ChronoStar

Version 1.0.0

Stephan Lorenzen

December 9, 2011

Abstract

ChronoStar is a program for analysis of time series data describing damped oscillations overlaid by a strong nonlinear trend. A typical but in no way exclusive application is the analysis of biological time series with decreasing circadian oscillations, e. g. bioluminescence data. The software can run both in a GUI environment with visualization and real-time adjustment of fitting parameters as well as in a batch mode to work on several files. The program is available for download free of charge for academic institutions on https://spideroak.com/browse/share/SciBioSoft/we-love-science. However, any publications using the software should cite the following paper:

S. Lorenzen, B. Maier, H. Herzel, and A. Kramer
ChronoStar: A software for analysis of time series with detrending,
Chronobiology International XXX:XXX–XXX (2011)

Contents

1 GUI mode

the GUI mode of the program is intended for everyday work and to 'play around' with your files. All settings can be changed in real time, and the graphical output gives you a feeling for the quality of your data. A typical screenshot of the application is shown in Fig. 1. The different areas of the program window are explained in the following section.

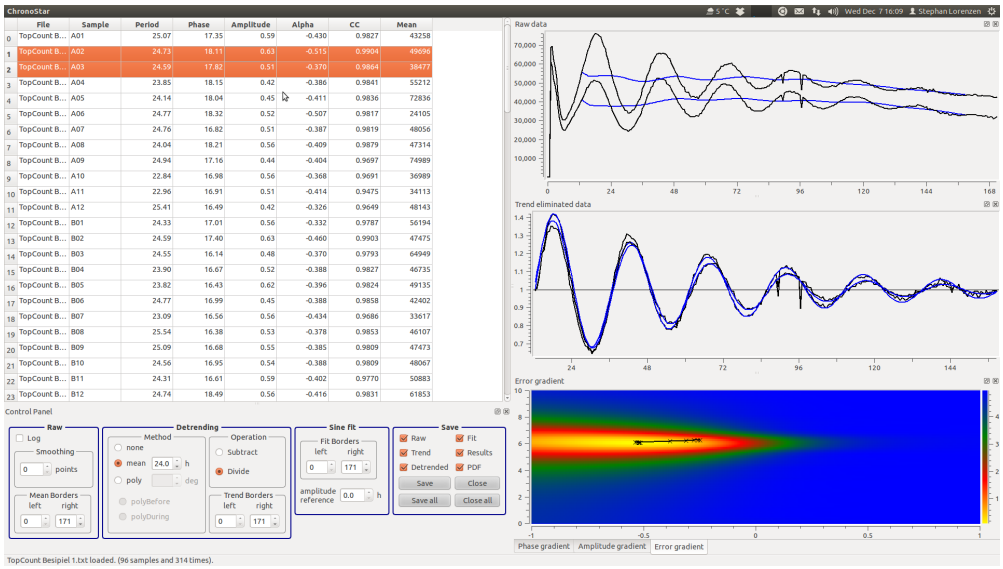

Figure 1: A screenshot of a *ChronoStar* session.

1.1 Areas of the *ChronoStar* Window

1.1.1 Data table

The upper right section of the screenshot (Fig. 2) shows the table with the data currently loaded in the application. The format of the table is variable (see section 1.2.3) to show exactly the items you are interested in. One of many rows of the table can be selected by clicking on them. Using Shift ⇧ lets you select a block of data, with Ctrl + Shift ⇧ you can select multiple noncontiguous rows. The selected data is plotted in the plot area (section 1.1.2) and the gradient of the last selected plot is displayed (section 1.1.4). Clicking on the upper left square of the table (where row and column headers cross) selects all data, pressing the Esc key closes the selected fields. By clicking on the table header, you can sort the data according to, e. g., amplitude or period.

	File	Sample	Period	Phase	Amplitude	Alpha	CC	Mean
0	TopCount B...	A01	25.07	17.35	0.59	-0.430	0.9827	43258
1	TopCount B...	A02	24.73	18.11	0.63	-0.515	0.9904	49696
2	TopCount B...	A03	24.59	17.82	0.51	-0.370	0.9864	38477
3	TopCount B...	A04	23.85	18.15	0.42	-0.386	0.9841	55212
4	TopCount B...	A05	24.14	18.04	0.45	-0.411	0.9836	72836
5	TopCount B...	A06	24.77	18.32	0.52	-0.507	0.9817	24105
6	TopCount B...	A07	24.76	16.82	0.51	-0.387	0.9819	48056
7	TopCount B...	A08	24.04	18.21	0.56	-0.409	0.9879	47314

Figure 2: The *table* section (upper part).

1.1.2 Plots

In the right area, you find two plot windows showing your raw data (Fig. 3) and the trend eliminated data. In the example, both plots contain two black curves and two blue fits, since two rows are selected in the table. The *raw data* plot contains a trend in blue. Note that, in the figure, the trend starts at 12h and ends 12h before the end of the recording. The reason is that the *running average* is only available in this area if you average over 24 hours. Accordingly,

the trend eliminated data in the second plot only covers this time span. If you select a different trend elimination method (e. g. polynomial fit), the trend covers your complete time range.

The lower plot contains both the trend eliminated data (black) and the fit calculated by *ChronoStar* (blue).

By using the mouse wheel, you can zoom in or out, the left mouse key lets you "grab" the curves so you can move around. By pressing Shift ⇧ and using the left mouse key, you can enlarge an area of the plot. Got lost by zooming and shifting around? No problem – pressing Shift ⇧ and the middle mouse key will bring you back one step in zoom history, Shift ⇧ and the right mouse key gives you the full overview again. You should play around with the zoom features a bit to get a feeling of how to use them.

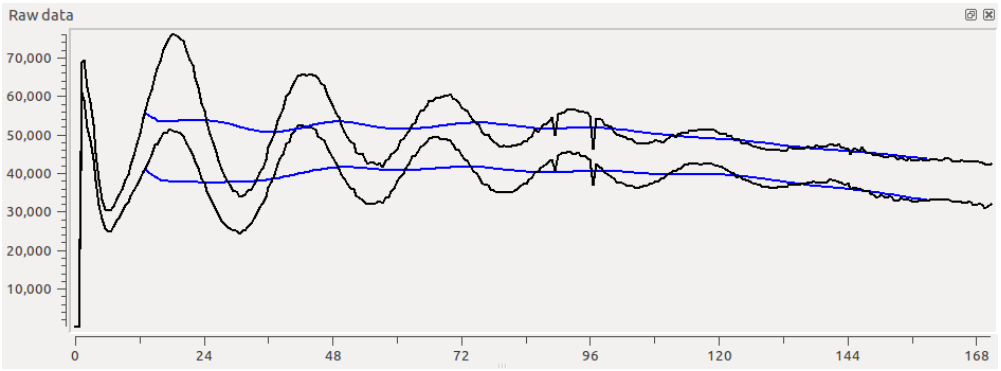

Figure 3: The *raw data* plot section.

1.1.3 Control Panel

The *Control panel* (Fig. 4) is your cockpit while you navigate *ChronoStar*. The input fields show the settings of the data currently selected in the table. Everything you change here is applied in real time to your *selected* data. The control panel consists of four parts described below.

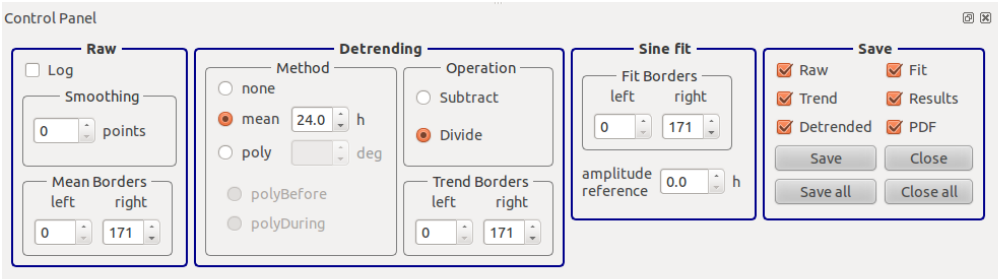

Figure 4: The *Control panel*.

Raw section This section lets you manipulate your raw data. You can take the logarithm of your data before processing it, or smooth your data by taking the mean of a window including neighboring points. The "Mean Borders" settings do not actually change your data but just determine the region for calculation of their mean value.

Detrending section The first step in processing your data is detrending it. If you do not want to detrend your data, choose "None". The standard method of detrending is by using a running average over one period length, i. e., about 24 hours (if your data is circadian). The size of the detrendiong window can be changed; however, keep in mind that by choosing an inappropriate

window size, part of the oscillations will be considered a "trend" and removed. Finally, you can detrend by a polynomial function whose degree you can select. If you do a polynomial fit, you can choose whether you want the polynomial fit done *before* fitting the damped sine (fit the trend, detrend, then fit the sine) or everything in one step, i. e. fit the polynomial *during* fitting the sine. The latter procedure usually gives better results and smaller errors.

In the *Operation* subsection, you can choose whether you want to *subtract* the trend from the data to yield detrended data with high amplitude centered around zero, or to *divide* by the trend to yield oscillations around one which amplitude shows the relative change of expression (in case your data is gene expression).

The third section, *Trend borders*, lets you select the region in which you want to fit the trend and detrend your data. Remember that, for **running average**, half a period length is used either side of the data to calculate the mean values, so, if you choose 24 hours period length and want the trend calculation to start at 20 hours, the actual trend and thus the detrended data starts at 32 hours.

Fit section This section determines which region of your data the damped sine is fitted to. Keep in mind that the fit can at most span the area included in the detrending (minus half a detrending period on either side, if you use *mean* detrending). Extending the fit borders further left or right does not have an effect on the fit. The last setting in this section, *amplitude reference*, is the reference point you want the amplitude to be reported at (see section 1.2.3); changing it has no effect on the fit.

Save section Here, you can determine which files are to be saved. The different kinds of poutput files are discussed in section 4, saving files is discussed in section 1.2.1.

1.1.4 Gradients

Often you want to know how good the minimization procedure was. Did you get stuck in a local minimum? How does the energy landscape of the period/damp space look like (see section 5)? This is what the *Gradient* plot (Fig. 5) is for. In dependence of the damping constant (ordinate) and phase (abscissa), the squared error of the fit is shown in a color coded mode (see color bar on the right). Smaller ("more yellow") values indicate a better fit. In the example plot, you can see a single *valley* in the energy landscape. The black trajectory indicates the steps taken during the minimization procedure. Obviously, you ended up in the right place.

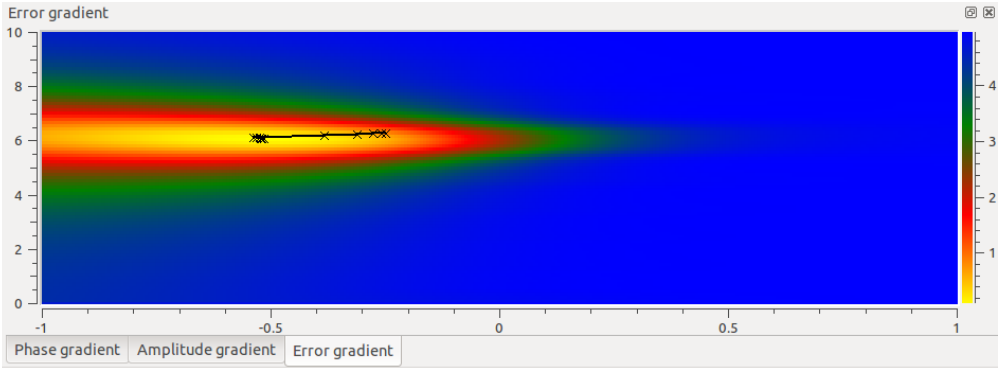

Figure 5: The *Gradient* panel showing a smooth minimization to a single minimum.

Two other gradients you can select in this area depict the amplitude and the phase landscape. Within all three gradients, you can navigate in a way similar to the plot windows (section 1.1.2): If you press Shift ⇑ , you can select a part with the left mouse key which will be enlarged, and you can move the plotted area by "grabbing" it with the mouse and moving it apart. Shift ⇑ and the right mouse key bring you back to a complete view.

Drawing the gradients is quite time consuming, especially if your gradient panel occupies a large space. For each selection of a new data set, a new gradient has to be calculated, which can be annoying if you play around a lot. If you feel that *ChronoStar* is slow, consider changing the **render velocity** (1.2.2) or closing the gradient panel completely (see section 1.1.5).

1.1.5 Moving, closing, and resizing sections

ChronoStar is very flexible in arranging your desktop and allows you to customize it to your favorite view. Besides resizing the different parts of the application window, you can also pull the panels away from the main window to let them become independent windows, or *dock* them to another area. You do not like the control panel at the bottom? Just grab it and dock it on the top. You like those fancy phase gradient plots? Make them a standalone window so you can maximize them to fill your whole screen! (Beware of the calculation time for the gradient, though...) Clicking on the ⊠ symbol in the upper right side of any panel will close it. Don't panic – all your closed panels can be brought back to life using the *Views* menu (1.2.2).

1.2 Program menu

1.2.1 File

In this menu, you can do all kinds of file operations: *Open* a new file, *Save* or *Close* your current selection or all currently opened files, and – in case you should ever want to do so – *Quit* the program. The *Open* menu will always remenber the last director you worked in, if you open a file the next time. The *Save* and *Close* options are also available in the control panel (section 1.1.3) where you also determine what kind of files you want to save. See section 4 for a description of the output files.

1.2.2 View

This menu lets you control the visibility of all the panels described above. Another important point is the adjustment of the *render velocity* (compare section 1.1.4). Small values mean better (but slower) rendering, larger values lead to a quick but not too beautiful gradient image. The meaning of the parameter is the size of the squares (in pixels) which are colored: the lowest possible number (1) colors each pixel separately, the highest value (10) colors squares of 10×10 pixels.

1.2.3 Format

The first three items in this menu let you arrange the format of the table displayed on screen, the table saved to file, and the parameters written to the PDF file upon saving (see section 4). Either of them opens a dialog where you can pick and reorder the items to display (Fig. 6).

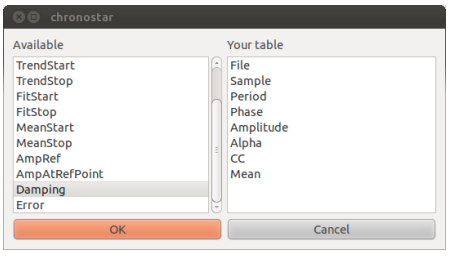

Figure 6: The Adjust format dialog.

On the left side of the dialog, you see all available items you can put in a table, on the right is your current choice. By grabbing items with the mouse, you can add them to your table (put them to the right) or delete them from your table (put them to the left). In the same way, you can change the order of the columns in the table. Be aware that, if you **change the format of the output file**, any other programs adopted to the old file format (e. g. scripts to plot something) might not work any more.

The choice to pick from is plentiful – table 1 describes the different items to pick from. The last item in the *Format* menu **resets all changes** you made to defaults. Be careful with this one! You loose the table formats you like, and *ChronoStar* will not remember the folder you like to work in.

File*	The file where your data comes from
Sample*	Sample number (coordinate if 96-well)
Period*	Period of the oscillations in *hours*
Phase*	Phase of the oscillations in *hours*
Amplitude*	Amplitude
AmpAtRefPoint	(damped) amplitude at reference point
AmpRef	Reference point in *hours*
Alpha*	Damping parameter (see section 5)
Damping	Decrease of amplitude per 24 hours ($e^{-\alpha}$)
Error	Sum of squared errors. This is actually minimized.
CC*	Correlation coefficient fit - experiment
Mean*	Mean experimental value (see also 1.1.3)
Method	Method used for detrending
Operator	Detrending operation (Subtraction/Division)
PolyDegree	Degree of the polynomial used for the fit
TrendStart, TrendStop	Area used for detrending (hours)
FitStart, FitStop	Area used for curve fitting (hours)
MeanStart, MeanStop	Area used for mean calculation (hours)

Table 1: Items to choose for table columns. Items marked by an asterisk are included in the default selection.

2 Batch mode

Sometimes, you have a bunch of data files you want to process without opening them manually or looking at them in the *ChronoStar*. Maybe you just want to calculate all these fits and process the output with some other program. Then, the *Batch mode* is what you want. It lets you work on *one file at a time* and writes its output to the disk, so you can write a script to process multiple input files. Whenever you provide a parameter to *ChronoStar*, it will start in batch mode and not open any GUI window. The parameters are explained in the following sections. All options are *paired*: you specify *what* you set and *which value* you set it to:

`-method none`	No detrending
`-method mean`	Detrending by running average. This is the **default**.
`-method polypre`	First detrending by polynomial fit, then sine fit.
`-method polydur`	Polynomial detrending and fit in one step.
`-oper sub`	The trend is subtracted from the data.
`-oper div`	The data is divided by the trend. This is the **default**
`-trendper 25`	Sets the trend period to 25 hours (**default: 24h**).
	Only has an effect if `-method mean` is set.
`-poly 5`	Sets the polynomial degree to 5 (**default: 4**).
	Only has an effect if `-method polypre` or `-method polydur` is set.
`-in myData.txt`	Loads the file `myData.txt`. Each run loads exactly one file.
`-outraw raw.txt`	Saves the raw data to `raw.txt`.
	Useful if you want to extract data from *LumiCycle* files.
`-outelim elim.txt`	Saves trend eliminated data to `elim.txt`.
`-outpara para.txt`	Saves fit parametes to `para.txt`.

An example program call would be
`chronostar -method polydur -in myData.txt -outpara para.txt`

3 Supported input file formats

Currently, *ChronoStar* supports three file formats. The file format is detected automatically upon openng the files. Example files of all types can be found on the website in the `examples` folder

3.1 TopCount

The *TopCount* format (Fig. 7) is a *tab-delimited* text file with data arranded in columns. The first column contains the time in *days*, subsequent columns contain luminescence data for an arbitrary number of experiments. Thus, each row contains the time (in days) and data for this time point for each experiment. If your data contains **exactly 97 experiments** (plus the mandatory time column), the second column will be ignored! This is due to a widely used special format for 96-well plates where an additional second row containing the temperature is added. If you have 96-well data stored in the usual way (i. e., time column and 96 data columns), everything will be fine. Beware if you have *97* experiments – use the LumiBox format or add a column with dummy data to get 98 data columns.

0	71776	68960	60872	83368	88088
0.022685	65240	69320	59032	80680	87256
0.045417	57536	62264	51840	71824	78000
0.068102	52920	58496	48560	66232	72992

Figure 7: Begin of a *TopCount* File with five data rows.

3.2 LumiBox

A *tab-delimited* file with an additional first line containing date and time in the format of Fig. 8 is treated as a *LumiBox* file. The first line (which is used to recognize the file format) is ignored, the first column is read as time in *hours*, subsequent columns contain the data. The number of data columns is arbitrary.

05.01.2011	11:19			
0	867621	1044027	1273674	876029
0.083	639131	874430	797365	713296
0.167	547260	764230	657847	615384

Figure 8: Begin of a *LumiBox* File with four data rows.

3.3 LumiCycle

LumiCycle is a binary format used by ActiMetrics.

4 Output file formats

Raw and trend eliminated data is both written as a *tab-delimited* file without headers with time in *hours* in the first column and data in subsequent columns. the parameter file is also tab delimited and consists of a header line and columns with Sample number, phase, period, amplitude, damping parameter, fit error and the parameters of the polynomial.

You can easily access and process this data with any other software like, e. g., `R`:

```
t    = read.table("raw.txt") [, 1]
elim = read.table("elim.txt")[,-1]
para = read.table("para.txt", header=TRUE)
attach(para)
MAXPOW = dim(para)[2]-7
ROW    = 1
fit = amp[ROW] * cos(2*pi*(t*24-pha[ROW])/per[ROW]) * exp(t*damp[ROW]) + 1
plot (t*24, elim[,ROW], xlab="Time [h]", ylab="Detrended data")
lines(t*24, fit)
```

5 The math behind *ChronoStar*