Visual ANalysis of timE SerieS dAta - *Drosophila* Activity Monitors (VANESSA-DAM) is a collection of useful tools to visualize and analyze Time series data obtained from *Drosophila* Activity Monitors (https://www.trikinetics.com/). The first in the series of tools is a shiny app for circadian rhythm analysis and visualization - VANESSA-DAM for circadian rhythm analysis (VANESSA-DAM-CRA). For any suggestions, questions, troubleshooting or customization, please contact arijitghosh2009@gmail.com.

VANESSA-DAM-CRA

VANESSA-DAM-CRA is dependent on Quentin Geissmann's rethomics family of packages - behavr, damr, ggetho, zeitgebr, for some analysis and visualization options. It offers several advantages over existing tools for circadian rhythm analysis from DAM systems, some mentionable ones are -

1. Analysis and visualization of multiple monitor files, genotypes, replicates together in a high-throughput manner.

2. Automatic period power detection through a range of periodogram methods.

3. Producing high-resolution publication-quality figures with a plethora of customization.

4. Data curation - automatic user-defined parameter-based removal of dead and arrhythmic individuals.

5. Individual wise CWT spectrograms and wavelet power plot.

6. Visual comparison among genotypes, replicates.

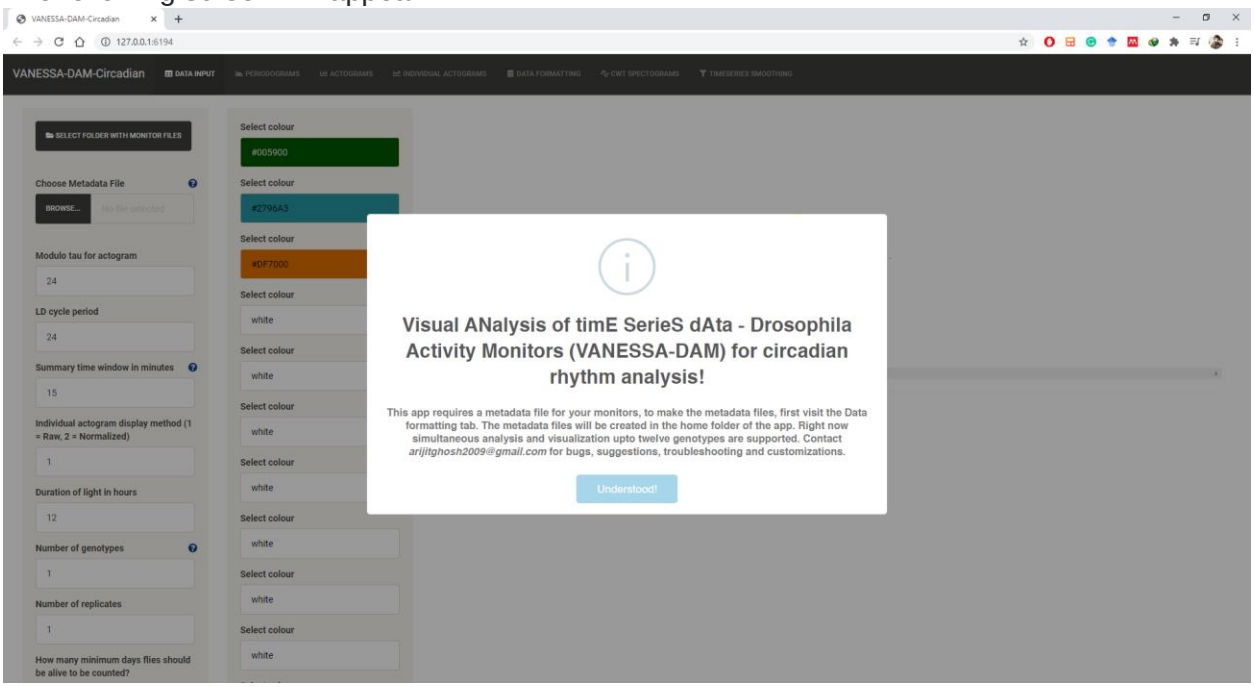
7. Timeseries filtering with kernel smoothing and Butterworth filters.

8. Reproducible code report so that you can generate the figures and analysis without the shiny app from RStudio directly.

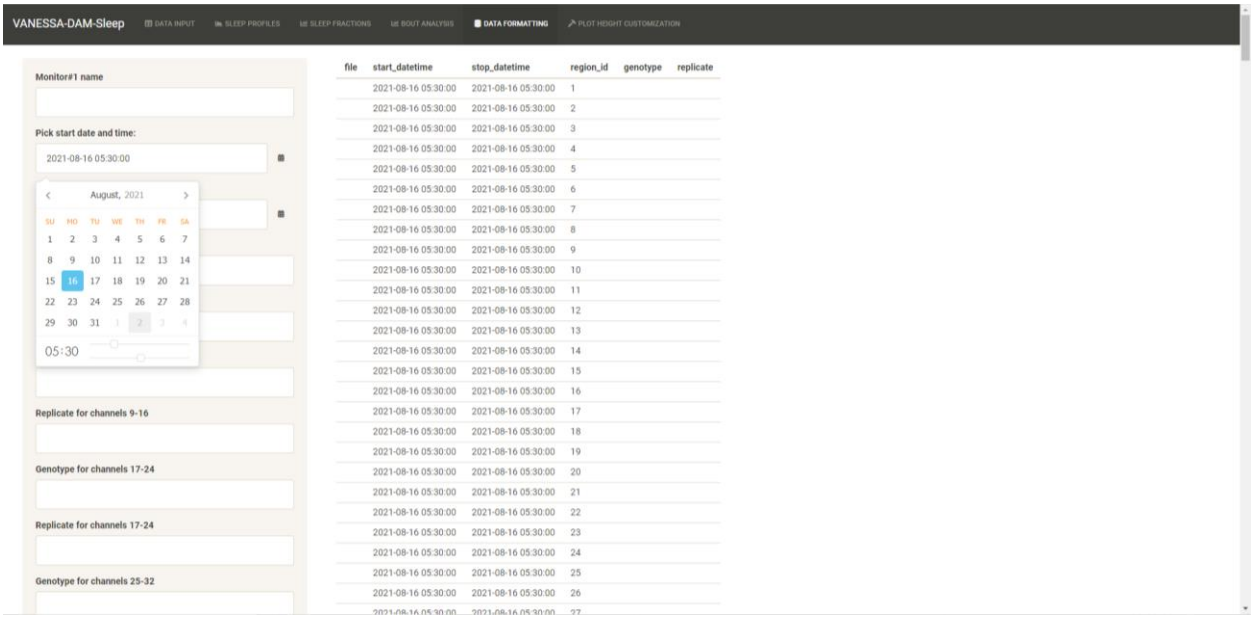
This short tutorial provided (*Easy tutorial to start using VANESSA-DAM-CRA.pdf*) here should be self-explanatory and should help beginners start using the app right away.

Easy tutorial to start using VANESSA-DAM-CRA

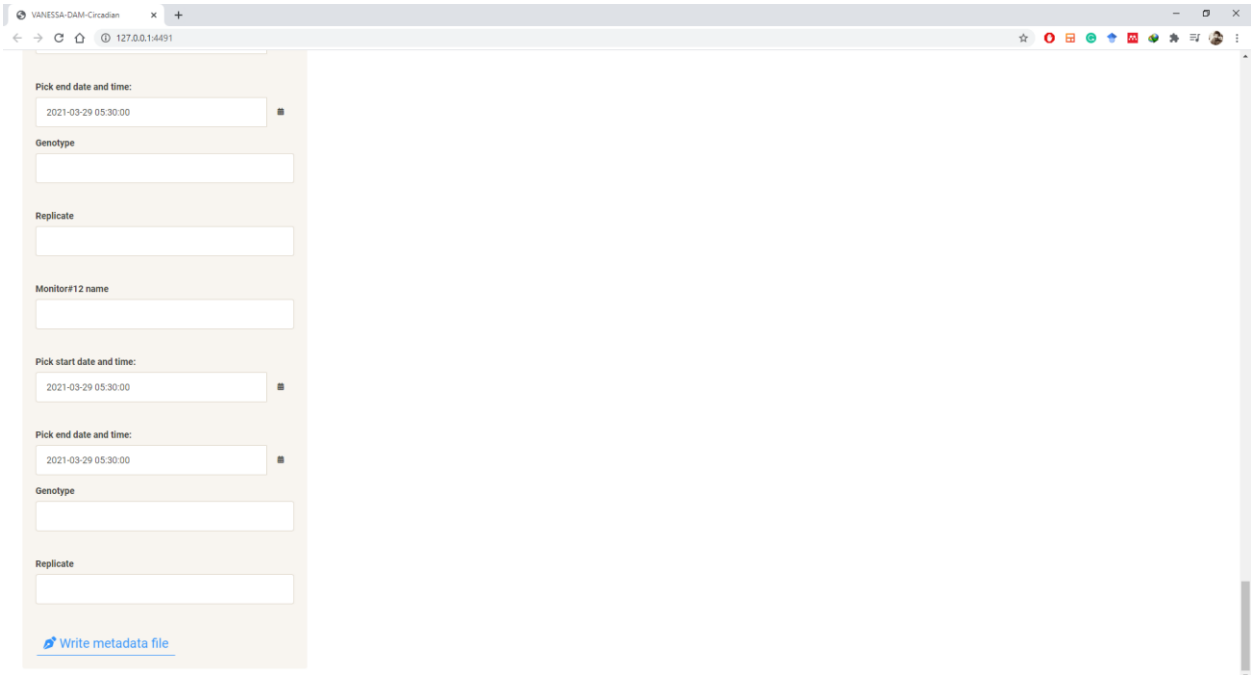
1. Install R (version ≥ 3.6.3) and RStudio (version ≥ 1.2) on your computer.
2. Download the zipped folder containing the app and unzip to a folder.
3. Open RStudio and Run the **startup.r** file to install all necessary packages to run the app.
   Note: If you have trouble installing the *damr* package, the source files are provided. Please go to RStudio > Tools > Install packages > Install from : Package Archive file and select the supplied *damr_0.3.7.tar.gz* file to install the package.
4. Set working directory as the home folder of the app (by pressing **ctrl+shift+h** OR by using the **setwd()** command) and then run the app by typing **shiny::runApp(launch.browser = T)** in the console.
5. The following screen will appear:



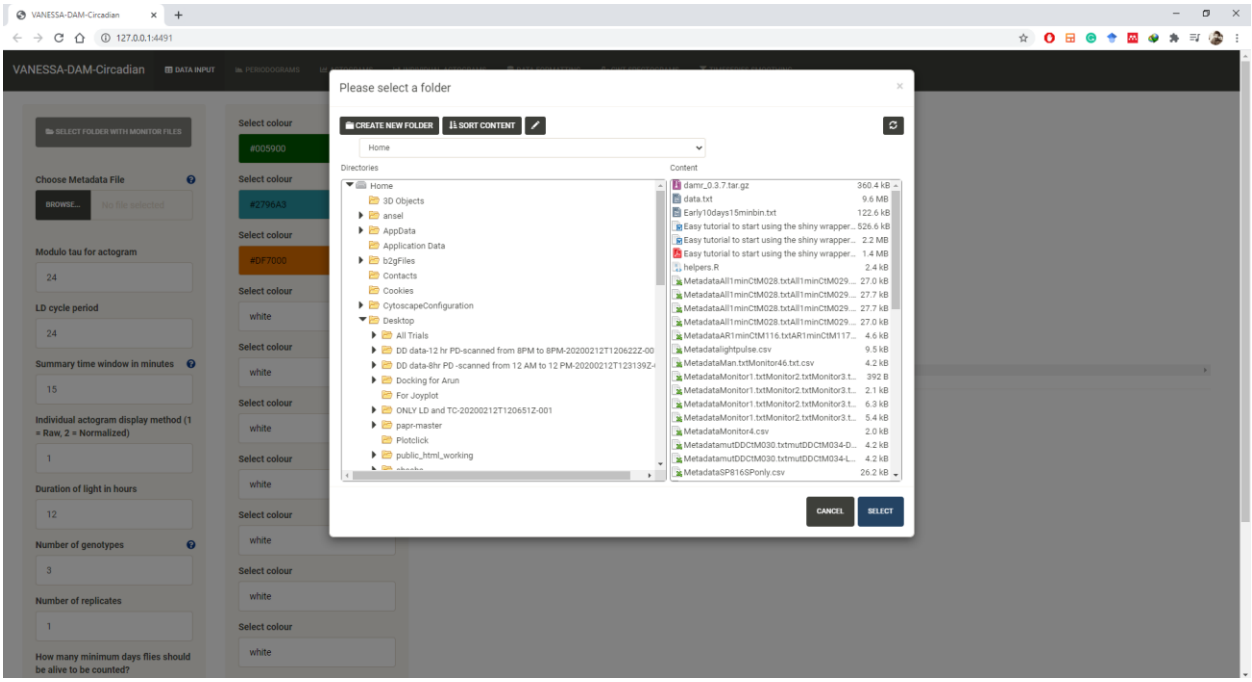
6. Click on **Understood!** button or press **Esc** to begin.
7. Assuming you have your Monitor files (*Monitor1.txt, Monitor2.txt, Monitor3.txt* provided with Genotypes *Early*, *Control* and *Late* respectively – check the ***Notes*** section for details of the data provided) from DAMScan in the home folder of the app, first job is to make a Metadata file (Provided as *MetadataMonitor1.txtMonitor2.txtMonitor3.txt.csv*).
8. In this case, we have 3 genotypes in 3 Monitor files, so enter 3 at the **Number of Genotypes** box. Then go to the **DATA FORMATTING** tab.



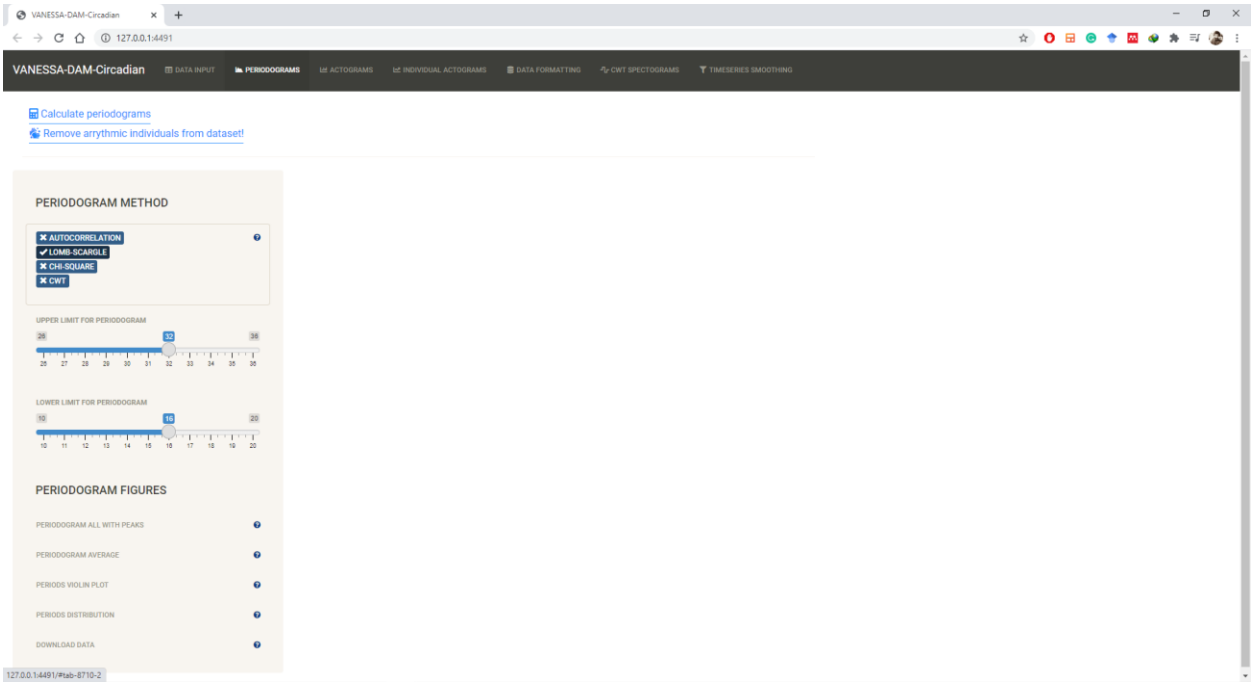
9. Enter name of your first monitor file in the **Monitor#1** name box, include extensions also (e.g., if the name of the monitor file Monitor1.txt, then enter the FULL name, including ".txt"). Fill the next boxes - **Start date and time (If your start date is 1st August 2020 and Lights-on time is 10 AM, then you write 2020-08-01 10:00:00), End date (If you start date is 10th August 2020, then you write 2020-08-10)** with your experimental details as shown above. Enter **Genotype** values and **Replicate number** (only if your experimental design has multiple replicates for all/some genotypes, else keep random number, it won't affect any calculation) in proper boxes for proper channels (marked as **Genotype for channels** 1-8 and **Replicate for channels** 1-8 etc.). A date and time picker has been added to facilitate easy entering of start and end date times without error in formatting. Each monitor can have maximum of 4 genotypes with 8 individuals.
Note: Creating the metadata file is the single most important step for using this app, any wrong information in the metadata file will wreak havoc on the results and the app may crash. DO NOT change information in the DATA FORMATTING tab once you have entered. For some unknown problem in shiny reactivity, this section does not handle deletion and change of values nicely. Alternatively, you can just crate your metadata file by changing data in the already provided metadata file in a good text editor (DO NOT USE EXCEL), like Notepad++. The metadata file serves as a base of information for your analysis too, and is good for record keeping of your experiments.



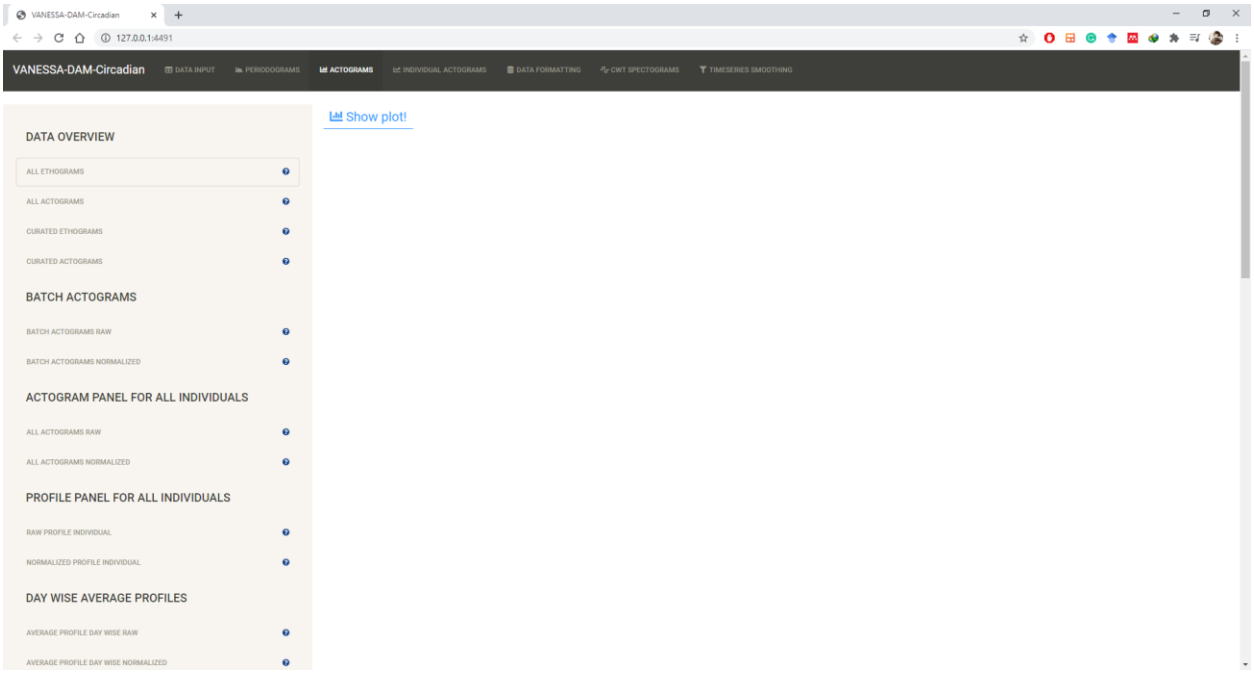
10. After filling in experimental details for all 3 monitors, go down of the page and click on the **Write metadata file** button, this will write the Metadata as a CSV file in the app home folder and is ready to be used. At this point the Metadata file will be stored in the folder set as your working directory.
11. Go back to the **DATA INPUT** tab. Enter choices one by one- **Duration of light in hours** in your experiment, **How many minimum days flies should be alive to be counted**. Also chose your **starting day** and **ending day** for analysis (**starting day** = 0 means first day in your data, so if you want to select first five days, **starting day** should be 0 and **ending day** should be 4). Enter **LD cycle period** in your experiment, **Modulo tau** for actogram can be changed for visualization purpose later, same goes for **Summary time window in minutes** and **Individual actogram display method**. You can change any of these parameters later while looking at plots, the plots will immediately update accordingly. For example, if you're looking at your actograms in 15 minutes bin, and you want to visualize them with 5 minutes bin, you just have to change the value in Summary time window in minutes box, similarly LD shading can also be changed. You are all set to start analysis now. **ONLY PROCEED FORWARD WHEN YOU HAVE FIXED ALL ANALYSIS PARAMETERS.** Select the folder where you have kept your monitor files to be analyzed by clicking the **Select Folder With Monitor Files** button and upload your Metadata file by clicking on the **BROWSE** button in the **Choose Metadata File** box. Press the Start **Calculations!** Button, when the calculations are done, you will be notified by a sound. Depending on how long your data is and how many monitors you are analyzing, this step (which involves curation also) will take anything between 5 seconds to 1 minute typically.



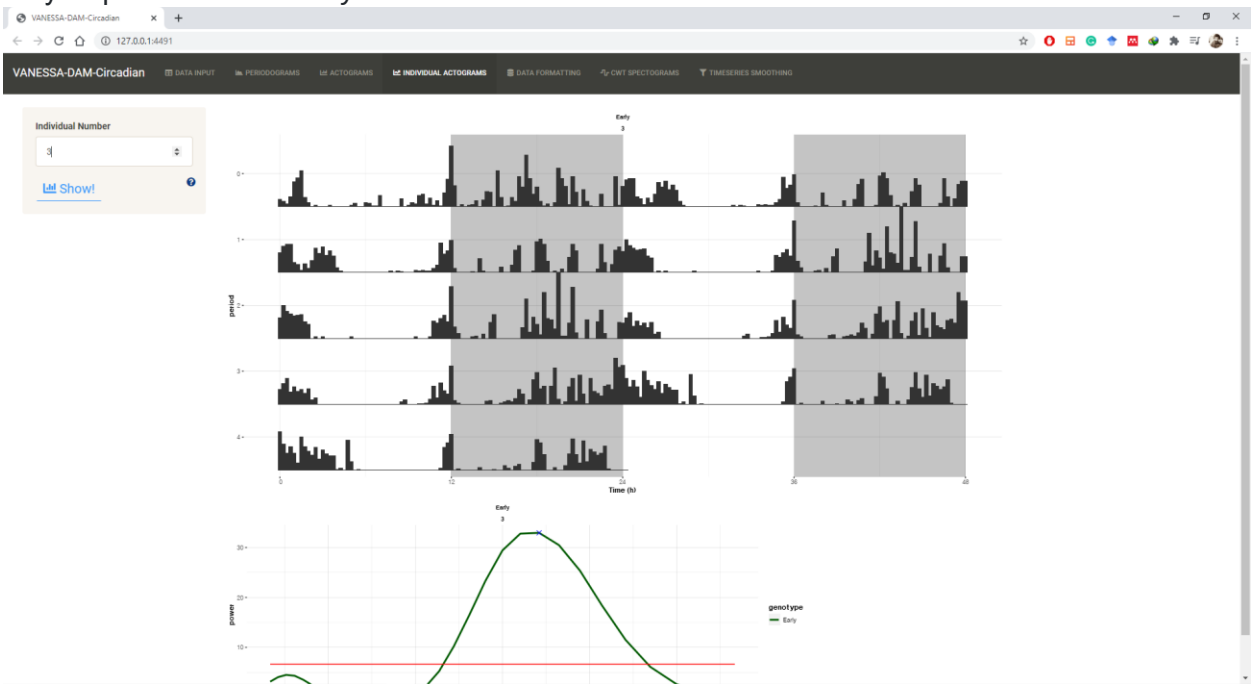
12. Go to the **PERIODOGRAMS** tab next. You will see the screen below. Choose **PERIODOGRAM METHOD** from four available methods, and choose **Upper limit for periodogram** and **Lower limit for periodogram**. Click on **Calculate periodogram** button to start calculating period of the individuals with the method you have chosen previously. If you click on the **Remove arrythmic individuals** from dataset! button next, all arrhythmic individuals will be removed from all calculations and plots. **Once you remove the arrhythmic individuals, ALL calculations and plots will be done with only rhythmic individual on curated data.** If you want to visualize the arrhythmic individuals also, do the analysis without clicking on the **Rhtymic arrhythmic individuals** button.



13. When you click on different buttons under **PERIODOGRAM FIGURES**, you will start seeing plots as mentioned in different side panels.
14. Go to the **Download data** tab to download a csv file for all period and power values. Different peaks of periodograms for all individuals are noted in the file. If you open the file in Excel or LibreOffice or any spreadsheet program, you can put a filter on the "peak" column which is the last column of the file. When you filter the data by "peak" and select "peak" as "1", it will show you the highest peak of the periodogram of all rhythmic individuals. You'll notice that this file will now show you these columns - a) id: start_datetime|Monitor name|Indivudual number, b) period: the period defined by the "peak" (if you select 1, it will show you the period of the highest peak in the respective periodogram), c) power: the peak value from the periodogram (in case of autocorrelation it's the peak autocorrelation value, i.e., rhythmicity index, in case of chi-sq, it's the power of the chi-sq periodogram, etc.), d) signif_threshold: significance threshold for the particular peak, e) p_value: the p_value is reported here for the particular peak, f) peak: different peaks in the periodogram (filter based on this).
15. Go to the **ACTOGRAMS** tab to start visualizing Actograms and Profiles as mentioned in different side tabs, feel free to explore.

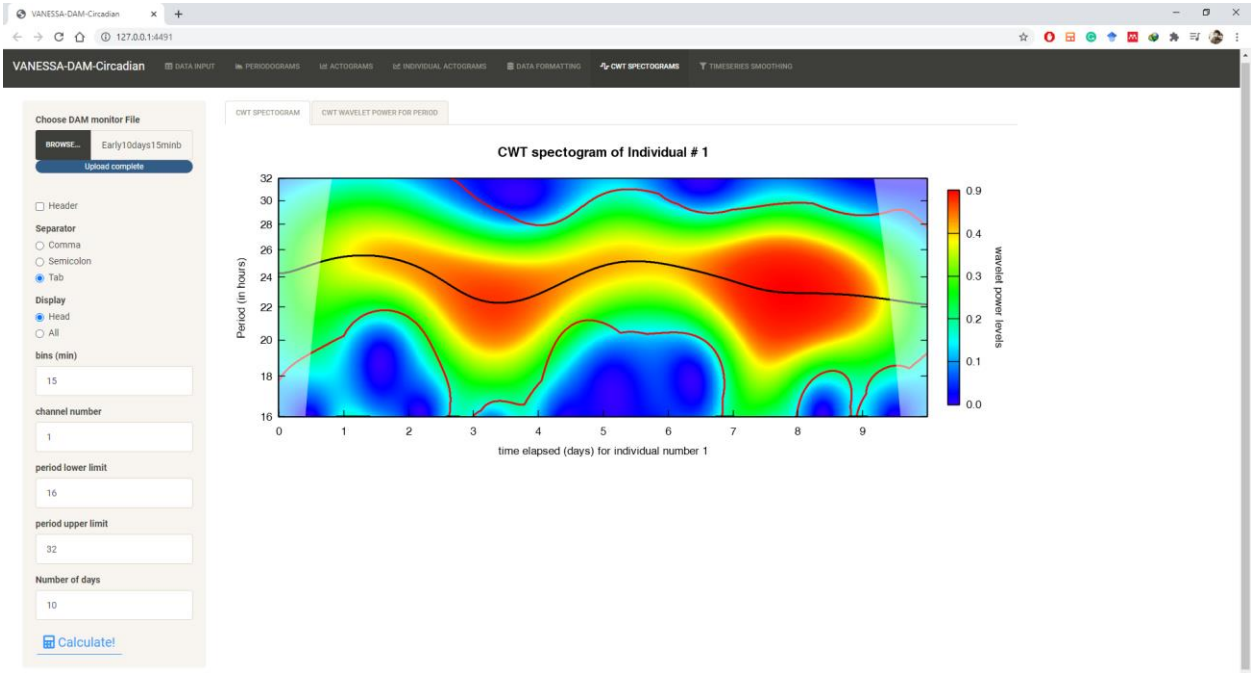


16. Go to the **INDIVIDUAL ACTOGRAMS** tab to see individual wise actograms (raw or normalized) for closer inspection of your data. Change **Individual Number** by pressing up or down arrow or use the buttons in the box to change individuals to look at and they'll update automatically.

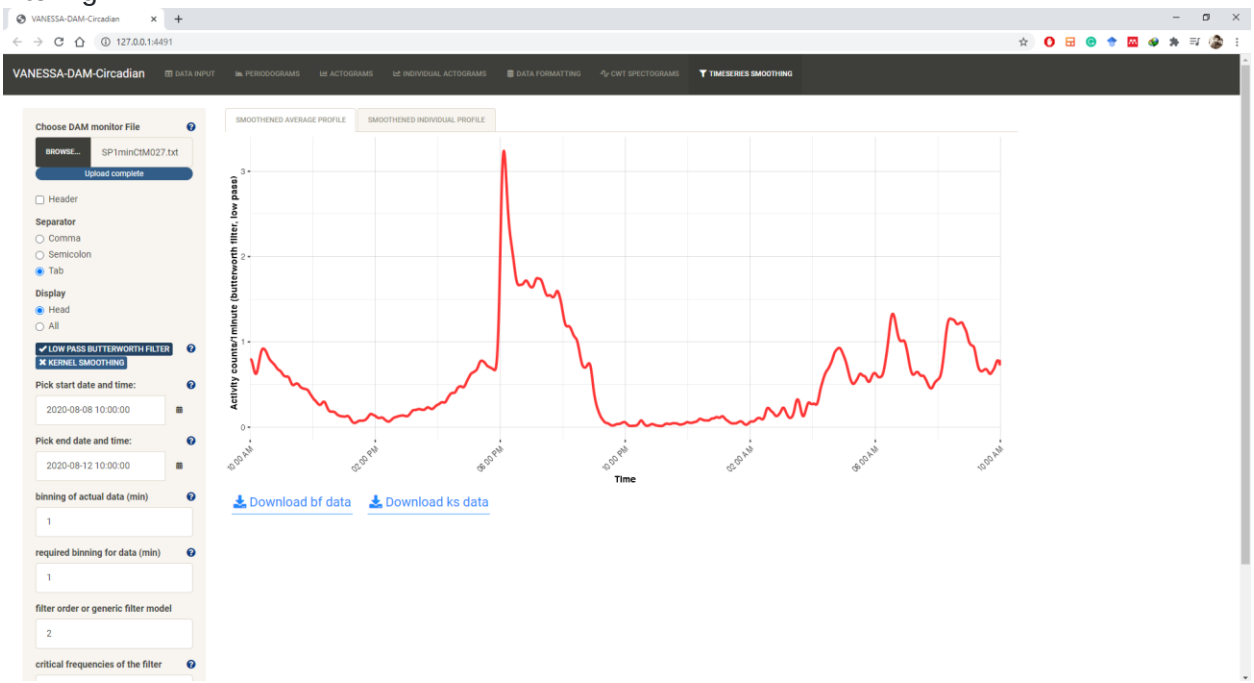


17. The next tab is **CWT SPECTOGRAMS**. This helps in constructing individual CWT spectrograms and plot ridge values and provides 95% CI. Useful for visualizing changing periods in case of long recordings and regime changes. Still in basic stage. After entering appropriate details in the **bins (min), period lower limit, period upper limit** and **Number of days** boxes, upload your Monitor file in the **Choose DAM monitor file** box and press the **Calculate!** Button. After calculations are done for the individual

chosen, a CWT spectrogram will be plotted. Change channel number to visualize different individual.



18. The last tab is **Timeseries Smoothing**. This tab has functions which are for day-to-day use while wrangling your data. You can re-bin your data, extract data from different dates, filter and smoothen them, and best of all, download the filtered/smoothened/extracted data. Choose appropriate options, and then visualize average profiles or individual profiles until you're satisfied with the smoothening or filtering.

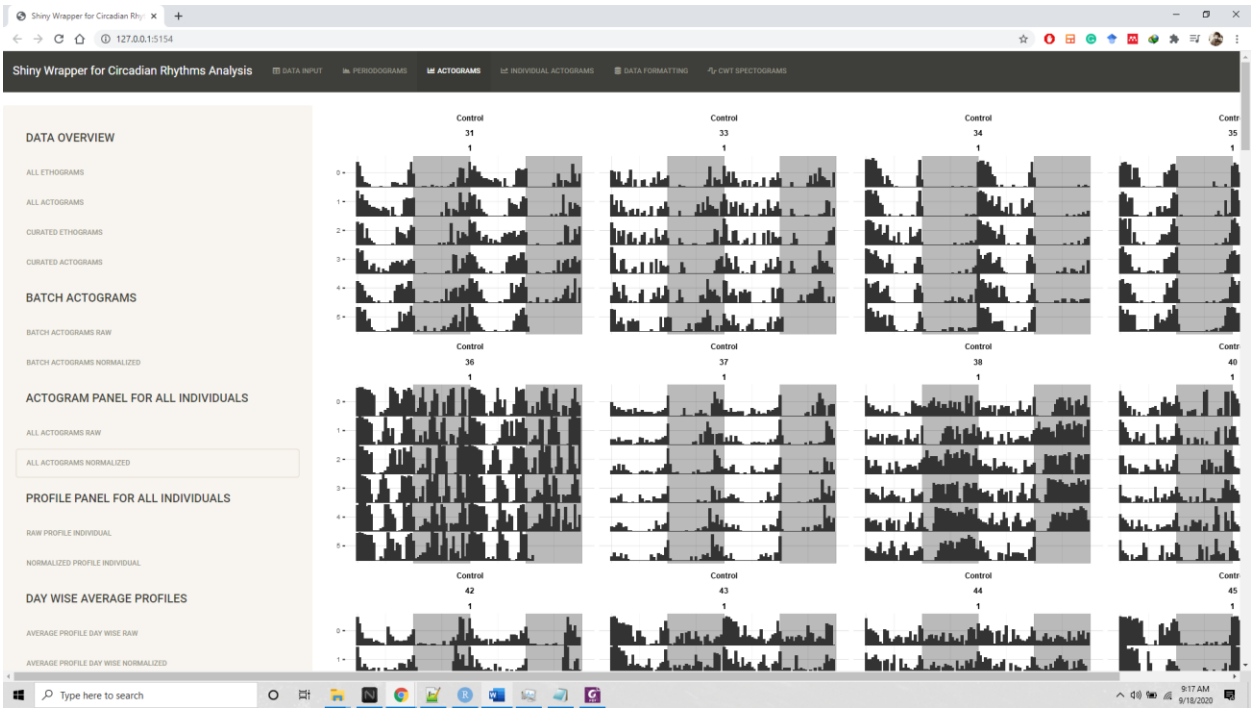


NOTES:

1. If you have different replicates, they will be plotted separately in **DAY WISE AVERAGE PROFILES, AVERAGE PROFILES** and **CIRCULAR AVERAGE PROFILES** tabs.

2. **RAW PROFILE ALL REPLICATE AVERAGE** and **NORMALIZED PROFILE ALL REPLICATE AVERAGE** tabs will average over ALL individuals in all replicates in a Genotype.

3. In faceted panels, like **ALL ACTOGRAMS RAW, ALL ACTOGRAMS NORMALIZED** etc., the values on top of each panel is as following: 1-Genotype name (Control in the following example), 2-Arbitrary serial ID assigned on raw data (31,33,34 etc. in example), 3-Replicate number (1 in example).



4. All images produced are high-resolution, can be copied onto clipboard, saved as png files and directly used.

5. For best results, use Monitor files with <5 minutes bin data. Also, cleaning up your Monitor file before using is desired, you don't necessarily have to subset your data by date, all data can be in the Monitor file, only the dates you specify in your Metadata file will be used, thus reducing hassle for the user.

6. If your run was in DD, please put **Duration of light in hours** as 23.99, as it does not 0 as input, it does not affect any calculations. Also, after you know your average period value in DD, you can change **Modulo tau for actogram** and **LD cycle period** accordingly, and plots will be updated accordingly.

7. The data provided is from a locomotor activity run with DAM2 system under LD12:12 (~750 lux light) for 8 days of three populations artificially selected for divergent chronotypes in our lab (as their names signify, early, control, and late). Each monitor file has 32 flies loaded onto it (standard format for DAM2 systems).

8. All plots can be resized from the last panel named **Plot height customization**. This is needed because when you remove arrhythmic individuals, the program won't automatically change dimensions of plot area, majorly only the "Height" parameter for the plots will have to be tweaked. It is also important to change the "Height" of actograms manually for achieving best visualization of your length of data.