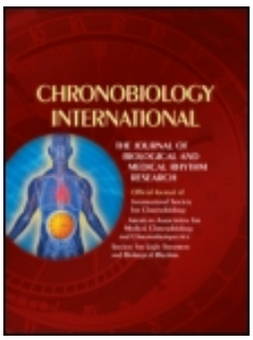


Taylor & Francis
Taylor & Francis Group

Chronobiology International

The Journal of Biological and Medical Rhythm Research

ISSN: 0742-0528 (Print) 1525-6073 (Online) Journal homepage: https://www.tandfonline.com/loi/icbi20

Inferential Statistical Method for Analysis of Nonsinusoidal Hybrid Time Series with Unequidistant Observations

Jose R. Fernández & RamÓN C. Hermida

To cite this article: Jose R. Fernández & RamÓN C. Hermida (1998) Inferential Statistical Method for Analysis of Nonsinusoidal Hybrid Time Series with Unequidistant Observations, Chronobiology International, 15:2, 191-204, DOI: 10.3109/07420529808998683

To link to this article: https://doi.org/10.3109/07420529808998683

Published online: 07 Jul 2009.

CHRONOBIOLOGY INTERNATIONAL, 15(2), 191–204 (1998)

INFERENTIAL STATISTICAL METHOD FOR ANALYSIS OF NONSINUSOIDAL HYBRID TIME SERIES WITH UNEQUIDISTANT OBSERVATIONS

Jose R. Fernández and Ramón C. Hermida
Bioengineering and Chronobiology Laboratories, ETSI de Telecomunicación, University of Vigo, Campus Universitario, Vigo, Spain

ABSTRACT

Most variables of interest in laboratory medicine show predictable changes with several frequencies in the span of time investigated. The waveform of such nonsinusoidal rhythms can be well described by the use of multiple components rhythmometry, a method that allows fitting a linear model with several cosine functions. The method, originally described for analysis of longitudinal time series, is here extended to allow analysis of hybrid data (time series sampled from a group of subjects, each represented by an individual series). Given k individual series, we can fit the same linear model with m different frequencies (harmonics or not from one fundamental period) to each series. This fit will provide estimations for $2m + 1$ parameters, namely, the amplitude and acrophase of each component, as well as the rhythm-adjusted mean. Assuming that the set of parameters obtained for each individual is a random sample from a multivariate normal population, the corresponding population parameter estimates can be based on the means of estimates obtained from individuals in the sample. Their confidence intervals depend on the variability among individual parameter estimates. The variance-covariance matrix can then be estimated on the basis of the sample covariances. Confidence intervals for the rhythm-adjusted mean, as well as for the amplitude-acrophase pair, of each component can then be computed using the estimated covariance matrix. The p-values for testing the zero-amplitude assumption for each component, as well as for the global model, can finally be derived using those confidence intervals and the t and F distri-

Received May 5, 1997; returned for revision September 12, 1997; accepted October 28, 1997.

Address correspondence to: Prof. José Ramón Fernández, Ph.D., Bioengineering and Chronobiology Laboratories, ETSI Telecomunicación, Campus Universitario, Vigo (Pontevedra) 36200, Spain.

butions. The method, validated by a simulation study and illustrated by an example of modeling the circadian variation of heart rate, represents a new step in the development of statistical procedures in chronobiology. (*Chronobiology International, 15(2), 191–204, 1988*)

Key Words: Multiple components analysis—Hybrid data—Population-mean cosinor—Periodic regression—Multifrequency modeling.

INTRODUCTION

Methods of linear least squares estimation have been designed for the detection of periodic components in sparse and noisy time series (as they are usually present in clinical situations involving patients). The approach is based on regression techniques, and as such, it is applicable to the analysis of unequidistant observations (1–3). Methods of least squares estimation presently in use include the single and the population-mean cosinor (1,3–7).

The single cosinor is a method applicable to longitudinal (single) biologic time series anticipated to be rhythmic with a given period (1,2,4,7). This procedure amounts to fitting to the data by least squares a cosine function of a fixed anticipated period. Thus, one obtains, for the period considered, an estimate of (1) the rhythm-adjusted mean or MESOR (midline estimating statistic of rhythm), defined as the average value of the rhythmic function (e.g., cosine curve) fitted to the data; (2) amplitude, defined as half the extent of rhythmic change in a cycle approximated by the fitted cosine curve (difference between the maximum and the MESOR of the fitted curve); and (3) acrophase, the lag from a defined reference time point (e.g., local midnight of the first day of measurement when the fitted period is 24h) of the crest time in the cosine curve fitted to the data (1,2,4,7). The goodness of fit is indicated by minimizing the sum of squares of the residuals from the analysis, that is, the differences between the actual measurements and the estimated functional form or best-fitting curve. Rhythm detection is sought by testing the null hypothesis of zero amplitude with an F-test (1). Provided that the residuals around the fitted curve are normally distributed, it can be shown (1) that estimates of the Cartesian projections of amplitude and acrophase follow a bivariate normal distribution. Conservative confidence intervals can thus be obtained for amplitude and acrophase separately (derived from a joint confidence ellipse). Nonconservative confidence intervals can also be computed for the amplitude and acrophase, as well as for the MESOR, by the use of their standard errors (1,8).

The single-cosinor method can be extended to allow inferences concerning a population. In order to summarize results obtained for different individuals belonging to a population (hybrid database, that is, data collected from a group of subjects, each represented by an individual time series), the rhythm parameters obtained for each individual by the single-cosinor procedure may serve as input for a population-mean cosinor for further quantification (1–3,5). The rhythm characteristics obtained by the single cosinor are then considered as imputations or first-order statistics. The population-mean cosinor, in turn, constitutes a second-order statistic, applied to derive confidence intervals for rhythm parameters pertaining to the whole population. The parameter estimates are based on the means of estimates obtained from individuals in the sample, and their confidence intervals depend on the variability among individual parameter estimates (1). The rejec-

tion of the zero-amplitude assumption by single cosinor ("rhythm detection") refers to the given data set and does not allow extrapolation to a population as a whole. As the name implies, the population-mean cosinor aims at extrapolation beyond the given sample to the population as a whole. If the sample is to be characterized without further inference to others in the population, a single cosinor is much more efficient; however, when inference is to be drawn on the basis of the sample for the population as a whole, the population-mean cosinor is indicated (1–3,5,9).

These methods for linear least squares modeling are only applicable within very restrictive conditions (1,3). Both the single (longitudinal data) and the population-mean (hybrid data) cosinor methods require that the data obtained can be reasonably well represented by a cosine curve, and nonsinusoidality limits the applicability of the method. For a meaningful rhythmometric analysis by cosinor, it is important, therefore, to determine the approximate sinusoidality of the data. This requires at least the inspection of the chronogram before application of the rhythmometric procedure and/or a mathematical test for sinusoidality. The lack of sinusoidality can be easily detected if the maximum value of the original signal occurs at an appreciably different time than that indicated by the acrophase and its confidence limits (10). Tests for normality and independence of residuals, as well as homogeneity of their variance, are also indicated. All these tests are usually incorporated into the computer program containing the cosinor (8,11). When one or more of these tests yields p-values less than a specified confidence level (usually .05), transformations or alternative models are to be considered (1,2). Users frequently do not take into consideration these restrictive conditions of the cosinor methods, which leads to numerous misuses (12).

In cases for which more than one period are statistically significant over the span of time investigated or when the waveform is nonsinusoidal, the use of a multiple components analysis to fit a model consisting of several cosine functions (harmonics or not from a fundamental period) is recommended (1,5,13). The method, however, has only been described for the analysis of longitudinal data (single time series). As explained above, one can go from the single to the population-mean cosinor to allow inferences for a group of subjects from fitting a linear model with only one period. We here followed a parallel line of thought in the development of a new inferential statistical method for analysis of nonsinusoidal hybrid time series with unequidistant observations, the so-called population multiple components method (14).

MULTIPLE COMPONENTS ANALYSIS

Analysis of Longitudinal Data

The multiple components procedure for analysis of longitudinal time series amounts to fitting to the data by least squares a function with several (m) fixed anticipated periods:

$$y_h = M + \sum_{j=1}^{m} A_j \cos(\omega_j t_h + \phi_j) + e_h \qquad h = 1, \ldots, N \tag{1}$$

where y_h is the observed value at time t_h; ω_j are the angular frequencies, that is, $2\pi/\tau_j$, where τ_j are the fitted periods; e_h is the residual from the analysis for the value y_h; and N is the number of observed values. Assuming $N > 2m + 1$ and that the residuals are independent, with zero mean and common variance, the linear least squares resolution

of this equation provides, for each fitted period, point and interval estimates of the amplitude (A_j in the equation) and the acrophase (ϕ_j in the equation) (1,2). Given each of the m periods, the model is fitted in its equivalent linear form after substituting A_j and ϕ_j by their Cartesian projections $\beta_j = A_j\cos\phi_j$ and $\gamma_j = -A_j\sin\phi_j$:

$$y_h = M + \sum_{j=1}^{m} \beta_j\cos(\omega_j t_h) + \sum_{j=1}^{m} \gamma_j\sin(\omega_j t_h) + e_h \qquad h = 1, \ldots, N \tag{2}$$

Rhythm detection is sought for each individual component, as well as for the overall fit of multiple components by testing the null hypothesis of zero amplitude with respective F-tests. The method also provides, for the overall fit of multiple components, point and interval estimates of the MESOR (M in the equation).

If all m frequencies in Eq. 2 are harmonics of a fundamental frequency, that is, if all fitted angular frequencies ω_j are integer multiples of ω_1, then the fitted curve is periodic with period $\tau_1 = 2\pi/\omega_1$, where τ_1 is the fundamental period. In such a case, the method of multiple components also provides three additional parameters not directly estimated from Eq. 2: the overall amplitude, defined as half the difference between the maximum and the minimum of the best-fitted curve in one fundamental period; the orthophase, defined as the lag from a defined reference time point of the crest time, within a fundamental period, in the curve of multiple components fitted to the data (10,13); and the bathyphase, defined as the lag from a defined reference time point of the time of the lowest value, within a fundamental period, in the curve of multiple components fitted to the data. The orthophase (similar to the acrophase of each component) and the bathyphase are usually given in negative angular degrees, with 360° = 1 fundamental period. In the case of a single-cosinor analysis, the orthophase is equivalent to the acrophase of the unique component. In the general nonsinusoidal case given by Eq. 2, the orthophase depends on all the amplitudes and acrophases. The orthophase is then calculated as the maximizing time point of Eq. 2. The computation of overall amplitude, orthophase, and bathyphase requires, then, that all periods in the curve fitted to the data must be harmonics of a fundamental period included within the time span investigated.

Analysis of Hybrid Data

Given k individual series from different subjects belonging to a population (hybrid data), we can fit the same model of multiple components given in Eq. 1 to each individual series using the same m anticipated periods. In so doing, each individual series is assumed to follow a multiple components curve with parameters given by Eq. 2, with the m periods known and fixed for all series in the sample. As a result, we obtain $2m + 1$ parameter estimates for each individual series, the MESOR, as well as the amplitude and acrophase of each component (or, equivalently, their Cartesian projections β_j and γ_j). The goal of the analysis is to make inferences concerning the population average (expected values) for each parameter in Eq. 2. This inference can be most simply done using a "second-order" analysis, for which confidence limits and statistical tests are based only on the variability among individual parameter estimates. This procedure is similarly used for deriving population parameters in the population-mean cosinor method (1).

In order to be able to develop a usable technique, we must assume that the individual parameter vectors $[M_i,\beta_{i1},\gamma_{i1}, \ldots, \beta_{im},\gamma_{im}]'$, $i = 1, \ldots, k$ are a random sample from a

multivariate normal population with mean $[M,\beta_1\gamma_1, \ldots, \beta_m,\gamma_m]'$ and covariance matrix Σ. The estimations for the population parameters are then given by

$$\hat{M} = \frac{1}{k}\Sigma_i M_i \qquad \hat{\beta} = \frac{1}{k}\Sigma_i \beta_i \qquad \hat{\gamma} = \frac{1}{k}\Sigma_i \gamma_i \tag{3}$$

The population estimates for the amplitude and acrophase of each component can be obtained by transforming the Cartesian coordinates β and γ given by Eq. 3 to polar coordinates.

The computation of confidence intervals is based on the variance-covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma_M^2 & \sigma_{M\beta_1} & \sigma_{M\gamma_1} & \cdots & \sigma_{M\gamma_p} \\ & & \cdots & \cdots & \\ & & & & \sigma_{\gamma_p}^2 \end{pmatrix} \tag{4}$$

with elements that can be estimated by

$$\hat{\sigma}_M^2 = \frac{1}{k-1}\Sigma_i (M_i - \hat{M})^2 \qquad \hat{\sigma}_{M\beta_1} = \frac{1}{k-1}\Sigma_i (M_i - \hat{M})(\beta_{i1} - \hat{\beta}_1) \qquad \text{etc.} \tag{5}$$

Then, a confidence interval for the MESOR is given by

$$\hat{M} \pm t_{1-\alpha/2}(k-1)\hat{\sigma}_M/\sqrt{k} \tag{6}$$

where $t_{1-\alpha/2}(k-1)$ is the $1-\alpha/2$ percentile of the t distribution with $(k-1)$ degrees of freedom.

It has been shown that (15)

$$\frac{k(k-2)}{2(k-1)}(\hat{\beta}_j - \beta_j, \hat{\gamma}_j - \gamma_j)(\hat{\Sigma}_j)^{-1}\begin{pmatrix} \hat{\beta}_j - \beta_j \\ \hat{\gamma}_j - \gamma_j \end{pmatrix} \tag{7}$$

follows an $F(2,k-2)$ distribution, where Σ_j is the submatrix of Σ corresponding to the jth component. This result can be used to determine a joint confidence region for the amplitude-acrophase pair of each component (1) and, therefore, to compute a p-value for testing the zero-amplitude assumption for that component.

A test for the null hypothesis that all amplitudes are simultaneously zero (zero global amplitude test) can be performed using (15)

$$\frac{k(k-2m)}{2m(k-1)}(\hat{\vec{\beta}} - \vec{\beta})'(\hat{\Sigma})^{-1}(\hat{\vec{\beta}} - \vec{\beta}) \tag{8}$$

distributed as an $F(2m,k-2m)$, where $\vec{\beta}$ is the vector of β's and γ's. This result implies that the minimum sample size in terms of number of individuals providing information from the population needed for a population multiple components analysis must be $k > 2m$, where m is the number of components in the model.

General Linear Model

All the results given above for analysis of either longitudinal or hybrid data can be easily extrapolated to the case of fitting a more general linear model consisting of several

cosine functions (as given by Eq. 1) and polynomial terms representing trends in the data. In such a case, Eq. 1 will become

$$y_i = M + \sum_{q=1}^{n} C_q t_i^q + \sum_{j=1}^{m} A_j \cos(\omega_j t_i + \phi_j) + e_i \qquad i = 1, \ldots, N \tag{9}$$

where C_q are the coefficients of the polynomial terms. The linear least squares resolution of this equation for any given longitudinal time series will provide estimates for $(2m + n + 1)$ parameters, that is, the mean level M, the n coefficients of the polynomial, and the amplitude and acrophase for the m sinusoidal components. Resolution of this model for longitudinal time series has been previously described (1).

Equations 3, 4, and 5 can be readily extrapolated, allowing a fit to the model from Eq. 9 for analyzing hybrid time series. A test for the null hypothesis that each of the coefficients C_q is zero can be performed using the t distribution. Confidence intervals for those coefficients can also be obtained from the t distribution following an approach equivalent to that described by Eq. 6 for the mean level M.

SIMULATION STUDY

In order to validate the method proposed here, we carried out a simulation study. We assumed that the population under investigation is characterized by a seven-dimensional normal distribution with mean values for the seven parameters given in the top portion of Table 1 and the variance-covariance matrix given in the bottom portion of Table 1. The seven parameters are those obtained from fitting, by multiple components analysis, a model with periods of 24h, 12h, and 8h to time series of subjects sampled from that population. In the proposed example, the amplitudes of each component are markedly different. Moreover, the statistical significance of the 8h component (by testing the zero-amplitude hypothesis) is highly questionable.

Following Eq. 2, the pair amplitude-acrophase for each of the three components was substituted by the corresponding Cartesian coordinates β and γ. Each individual from the population can then be represented by a seven-dimensional vector (a sample from a Gaussian variable) that includes the seven parameters obtained by multiple components analysis after fitting Eq. 2 by linear least squares. Each simulation consisted of generating 100 random values for this seven-dimensional vector, that is, generating a sample of 100 potential individuals from the population with parameters given in Table 1. The variance-covariance matrix given by Eq. 4 was computed for the random sample of 100 individuals in the simulation by using Eq. 5. The population parameters (MESOR and the amplitude and acrophase of each of the three sinusoidal components, as well as their 95% confidence intervals) of this artificially generated sample were computed following Eqs. 3, 6, and 7.

Since the real values of the population parameters (given in the top of Table 1) are known, we can check if those true values are within the 95% confidence intervals computed for the corresponding estimated parameters for the sample of 100 individuals obtained from the population multiple components method. By repeating this whole procedure many times, we can calculate how many times the true values at the top of Table 1 are inside the 95% confidence intervals of the estimated parameters. If the population multiple components approach works satisfactorily, we should expect that in about 95%

Table 1. Simulation Study: True Values for the Population Parameters[a]

Mean values in polar coordinates

Period	MESOR	Amplitude	Acrophase
24.00		12.15	−245
12.00		5.34	−302
8.00		0.68	−258
Global	116.8		

Variance-covariance matrix

MESOR	β_{24}	γ_{24}	β_{12}	γ_{12}	β_{8}	γ_{8}
28.613						
−0.998	7.258					
−5.103	−0.471	10.915				
−0.834	−0.254	−1.420	6.111			
−5.214	2.770	−0.336	0.797	9.689		
0.295	−1.229	1.567	−0.609	−1.419	5.036	
0.410	−1.305	−0.199	0.405	−0.456	−0.349	3.840

[a]We assumed that the population is characterized by a seven-dimensional normal distribution with mean values given in the top portion and variance-covariance matrix given in the bottom portion. The seven parameters are those obtained from fitting, by multiple components analysis, a model with periods of 24h, 12h, and 8h. β and γ are the Cartesian projections of amplitude and acrophase, respectively, for each of the three components. The acrophase is given in angular degrees, with 360° = 24h (the fundamental period).

of the replicates the 95% confidence intervals would include the true parameters. Table 2 provides the results of this simulation study.

For a total number of 1000 replicates of the simulation procedure and an expected confidence level of 95%, Table 2 gives the coverage for each of the seven parameters in the linear model. The term *coverage* stands for the total number of replicates in which the real parameter (Table 1, top) was inside the confidence interval of the estimated parameter. Results from Table 2 indicate that the expected coverage (950) was mostly obtained for all but one of the parameters in the model. The worst results characterized the acrophase of the 8h component. By looking to the true parameters at the top of Table 1, we must realize that the amplitude of this 8h is very small, almost 20 times smaller than the amplitude of the 24h component, and about 8 times smaller than the amplitude of the 12h component. Accordingly, as pointed out above, we should be concerned about the statistical significance of the 8h component in the model. The lack of coverage for the acrophase of this 8h component from the application of the population multiple components method (Table 2) could corroborate the lack of significance for this component in the model (amplitude not statistically different from zero). In fact, relatively large confidence intervals for an acrophase obtained by linear least squares estimation usually corroborate similar conclusions (12). With these qualifications notwithstanding, results in Table 2 validate the proposed method for population multiple components analysis.

Table 2. Simulation Study: Results After 1000 Replicates[a]

Parameter	MESOR	A_{24}	ϕ_{24}	A_{12}	ϕ_{12}	A_8	ϕ_8
Coverage	954	946	942	945	964	962	905

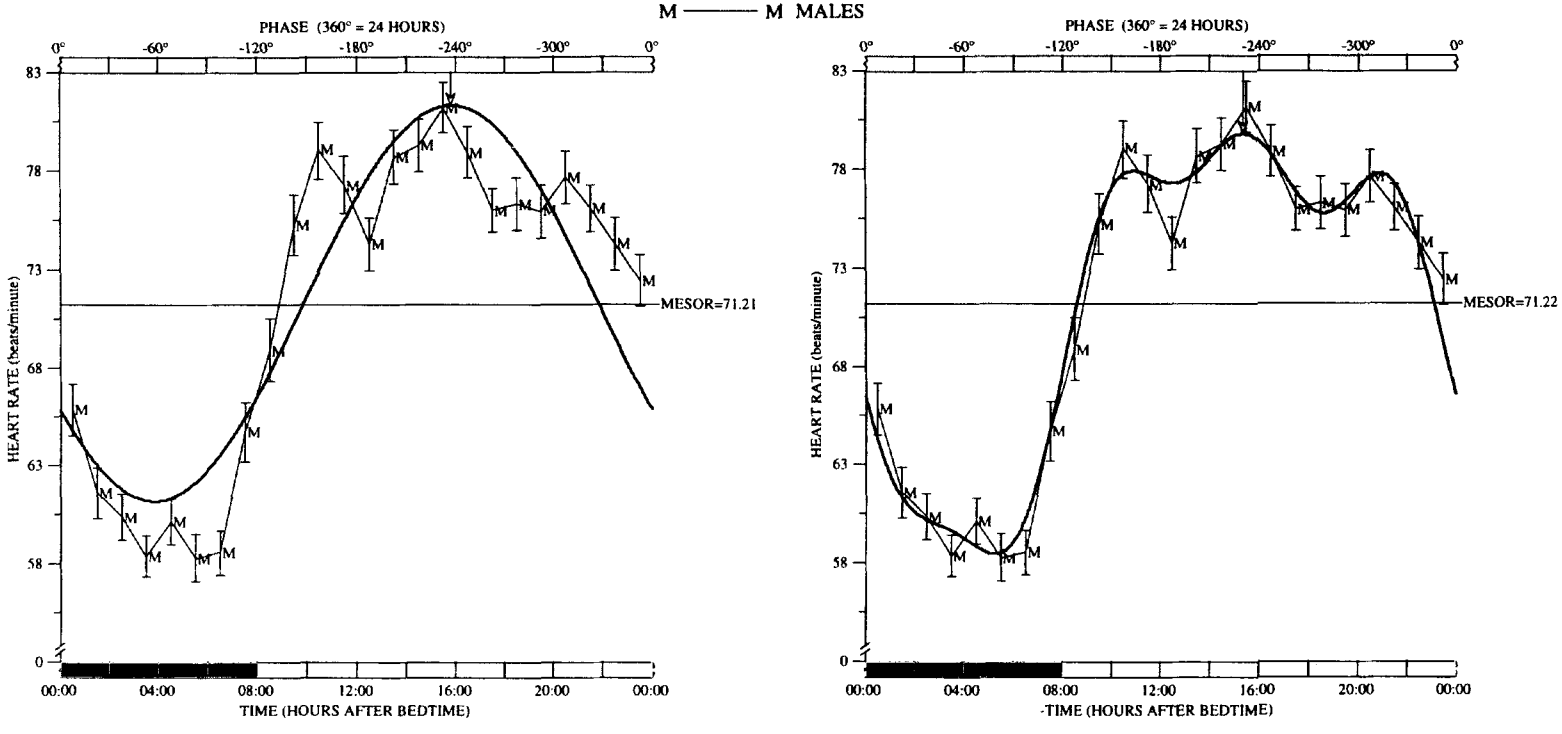
[a]Coverage indicates the total number of replicates in which the real parameter given in the top of Table 1 was inside the 95% confidence interval of the estimated parameter obtained by population multiple components analysis. The parameters are those obtained from fitting a multiple components model with periods of 24h, 12h, and 8h. We should expect that in about 95% of the replicates the 95% confidence intervals would include the true parameters (expected coverage is 950).

EXAMPLE

Despite the validation of the method here proposed by simulation studies, a better understanding of its differences with respect to the frequently used population-mean cosinor method could be obtained by visualizing an illustrative example. The data summarized in this example were sampled from a total of 50 clinically healthy men, 20 to 24 years of age, as part of a larger protocol designed to derive time-specified reference limits for the circadian variability of blood pressure and heart rate (16–18). The systolic, mean arterial, and diastolic blood pressures and heart rate of each subject were automatically monitored every 30 minutes for at least 48h with an ABPM-630 (Colin Medical Instruments, Inc., San Antonio, Texas) device. The subjects maintained their usual diurnal waking (~08:00 to ~24:00 for most subjects) and nocturnal resting routines during sampling, following everyday life conditions with minimal restrictions: they were told to follow a similar schedule during the days of sampling and to avoid the use of medication for the duration of the trial, as well as for at least 2 weeks prior to entering the study. The blood pressure cuff was worn on the nondominant arm. Cuff size was determined by upper arm circumference. During monitoring, each subject maintained a diary regarding information about their activity cycle, dietary consumption, physical activity, emotional state, and other external or internal stimuli possibly affecting blood pressure. In this example, we only describe the circadian variability of heart rate.

Original oscillometric data from each heart rate series were first synchronized according to the rest-activity cycle of each individual by recomputing all times of sampling in hours from bedtime to avoid differences among subjects in actual times of daily activity and to express results in circadian time rather than in less meaningful clock hours (17,18). After synchronization, heart rate values were edited according to commonly used criteria for the removal of outliers and measurement errors (19,20). The remaining data were analyzed using an updated version of Chronolab (8), a software package for biological signal processing by least squares estimation. In particular, data were analyzed, for comparative purposes, by population-mean cosinor, as well as by the proposed method, for population multiple components analysis.

Figure 1 represents, on both the left and on the right, the same circadian population chronogram (display of data as a function of time), with hourly means and standard errors of data computed as follows. First, hourly means were computed from each individual series after stacking all data sampled during a 48h monitoring span in only one



POPULATION MEAN COSINOR — FITTED PERIOD: 24.0 hours

KEY	#Sub.	p	MESOR	C.I.	AMP	C.I.	ACR	C.I.
M ——M	50	<0.001	71.21	(69.4, 73.1)	10.06	(9.1, 11.1)	-237.5°	(-233°,-242°)

POPULATION MULTIPLE COMPONENTS RHYTHMOMETRY
FITTED COMPONENTS: 24.0, 12.0, 6.0 hours

KEY	#Ind.	p	MESOR	C.I.	AMP	ORTHOPH	BATHYPH
M ——M	50	<0.001	71.22	(69.4, 73.1)	10.65	-231°	-79°

FIGURE 1. Circadian variation of heart rate in clinically healthy men sampled by 48h ambulatory monitoring modeled by population-mean cosinor (left) and by population multiple components (right). Each chronogram represents hourly means and standard errors of data sampled from 50 clinically healthy men, 20–24 years of age. The curve represented in each chronogram corresponds to the fitted model obtained by population-mean cosinor (left) and by population multiple component analysis (right), with corresponding characteristics given in the table below each chronogram. The arrows from the upper horizontal axis indicate the circadian acrophase (left) or orthophase (right).

idealized 24h span (given the highly statistically significant rhythm with a period of 24h demonstrated in all 50 series of heart rates studied). In a second step, the average of those individual means at each interval was computed, averaging across the total number of series for the population. The lower horizontal axis represents circadian time in hours after bedtime; the average resting span is indicated by the dark bar in the lower horizontal axis. The cosine curve represented in the chronogram on the left of Fig. 1 corresponds to the best-fitted model obtained by population-mean cosinor applied to all heart rate values (not just to the hourly means). The arrow from the upper horizontal axis in this left chronogram indicates the circadian acrophase of heart rate. The values of acrophase, in angular degrees, were computed taking bedtime as the circadian reference time point. The characteristics of the circadian rhythm with their confidence intervals are provided in the table below the chronogram, as well as in Table 3 (top). The table also includes information on the number of series analyzed, as well as on the *p*-value from testing the zero-amplitude hypothesis.

The nonsinusoidal curve represented in the chronogram on the right of Fig. 1 corresponds to the best-fitted model obtained by population multiple components analysis using three components with periods of 24h, 12h, and 6h, respectively (all harmonics of a fundamental period of 24h). Even when other ultradian components can be demonstrated as statistically significant in a small percentage of the subjects investigated, this model with only three components describes quite well the nonsinusoidal waveform of circadian variability in heart rate. The arrow from the upper horizontal axis in this right chronogram represents the circadian orthophase of heart rate. The characteristics of this three-component model are provided in the table below the chronogram. This table is complemented by the information of Table 3 (bottom), which provides the parameters for each of the three components in the model fitted to the data. Differences between the

Table 3. Time Structure of Circadian Variation of Heart Rate in Clinically Healthy Men, 20–24 Years of Age, Sampled by 48h Ambulatory Monitoring[a]

Population-mean cosinor

Period	*p*	MESOR	SE	Amp.	SE	CI Amp.	Acr.	SE	CI Acr.
24.000	<0.001	71.21	0.92	10.06	0.50	(9.1, 11.1)	−237.5°	2.4°	(−233°, −242°)

Population multiple components analysis

Period	*p*	MESOR	CI MESOR	Amp.	CI Amp.	Acr.	CI Acr.
24.000	<0.001			10.11	(9.1, 11.1)	−237.4°	(−233°, −242°)
12.000	<0.001			3.63	(2.9, 4.4)	−307.2°	(−296°, −318°)
6.000	<0.001			1.84	(1.1, 2.5)	−218.3°	(−200°, −236°)
Overall	<0.001	71.22	(69.4, 73.1)	10.65	Orthophase = −231°		Bathyphase = −79°

[a] *p* = *p* value from testing the zero-amplitude (Amp.) assumption; MESOR = midline estimating statistic of rhythm, a rhythm-adjusted mean; CI = confidence interval; Acr. = acrophase; acrophase, orthophase, and bathyphase all given in angular degrees, with 360° = 24h, and computed taking bedtime as the circadian reference time point.

cosine curve on the left and the nonsinusoidal model on the right obtained by two different methods for modeling hybrid time series can be readily observed from the graphic representation.

DISCUSSION

The fit of multiple components that are statistically and biologically significant accounts for nonsinusoidal waveforms and provides parameters characterizing them. When the data are nonsinusoidal, the least squares fit of a cosine curve may be used for rhythm detection, although this approach may not be as powerful as the simultaneous fit of all statistically significant components. The p-value obtained in testing the zero-amplitude assumption in fitting a unique component by single cosinor should thus be regarded as reflecting whether the data are better approximated by a cosine curve than by a horizontal line (1,2). Similar qualifications apply to the use of the population-mean cosinor method.

When more than one period are statistically significant over the span of time investigated or when the waveform is nonsinusoidal, the use of a multiple components analysis to fit a model consisting of several cosine functions (harmonics or not from one fundamental period) is recommended (1,10,13,21). The use of linear least squares methods with only one component (single and population-mean cosinor) should then be much more restricted than in general real practice, in which misuses are numerous (12). The sampling scheme could, however, limit the possible methods to be used. Too many clinical trials are frequently designed to obtain no more than five or six data for the variable of interest sampled over the time span covering the fundamental period investigated (say, 24h). Since the application of Eq. 1 needs a sample size $N > 2m + 1$, the described sampling rate would allow fitting a multiple model with no more than two components. In those cases or when, for instance, the resting and activity spans (or light-dark schedules) have different durations, one needs to realize that modeling the data with a unique cosine curve (and thus assuming the same duration for the above-the-MESOR and the below-the-MESOR spans) could only serve as a first approximation. The results should then be provided for descriptive or comparative purposes without further inference. The conditions specified above with respect to sample size needed to fit a model with m components ($k > 2m$ and $N > 2m + 1$) are the minimal mathematical requirements. Nevertheless, if the user wants reliable statistics, it is advisable to take higher values for k and N.

Advantages of using multiple components analysis have been limited by the fact that the method was only applicable for the analysis of longitudinal time series (1,2). Results from multiple components analysis obtained by fitting Eq. 2 to data sampled from an individual are dependent on the longitudinal time series at hand, characterize that given individual, and cannot be extrapolated to the population as a whole. The use of this approach for analysis of hybrid data has restrictions similar to the use of the single-cosinor method for simultaneously analyzing data sampled from several individuals (1). In particular, the parameters estimated from this inappropriate approach characterize only the individuals in the sample, and they cannot be extrapolated or inferred to characterize the population. An equivalent to the population-mean cosinor method (for analysis of short, sparse, and non-equidistant hybrid time series) with the fit of multiple components, allowing inference to the population, has been described here and tested by the results of a simulation study.

The method for population multiple components analysis presented here represents a new step in the development of statistical procedures in chronobiology. The approach has few limitations inasmuch as the usual restricting hypotheses such as equidistant sampling were not assumed in deriving Eqs. 3–8. Apart from the estimated values of the amplitude and acrophase for each component and the MESOR, with their respective confidence intervals, the method also provides estimates for parameters not included in Eq. 2, but highly useful for describing the general waveform of nonsinusoidal variability. Those parameters are the global amplitude, characterizing half of the total predictable change in the studied variable along the fundamental period; the orthophase, characterizing the time of the highest value, within the fundamental period, in the curve of multiple components fitted to the data; and the bathyphase, characterizing the time of the lowest value in the fitted curve.

The model resulting from fitting Eq. 9 to either single or population time series is not periodic since the polynomial trend is included. Such a situation, although unusual, is known in chronobiology or general chronopharmacology. For instance, predictable trends during gestation have been shown to characterize blood pressure and heart rate in both healthy and complicated pregnancies (22–25). Ultradian, circadian, circaseptan, and circatrigintan components in blood pressure have been described during gestation (26), along with a second-order polynomial trend found as statistically significant in normotensive pregnant women. The amplitude of these infradian components, free running from the social week and from the menstrual cycle in amenorrhea, as well as the coefficients of the polynomial trends, seem to be markedly different between healthy pregnant women and those who developed preeclampsia (26). The need of using the general model given by Eq. 9 becomes more frequent as one extends the time span of observation. Trends in variables of clinical interest not apparent when analyzing data sampled during two or just a few days become highly statistically significant when one takes aging into account by studying the same subjects for longer spans, even for years (9,27,28).

Apart from the restrictions concerning sample sizes given above, some other considerations must be taken into account when using the method proposed here. First, following a parallel line of thought as if comparing the single- and the population-mean cosinor methods, the population multiple components procedure assumes that all individual time series in the sample are well modeled by the same multiple components model that, therefore, should include the same anticipated periods. When analyzing longitudinal time series, the problem of selecting those anticipated relevant periods is not trivial (21). In the analysis of hybrid data, the problem becomes even more complicated. Second, there is not, at the moment, any method for comparing parameters estimated from the population multiple components models. For analysis of hybrid time series, a parameter test (the so-called Bingham test) has only been described for comparison of parameters estimated from the population-mean cosinor method (1). In the multiple components analysis, the comparison of orthophases is not trivial (1,10). The problem has just recently been solved for comparison of parameters obtained from the multiple components analysis of longitudinal time series (29). Extrapolation of this nonparametric method to the comparison of parameters obtained from analysis of hybrid data needs further investigation. Finally, the population multiple components procedure proposed here as the population-mean cosinor method only takes into account the among-individuals variability, assuming that this variance is considerably greater than the within-individuals variability. Precautions should be taken before using the proposed method when that assumption cannot be met.

ACKNOWLEDGMENTS

This research was supported in part by grants from Dirección General de Investigación Científica y Técnica, DGICYT, Ministerio de Educación y Ciencia (PB92-1111 and PB93-0372); Consellería de Educación e Ordenación Universitaria, Xunta de Galicia (XUGA-32202-B97, XUGA-32201-B95, and XUGA-32205-B95); by Química Farmacéutica Bayer S.A.; and by Vicerrectorado de Investigación, Universidad de Vigo.

REFERENCES

1. Bingham C, Arbogast B, Cornélissen G, Lee JK, Halberg F. Inferential statistical methods for estimating and comparing cosinor parameters. *Chronobiologia* 1982; 9:397–439.
2. Hermida RC. Chronobiologic data analysis systems with emphasis in chronotherapeutic marker rhythmometry and chronoepidemiologic risk assessment. In: Scheving LE, Halberg F, Ehret CF, eds. *Chronobiotechnology and chronobiological engineering*. Dordrecht, The Netherlands: Martinus Niijhoff, NATO ASI Series No. 120; 1987:88–119.
3. Nelson W, Tong YL, Lee JK, Halberg F. Methods for cosinor rhythmometry. *Chronobiologia* 1979; 6:305–23.
4. Halberg F. Chronobiology. *Annu Rev Physiol* 1969; 31:675–725.
5. Halberg F. Chronobiology: methodological problems. *Acta Med Rom* 1980; 18: 399–440.
6. Halberg F, Tong YL, Johnson EA. Circadian system phase—an aspect of temporal morphology; procedures and illustrative examples. Proc Internat Congress of Anatomists. In: *The cellular aspects of biorhythms. Symposium on biorhythms*. Berlin: Springer-Verlag, 1967:20–48.
7. Halberg F, Johnson EA, Nelson W, Runge W, Sothern R. Autorhythmometry—procedures for physiologic self-measurements and their analysis. *Physiol Teacher* 1972; 1:1–11.
8. Mojón A, Fernández JR, Hermida RC. Chronolab: an interactive software package for chronobiologic time series analysis written for the Macintosh™ computer. *Chronobiol Internat* 1992; 9:403–12.
9. Haus E, Nicolau GY, Lakatua D, Sackett-Lundeen L. Reference values for chronopharmacology. In: Reinberg A, Smolensky M, Labrecque G, eds. *Annual review of chronopharmacology*, Vol. 4. Oxford, England: Pergamon Press, 1988:333–424.
10. Tong YL, Nelson W, Sothern RB, Halberg F. Estimation of the orthophase (timing of high values) on a non-sinusoidal rhythm, illustrated by the best timing for experimental cancer chronotherapy. *Proc 12th Internat Conf Internat Soc Chronobiol*, 1977:765–69.
11. Mattes A, Witte K, Hohmann W, Lemmer B. Pharmfit—a nonlinear fitting program for pharmacology. *Chronobiol Internat* 1991; 8:460–76.
12. De Prins J, Waldura J. Sightseeing around the single cosinor. *Chronobiol Internat* 1993; 10:395–400.
13. Ayala DE, Hermida RC, García L, Iglesias T, Lodeiro C. Multiple component analysis of plasma growth hormone in children with standard and short stature. *Chronobiol Internat* 1990; 7:217–20.
14. Fernández JR, Hermida RC. Estimation of population parameters from rhythmome-

tric models with multiple components [abstract]. *Biol Rhythm Res* 1995; 26:388.
15. Johnson RA, Wichern DW. *Applied multivariate statistical analysis*. New York: Prentice-Hall, 1982.
16. Hermida RC, Ayala DE, Fernández JR, Mojón A. Time-specified reference limits for ambulatorily monitored blood pressure in clinical health. *Biomed Instrum Technol* 1993; 27:235–43.
17. Hermida RC, Fernández JR. Computation of time-specified tolerance intervals for ambulatorily monitored blood pressure. *Biomed Instrum Technol* 1996; 30:257–66.
18. Hermida RC, Mojón A, Fernández JR, Ayala DE. Computer-based medical system for the computation of blood pressure excess in the diagnosis of hypertension. *Biomed Instrum Technol* 1996; 30:267–83.
19. Cornélissen G. Instrumentation and data analysis methods needed for blood pressure monitoring in chronobiology. In: Scheving LE, Halberg F, Ehret CF, eds. *Chronobiotechnology and chronobiological engineering*. Dordrecht, The Netherlands: Martinus Nijhoff, 1987:241–61.
20. Staessen J, Fagard R, Lijnen P, Thijs L, Vaa Hoof R, Amery A. Ambulatory blood pressure monitoring in clinical trials. *J Hypertens* 1991; 9(Suppl 1):s13–s19.
21. De Prins J, Hecquet B. Data processing in chronobiological studies. In: Touitou Y, Haus E, eds. *Biological rhythms in clinical and laboratory medicine*. Berlin: Springer-Verlag, 1992:90–113.
22. Ayala DE, Hermida RC. *Circadian time structure of cardiovascular characteristics in human pregnancy. Chronobiol Internat* 1993; 10:128–36.
23. Ayala DE, Hermida RC, Cornélissen G, Brockway B, Halberg F. Heart rate and blood pressure chronomes during and after pregnancy. *Chronobiologia* 1994; 21: 215–25.
24. Hermida RC, Ayala DE, Mojón A, Fernández JR, Silva I, Ucieda R, Iglesias M. High sensitivity test for the early diagnosis of gestational hypertension and preeclampsia. I. Predictable variability of cardiovascular characteristics during gestation in healthy and hypertensive pregnant women. *J Perinatal Med* 1997; 25(1): 101–9.
25. Ayala DE, Hermida RC, Mojón A, Fernández JR, Silva I, Ucieda R, Iglesias M. Blood pressure variability during gestation in healthy and complicated pregnancies. *Hypertension* 1997; 30(3):611–18.
26. Ayala DE, Hermida RC. Multifrequency infradian variation of blood pressure during and after human pregnancy. *Chronobiol Internat* 1995; 12(5):333–44.
27. Kanabrocki EL, Sothern RB, Scheving LE, Halberg F, Pauly JE, Greco J, Nemchausky BA, DeBartolo M, Kaplan E, McCormick JB, Olwin JH, Marks GE, Bird T, Redmond DP, Graeber RC, Ferrara A, Hrushesky WJM. Ten-year-replicated circadian profiles for 36 physiological, serological and urinary variables in healthy men. *Chronobiol Internat* 1988; 5:237–84.
28. Kanabrocki EL, Sothern RB, Scheving LE, Vesely DL, Tsai TH, Shelstad J, Cournoyer C, Greco J, Mermall H, Nemchausky BM, Bushnell DL, Kaplan E, Kahn S, Augustine G, Holmes E, Rumbyrt J, Sturtevant RP, Sturtevant F, Bremner F, Third JLHC, McCormick JB, Mudd CA, Dawson S, Sackett-Lundeen L, Haus E, Halberg F, Pauly JE, Olwin JH. Reference values for circadian rhythms of 98 variables in clinically healthy men in fifth decade of life. *Chronobiol Internat* 1990; 7(5/6): 445–61.
29. Fernández JR. Nonparametric method for comparison of parameters from rhythmometric models with multiple components [abstract]. *Biol Rhythm Res* 1995; 26:387.