14) = 0.9417$, we are able to detect what appears to be a faint El Niño signal by borrowing information from nearby locations.

5.2. Physiological signals data

In the intensive care unit (ICU), monitors are used to provide information about patients in numerical and waveform formats, such as arterial blood pressure (ABP), electrocardiography (ECG), and fingertip plethysmograph (PLE). These physiological signals are often severely corrupted by noise, which makes the cardiac cycle estimation less accurate in the diagnosis and tracking of medical conditions. Thus, it results in a high incidence of false alarms from ICU monitors (Li et al., 2008).

The frequency of the cardiac cycle is described by the heart rate, which is the number of heartbeats per unit of time, typically defined as beats per minute. Therefore, the signals from ABP, ECG, and PLE related to heartbeats all show a certain periodicity. The behavior of blood pressures is an important indicator of human health. We are interested in studying its periodicity, in particular, identifying the period as well as a typical blood pressure cycle, which is not necessarily the cycle between two heartbeats.

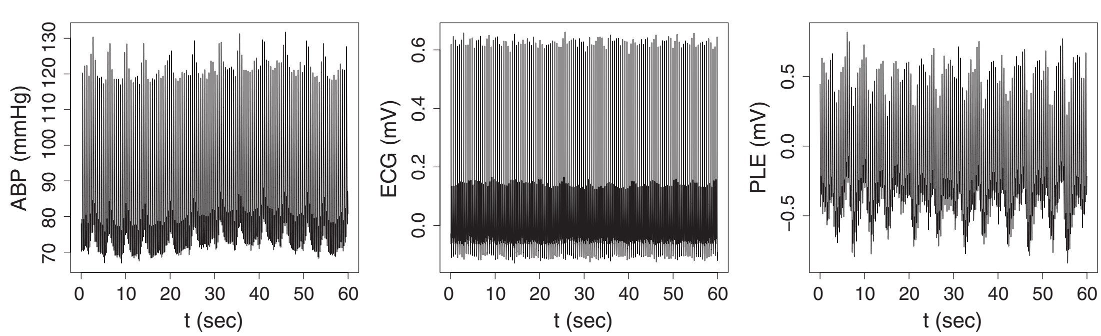

Figure 6. Time-series plots of the three signals: the arterial blood pressure (ABP) with unit mmHg, electrocardiography (ECG), and fingertip plethysmograph (PLE) with unit millivolt (mV).

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

The ISI's Journal for the Rapid
Dissemination of Statistics Research

We randomly select one patient from the Multiparameter Intelligent Monitoring in Intensive Care database (from http://www.physionet.org/physiobank/database/mimicdb/) with three signals: ABP (Y_1), ECG (Y_2), and PLE (Y_3). The description of this database can be found in Moody & Mark (1996) and Goldberger et al. (2000). The monitors take 500 samples per second, and thus measurements are available every 0.002 s. We analyze the record of the three signals of total length 1 min, that is, the sample size is $N = 30,000$. The three sequences are shown in Figure 6, and the sample correlations are $\rho_{12} = 0.1183$, $\rho_{13} = 0.3944$, and $\rho_{23} = 0.0001$.

The CV method for one sequence proposed by Sun et al. (2012) yields a period estimate of $\hat{p} = 475 \times 0.002 = 0.95$ s for all three sequences separately, which means that within 1 min, such a periodic event happens $60/0.95 = 63.15$ times. Suppose we only have shorter sequences with a sample size $n = 1500$. For the ABP sequence, in this case, the CV method gives a period estimate of $\hat{p}_1 = 708 \times 0.002 = 1.42$ s, that is, 42.37 cycles; see the left panel of Figure 7 for the CV curve. Notice that the CV curve is also locally minimized at 0.95 s.

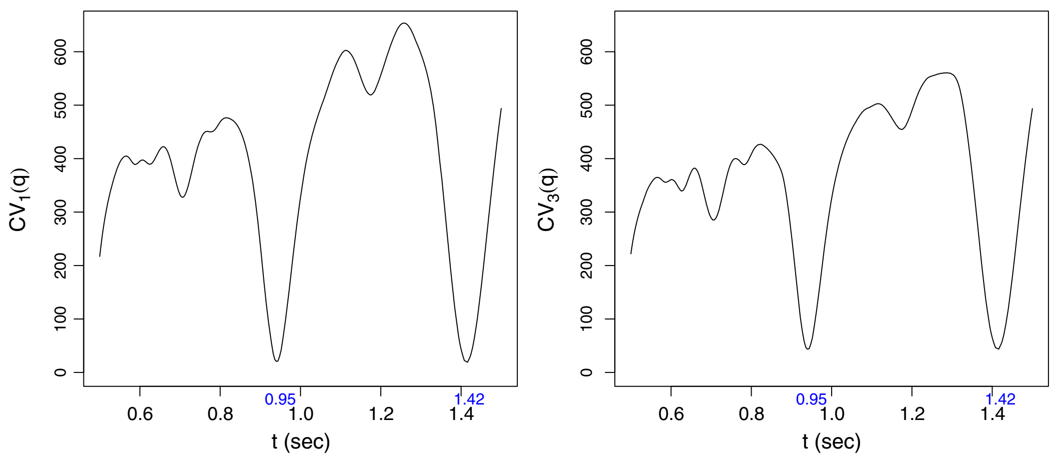

Figure 7. Left panel: the plot of the CV curve for estimating the period of blood pressures from ABP sequence only. Right panel: the plot of the CV curve for estimating the period of blood pressures using ABP, ECG, and PLE sequences.

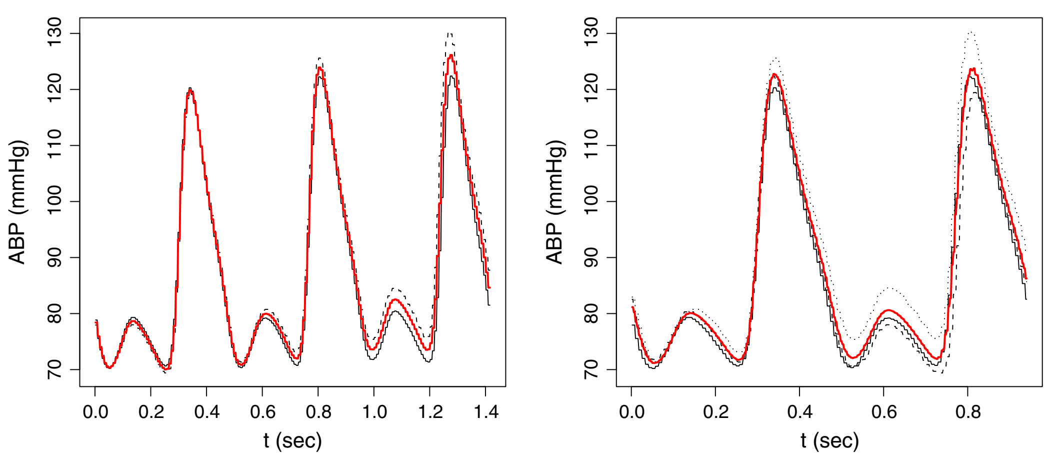

Figure 8. Left panel: a plot of the cycles of data corresponding to a period of 1.42 s superimposed on each other. Right panel: a plot of the cycles of data corresponding to a period of 0.95 s superimposed on each other. The estimated means are connected by the red lines.

The ISI's Journal for the Rapid
Dissemination of Statistics Research

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

The plot of the cycles of data superimposed on each other for $\hat{p}_1 = 1.42$ s is in the left panel of Figure 8, which contains three peaks. Interestingly, if only one of ECG or ELP is used with ABP, the period estimate does not change, but if we use both with ABP, $\hat{p}_1$ becomes $470 \times 0.002 = 0.94$ s. In other words, there are 63.83 cycles within 1 min, which is much closer to the estimate from the whole series of ABP; see the right panel of Figure 7 for the CV curve. A plot of the cycles of data superimposed on each other for $\hat{p}_1 = 0.95$ s is in the right panel of Figure 8, which includes two peaks. By improving the period estimation of blood pressures as well as the periodic mean estimation, the cardiac cycle between two maximum blood pressures can also be better characterized.

6 Discussion

In this paper, we have proposed a CV period estimator when two or more equally spaced periodic sequences are available. Using an idea similar to the CV method for one sequence, a leave-out-one-cycle version of CV is used to compute an average squared prediction error given a particular period, or cycle length. The multivariate CV method uses the conditional means, that is, conditional on other correlated sequences, to predict the left-out cycle in CV. In this way, the period estimation for a sequence X has been improved by borrowing information from other sequences with which X is correlated. In theory, we argue heuristically that the asymptotic behavior of the bivariate CV is the same as that of the CV for one sequence. In our simulation studies, however, it was shown that for finite samples, the better the periods of the other correlated sequences were estimated, the more substantial improvements could be obtained in estimating the period of interest. We have also shown that more correlated sequences lead to more improvements, although the improvement from one to two sequences is apparently larger than improvements from two to three, and so on. In addition to the CV method, we also considered a model selection criterion, AIC, to estimate the period for multiple periodic sequences. Similarly, for finite samples, a simulation study showed an improvement in period estimation from using information in a correlated sequence. The asymptotic properties of the AIC method need further exploration.

Appendix A

A.1. Deriving MLEs of sequence of means

For the bivariate case, suppose we observe bivariate sequences from the following model:

$$\begin{cases} X_t = \mu_{1,t} + \varepsilon_{1,t}, \\ Z_t = \mu_{2,t} + \varepsilon_{2,t}, \end{cases}$$

where $\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}_t \sim$ i.i.d. $N_2(\mathbf{0}, \mathbf{\Sigma})$, and $\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. The likelihood is of the form

$$L = \frac{1}{1-\rho^2}\left\{\frac{1}{\sigma_1^2}\sum_{t=1}^{n}(X_t - \mu_{1,t})^2 + \frac{1}{\sigma_2^2}\sum_{t=1}^{n}(Z_t - \mu_{2,t})^2 - \frac{2\rho}{\sigma_1\sigma_2}\sum_{t=1}^{n}(X_t - \mu_{1,t})(Z_t - \mu_{2,t})\right\}. \quad \text{(A.1)}$$

Let q_1 and q_2 be the period candidates for X_t and Z_t, respectively, and $m_i = \mu_{2,i+q_2} = \mu_{2,i+2q_2} = \cdots$ for $i = 1, \ldots, q_2$, $\eta_i = \mu_{1,i+q_1} = \mu_{1,i+2q_1} = \cdots$ for $i = 1, \ldots, q_1$. First, minimize

$$\frac{1}{\sigma_2^2}\sum_{t=1}^{n}(Z_t - \mu_{2,t})^2 - \frac{2\rho}{\sigma_1\sigma_2}\sum_{t=1}^{n}(X_t - \mu_{1,t})(Z_t - \mu_{2,t}) \quad \text{(A.2)}$$

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

The ISI's Journal for the Rapid Dissemination of Statistics Research

with respect to $m_1, \ldots, m_{q_2}$. Define $Z_{q_2ij} = Z_{i+(j-1)q_2}$, where $i = 1, \ldots, q_2$ and k_{q_2i} is the largest integer such that $i + q_2k_{q_2,i} \leqslant n$. Then (A.2) can be written as

$$A = \frac{1}{\sigma_2^2}\sum_{i=1}^{q_2}\sum_{j=1}^{k_{q_2i}}(Z_{q_2ij} - m_i)^2 - \frac{2\rho}{\sigma_1\sigma_2}\sum_{i=1}^{q_2}\sum_{j=1}^{k_{q_2i}}(X_{q_2ij} - \mu_{1,q_2ij})(Z_{q_2ij} - m_i).$$

The first derivative

$$\frac{\partial A}{\partial m_i} = -\frac{2}{\sigma_2^2}\sum_{j=1}^{k_{q_2i}}(Z_{q_2ij} - m_i) + \frac{2\rho}{\sigma_1\sigma_2}\sum_{j=1}^{k_{q_2i}}(X_{q_2ij} - \mu_{1,q_2ij}) = 0$$

gives

$$m_i = \bar{Z}_{q_2i} - \rho\frac{\sigma_2}{\sigma_1}(\bar{X}_{q_2i} - \bar{\mu}_{1,q_2i}), \tag{A.3}$$

where $\bar{Z}_{q_2i}$ is the average of $Z_{q_2i\ell}$, $\ell = 1, \ldots, k_{q_2i}$. Then plug (A.3) into Equation (A.1),

$$\begin{aligned} L_1 =& \frac{1}{\sigma_1^2}\sum_{i=1}^{q_1}\sum_{j=1}^{k_{q_1i}}(X_{q_1ij} - \eta_i)^2 + \frac{1}{\sigma_2^2}\sum_{i=1}^{q_2}\sum_{j=1}^{k_{q_2i}}(Z_{q_2ij} - \bar{Z}_{q_2i})^2 \\ &- \frac{\rho^2}{\sigma_1^2}\sum_{i=1}^{q_2}k_{q_2i}(\bar{X}_{q_2i} - \bar{\mu}_{1,q_2i})^2 - \frac{2\rho}{\sigma_1\sigma_2}\sum_{i=1}^{q_2}\sum_{j=1}^{k_{q_2i}}(X_{q_2ij} - \mu_{1,q_2ij})(Z_{q_2ij} - \bar{Z}_{q_2i}). \end{aligned}$$

We then have $\partial\mu_{1,q_2ij}/\partial\eta_r = 1$, if $i + (j-1)q_2 = r + (\ell - 1)q_1$ for some ℓ, otherwise 0, and $\partial\bar{\mu}_{1,q_2i}/\partial\eta_r = N_{ir}/k_{q_2i}$, where N_{ir} is the number of times $i + (j-1)q_2 = r + (\ell-1)q_1$ when j ranges from 1 to k_{q_2i} and ℓ ranges from 1 to k_{q_1r}. Then

$$\frac{\partial L_1}{\partial\eta_r} = -\frac{2}{\sigma_1^2}k_{q_1r}(\bar{X}_{q_1r} - \eta_r) + \frac{2\rho^2}{\sigma_1^2}\left(\sum_{i=1}^{q_2}\bar{X}_{q_1i}N_{ir} - \sum_{i=1}^{q_2}\frac{1}{k_{q_2i}}\sum_{k=1}^{q_1}N_{ir}N_{ik}\eta_k\right) + \frac{2\rho}{\sigma_1\sigma_2}S_{q_1,q_2,r},$$

where $S_{q_1,q_2,r} = \sum_{i=1}^{k_{q_1r}}(Y_{q_1rj} - \bar{Y}^{q_2}_{q_1rj})$. We can see that $\partial L_1/\partial\eta_r = 0$ if and only if

$$k_{q_1r}(\bar{X}_{q_1r} - \eta_r) + \rho^2\sum_{i=1}^{q_2}\frac{1}{k_{q_2i}}\sum_{k=1}^{q_1}N_{ir}N_{ik}\eta_k = \rho^2\sum_{i=1}^{q_2}\bar{X}_{q_2i}N_{ir} + \rho\frac{\sigma_1}{\sigma_2}S_{q_1,q_2,r}.$$

This yields a set of q_1 linear equations. Arrange them so that we can solve a smaller system, that is, $q_1 < q_2$. The coefficient of η_r in the r-th equation is

$$-k_{q_1r} + \rho^2\sum_{i=1}^{q_2}\frac{1}{k_{q_2i}}N_{ir}^2,$$

and the coefficient of η_k ($k \neq r$) is

$$\rho^2\sum_{i=1}^{q_2}\frac{1}{k_{q_2i}}N_{ir}N_{ik}.$$

The equations are $\mathbf{B}_{q_1,q_2}\boldsymbol{\eta} = \mathbf{b}_{q_1,q_2}$, where $\mathbf{B}_{q_1,q_2}(j,k) = \rho^2\sum_{i=1}^{q_2}\frac{1}{k_{q_2i}}N_{ij}N_{ik} - k_{q_1j}I(j-k)$ for $I(0) = 1$, 0 otherwise, and

$$\mathbf{b}_{q_1,q_2}(j) = \rho^2\sum_{i=1}^{q_2}\bar{X}_{q_2i}N_{ij} + \rho\frac{\sigma_1}{\sigma_2}S_{q_1,q_2,j} - k_{q_1,j}\bar{X}_{q_1j}.$$

When $q_1 = q_2$, the MLEs of $\eta_1, \ldots, \eta_{q_1}$ and $m_1, \ldots, m_{q_2}$ are the usual stacked means. □

The ISI's Journal for the Rapid Dissemination of Statistics Research

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

A.2. Proof of consistency of mean estimator for one-series method

Consider observations $Y_1, \ldots, Y_n$ from $Y_t = \mu_t + \varepsilon_t$, $t = 1, \ldots, n$, where $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. $N(0, \sigma^2)$, and the sequence of means $\mu_1, \mu_2, \ldots$ is periodic with (smallest) period p_0. Let $\hat{\boldsymbol{\mu}}_{q,n}$ be the vector of stacked mean estimates of $\boldsymbol{\mu}_n = (\mu_1, \ldots, \mu_n)$ that correspond to assuming that the period is q. For example, if $n = 10$ and $q = 3$, $\hat{\boldsymbol{\mu}}_{q,n}$ has the form $\left(\hat{\mu}_1^3, \hat{\mu}_2^3, \hat{\mu}_3^3, \hat{\mu}_1^3, \hat{\mu}_2^3, \hat{\mu}_3^3, \hat{\mu}_1^3, \hat{\mu}_2^3, \hat{\mu}_3^3, \hat{\mu}_1^3\right)$. Let $\hat{p}$ be the Sun et al. (2012) (SHG) estimator of p_0, and, for simplicity of notation, let $\hat{\boldsymbol{\mu}}_n = (\hat{\mu}_1, \ldots, \hat{\mu}_n) = \hat{\boldsymbol{\mu}}_{\hat{p},n}$. Under the conditions of SHG, we will show that $\hat{\boldsymbol{\mu}}_n$ is $\sqrt{n}$-consistent for $\boldsymbol{\mu}_n$ in the sense that

$$\delta_n = \max_{1 \leqslant i \leqslant n} |\hat{\mu}_i - \mu_i| = O_p\left(n^{-1/2}\right).$$

Two facts are key in proving consistency of $\hat{\boldsymbol{\mu}}_n$. First, although $\hat{p}$ is not a consistent estimator of p_0, SHG show that, asymptotically, the support of $\hat{p}$ is $p_0, 2p_0, \ldots$. The second key is that, for any j, $\hat{\mu}_k^{jp_0}$ is $\sqrt{n}$-consistent for μ_k, $k = 1, \ldots, jp_0$, because each $\hat{\mu}_k^{jp_0}$ is a sample mean of approximately $n/(jp_0)$ i.i.d. random variables. So with probability tending to 1, $\hat{\boldsymbol{\mu}}_n$ is one of $\hat{\boldsymbol{\mu}}_{p_0,n}, \hat{\boldsymbol{\mu}}_{2p_0,n}, \ldots$, any of which is $\sqrt{n}$-consistent.

We need to show that $\sqrt{n}\delta_n$ is bounded in probability. This means that for each $\varepsilon > 0$, we must find M_ε and n_ε such that $P(\sqrt{n}\delta_n \geqslant M_\varepsilon) \leqslant \varepsilon$ for all $n > n_\varepsilon$. Let E_n be the event that $\hat{p}$ is in the set $\{p_0, 2p_0, 3p_0, \ldots\}$. We have

$$P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon\right) = P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap E_n\right) + P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap E_n^c\right) \leqslant P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap E_n\right) + P\left(E_n^c\right). \tag{A.4}$$

SHG show that $P(E_n^c) \to 0$ as $n \to \infty$, and so obviously there exists $n_{1\varepsilon}$ such that $P(E_n^c) \leqslant \varepsilon/2$ for all $n > n_{1\varepsilon}$.

SHG minimize the CV criterion over a set $\{1, \ldots, M_n\}$, where $M_n = o(\sqrt{n})$. Let m be an integer such that $1 \leqslant mp_0 < M_n$. For m_n the largest integer such that $m_n p_0 \leqslant M_n$, we have

$$\begin{aligned} P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap E_n\right) &= P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap \hat{p} \in \{p_0, 2p_0, \ldots, mp_0\}\right) + P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap \hat{p} \in \{(m+1)p_0, \ldots, m_n p_0\}\right) \\ &\leqslant P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap \hat{p} \in \{p_0, 2p_0, \ldots, mp_0\}\right) + P(\hat{p} \in \{(m+1)p_0, \ldots, m_n p_0\}). \end{aligned} \tag{A.5}$$

The theorem of SHG entails that $P(\hat{p} \in \{(m+1)p_0, \ldots, m_n p_0\})$ can be made arbitrarily small by taking m sufficently large. Therefore, there exist m_ε and $n_{2\varepsilon}$ such that for all $n > n_{2\varepsilon}$

$$P(\hat{p} \in \{(m_\varepsilon+1)p_0, \ldots, m_n p_0\}) \leqslant \frac{\varepsilon}{4}. \tag{A.6}$$

Turning to the other probability on the right-hand side of (A.5),

$$\begin{aligned} P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap \hat{p} \in \{p_0, 2p_0, \ldots, m_\varepsilon p_0\}\right) &= \sum_{j=1}^{m_\varepsilon} P\left(\sqrt{n}\delta_n \geqslant M_\varepsilon \cap \hat{p} = jp_0\right) \leqslant \sum_{j=1}^{m_\varepsilon} P\left(\sqrt{n} \max_{1 \leqslant k \leqslant jp_0} |\hat{\mu}_k^{jp_0} - \mu_k| \geqslant M_\varepsilon\right) \\ &\leqslant \sum_{j=1}^{m_\varepsilon}\sum_{k=1}^{jp_0} P\left(\sqrt{n}|\hat{\mu}_k^{jp_0} - \mu_k| \geqslant M_\varepsilon\right) \leqslant \sum_{j=1}^{m_\varepsilon}\sum_{k=1}^{jp_0} \frac{n}{M_\varepsilon^2} \frac{\sigma^2}{[n/(jp_0)]}, \end{aligned} \tag{A.7}$$

where $[x]$ denotes the largest integer not greater than x. The last inequality simply makes use of Markov's inequality and the fact that $\hat{\mu}_k^{jp_0}$ is a sample mean of at least $[n/(jp_0)]$ i.i.d. random variables.

It is straightforward to check that the right-hand side of (A.7) is bounded by $n\sigma^2 p_0^2 m_\varepsilon^3/[(n - m_\varepsilon p_0)M_\varepsilon^2]$. Obviously, there exists $n_{3\varepsilon}$ such that for all $n > n_{3\varepsilon}$, $n/(n - m_\varepsilon p_0) < 3/2$. Therefore, we need $(3/2)\sigma^2 p_0^2 m_\varepsilon^3/M_\varepsilon^2 \leqslant \varepsilon/4$, or

$$M_\varepsilon \geqslant \frac{\sqrt{6}\sigma p_0 m_\varepsilon^{3/2}}{\sqrt{\varepsilon}}. \tag{A.8}$$

(wileyonlinelibrary.com) DOI: 10.1002/sta4.28

The ISI's Journal for the Rapid Dissemination of Statistics Research

Once m_ε is determined, we can obviously choose M_ε to satisfy (A.8), and so, putting together (A.4)–(A.7), for all $n > \max(n_{1\varepsilon}, n_{2\varepsilon}, n_{3\varepsilon})$ and M_ε satisfying (A.8), $P(\sqrt{n}\delta_n \geqslant M_\varepsilon) \leqslant \varepsilon$, which completes the proof. □

Acknowledgement

The work of Professor Hart was supported by NSF grant DMS 1106694.

References

Akaike, H (1973), 'Information theory and an extension of the maximum likelihood principle', in Petrov, BN & Csake, F (eds), *Second International Symposium on Information Theory*, Akademiai Kiado, Budapest, 267–281.

Canova, F (1996), 'Three tests for the existence of cycles in time series', *Ricerche Economiche*, **50**, 135–162.

Genton, MG & Hall, P (2007), 'Statistical inference for evolving periodic functions', *Journal of the Royal Statistical Society Series B*, **69**, 643–657.

Goldberger, AL, Amaral, L, Glass, L, Hausdorff, JM, Ivanov, P, Mark, RG, Mietus, JE, Moody, GB, Peng, C & Stanley, HE (2000), 'PhysioBank, physioToolkit, and physioNet: components of a new research resource for complex physiologic signals', *Circulation*, **101**, 215–220.

Hall, P (2008), 'Nonparametric methods for estimating periodic functions, with applications in astronomy', in Brito, P (ed.), *COMPSTAT 2008: Proceedings in Computational Statistics*, Physika-Verlag, Springer, Heidelberg, 3–18.

Hall, P & Li, M (2006), 'Using the periodogram to estimate period in nonparametric regression', *Biometrika*, **93**, 411–424.

Hall, P, Reimann, J & Rice, J (2000), 'Nonparametric estimation of a periodic function', *Biometrika*, **87**, 545–557.

Hall, P & Yin, J (2003), 'Nonparametric methods for deconvolving multiperiodic functions', *Journal of the Royal Statistical Society Series B*, **65**, 869–886.

Li, T (2012), 'Detection and estimation of hidden periodicity in asymmetric noise by using quantile periodogram', in *Proceedings of 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IBM TJ Watson Res. Center, Yorktown Heights, NY, USA, 3969–3972.

Li, Q, Mark, RG & Clifford, GD (2008), 'Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter', *Physiological Measurement*, **29**, 15–32.

Moody, GB & Mark, RG (1996), 'A database to support development and evaluation of intelligent intensive care monitoring', *Computers in Cardiology*, **23**, 657–660.

North, GR, Genton, MG & Chylek, P (2012), 'Using ice core data to constrain climate sensitivity', Manuscript.

Quinn, BG & Thomson, PJ (1991), 'Estimating the frequency of a periodic function', *Biometrika*, **78**, 65–67.

Rice, JA & Rosenblatt, M (1988), 'On frequency estimation', *Biometrika*, **75**, 477–484.

Sun, Y, Hart, JD & Genton, MG (2012), 'Nonparametric inference for periodic sequences', *Technometrics*, **54**, 83–96.

Torrence, C & Webster, P (1999), 'Interdecadal changes in the ENOS–Monsoon system', *Journal of Climate*, **12**, 2679–2690.

Walker, AM (1971), 'On the estimation of a harmonic component in a time series with stationary independent residuals', *Biometrika*, **58**, 21–36.