Features of a Chronomics Analysis Toolkit (CATkit)

Lee Gierke C, Cornelissen G

University of Minnesota, Minneapolis, MN USA

UNIVERSITY OF MINNESOTA
Driven to Discover℠

Introduction

Chronomics, and the study of biological rhythms is a rapidly growing area of study. As of March 2015, a search of PubMed returned 73,672 journal entries containing the keyword "circadian" – an increase by half over the same search in 2006. As biologists, biochemists and others transition to the study of the role of biological oscillators, and how this vast network modulates genetic, molecular, physiological, and behavioral rhythms, there is increasing need for accessible tools to characterize rhythmicity, not only in general terms, but in quantitative terms. The Chronomics Analysis Toolkit (CATkit), an R package for analysis of periodicities in time series, is a free and open source suite of tools, especially suited to the often scarce, frequently noisy, biological data. R is a widely used environment for statistical programming, freely available from the Comprehensive R Archive Network, and runs on a wide variety of UNIX, Mac and Windows platforms. CAT provides visualization tools. CAT Cosinor gives quantitative assessment, by cosinor, of mean, amplitude and phase at an assumed period (or periods), with a measure of uncertainty for each parameter.

CAT Visualization Tools

Visualization tools allow inspection of raw data and a generalized characterization of periodicities present in the data. Inspection is an important step in determining whether data meet key assumptions for subsequent analyses, in that non-normal distributions, trends or other non-stationarities, and gaps in the data can signal a need for further or specialized treatment. CAT tools include: a moving average; an actogram at any selected period; autocorrelation and crosscorrelation; and a periodogram by Fourier analysis, Figures 1-4. Data must be equidistant, but gaps are interpolated linearly. Data can also be binned.

CAT Cosinor

An array of cosinor-based techniques provides quantitative results. Cosinor does not require equidistant data. Plots display the raw data, and the model parameters: MESOR, Amplitude, Acrophase, percent rhythm (PR) and P-value. In addition to the basic **single-component cosinor**, there is a periodogram-like **spectrum analysis** performed by cosinor, with the advantage that it is able to get estimates at frequencies intermediate to Fourier frequencies. Where data are non-stationary, a progressive analysis through successive sections of the data, called a **chronobiologic serial section**, identifies changing rhythm dynamics over time. A **gliding spectrum** generates a heat map of amplitudes over time and frequency for a 3-dimensional visualization of changes, Figures 5-8.

The **multiple-component cosinor** can model complex signals consisting of multiple sinus curves, Figure 9. Many physiological processes are better modeled by multiple cosines, than single. The **serial section** for non-stationary data can be performed with the multiple-component cosinor, as well.

CAT Visualization Output

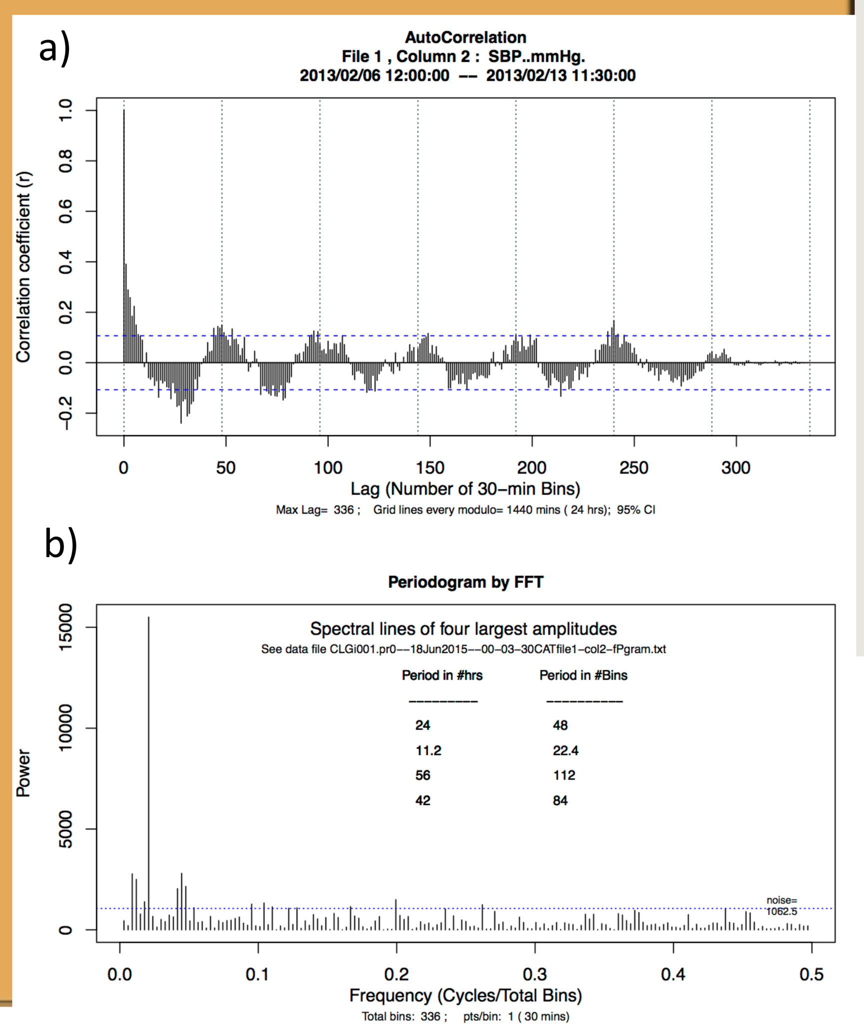

Figure 3: a) Autocorrelation of SBP with a maximum lag of N-1 shows peaks approximately every 24 hours. Horizontal dashed lines are the 95% CI. b) Periodogram of SBP also shows a peak at 24 hours, with additional peaks. Horizontal dotted line is the estimated white noise level, assuming normality.

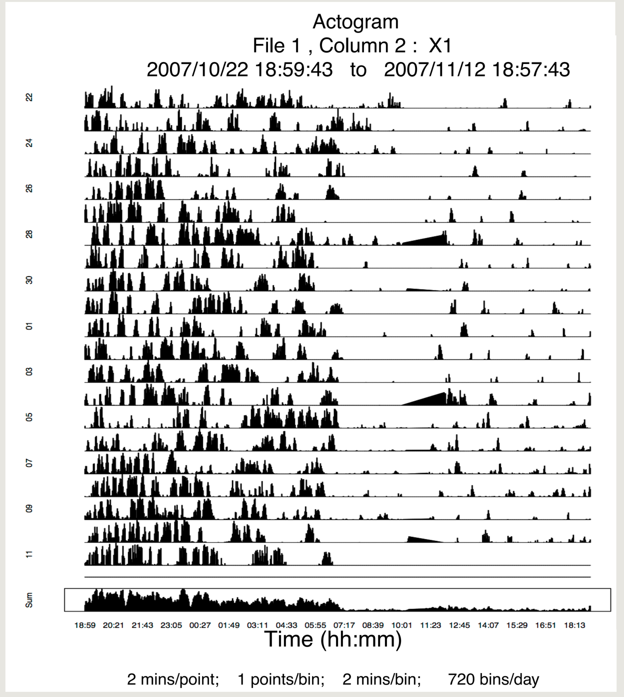

Figure 1: Actogram of murine activity data. A ragged, but distinct vertical alignment can be seen between high nightly activity levels, and lower daytime levels. Ramps caused by interpolation of large gaps are visible on days 7, 14, and 20, all occurring around 11:00 a.m. The bottom row sums all rows to highlight daily patterns.

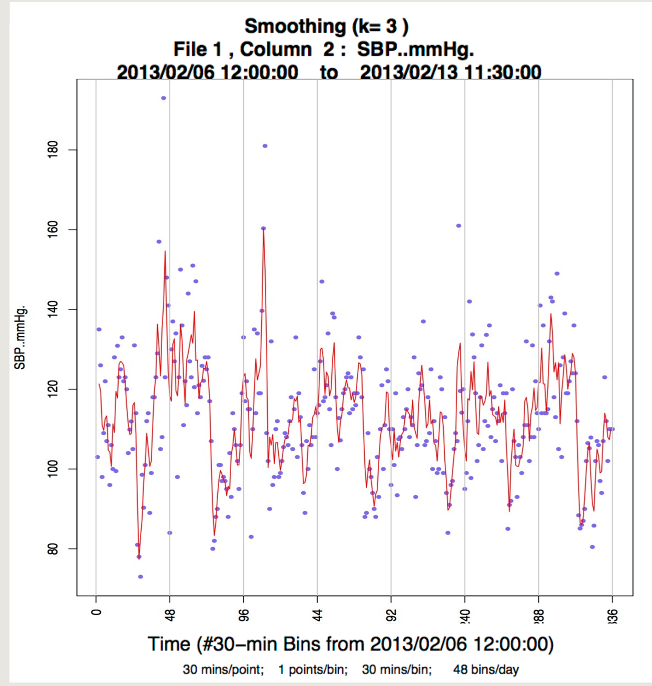

Figure 2: Smoothing – a moving average of $2k+1$ points reduces jitter, making the general pattern more visible. Seven days of systolic blood pressure (SBP) data (dots) are overlaid by a moving average (red line). File, column, date range and bin count are displayed (as on all graphs). k is selectable by parameter.

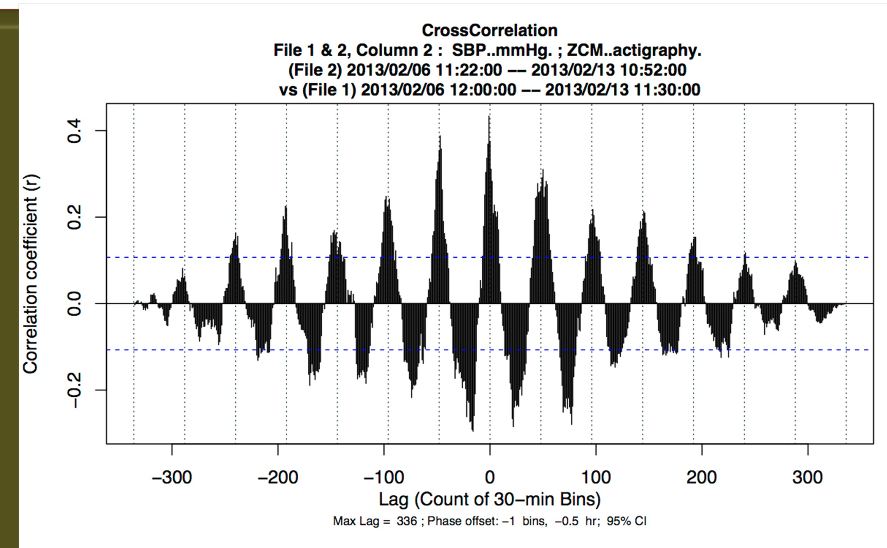

Figure 4: Crosscorrelation of 7 days of SBP and actigraph (ZCM) data show persistent peaks every 48 bins, or 24 hours, confirming a 24-hour periodicity. The central peak at 0 indicates the two time series, SBP and ZCM (a measure of activity) are in phase. A central peak shifted right would indicate file 1 was peaking after file 2.

CAT Cosinor Output

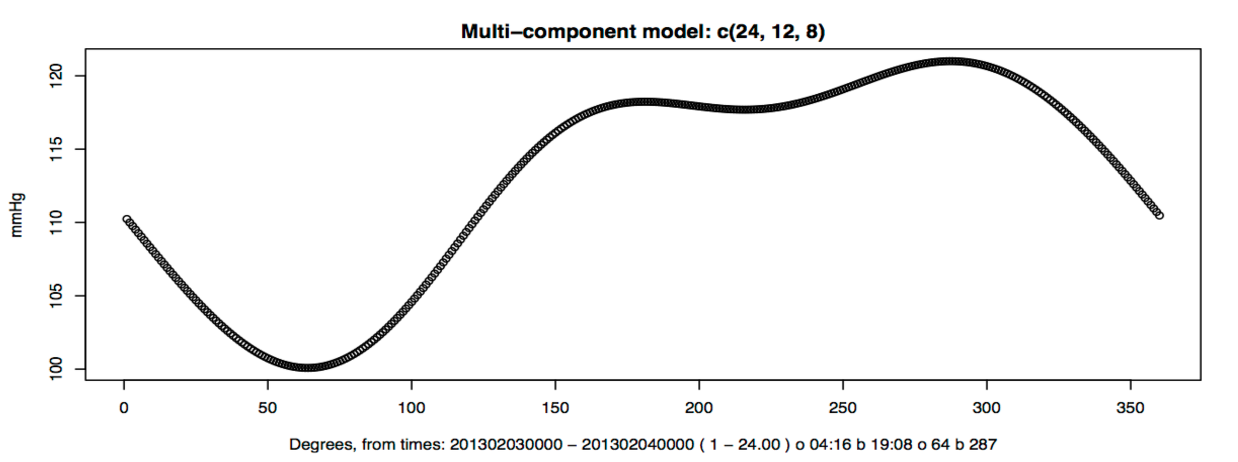

Figure 9: Multi-component cosinor modeling one cycle of a seven-day SBP record: three cosines make up this curve, with periods of 24, 12 and 8 hours.

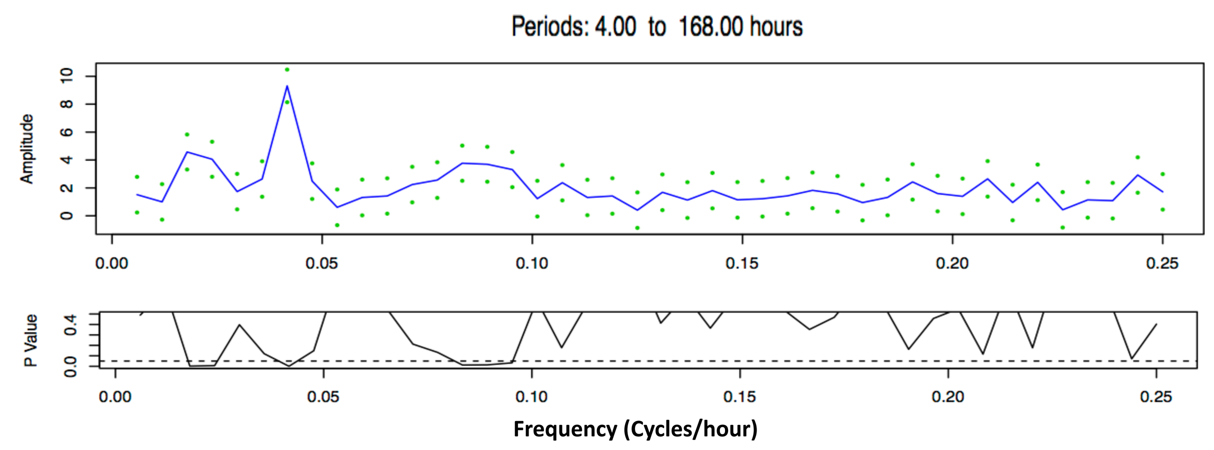

Figure 6: Summary graphs for Cosinor Least Squares Spectrum: Amplitude (dots give standard error) and the P–value for a zero-amplitude hypothesis test, across frequencies. Dashed line corresponds to P=0.05. Fundamental period of 168 hours set by user – although there is only 166.5 actual hours of data.

Rhythmometric Summary

Err	Y	Time Pts	hours from RefDateTime	#Pts	Period in hrs	PR	P	Mesor	Mesor SE	Amp	Amp SE	PHI	PHI SE
	4	201302061200 - 201302131030	0.0 - 192.0	1 - 327 (327)	24.00	16.5800	<.001	112.88220	0.8249	9.31129	1.169	-245.2	7.2
	4	201302061200 - 201302131030	0.0 - 192.0	1 - 327 (327)	12.00	2.7049	0.0118	112.91203	0.8909	3.77034	1.260	-292.8	19.1
	4	201302061200 - 201302131030	0.0 - 192.0	1 - 327 (327)	8.00	0.0312	0.9506	112.87997	0.9032	0.40512	1.273	-86.5	181.3

Data Summary

Err	Y	Time Pts	hours from RefDateTime	#Pts	N	Low	High	Mean	Median	Mode	dT	s.d.
	4	201302061200 - 201302131030	0.0 - 192.0	1 - 327 (327)	327	73.00	193.00	112.8746	112.00	114	306.00	16.28107

Figure 5: CAT Cosinor Rhythmometric and Data Summaries for single trial periods of 24, 12 and 8 hours.

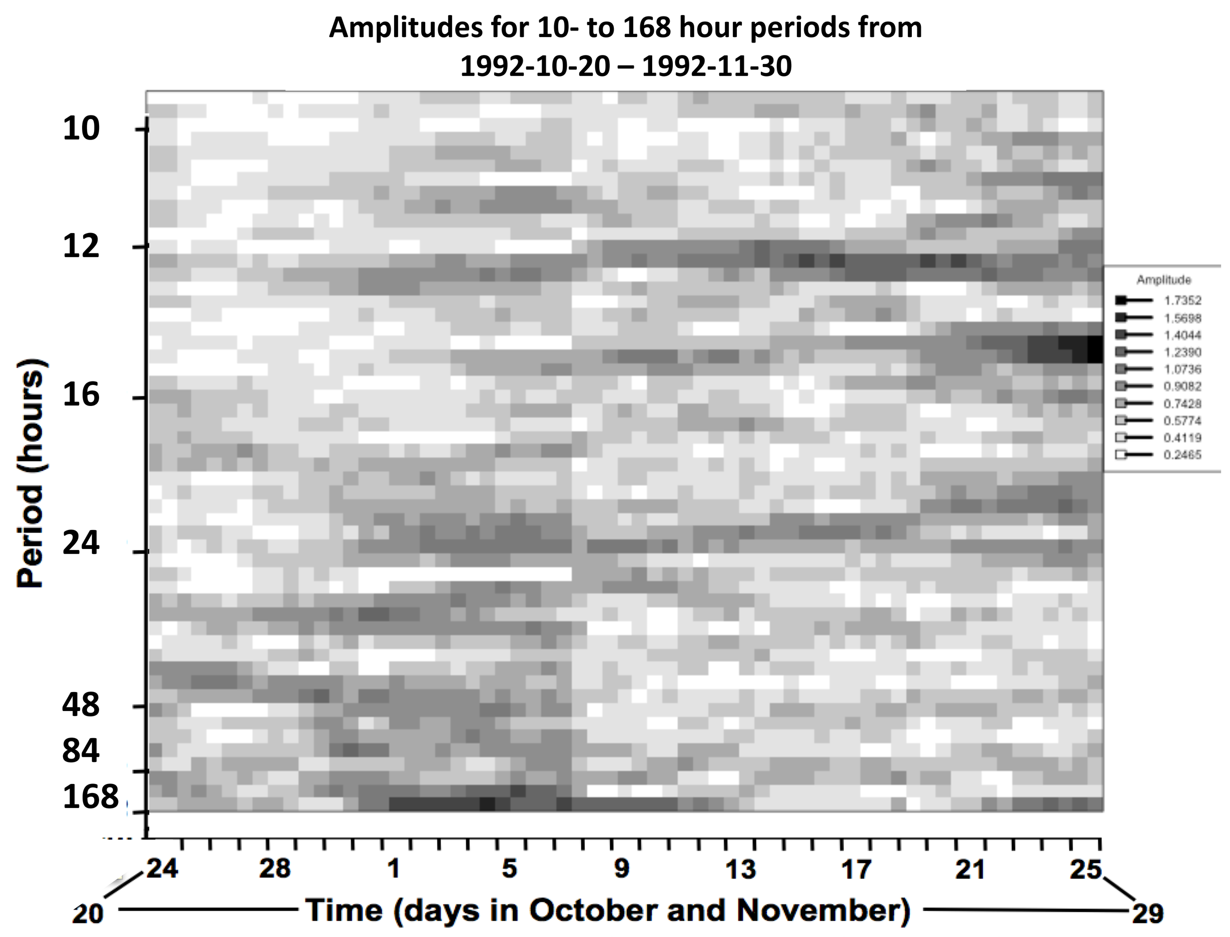

Figure 8: CAT Cosinor Gliding spectrum analysis of heart rate over the first 40 days of a newborn. At birth, several periods longer than 24 hours are present (dark areas), gradually changing to predominantly 24- and 12-hour periods. A 252-hour interval is moved by 12-hour increments, for the full 40 days. Data from Y Watanabe.

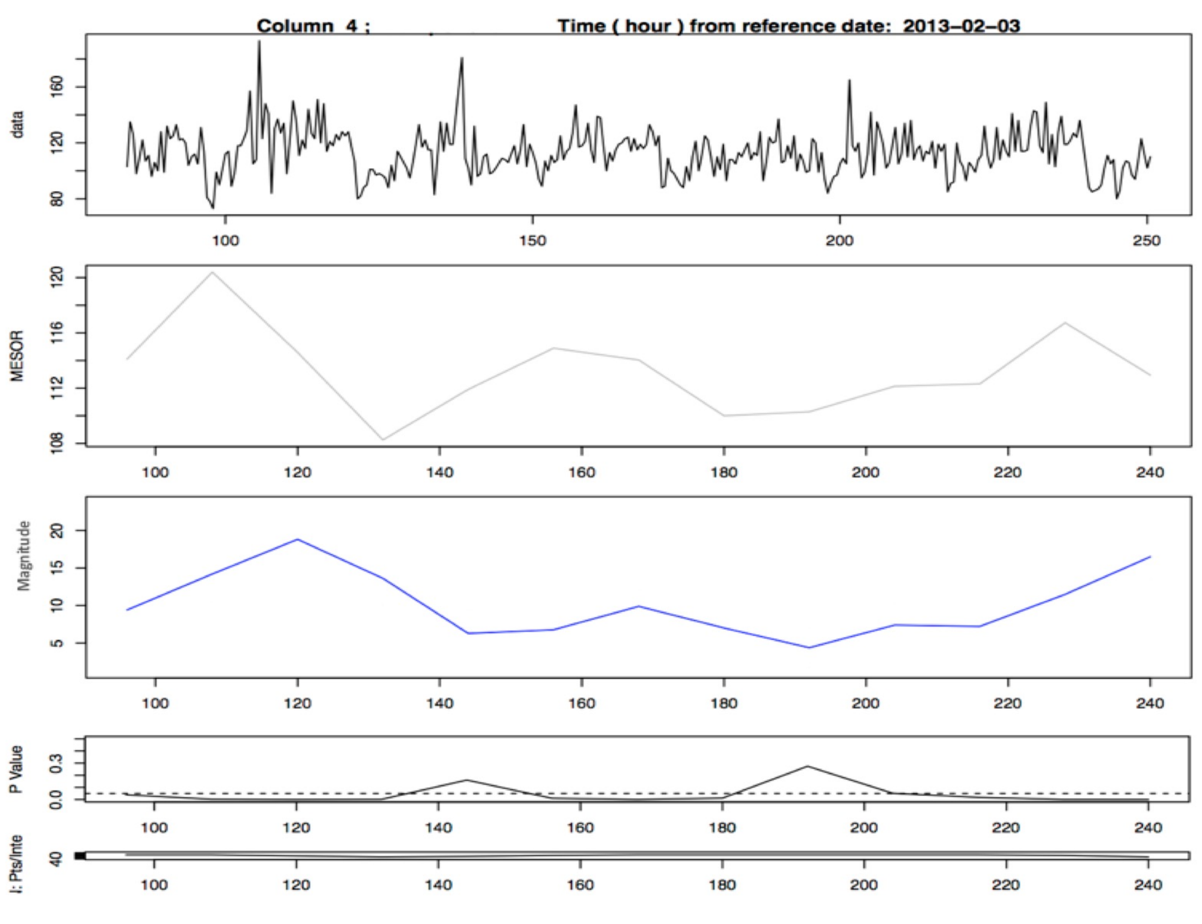

Figure 7: Serial Section: A multiple component cosinor is performed progressively across 240 hours of SBP data. Plots of Data, MESOR, Magnitude, P-value, and N data points per span, at each successive 24-hour span, incremented by 12 hours. Plots of Bathyphase and Orthophase are not shown.

Conclusion

The R environment affords CATkit and its functions extensibility through scripting options, direct source edits and a large library of R functions. Development of CATkit continues, and additional functions, such as population-mean cosinor, are planned. Detailed instructions for running CATkit, with examples are available at: http://z.umn.edu/CATkit. Further information can also be found regarding fundamental assumptions behind the cosinor calculation, use and interpretation of CATkit functions, and time-series analysis resource material. Quantitative rhythm characterization is a fundamental tool for studying biological rhythms. Reliable, easy to use tools, and adequate support for biologists moving in this direction are needed to facilitate advancement of the field.

Funded by University of Minnesota Retirees Association Professional Development Grant