Biological Rhythm Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/nbrr20

A Simple Approach for the Computation of Multiple Periodicities in Biological Time Series

Anoop V. Rao & Vijay Kumar Sharma

Available online: 09 Aug 2010

To cite this article: Anoop V. Rao & Vijay Kumar Sharma (2002): A Simple Approach for the Computation of Multiple Periodicities in Biological Time Series, Biological Rhythm Research, 33:5, 487-502

To link to this article: http://dx.doi.org/10.1076/brhm.33.5.487.13933

Biological Rhythm Research
2002, Vol. 33 No. 5. pp. 487–502

0165-0424/3305-487$16.00

A Simple Approach for the Computation of Multiple Periodicities in Biological Time Series

Anoop V. Rao[1]* and Vijay Kumar Sharma[2]

[1]Kasturba Medical College, Mangalore, India; [2]Chronobiology laboratory, Evolutionary and Organismal Biology Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore, India

Abstract

We have described a simple approach for the analysis and isolation of multiple periodicities from a biological time series. For the estimation of the periodicities, we used simulated data and data from ongoing experiments in our laboratory. Two time series were simulated, one which consisted of only white noise and the other consisted white noise along with periodicities of 6, 11, 17 and 23 h, to demonstrate that our method can successfully isolate multiple patterns in a time series. Our method of analysis is objective, simple, flexible and adaptive since it distinctly delineates the individual contribution from an overlap of multiple periodicities. The key features of our method are: (i) identification of a reliable phase reference point, (ii) scanning the time series using a moving window in increments, (iii) use of Siegel's modification of Fisher's method to detect significant periodicit(y)ies in the time series. The use of window sizes of increasing length to examine the time series elegantly reduces noise while identifying periodicities that are otherwise not apparent. Finally, the periodogram can be smoothed in order to normalize the contribution by attendant frequency components within the waveform. A minimum critical value for relative contribution of various frequencies was calculated to delineate the periodicities that contributed significantly to the time series. We executed this method of time series analysis using MS Excel and C.

Keywords: Time series, Fisher's test, periodogram, critical points.

Address correspondence to: Vijay Kumar Sharma, Chronobiology laboratory, Evolutionary and Organismal Biology Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, P.O. Box 6436, Jakkur, Bangalore 560064, India. Tel: +91-80-8462750 (Extn. 2258) (W) or +91-80-8564671 (H); Fax: +91-80-8462766; E-mail: vsharma@jncasr.ac.in
*Current Address: 56-255, Biological Engineering Division, Massachusetts Institute of Technology, 77, Massachusetts Avenue, Cambridge, MA 02139

Introduction

Oscillatory phenomenon in various time domains is a commonly and conspicuously observed phenomenon in biological systems. They are known to occur at several levels of organizational hierarchy; from molecular to populational (Takahashi & Zatz, 1982; Vanden Driessche, 1989; Takahashi, 1991; Pedersen & Johnsson, 1994; Rosato et al., 1997). Body temperature, hormone levels, immune cells and their response (De Boer et al., 1993), sleep-wake cycle, biochemical cycle (Goldbeter & Decroly, 1983), cell division cycle (Smaaland, 1996), are a few well-studied examples (Robertson & Takahashi, 1983; Ishida et al., 1999). Identification and precise quantification of such fluctuations in biological systems constitute an important step in order to relate to and understand its regulatory mechanisms. Amplitude and period are two important parameters that characterize such oscillations which are represented as a time series. A time series can be continuous (observations made continuously over time) or discrete (observations made only at certain times); stationary (data that fluctuate around a constant value) or non-stationary (a series having parameters of the cycle i.e., length, amplitude or phase, change over time); deterministic (data can be predicted exactly) or stochastic (data are only partly determined by past values and successive values have to be described with a probability distribution).

While estimating the parameters of biological oscillations, several authors have used empirical procedures that include eye fitting, approximations to sine curves, spectral analysis (Cambras & Diez-Noguera, 1988), Fast Fourier transform (FFT) (Cambras & Diez-Noguera, 1988; Araujo & Marques 1996; Wang & Brown, 1996), maximum entropy spectral analysis (Dowse & Ringo 1989; Mormont et al., 1996), autocorrelation, Enrights (chi square) periodogram method (Siegel, 1980; Araujo & Marques, 1996), Sokolove-Bushell's Q statistic (Refinetti, 1991), linear regression of onsets, inter onset averaging, acrophase counting and semi parametric periodic spline function (Araujo & Marques, 1996). Mathematical models of a given biological rhythm (Edmunds Jr, 1983; Diez-Noguera, 1994; Pedersen & Johnsson, 1994; Scheper et al., 1999), are close approximations of the temporal process. Geometric methods such as phase plane plot aid in visual representation of the dynamics (Li & Goldbeter, 1989; Goldbeter & Moran, 1988). These methods are based on several assumptions and prerequisites such as regular sampling, prior knowledge of range of period, detrending and so forth, and between them, the results are often inconsistent. The crux of using an unambiguous method for analyzing biological time series data lies in the fact that the period derived by a method may considerably alter our perception and inference of the underlying biological rhythm and therefore, the choice of an appropriate technique is crucial. Several authors have tried to compare the available methods of time series analysis (Refinetti, 1993; Klemfuss & Clopton, 1993; Mormont et al., 1996). Some of their results indicate the necessity to collate several methods for computation of the period. However, the choice of the method of analysis depends considerably on the nature of data such as its length, multiplicity of frequency components, noise, effects of smoothening apart from experimental and sampling error.

Often, sampling a biological rhythm yields data sets with no apparent temporal

pattern. In such circumstances, the analysis needed to extract a pattern in the oscillation can be simplified if there is a simple and unambiguous mathematical approach to this problem. In this paper, we describe our effort towards developing a simple method that does not warrent any preconditions such as prior knowledge of period, scale of measurement, regularity of sampling, detrending and continuity of data, imperative.

Mathematical Approach

Our aim was to derive a simple mathematical procedure to isolate and quantify the significant periodicit(y)ies in a biological time series. Consider a variable $y(i) = y_1, y_2, y_3 \ldots y_n$, that is sampled at time, $t(i) = t_1, t_2, t_3 \ldots t_n$. A time series plot can be obtained by joining the points, (t_1,y_1), (t_2,y_2), $(t_3,y_3) \ldots (t_n,y_n) \ldots$.

The waveform is characterized by determining the time of occurrence of the extremes in amplitude (the 'local' maxima and minima). Let y_k, y_l and y_m define three successive amplitudes of the waveform, sampled at time t_k, t_l and t_m, respectively. The inequality, $y_k < y_l$, *and* $y_l > y_m$ holds true for all the local maxima (for three successive points). Let y_p, y_q and y_r define three successive amplitudes in the waveform. The inequality, $y_p > y_q$ *and* $y_r > y_q$ is true for all the local minima (for three successive points). Similarly, a local maxima/minima can be determined for an odd number of successive points along the evolution of a time series. Therefore, in principle, if a local maxima/minima exists, it is located halfway along an odd number of successive points in the time series. Therefore, if we consider using a window of size 'w' = 5 for scanning a time series that spans from t_1,y_1 to t_7,y_7, then the maxima/minima for the overlapping points t_1,y_1–t_5,y_5, t_2,y_2–t_6,y_6, t_3,y_3–t_7,y_7 could be t_3,y_3 or t_4,y_4 or t_5,y_5, respectively, or there could be none at all if the maxima/minima does not lie right in the middle for that segment of time series.

The x coordinate of each of these maxima and minima represents its occurrence with respect to time and we called it the 'time tag'. A time series therefore yielded two sets of time tags, one for each maxima and minima. In principle, the time series can be resolved to an array of time tag outputs that represents the occurrence of the critical temporal events (maxima/minima) over several time scales (increasing window sizes). The absolute value of differences between time of occurrence of successive maxima or minima was used as an appropriate means of estimation of the possible period of the signal. We relied on these numerical differences for further analysis. A critical aspect of our method is that a moving window scans the entire time series and this process was iterated over increasing window size, 'w'. To begin with, the entire length of the time series was scanned with a window size of 3 (three points at a time) until the end of the time series. Window size of 3 is minimum, since fewer than three points cannot be examined for locating local minima. The same process was then iterated over window size of 5, 7 and so on, awaiting that size of the window that yielded a single, global maxima. In principle, if a significant pattern is present in a time series, it should be detectable consistently as the window size is increased. It is worthwhile to note that as the window size is increased (and a longer

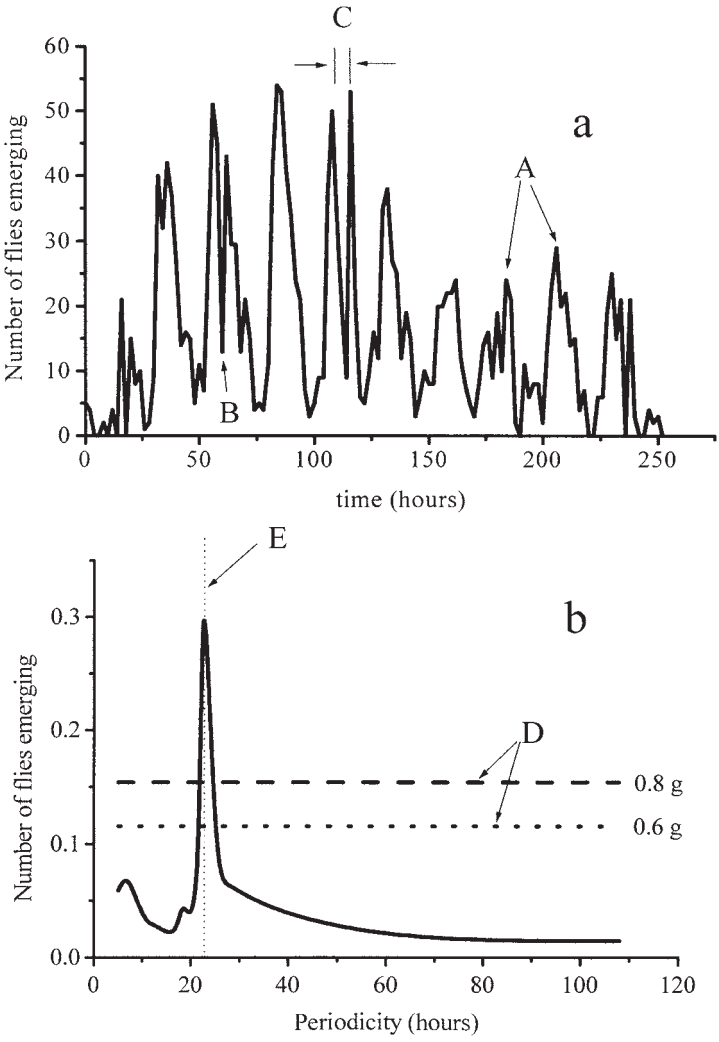

Figure 1a–b. A Step-wise approach for the computation of significant periodicit(y)ies in the time series. *Step 1*: As a first step the critical points, maxima (A) or minima (B) were located along the evolution of the time series. This operation was repeated with a dynamically scaled window. *Step 2*: The differences between successive maxima or minima (C) were computed to estimate the cycle to cycle periods. *Step 3*: The normalized periodogram was examined for the critical frequency (D) using Siegels' modification of the Fisher test to locate the significant frequency (E).

segment of time series is examined) the local maxima/minima lose their connotation as they might be transcended by other maxima/minima (Fig. 2a–c).

In summary, the calculation of differences between successive phase reference points (maxima/minima) in the time series was repeated for each time tag output

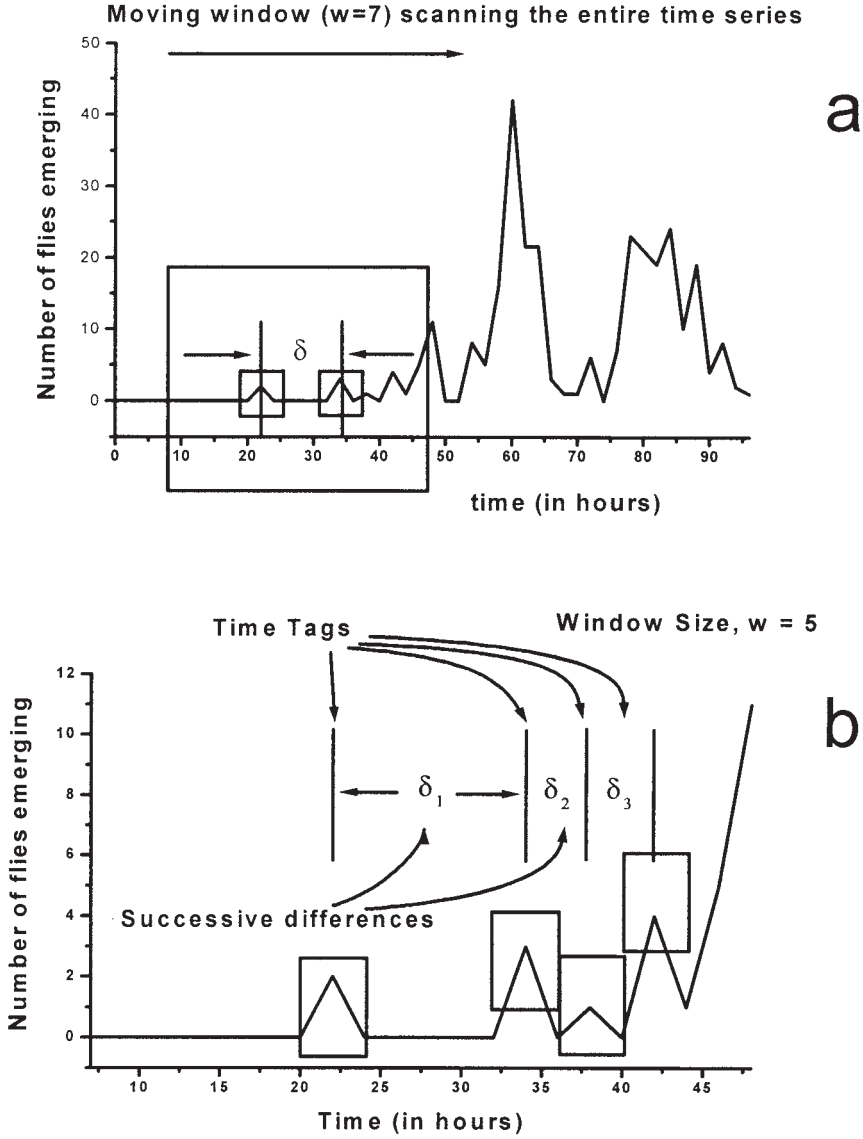

Figure 2a–c. (a) A moving window of ($w = 7$ in above example), scanning the time series starting from window size $w = 3$. It locates the maximum (or minimum) only when it is the middle of the window but not otherwise. Therefore, window size is always an odd number. A portion of time series is magnified for clarity and annotated in Figure 2b. (b) The time tags denote x intercept of the maxima/minima. Successive differences between time tags (δ_1, δ_2, δ_3, . . . δ_n) are computed. The process is iterated for window size $w = 3$ to that value of w that results in one global maxima/minima. In this example, notice that using a window size $w = 5$ results in four local maxima. (c) Increasing the window size to $w = 7$ decreases the number of maxima detected (e.g., maxima 3, when $w = 5$ is not detected when $w = 7$).

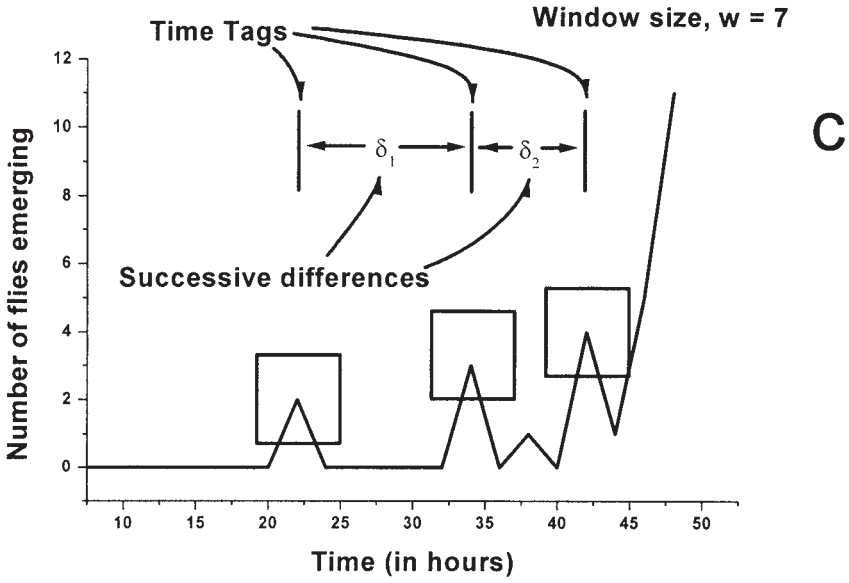

Figure 2a–c. Continued

starting from a window size of 3 until that window size which yields a global maxima/minima. The global maxima or minima is actually a numerical artifact of the biological time series and does not impart any significance to the time series or the computation, otherwise. Differences of time tags across all window sizes were clustered together and represented as a periodogram. The periodogram was normalized by representing the magnitude of the difference (that represents the period) as x-coordinate and the relative contribution of each frequency as the y coordinate. The likelihood that such a difference is a plausible significant frequency was augmented when it was sustained over increments in window size (i.e., several time scales). Among others, this repetition precludes chance as a factor in yielding a 'difference' and its subsequent classification as a possible frequency. In contrast, those differences that are ill sustained typically arise due to minor fluctuations that exist over shorter time scales. Noise is predominant with a smaller window size. Therefore, increasing the window size is accompanied by a reduction in noise along with an emergence of those differences that are most likely to represent the significant periodicities.

Short-term fluctuations in biological rhythms can result in differences that may be numerically close yet discrete and independent. Their independence is arbitrary; therefore the periodogram may not truly represent the original pattern due to a leakage of frequencies (translation of power from a discrete spectral peak to adjacent frequencies above and below the true frequency). This usually arises due to discontinuous sampling of the data. A weighted moving average of the periodogram could

be used as a remedial measure to obviate the leakage of frequencies and validate their contribution.

Critical frequency

Fisher's test is considered as a powerful and decisive test in detecting the periodicities that exist at a single frequency (Fisher, 1929). However, this method may not be suitable for complex biological rhythms with coexistent periodicities. For identifying significant frequencies in a time series with multiple periodicities, Siegel's modification of Fisher's test is used (Siegel, 1980). We delineated the frequency components of the periodicity by defining a critical frequency $g = (1 - 0.05^{1/(n-1)})$ where n is the number of independent differences. This value g is then expressed as percentage and is depicted by a line, $y = g$, parallel to the x-axis. All the frequencies above g were considered significant. Siegel's method for estimation of significant contribution has been used in many publications in chronobiology (Sheeba et al., 1999; Koilraj et al., 1999; Koilraj et al., 2000; Sheeba et al., 2001a,b).

Steps of the program

We first examined the time series for the maxima and minima as described. The time tag output file of a set of maxima was then processed to compute absolute values of their successive differences. This operation was repeated over all odd values from $w = 3$ to that value of w which elicits one global extreme. The frequency of occurrence of each difference was transformed into a periodogram that carries the option of being normalized by using a weighted moving average in order to justify the contribution of individual periodicities (periodogram). For the points t_1,y_1, t_2,y_2, t_3,y_3 . . . the normalized x coordinate for the point (t_1,y_1) for a window of three points, is calculated by the equation:

$$\left\{\frac{t_1y_1 + t_2y_2 + t_3y_3}{y_1 + y_2 + y_3}\right\}$$

The critical frequency was computed using Siegel's modification factor λ, in the equation $g = \lambda\,(1 - [0.05^{1/(n-1)}])$; where λ can attain values 0.6 or 0.8 (Siegel, 1980). The value of λ is taken based on the number of distinct periodicities constituting the periodogram. Periodicites that contributed more than the critical frequency were considered significant.

Relevant portion of the C code

Relevant portions of the C code, which we wrote to execute the program, are given below.

```
int data [MAX]; int diff [MAX];
float diff2 [MAX]; float i2 [MAX];
```

```
int lastWindow = 0; int lastMaxIndex = 0;
void CheckMax (int window, int base)
{int i; for (i = 1; i <= window/2; i++)
if ((data[base+window/2-i] >= data[base + window/2]) ||
(data[base+window/2+i] >= data[base + window/2]))
return; if (lastWindow ! = window)
{lastWindow = window; lastMaxIndex = base + window/2;
return;} else {diff [base + window/2 - lastMaxIndex]
++;
lastMaxIndex = base + window/2;}}
float GetSiegel (int i)
{float Siegel[50] = {.684, .684, .684, .684, .684, .616,
      .561, .516, .477, .445, .445, .445, .445, .445,
      .335, .335, .335, .335, .335, .270, .270, .270,
      .270, .270, .228, .228, .228, .228, .228, .198,
      .198, .198, .198, .198, .175, .175, .175, .175,
      .175, .131, .131, .131, .131, .131, .131, .131,
      .131, .131, 131};
if (i <= 0) return 1; if (i >= 50)
return 0.131; return Siegel[i];}
int main(int argc, char *argv[])
{FILE *fptr = fopen (argv[1], "r");
int i, j, windowSize;
int sizeData, numFreq, SigmaFreq;
for (i = 0; i < MAX; i++) {diff[i] = 0; data[i] = 0;}
i = 0; do {if (feof (fptr)) break;
fscanf (fptr, "%d", &data[i]); if (feof (fptr))
break; i++;} while (1);
sizeData = i; fclose (fptr);
for (windowSize = 3; windowSize < sizeData; windowSize
+ = 2)
{for (j = 0; j <= sizeData-windowSize; j++)
CheckMax (windowSize, j);}
numFreq = 0; diff2 [0] = diff [0];
diff2 [sizeData – 1] = diff [sizeData–1];
if (diff2 [0] > 0) numFreq++;
if (diff2 [sizeData–1] > 0) numFreq++; SigmaFreq = 0;
for (i = 1; i < sizeData – 1; i++)
{i2 [i] = diff [i–1] * (i–1) + diff [i] * i + diff
[i+1] * (i+1);
if ((diff[i–1] + diff[i] + diff[i+1]) > 0)
i2 [i] = i2 [i] / (diff[i–1] + diff[i] + diff[i+1]);
else i2 [i] = i; diff2 [i] = (diff [i–1] + diff[i] +
diff [i+1]) / 3;if (diff [i] > 0) numFreq++;
SigmaFreq += diff2 [i];}
```

```
SigmaFreq += diff2 [0] + diff2[sizeData–1];
printf ("SigmaFreq = %d, numFreq = %d\n", SigmaFreq,
numFreq);
printf ("Cutoff freq = %f\n", 0.6 * GetSiegel(numFreq)
* SigmaFreq);printf ("Smoothened frequencies are:
\n\n");
for (i = 1; i < sizeData–1; i++)
if (diff2 [i] > 0) printf ("%f \t %f\n", i2 [i], diff2
[i]);
printf ("\n\n\n Raw frequencies are \n\n\n");
for (i = 1; i < sizeData–1; i++)
if (diff [i] > 0) printf ("%d \t %d\n", i, diff [i]);
printf ("\n\n\n Significant frequencies are \n\n\n");
for (i = 1; i < sizeData–1; i++) if (diff [i] > 0.6 *
GetSiegel(numFreq) * SigmaFreq) printf ("%f \t %f\n",
i2 [i], diff2 [i]); getchar(); return 0;}
```

Examples of data analyzed

We analyzed the experimental data from oviposition (egg laying) and eclosion rhythm (adult flies emerging from the pupal case) in the fruit fly, *Drosophila melanogaster* (Figs. 3, 4) and simulated data (Fig. 5a–b).

The simulated time series data were either white noise (Fig. 5a1) or a time series obtained by overlapping time series of periodicities 7, 11, 17 and 23 h along with white noise (Fig. 5b1).

The critical frequency was incorporated in the periodogram. For the x-axis to represent the true periodicity, the time scale of this periodogram has to be rescaled to correspond to the time interval of sampling. The periodogram of the experimental data shows a significant pattern of 24 h present in the time series (Figs. 3, 4). The periodogram of the simulated data revealed four periodicities of value 6, 11, 17 and 23 h, contributing significantly in the time series (Fig. 5b2). When a time series consisting of white noise was used in our analysis (Fig. 5a1), the periodogram (Fig. 5a2) revealed significant contribution from very short periodicities of magnitude 1 and 2 h. The analysis of the simulated time series using multiple periodicities along with white noise (Fig. 5b1), successfully isolated all of the periodicities (Fig. 5b2). However, when a time series obtained by pooling two time series of equal amplitude, one with shorter periodicity and longer periodicity, was analyzed, the periodogram revealed that the shorter periodicity always contributed significantly. The longer periodicities were identified as significant patterns only when their amplitude was higher than the shorter periodicity.

Discussion

In analyzing a biological time series, several approaches have been adopted. The conclusions drawn thereafter depend very much on its mathematical basis. In our

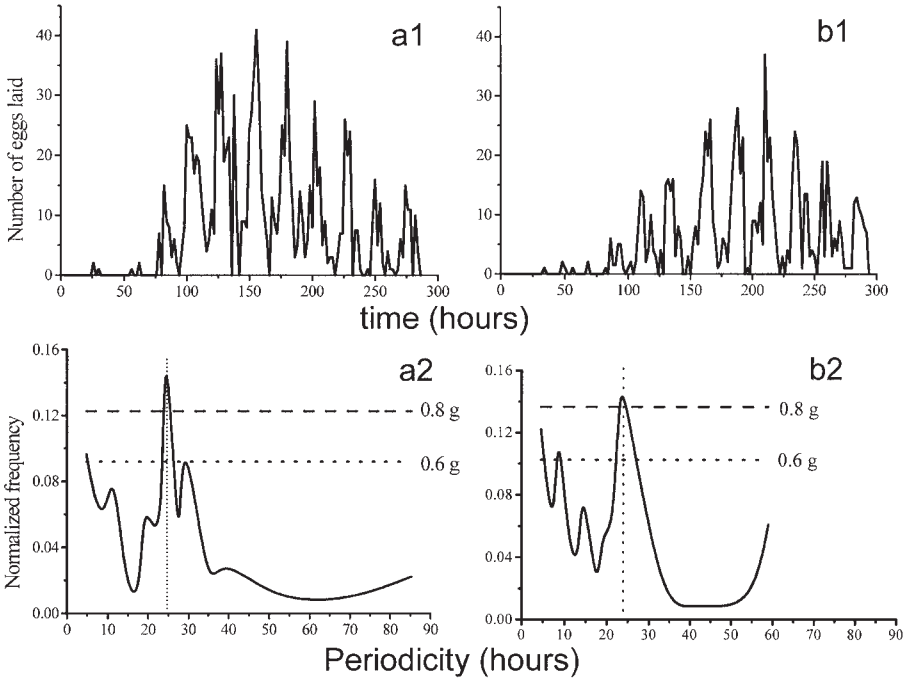

Figure 3a1, a2, b1, b2. Representation of the time series data of *D. melanogaster* oviposition rhythm. The data was collected at fixed intervals of 2 h. The number of eggs is indicated on the Y axis. The time series (a1 and b1) does not explicitly exhibit a pattern. The normalized periodograms (a2 and b2) are used to delineate significant periodicit(y)ies within this time series.

approach, we have tried to overcome some of the shortcomings of the conventional methods of time series analysis and tried to address some issues that considerably influence the outcome of the data analysis. We believe that the problem of aliasing can be avoided by using this method where we use sampling of data at a sampling interval much shorter than the expected period. Experimental data can be periodic, may have noise that obscures periodicity or may have white noise alone. Analysis of periodic data is relatively simple when compared to periodic data in association with noise. In the presence of noise, as we increase '*w*', the number of 'local' maxima/minima decreases. This decrease persists as an increasing number of (odd) successive points which is examined for the presence of a local maxima/minima. Eventually, we arrive at a value of '*w*', which yields one 'global' maxima/minima for the entire time series; a result obtained as an artifact of variations in amplitude.

Several methods are used to analyze biological time series and most of them have several limitations. A Fourier Transform converts equally spaced information from

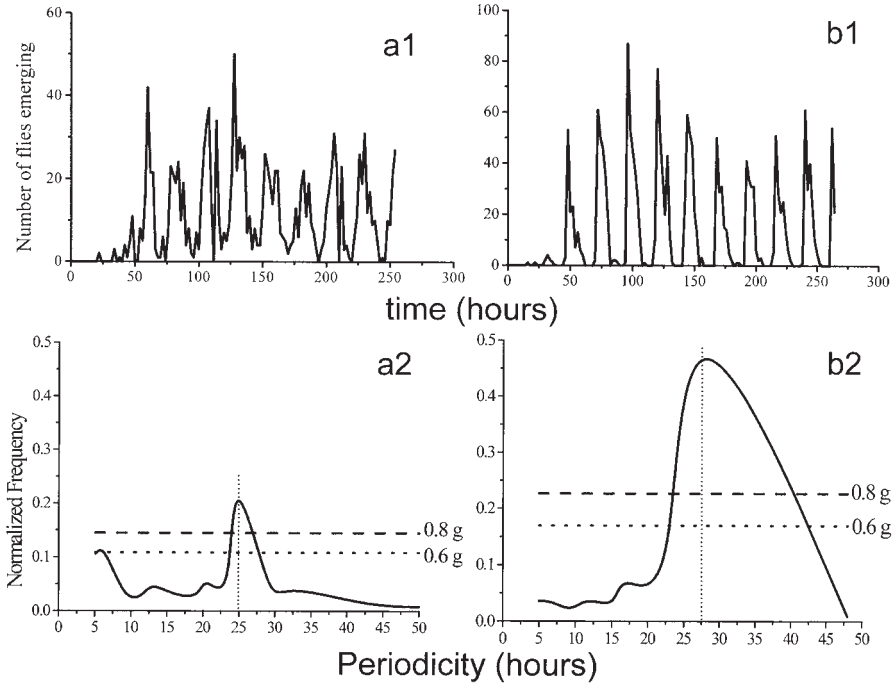

Figure 4a1, a2, b1, b2. Representation of the time series data of *D. melanogaster* eclosion rhythm. The data was recorded at fixed intervals of 2 h. The number of enclosed flies was plotted on the Y axis. The time series (a1 and b1) does not explicitly exhibit a pattern. The normalized periodograms (a2 and b2) are used to delineate significant periodicit(y)ies within this time series.

the time domain to the frequency domain, and vice versa (Van Dongen et al., 1999). It has been used extensively by biologists for the detection and analysis of frequency components of their interest. However, this method cannot be used when the sampling of biological variables is irregular (Van Dongen et al., 1999). Spectral analysis is used when there is no prior knowledge to suspect a specific period. Like the Fourier transform, it relies on a regularly sampled time series to be decomposed into a sum of trigonometric periodic functions with different frequencies, amplitudes and phases (Van Dongen et al., 1999). Another disadvantage of the spectral analysis method is that preprocessing of the data such as insertion of means for missing data points or removal of a linear trend must be performed in order to obtain reliable estimate of the period. Harmonic analysis includes various techniques, in which the time series is decomposed into a number of periodic components of sinusoidal form. Nevertheless, before using a harmonic analysis technique, any trend in the data needs to be removed. Lomb Scargle Periodogram uses linear least-squares regression method for

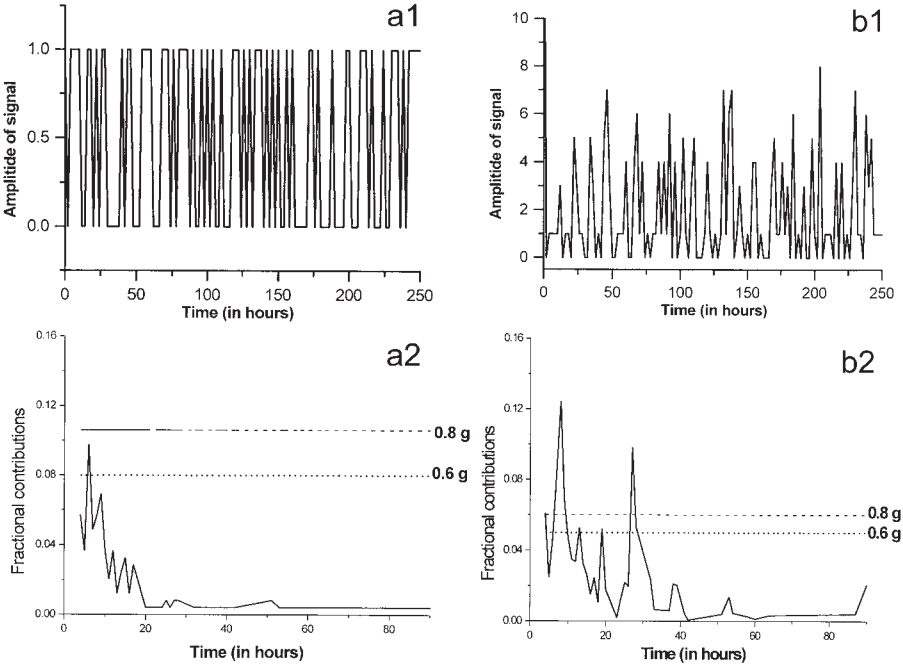

Figure 5a–b. (a1) Time series data consisting of white noise and its periodogram (a2), revealing significant contributions from very short periodicities. (b1) Simulated time series data obtained by pooling five time series data of periodicities 6, 11, 17 and 23 h. A 250-h long time series was simulated and then subjected to analysis using our method. (b2) Periodogram for the time series data simulated using four periodicities (b1). The periodogram reveals significant contributions from all of the periodicities.

transforming unevenly spaced data to a sine or cosine series of different frequencies. However, this periodogram has to be normalized for better representation of the original time series (Scargle, 1982; Ruf, 1999). The autocorrelation function may also be used to detect deterministic components masked by a random background (Blackman & Tukey, 1958). Nevertheless, noise has to be reduced for better assessment of the period using the autocorrelation function. Sine curve fits involve decomposition of the waveform and its representation as a sine wave (Blackman and Tukey, 1958; Hickey, 1984). While biological oscillations can be periodic, not all of them can be expressed as a sine wave (Mercer, 1960). Cosinor method is a frequently used parametric method that provides a least-squares fit and consists of one or more cosine curves with one or more periods expected to represent the data (Nelson et al., 1979; Bingham et al., 1984). The chi-square method for computing the periodicities determines the intensity (power) of rhythms in the time series for a range of periods (Enright, 1965; Ruf, 1999). In this method, significant frequencies are identified on the basis of a threshold. However, the chi-square test does not take into account the ordered structure of the data, and it does not distinguish between irregular fluctua-

tion and a smooth cyclical pattern (James, 1976; Walter, 1977). It is also designed for equally sampled data and not unequally spaced time series (Ruf, 1999). The Maximum Entropy Spectral Analysis (MESA), a parametric method, also makes use of the autocorrelation function that is approximated by an autoregressive process (Dowse & Ringo, 1989). A study found MESA to be no better than FFT or autocorrelation in estimating period (Klemfuss & Clopton, 1993). Noise is inherent in biological time series. To reduce noise, preprocessing (smoothening) of the data is frequently resorted to. While doing so there may be a loss of vital data points at the cost of obtaining regularity of the data structure. Eye fitting, a commonly used procedure in biological time series analyses, is an empirical estimation of the periodicity that is limited in terms of repeatability.

With our method, we satisfactorily analyzed data without smoothening. While computing the periodicities, we assigned a scale that was independent of the original sampling time and interval. This aids the evaluation of any length of sampled time series, irrespective of the scale of its measurement. This makes our analysis flexible for the elucidation of multiple periodicities that might exist in a biological time series. Our method focuses on the temporal organization of the rhythms and, consequently, the variation in the amplitude of the waveform does not influence the outcome of our computation. However, the detection of the critical points is of utmost importance. Erroneous or missing data may alter the periodogram significantly, if it were to correspond with a nominal peak. Missing data elsewhere along the time series may be accompanied by a flattening of the periodogram. Computing successive differences between all the time tags and using an overlapping window with a variable size offers an elegant solution for minimizing the errors due to missing data. The margin of error of estimation of the resultant periodicities is of the order of the sampling interval (regularly sampled data) or lesser than the largest sampling interval (for irregularly sampled data). Periodicities may be estimated by normalizing this periodogram by computing a weighted moving average. The weight for each case is proportional to its contribution to the cumulative frequency. To account for inherent biological variability we can use a window of three points to compute the weighted moving average, for this operation (Sollberger, 1965). Siegel's modification of the Fisher's test incorporates a variable multiplication factor, λ to distinguish significant frequenc(y)ies that would have otherwise been neglected.

If two or more successive observations are equal and they otherwise correspond to a nominal 'peak', our method fails to detect it as we rely on the determination of a 'local' peak for an odd set of points. Nevertheless, the user has the option of computing the periodicities using the local minima instead of the maxima. However, in a case where the waveform has no apparent maxima/minima, one can estimate periodicities by incorporating alternate points of reference, for example the points of inflection. Hence, the algorithm for the detection of critical points should then be modified accordingly.

Our method worked very well with time series with single frequency embedded in white noise. It successfully identified the pattern present in the time series without any smoothing or detrending. Our method isolated all of the frequencies from a pooled time series consisting of several time series with multiple frequencies em-

bedded in white noise. However, the analyses of white noise alone resulted in a periodogram, which revealed significant contribution from very short periodicities. This demonstrates that our methods works equally well with experimental data with single periodicity and simulated time series data with multiple periodicities.

Conclusions

Our method can be used to analyze time series data that has been sampled at irregular intervals. It is simple but lucid, flexible and unambiguous since each of the few operations has a simple mathematical basis. It is an objective method of estimation of the period with an option for the user to refine the periodogram by computing a weighted moving average. The method can also be tailored to use several other points of reference in a time series for computation of periodicities; for example maxima/minima (as we have described) or other points of inflection. In summary, our approach is a simple means of decomposing time series with compound periodicities without any prior assumption of the pattern(s) present. No pre-processing, detrending of data, is required to identify the significant pattern(s) present in the data. We can analyze time series of any length irrespective of the scale of measurement. This method can also be used to analyze irregularly sampled time series data and also time series data with missing points. We also feel that this method is flexible, because one can analyze time series data with various waveforms by using well-defined phase reference points.

Acknowledgements

We thank the Indian Academy of Sciences for its fellowship to AVR and the Jawaharlal Nehru Centre for Advanced Scientific Research for providing excellent infrastructure and experimental facility. We thank Srikanth BN for writing the C code to excute the program and Prof. M.K. Chandrashekaran for his encouragement given during the work. A portion of this work was funded by Indian National Science Academy, New Delhi.

References

Araujo JF, Marques N (1996): Circadian and ultradian rhythms of drinking behavior of albino rats maintained in constant darkness. *Braz J Med Biol Res 29*: 1369–1372.

Bingham C, Cornelissen G, Halberg E, Halberg F (1984): Testing period for single cosinor: extent of human 24-h cardiovascular 'synchronization' on ordinary routine. *Chronobiologia 11*: 263–274.

Blackman RB, Tukey JW (1958): Measurement of power spectra. *Bell Sys Tech Journal 37*: 185–282; 485–569.

Cambras T, Diez-Noguera A (1988): Generational variability in the patterns of motor activity circadian rhythms in the rat. *A Rev Esp Fisiol 44*: 243–246.

De Boer RJ, Perelson AS, Kevrekidis IG (1993): Immune network behavior–II. From oscillations to chaos and stationary states. *Bull Math Biol 55*: 781–816.

Diez-Noguera A (1994): A functional model of the circadian system based on the degree of intercommunication in a complex system. *Am J Physiol 267*: R1118–1135.

Dowse HB, Ringo JM (1989): The search for hidden periodicities in biological time series revisited. *J Theor Biol 139*: 487–515.

Edmunds LN Jr (1983): Chronobiology at the cellular and molecular levels: models and mechanisms for circadian timekeeping. *Am J Anat 168*: 389–431.

Enright JT (1965): The search for rhythmicity in biological timeseries. *J Theor Biol 8*: 426–468.

Fisher RA (1929): Tests for significance in harmonic analysis. *Proc Royal Soc A 125*: 54–59.

Goldbeter A, Decroly O (1983): Temporal self-organization in biochemical systems. *Am J Physiol 245*: R478–483.

Goldbeter A, Moran F (1988): Dynamics of a biochemical system with multiple oscillatory domains as a clue for multiple modes of neuronal oscillations. *Eur Biophys J 15*: 277–287.

Hassnaoui M, Pupier R, Rehailia M (2000): A concordance method for analyzing categorical time series. An application for the search of periodicities. *Biol Rhythm Res 31*: 177–201.

Hickey DS, Kirkland JL, Lucas SB, Lye M (1984): Analysis of circadian rhythms by fitting a least squares sine curve. *Comput Biol Med 14*: 217–223.

Ishida N, Kaneko M, Allada R (1999): Biological clocks. *Proc Natl Acad Sci*, USA *96*: 8819–8820.

James WH (1976): The power of a test for seasonality of births with reference to schizophrenia. *Br J Psychiatry 129*: 94–95.

Klemfuss H, Clopton PL (1993): Seeking tau: a comparison of six methods. *J Interdiscip Cycle Res 24*: 1–16.

Koilraj J, Marimuthu G, Sharma VK (1999): Circadian rhythm in the locomotor activity of a surface-dwelling millipede *Syngalobolus sp. Biol Rhythm Res 30*: 529–533.

Koilraj J, Sharma VK, Marimuthu G, Chandrashekaran MK (2000): Presence of circadian rhythms in the locomotor activity of a typical cave-dwelling millipede *Glyphiulus cavernicolus*. *Chronobiol Internat 17*: 757–765.

Li Y, Goldbeter A (1989): Oscillatory isozymes as the simplest model for coupled biochemical oscillators. *J Theor Biol 138*: 149–174.

Mercer DMA (1960): Analytical methods for the study of periodic phenomena obscured by random fluctuations. *Cold Spring Harb Symp Quant Biol 25*: 73–86.

Mormont MC, De Prins J, Levi F (1996): Study of circadian rhythms of activity by actometry: preliminary results in 30 patients with metastatic colorectal cancer. *Pathol Biol 3*: 165–171.

Nelson W, Tong YL, Lee JK, Halberg F (1979): Methods for cosinor-rhythmometry. *Chronobiologia 6*: 305–323.

Pedersen M, Johnsson A (1994): A study of the singularities in a mathematical model for circadian rhythms. *Biosystems 33*: 193–201.

Refinetti R (1991): Use of chi square periodogram in the analysis of estrous rhythmicity. *Int J Biomed Comput 27*: 125–132.

Refinetti R (1993): Laboratory instrumentation and computing: comparison of six methods for the determination of the period of circadian rhythms. *Physiol Behav 54*: 869–875.

Robertson LM, Takahashi JS (1983): Circadian clock in cell culture: I. Oscillation of melatonin release from dissociated chick pineal cells in flow-through microcarrier culture. *J Neurosci 8*: 12–21.

Rosato E, Piccin A, Kyriacou CP (1997): Circadian rhythms: from behaviour to molecules. *Bioessays 19*: 1075–1082.

Ruf T (1999): The lomb scargle periodogram in biological rhythm research: analysis of incomplete and unequally spaced time series. *Biol Rhythm Res 30*: 178–201.

Scargle JD (1982): Studied in astronomical time series analysis: II. Statistical aspects of spectral analysis of unevenly spaced data. *Ap J 263*: 835–853.

Scheper T, Klinkenberg D, Pennartz C, van Pelt J (1999): A mathematical model for the intracellular circadian rhythm generator. *J Neurosci 19*: 40–47.

Sheeba V, Sharma VK, Chandrashekaran MK, Joshi A (1999): Persistence of *Drosophila* eclosion rhythm after 600 generations in an aperiodic environment. *Naturwissenschaften 86*: 448–449.

Sheeba V, Chandrashekaran MK, Joshi A, Sharma VK (2001a): A case for multiple oscillators controlling different rhythms in *Drosophila melanogaster*. *J Insect Physiol 47*: 1217–1225.

Sheeba V, Chandrashekaran MK, Joshi A, Sharma VK (2001b): Persistence of oviposition rhythm in individuals of *Drosophila melanogaster* reared in an aperiodic environment for several hundred generations. *J Exp Zoology 290*: 541–549.

Siegel AF (1980): Testing for periodicity in a time series. *J Am Stat Assoc 75*: 345–348.

Smaaland R (1996): Circadian rhythm of cell division. *Prog Cell Cycle Res 2*: 241–266.

Sollberger A (1965): *Biological Rhythm Research.* Amsterdam, Elsevier, pp. 187–188.

Takahashi JS (1991): Circadian rhythms: From gene expression to behavior. *Curr Opin Neurobiol 1*: 556–561.

Takahashi JS, Zatz M (1982): Regulation of circadian rhythmicity. *Science 217*: 1104–1111.

Van Dongen HP, Olofsen E, VanHartevelt JH, Kruyt EW (1999): Searching for biological rhythms: peak detection in the periodogram of unequally spaced data. *J Biol Rhythms 14*: 617–620.

Vanden Driessche T (1989): The molecular mechanism of circadian rhythms. *Arch Int Physiol Biochim 97*: 1–11.

Walter SD (1977): The power of a test for seasonality. *Br J Prev Soc Med 31*: 137–140.

Wang Y, Brown MB (1996): A flexible model for human circadian rhythms. *Biometrics 52*: 588–596.