CHRONOBIOLOGY INTERNATIONAL, 19(4), 793–803 (2002)

CICLOGRAMA: A TOOL FOR DETECTION OF RHYTHMICITIES IN SLEEP/WAKE CYCLES

L. Diambra,* J. R. Lopes, L. Menna-Barreto, and R. Rigolino

GMDRB, Departamento de Fisiologia e Biofísica, ICB, Universidade de São Paulo, Av. Prof. Lineu Prestes 1524, ICB1 05508-900, São Paulo, SP, Brazil

ABSTRACT

The Fourier spectral analysis of binary time series (or rectangular signals) causes methodological problems, due to the fact that it is based on sinusoidal functions. We propose a new tool for the detection of periodicities in binary time series, focusing on sleep/wake cycles. This methodology is based on a weighted histogram of cycle durations. In this paper, we compare our methodology with the Fourier spectral analysis on the basis of simulated and real binary data sets of various lengths. We also provide an approach to statistical validation of the periodicities determined with our methodology. Furthermore, we analyze the discriminating power of both methods in terms of standard deviation. Our results indicate that the Ciclograma is much more powerful than Fourier analysis when applied on this type of time series. (*Chronobiology International, 19(4), 793–803, 2002*)

Key Words: Spectral analysis; Binary data; Sleep/wake cycle; Ciclograma; Method; Time series analysis

INTRODUCTION

From a chronobiological point of view, the physiological and behavioral states of a living organism change as time goes by, forming a sequence of patterns that repeat themselves periodically or nearly periodically. In this sense, an important

*Corresponding author. E-mail: diambra@fisio.icb.usp.br

physiological process is the alternance of states of wakefulness and sleep, forming a discrete binary signal. Any binary process can be described by two values $\{-1, 1\}$ representing single events. In sleep/wake cycle, it is thus possible to approximate the underlying time series by a combination of different basis functions. A well-known algorithm is the spectral analysis (SA), based on Fourier transform, which decomposes the signal into weighted sum of sinusoidal functions whose frequencies are multiples of the fundamental frequency.[1,2] Notwithstanding its helpful use in many circumstances, when this protocol is applied on a binary time series some problems come up due to rectangular signals with unequal periodicities, which cannot be easily reproduced by sinusoidal functions.[3]

There are many situations where we do not need a complete representation of the time series, but rather a quantitative characterization that allows some interpretation of the time series. In chronobiology, periodicity is one of the most important properties for a biological interpretation. For sinusoidal functions, periodicity is defined by zero crossing at equidistant points in time. Thus, frequency is defined as one half the number of zero crossings per unit of time.[4] However, a signal may show rhythmic but not strictly periodic behavior, as many real-life phenomena that exhibit a nearly periodic characteristic.[5,6] Such data do not fulfill the criterion of equidistant zero crossings. These considerations imply that the terms, periodicity and frequency, in the SA sense, are not appropriate to describe rhythmicities of the signal.[3] Moreover, two subjects with different durations of sleep (or wake) episodes but with the same mid-sleep time have different periodograms, according to simulations performed by Louzada.[7]

Therefore, much effort has been devoted to the task of developing suitable tools that are able to represent the relative power of each rhythmic pattern (or component, in the SA sense).[8–11] In this paper, we propose a new method, termed Ciclograma, with which one is able to compute a histogram of the distribution of durations of sleep/wake cycles present in a given time series.[12] By this histogram, one can detect the periodic components present in a time series and their relative contribution to the overall rhythmicity. We have compared the rhythmicity descriptive power of the Ciclograma and the Fourier periodogram in two simulated time series. We have also analyzed the discrimination power of the Ciclograma and the Fourier periodogram in the task to distinguish subjects with different levels of variability in the wake and sleep onset time. Finally, we have applied our technique for the characterization of the ontogeny of rhythmicity in human data.

METHODS AND MATERIALS

Methods

Our technique is based on building a histogram of the distribution of intervals between successive awakenings and successive sleep onsets with some modifications. We realized the fact that consecutive sleep/wake cycles form another broader cycle whose duration is equal to the sum of length of time of each

cycle. This *composing* procedure is shown schematically in Fig. 1. Thus, we considered the duration of each sleep/wake cycle as well as the duration of the cycles formed by two or more consecutive cycles in the counting procedure. In order to obtain a better description of polyphasic patterns, we extended the composing procedure to cycles formed by eight consecutive sleep/wake cycles.

The algorithm counts the number of all these cycles within consecutive bins. The size of bins b is an arbitrary choice. In this work, we set $b = 30$ min (i.e., from 0 to 0.5 h, from 0.5 to 1 h, from 1 to 1.5 h, and so on until the bin from 24.5 to 25 h). After the counting procedure, we normalized the histogram in the following manner: i) There are more short intervals than long intervals during a day (e.g., there are 24 intervals of 1 h in a day, but only four intervals of 8 h). Consequently, short intervals are more probable than long intervals. Assigning a different weight to each bin can circumvent this bias. Thus, the bin corresponding to cycles between 0.5 and 1 h is weighted by a factor 1/24, while the bin corresponding to cycles whose durations last between 7.5 and 8 h is weighted by a factor 1/4. ii) After setting the weights, we normalized the amplitudes Q_w such that $\sum_w Q_w = 1$. These amplitudes express the weighted fraction of the total number of cycles whose durations lie in the interval $\{w - b, w\}$. The final result of this procedure can be depicted as a histogram of duration of cycles, where the height in each bin represents the power or amplitude of each component (or rhythmic pattern) present in the time series. In other words, Q_w represents the relative contribution of the component w for the overall rhythmicity.

Analysis of Significance

After histogram estimation, an analysis of significance became mandatory in order to eliminate some spurious periodicities. The approach consists of specifying a well-defined null hypothesis, as if the data were derived from a

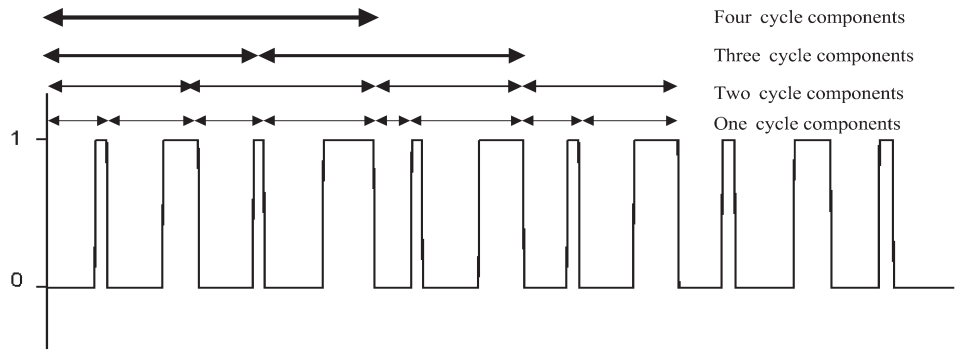

Figure 1. Schematic diagram of Ciclograma counting procedure. Arrows at level 1 indicate the temporal intervals between successive sleep onsets. Arrows at level 2 indicate the broad cycles that are formed by two consecutive cycles, and so forth for the higher levels. The graph shows the composing procedure to cycles formed up to four consecutive sleep/wake cycles.

stochastic process. The second step was to compute the histogram of the duration of cycles of this process and for the original data. Finally, we tested the null hypothesis against the observations. In this paper, our null hypothesis is that the data were derived from a stochastic process with exactly the same statistical distribution of the original data. This null hypothesis was tested by firstly randomly shuffling the binary values of the original time series. With this operation, the new time series loses the temporal order but preserves the statistical distribution properties of the original time series. Then, we determined the distribution of the histogram for an ensemble of surrogate time series, which are only different realizations of the hypothesized stochastic process. In this paper, we have worked with two different sizes of ensembles: 100 and 500 surrogate time series, respectively. After that, we computed numerically, the average histogram and the corresponding standard deviation from the ensemble of surrogate time series. This can be reliably done because we have many realizations of the null hypothesis. Finally, rather than estimate error bars on the histogram of the original data, we put error bars on the average histogram. These error bars were estimated from the standard deviation. By this way, we avoided analytical derivations that could be difficult if not impossible.

In order to compute the significance, we considered Q_w and H_w^i, the amplitudes corresponding to the rhythmic patterns whose durations lie in the interval $\{w - b, w\}$ computed from the original time series and from the ith surrogate data generated under the null hypothesis H, respectively. Now we define, as measure of significance, the statistic $S = |Q_w - \mu_w^H|/\sigma_w^H$, where μ_w^H and σ_w^H denote the mean and the standard deviation of the distribution of H_w. The significance S is a dimensionless quantity, and we can express it as unit of sigmas. Numerical experiments indicated that if the distribution H_w^i is Gaussian, then the p-value is given by $p = \mathrm{erfc}(S/\sqrt{2})$.[13] This is the probability of observing a significance S or larger if the null hypothesis is true. The significance level at 2.81 standard deviations ($p = 0.001$) is plotted in the histogram of cycle duration as a gray line. Thus, when the amplitude, Q_w crosses the gray line upwards, we can say that the rhythmic pattern whose duration lies in the interval $\{w - b, w\}$ is significant with $p < 0.001$.

Data

Our procedure has been applied to simulated and real binary data sets of various lengths. Simulated data were derived from signals with defined cyclic patterns, where the beginning and ending of each event were perturbed with some level of noise. In Fig. 2, two simulated binary time series are shown. On the left, 50 d of monophasic patterns are displayed. In this case, the data are simulating a subject whose wake-up time has a Gaussian distribution with a mean at 06:00h in the morning and a standard deviation of 30 min, while sleep time shows a mean at 22:00h and a standard deviation of 60 min. On the right, 50 d of biphasic patterns

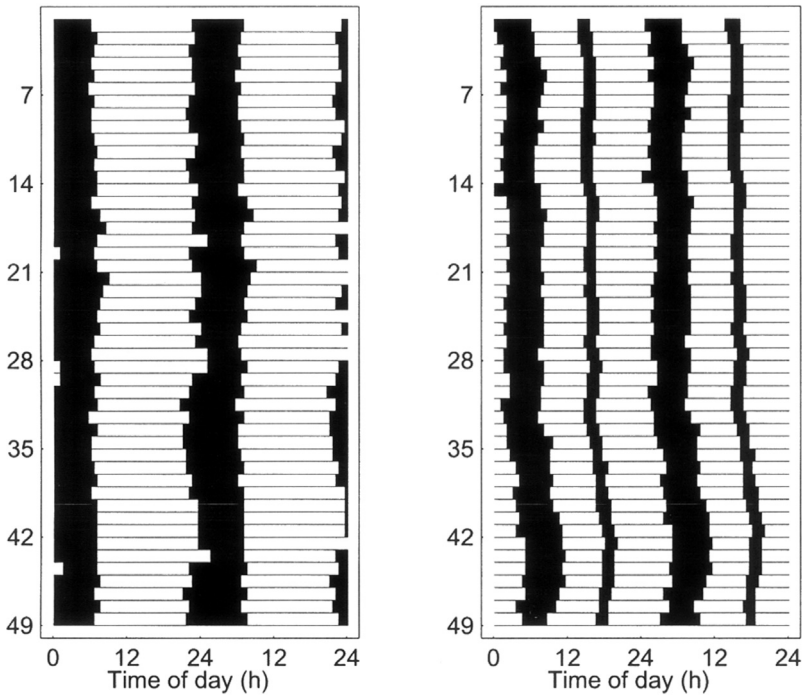

Figure 2. Double actograms show 50 d (on vertical axis). Left: Simulated monophasic sleep pattern (standard deviation of 60 min from the sleep time and 15 min from the wake-up time). Right: Simulated biphasic sleep pattern (standard deviations of 30 min both to sleep and wake-up time, and a nap around 2 h centered at 14:00h).

are displayed. In this case, the data are stimulating a subject with wake-up and sleep onset time with a Gaussian distribution of 30 min of standard deviation, and means at 06:00 and at 24:00h, respectively. In addition, there is a nap that lasts around 2 h, centered at 14:00h and with 10 min of standard deviation both at the beginning and at the ending of the nap.

In the case of the real time series, the data for the wake-up and sleep onset time were recorded by the mother of the child. Then the data were arranged in 10min bins, each bin representing sleep (-1) or wake (1) states. The subject did not show any evidence of neurological disorder and other serious physiologic disturbances. The data were collected for 128 consecutive weeks beginning from the first days of life.

RESULTS

To analyze periodicities in the sleep/wake cycles, we applied the Ciclograma procedure to several situations. For comparison, we also applied the fast Fourier transform (FFT) periodogram. As a first example, we considered the simulated

time series whose actogram is shown in Fig. 2 on the left side. This sleep/wake time series is mimicking a monophasic pattern. On the left of Fig. 3, the histogram of cycle duration is depicted with the corresponding significance line in thick gray computed on 100 surrogated time series. One can see that the bulk of the rhythmic patterns lies in the region of 24 h, while the width of the 24h peak is related to the variability of the rhythmic pattern. The Fourier periodogram, from the same data, is shown on the right side of Fig. 3. In this case, several artifactural peaks appear, which do not correspond to sleep/wake episodes, and the variability of the rhythmic pattern is not apparent. In the following example, we applied the same procedure to the biphasic rhythm shown on the right side of Fig. 2, simulating napping episodes. Figure 4 shows the histogram of cycle duration produced by the Ciclograma (left) and the FFT periodogram (right). In this case, the Ciclograma shows, besides the circadian component, four small significant peaks corresponding to rhythmic patterns of duration around 9 h (corresponding to events beginning approximately at 06:00h and ending at 15:00h), 11 h (beginning approximately at 13:00h and ending at 24:00h), and 13 h (beginning approximately at 00:00h and ending approximately at 06:00h of the next day). The FFT periodogram shows more power for the period corresponding to 12 h than for the period corresponding to 24 h. This fact can develop from the decomposition of the signal in sinusoidal functions by Fourier transform rather than from the data itself. These examples show that the Ciclograma provides better description of the periodicities than the classical FFT periodogram.

We have also studied the relative power corresponding to the rhythmic pattern of 24 h for different values of standard deviation in the wake-up time and sleep onset. Again, we considered simulated time series that mimic monophasic and biphasic subjects. We defined the variability of a subject as the sum of standard deviations of both the wake-up and the sleep times. It is expected that the subjects with different variabilities are associated with different values of the relative power corresponding to the rhythmic pattern of 24 h. In this sense, subjects with greater variabilities are associated with a

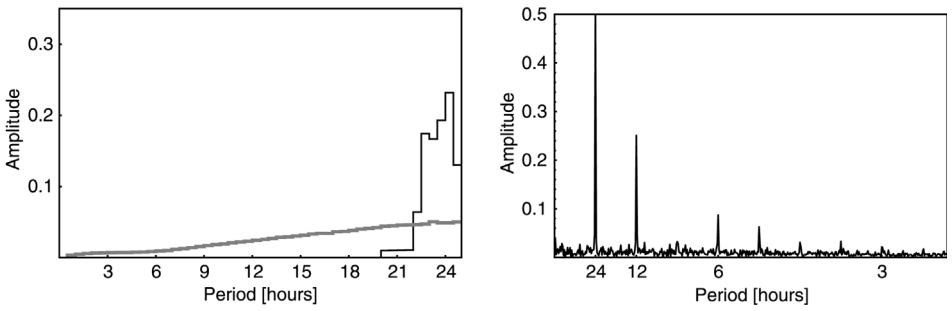

Figure 3. Histogram of cycle durations produced by the Ciclograma (left) and FFT periodogram (right) for the simulated data shown in the left side of Fig. 2. The gray line on the histogram is the significance at level $p = 0.001$, computed over 100 surrogated time series.

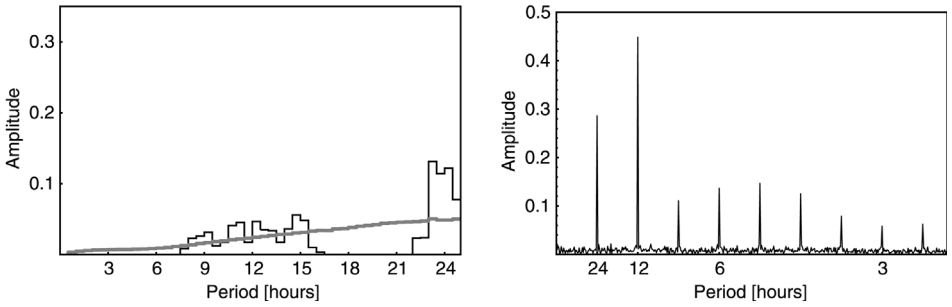

Figure 4. Histogram of cycle durations produced by the Ciclograma (left) and FFT periodogram (right) for the simulated data shown in the right side in Fig. 2. The gray line on the histogram is the significance at level $p = 0.001$, computed over 100 surrogated time series.

smaller circadian power. In order to analyze the discriminating power of the Ciclograma when comparing subjects with different variabilities, we performed the following computation: we generated an ensemble of 50 sleep/wake time series for each value of variability and then computed the relative power corresponding to rhythmic patterns with durations between 23.5 and 24.5h of each time series. In order to make a comparison with classical analysis, we also performed the same procedure with the FFT periodogram. In Fig. 5, we show the mean value of the circadian power in the monophasic case corresponding to different values of the variability obtained by the Ciclograma (solid bars), and the corresponding circadian power obtained by the Fourier analysis (open bars). The standard deviations are represented by vertical error bars. Our results

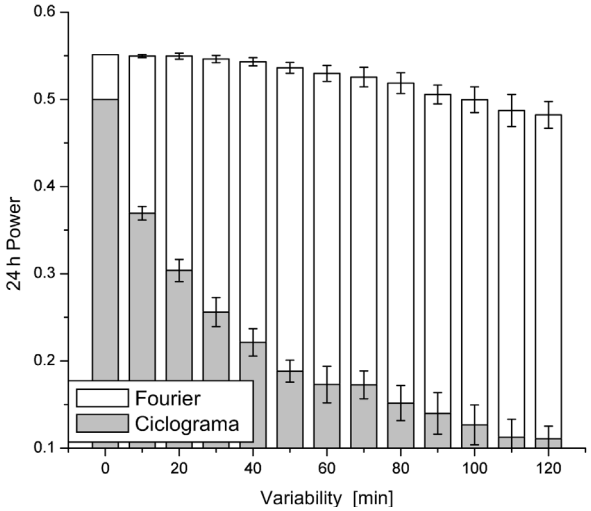

Figure 5. Relative power corresponding to the circadian cycle vs. the variability of the wake-up time and the sleep time, in a monophasic case. The solid bars show the relative power calculated with the Ciclograma and the open bars show the relative amplitude computed with the FFT periodogram.

indicate that the Ciclograma produces clearer discrimination than the FFT periodogram when comparing subjects with different variabilities. In Fig. 6, we depict the mean value of the circadian power in the biphasic case (the standard deviation is represented by the vertical error bars) corresponding to different values of the variability obtained by the Ciclograma (solid bars) and by the FFT analysis (open bars). In this case, the discrimination power of the Ciclograma is better than the classical technique. The discrimination task is more difficult for the biphasic case than for the monophasic case.

As a final example, we have also applied the Ciclograma to a real time series. In Fig. 7, we depict a time-period diagram computed by the Ciclograma for the first 128 wk of life of a child. In this case, we applied the Ciclograma for each week and displayed the relative amplitudes in a gray scale. This time-period diagram shows the development of the rhythmic patterns during the first 128 wk. One can see that the first 25 wk on the diagram show ultradian rhythms of various durations. Only around the circadian rhythm there is power. Between the 25th and the 100th wk of life, there are besides the 24h rhythm, ultradian components of 9 and 15 h. The same data were analyzed with FFT periodogram technique, which was subject of a former publication[14]. In Fig. 8, the histogram of cycle durations computed by Ciclograma for the 54th week from birth, is shown related to the diagram shown in Fig. 7. In this case, the corresponding significance line at $p = 0.001$ is computed on 500 surrogated time series. Shorter time series require larger ensembles of surrogated time series in order to guarantee good statistics. The polyphasic characteristic is evident.

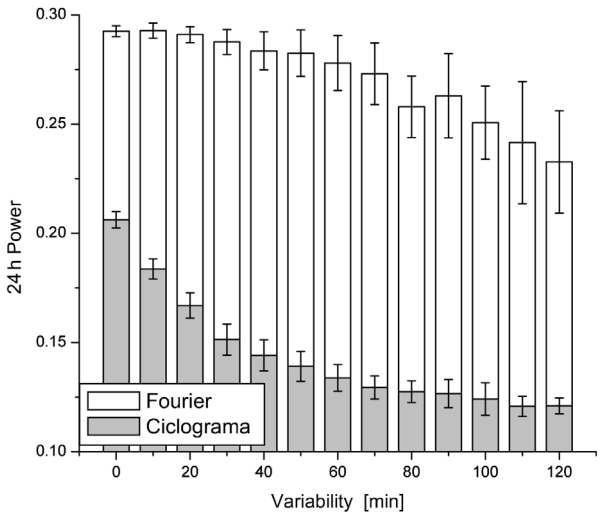

Figure 6. Relative power corresponding to the circadian cycle vs. the variability of the wake-up time and the sleep time, in a biphasic case. The solid bars show the relative power calculated with the Ciclograma, and the open bars show the relative amplitude computed with the FFT periodogram.

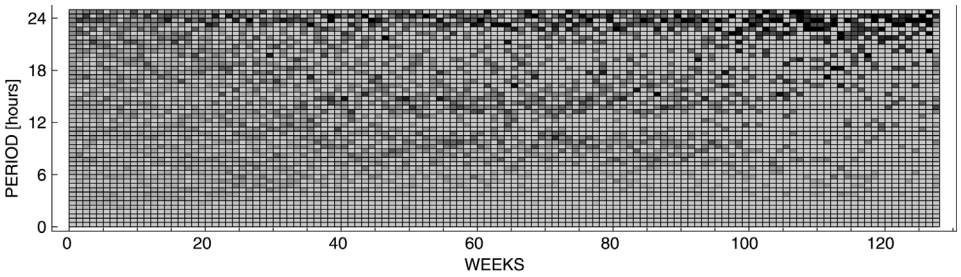

Figure 7. Time-period histogram of cycle duration computed weekly by Ciclograma for the first 128 wk of a child.

CONCLUSIONS AND DISCUSSION

In this paper, we propose a new tool, based on a modified histogram of cycle durations to describe the rhythmicity of binary signals. We have illustrated its descriptive power in several applications concerning simulated and real sleep/wake data. We also provide an approach to statistical validation of the periodicities determined with our methodology. The relative contribution of a specific rhythmic pattern for the overall rhythmicity is given by the amplitude of the Ciclograma, i.e., the height of the histogram Q_w. Thus, the amplitude Q_w expresses the probability of observing a sleep/wake cycle whose duration lies in the interval $\{w - b, w\}$. In order to make comparison with the FFT periodogram, we have compared the amplitude Q_w with the square weight, associated with the period w, which is present in the sum of sinusoidal functions that represents the binary signal. Our results suggest that with Ciclograma we are able to detect the rhythmic components present in binary time series and their relative contribution to overall rhythmicity better than the FFT periodogram.

Moreover, the width in the 24h peak is related to the variability of circadian pattern. In the case of the FFT periodogram, this variability is not evident. Thus,

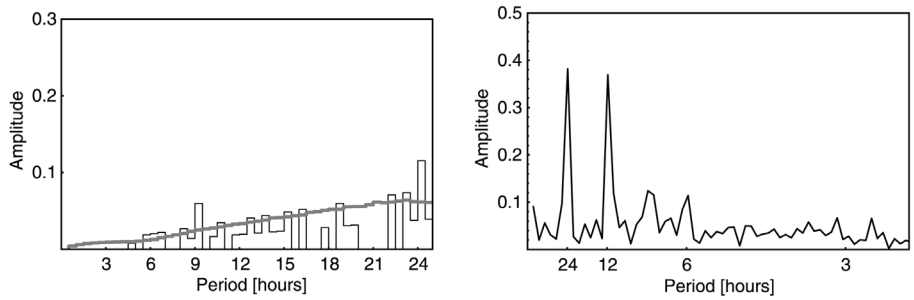

Figure 8. Histogram of cycle duration computed by Ciclograma (left) and Fourier periodogram (right) for 54th wk of the child shown in Fig. 7. The gray line on the histogram is the significance at level $p = 0.001$, computed over 500 surrogated time series.

we can say that the Ciclograma gives a better description of the periodicities than the classical methods.

We have also compared the descriptive power for rhythmicity of the Ciclograma and FFT periodogram in simulated time series. We have analyzed the discrimination power of the Ciclograma and FFT periodogram in the task of distinguishing subjects with different levels of variability in the wake-up and sleep onset time. Finally, we have applied our technique to the characterization of the human ontogenetic development of the sleep/wake cycle.

Furthermore, the Fourier techniques can produce aliasing effects. This phenomenon spuriously moves all power spectra, which lie outside of the frequency range $-f_s/2 < f < f_s/2$, where f_s is the sampling frequency, into that range. Thus, the considerable spectral power of the high frequencies in binary signals will be spuriously moved to the frequency range of our interest. On the other hand, the Ciclograma approach avoids these undesirable effects.

Spectral analysis techniques are based on the description of the time series by the sum of sinusoidal functions rather than by an effective characterization of the rhythmicity. Thus, we believe that the terms periodicity and frequency, in the SA sense, are not appropriate to describe rhythmicities of binary signals, and the results based on SA of binary signals should be interpreted bearing in mind the signal decomposition environment rather than the rhythmic content, as in Ciclograma. Thus, the Ciclograma approach, for the estimation of the relative contribution of each rhythmic pattern to the overall rhythmicity, seems to offer promising perspectives.

ACKNOWLEDGMENTS

The authors acknowledge fruitful discussions with Antoni Diez-Noguera and F. Louzada. L. Diambra acknowledges the financial support of FAPESP #99/07186-3 (Brazil), J. R. Lopes acknowledges the financial support of CAPES (Brazil), L. Menna-Barreto acknowledges the financial support of FAPESP (Brazil) and R. Rigolino acknowledges the financial support of FAPESP #01/03184-8 (Brazil).

REFERENCES

1. Hernandez, G. Time Series, Periodograms and Significance. J. Geophys. Res. **1999**, *104*, 10355–10368.
2. Blackman, R.B.; Tukey, J.W. *The Measurements of Power Spectra*; Dover Publications Inc.: New York, 1958.
3. Kowalski, A.; Musial, F.; Enck, P.; Kalveram, K.T. Spectral Analysis of Binary Time Series: Square Waves Vs. Sinusoidal Functions. Biol. Rhythm Res. **2000**, *31*, 481–498.
4. Harmuth, H.F. *Transmission of Information by Orthogonal Functions*; Springer: Berlin, 1969.

5. Kanjilal, P.P.; Bhattacharya, J.; Saha, G. Robust Method for Periodicity Detection and Characterization of Irregular Cyclical Series in Terms of Embedded Periodic Components. Phys. Rev. E **1999**, *59*, 4013–4025.
6. Tong, H. *Non-linear Time Series*; Oxford University Press: New York, 1990.
7. Louzada, F. Personal communication.
8. Sokolove, P.G.; Bushell, W.N. The Chi Square Periodogram: Its Utility for the Analysis of Circadian Rhythms. J. Theor. Biol. **1978**, *72*, 131–160.
9. Brillinger, D.R. Asymptotic Distribution of Whittaker Periodogram and a Related Chi-Squared Statistic for Stationary Processes. Biometrika **1974**, *61*, 419–422.
10. Euright, J.T. Search for Rhythmicity in Biological Time Series. J. Theor. Biol. **1965**, *8*, 426–468.
11. Menna-Barreto, L.; Benedito-Silva, A.A.; Marques, N.; Andrade, M.M.; Louzada, F. Ultradian Components of the Sleep–Wake Cycle in Babies. Chronobiol. Int. **1993**, *10*, 103–108.
12. Bueno, C.; Diambra, L.; Menna-Barreto, L. Sleep–Wake and Temperature Rhythms in Preterm Babies Maintained in a Neonatal Care Unit. Sleep Res. Online **2002**, in press.
13. Theiler, J.; Eubank, S.; Longtin, A.; Galdrikian, B.; Farmer, J.D. Testing for Nonlinearity in Time Series: The Method of Surrogate Data. Physica D **1992**, *58*, 77–94.
14. Menna-Barreto, L.; Isola, A.; Louzada, F.; Benedito-Silva, A.A.; Mello, L. Becoming Circadian: A One-Year Study of the Sleep–Wake Cycle in Children. Braz. J. Med. Biol. Res. **1996**, *29*, 125–129.

Received October 21, 2001
Returned for revision November 21, 2001
Accepted March 18, 2002