

# FSLRetho2 Documentation

An R-Shiny application designed to streamline analysis of behavioral datasets generated in the lab with DAM <sup>1</sup> and Ethoscope <sup>2</sup> devices, without the need to know R.

- Analyse your data remotely, without the need of having the behavioral dataset downloaded in your local computer.
- No need to write any code!
- User friendly error messages for most common mistakes.
- Consistent analyses for easy replication of results.

## Generate and upload a metadata table

### Generate metadata

Provide fslretho2 with information about your experiments in the form of a metadata file in .csv (comma separated value) format. A valid metadata table will have one row per fly. Columns are separated by ,. Information about which flies to look for is **required**. Extra information about the flies that you want to use in your analysis (sex, age, genotype, treatment, ...) is **optional** (but still you must put it as extra columns in the metadata if you are to use it in the analysis).

You can generate a metadata table using Google sheets or Excel.

*Please be very careful with columns storing date and/or time. These programs tend to change the format entered by the user, which can lead to invalid metadata when uploaded to fslretho2.*

Required fields in **ethoscope**:

- **region\_id**: Position of the fly in the arena. Ex 2
- **machine\_name**: Name of the ethoscope. Ex ETHOSCOPE\_007
- **date**: Day when the tracking was started in format YYYY-MM-DD. Ex 2020-07-06

---

<sup>1</sup> "TriKinetics." <https://trikinetics.com/>. Accessed 6 Jul. 2020.

<sup>2</sup> "Ethoscopes: An open platform for high-throughput ... - PLoS." 19 Oct. 2017, <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2003026>. Accessed 6 Jul. 2020.

- **reference\_hour**: GMT time in hours when the lights turn on in the incubator.  
Example: 12 for an incubator that turns lights on at 2/1 PM (summer/winter)

Example ethoscope metadata:

region_id	machine_name	date	reference_hour	motor	genotype	age low	age high
1	ETHOSCOPE_007	2020-06-29	12	N	DGRP441	4	5
2	ETHOSCOPE_007	2020-06-29	12	N	DGRP441	4	5
3	ETHOSCOPE_007	2020-06-29	12	N	DGRP441	4	5
4	ETHOSCOPE_007	2020-06-29	12	N	DGRP441	4	5
5	ETHOSCOPE_007	2020-06-29	12	N	DGRP441	4	5
6	ETHOSCOPE_007	2020-06-29	12	N	DGRP441	4	5

Required fields in **DAM**:

- **region\_id**: Position of the fly in the arena. Ex 2
- **file**: Monitor.txt file. Ex Monitor12.txt
- **start\_datetime**: Day when the experiment was started in format YYYY-MM-DD followed by the time when the incubator turned on, in format HH:MM:SS
- **stop\_datetime**: Day when the experiment was stopped in format YYYY-MM-DD followed by the time when the incubator turned on, in format HH:MM:SS

Example DAM metadata:

region_id	file	start_datetime	stop_datetime	genotype
1	Monitor12.txt	2020-04-27 08:00:00	2020-05-04 08:00:00	iso31
2	Monitor12.txt	2020-04-27 08:00:00	2020-05-04 08:00:00	iso31
3	Monitor12.txt	2020-04-27 08:00:00	2020-05-04 08:00:00	iso31
4	Monitor12.txt	2020-04-27 08:00:00	2020-05-04 08:00:00	iso31

## Upload metadata

Go to the **Load** tab in fslretho2 and you will find this screen. Provide your metadata file in csv format in the correct input slot ([Browse...](#)). You can either click on Browse and select your file or drag and drop it. After that, click on the [Submit](#) button below the slot you used.

FSLRetho2

Scoring parameters Binning parameters Animal groups

Loaded dataset: NONE

Currently uploaded dataset (taken from the metadata name)

Ethoscope Input

Browse... No file selected **Ethoscope metadata input**

/ethoscope\_data/results **Path to ethoscope database**

Submit

Dam Input

Browse... No file selected **DAM metadata input**

/DAM\_data/results **Path to DAM database**

Submit

FSLRetho2 receives and processes your metadata when you press "Submit"

It will

I) **Load** the data specified in your metadata

II) **Score** the activity and return a machine-generated behavioral label

If any errors are raised, you will see them on the bottom. If there's an error and the program cannot handle it, the whole screen will turn grey. Contact Antonio then!

FSLRetho2

Scoring parameters Binning parameters Animal groups

Loaded dataset: aci\_metadata - Jiekun\_screening.csv

Ethoscope Input

Browse... No file selected

/ethoscope\_data/results

Submit

Dam Input

Browse... metadata\_duplicates.csv

Upload complete

/DAM\_data/results/2020-07-06

Submit

Error: 'metadata' must not contain repeated rows: Rows # 5 are repeated

FSLRetho2

Scoring parameters Binning parameters Animal groups

Loaded dataset: aoj\_metadata - Jiekun\_screening.csv

Welcome

Load

Sleep analysis

Text editor

Revise metadata

### Ethoscope Input

Browse... laptop\_test2.csv  
Upload complete

/ethoscope\_data/results

Submit

### Dam Input

Browse... metadata\_duplicates.csv  
Upload complete

/DAM\_data/results/2020-07-06

Submit

Error : "metadata" must contain columns machine\_name date reference\_hour

If the upload is successful, the name of the file will appear on the top right corner

FSLRetho2

Scoring parameters Binning parameters Animal groups

Loaded dataset: aoj\_metadata - Jiekun\_screening.csv

Welcome

Load

Sleep analysis

Text editor

Revise metadata

### Ethoscope Input

Browse... No file selected

/ethoscope\_data/results

Submit

### Dam Input

Browse... aoj\_metadata - Jiekun\_screeni  
Upload complete

/DAM\_data/results/2020-07-06

Submit

Your metadata was successfully processed and at least a matching entry in the behavioral database was found. You are then ready for the analysis step.

# Analyze

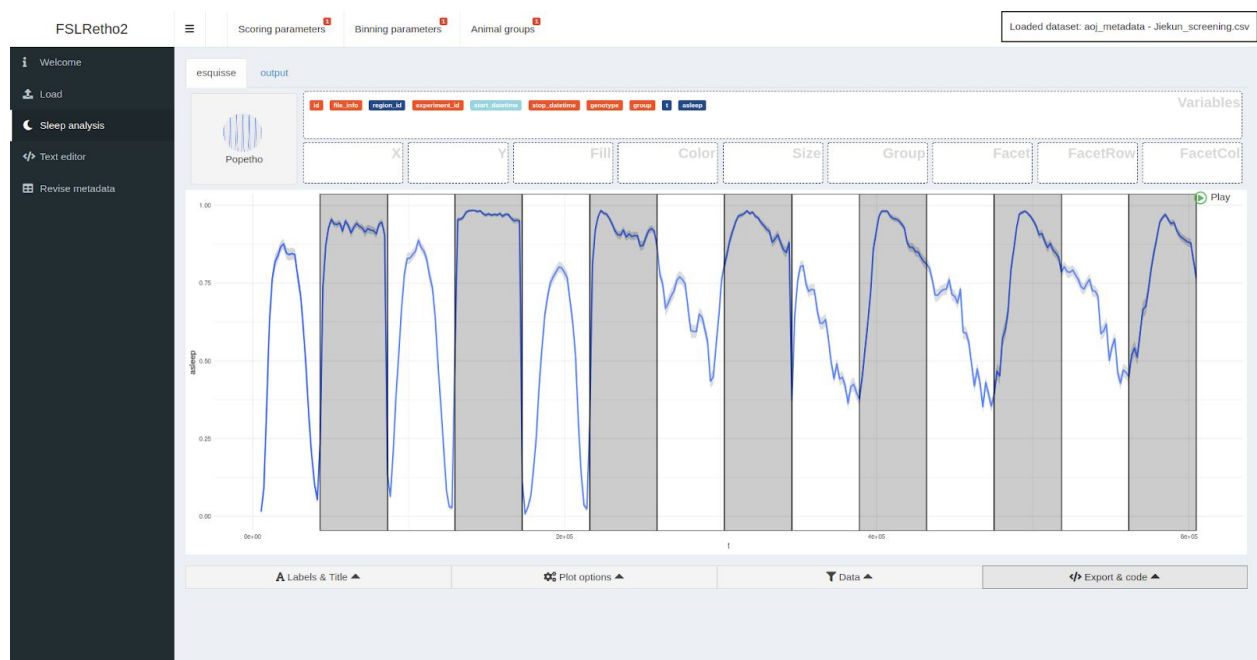
## Amount of sleep

fslretho2 follows the rethomics workflow for the analysis of behavioral data:

In an experiment with N flies

- Input data has one row per fly and observation. This is N rows per minute in DAM and N rows per frame in ethoscope.
- Score several datapoints within a window of predefined length. The result is 1 behavioral label, in this case **moving (TRUE)** or **not moving (FALSE)**
- Apply the 5 minute rule to consider significantly long bouts of non movement as asleep and discard non significant ones. The result is a new (derived) behavioral label **asleep (TRUE)** or **not asleep (FALSE)**
- Group several of such windows into bins of a greater predefined length and compute a summary statistic of the asleep label. The default is the mean for F=30 minutes, which has the effect of computing the fraction of windows where the fly is scored as asleep
- Plot this asleep vs time series with frequency F

Such a plot is created automatically the moment the **Analysis** tab is clicked. For the DAM dataset shown above:



## Customize the analysis: scoring

Sensible default parameter values are used to generate the plot above. However, the user has the power to change the following parameters:

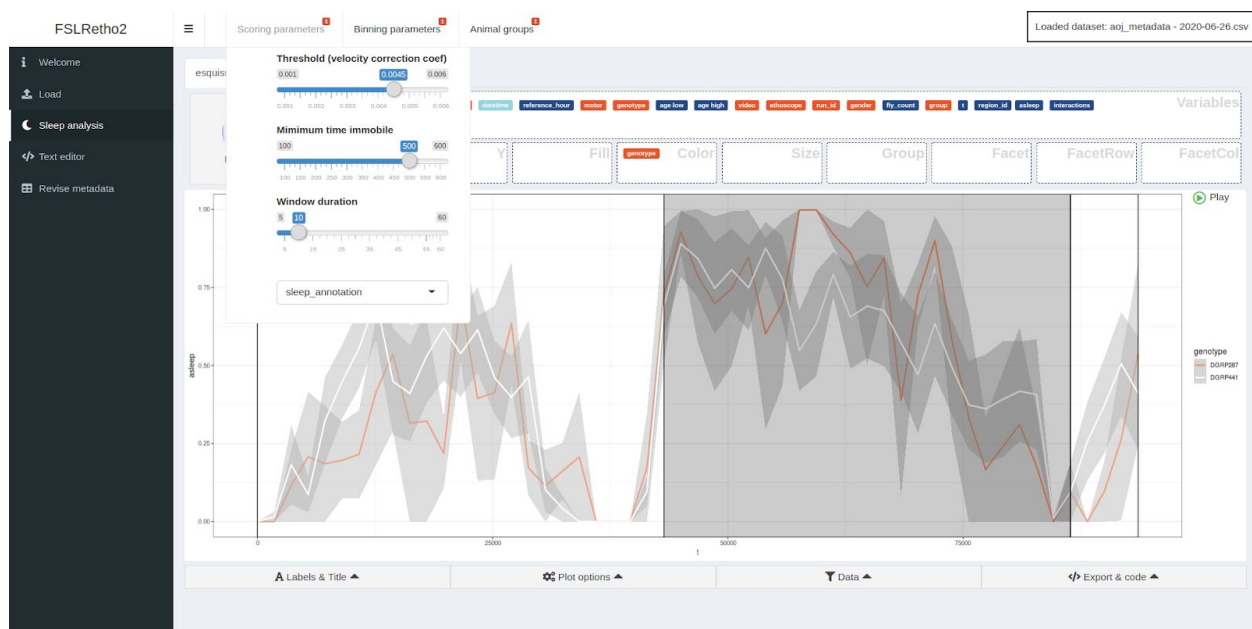
(ethoscope only) **velocity\_correction\_coef**: maximal distance crossed by a fly between two consecutive frames that is still considered non movement. The distance is defined in terms of the width of the ROI, so a value of the parameter of 0.004 means 0.4% of the length of the tube.

The **higher** this value, the **higher** the quantified amount of sleep.

(ethoscope only) **window\_duration**: length of the predefined windows that define the behavioral unit datapoint. The default is 10 seconds.

**minimum\_time\_immobile**: minimal duration of a non movement bout before it is considered a sleep bout. The default is 300 seconds (5 minute rule). The **higher** this value, the **lower** the quantified amount of sleep.

Here the same dataset is analyzed using `velocity_correction_coef=0.0038` and `velocity_correction_coef=0.0045`.



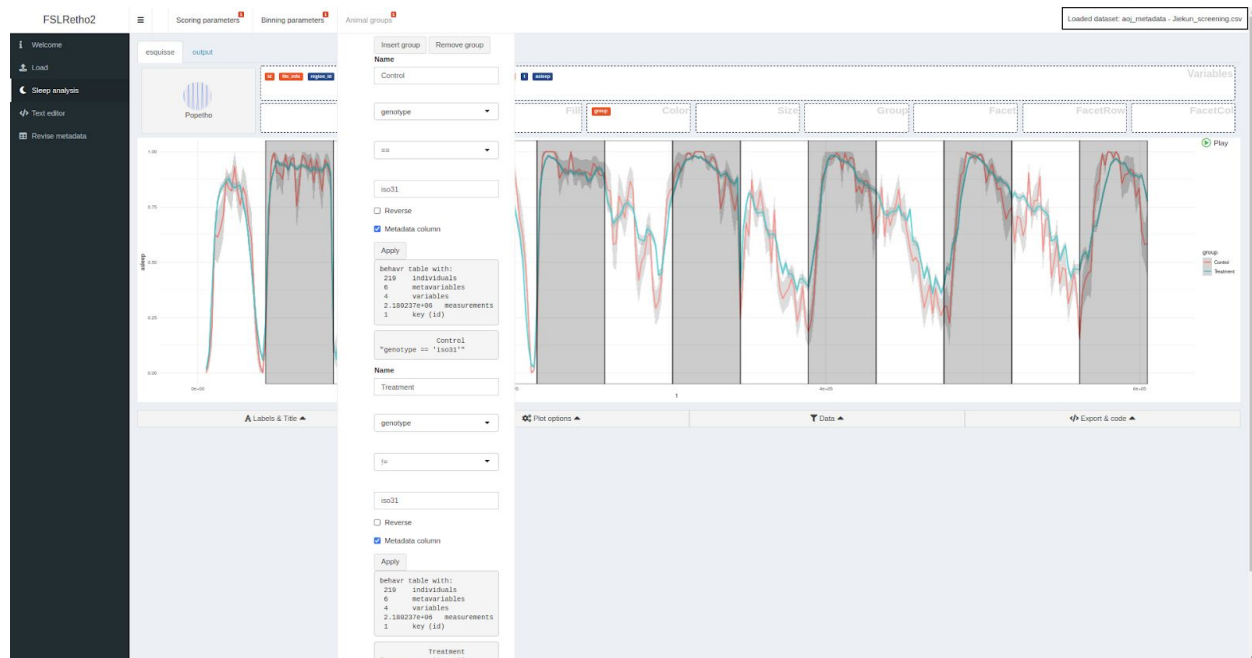
## Customize the analysis: binning

Besides controlling the scoring process, the user can also tweak how the binning of the time series is performed. More concretely, the user can change:

- The size of the bins (default of 30 minutes). The **longer** the bin, the **smoother** the shape of the plot
- The summary statistic that aggregates all the data in the bin into a single number. Besides mean, other statistics like min or max can be used. Moreover, P(Doze) and P(Wake) are supported. P(Doze) is postulated to be a proxy for sleep pressure, while P(Wake) is a proxy of sleep depth.<sup>3</sup>

## Customize the analysis: grouping

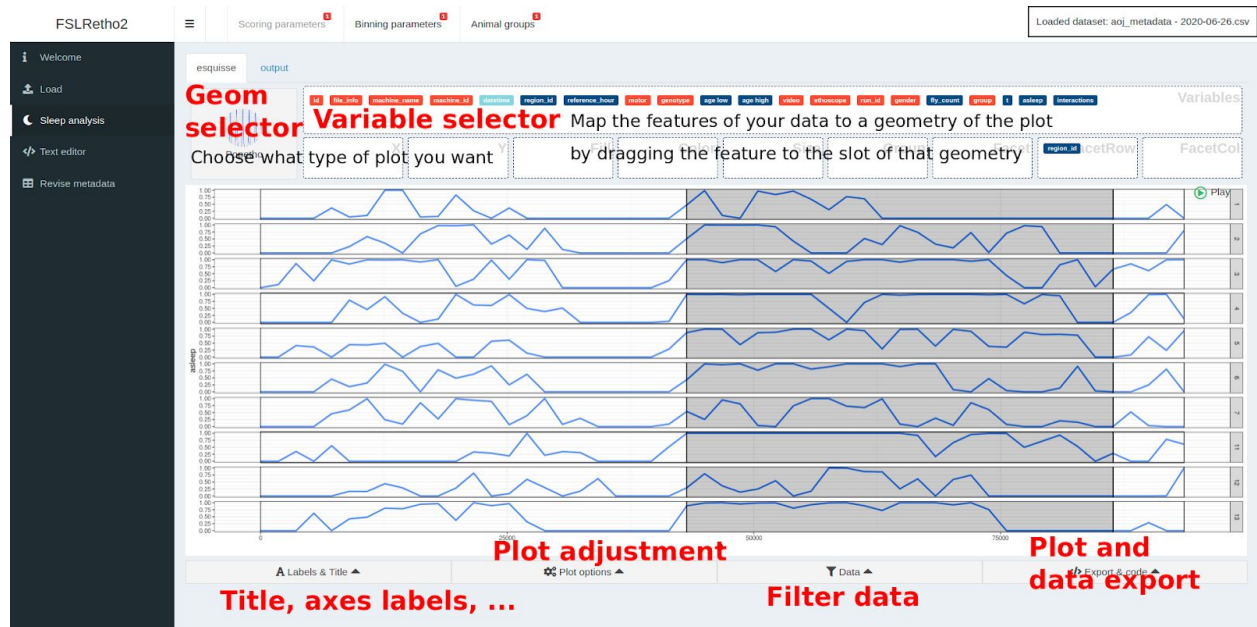
It is also possible to define groups of flies based on data or metadata variables, without the need of defining them in the metadata. For instance, one could split the animals in two groups: treatment, given by a group of genotypes, and control, made up by just iso31 flies. Such a scheme can be implemented like so:



## Advanced plotting

Esquisse allows the user to graphically customize the plot

<sup>3</sup> "Covert sleep-related biological processes are revealed by ...." 17 Apr. 2020, <https://www.pnas.org/content/117/18/10024>. Accessed 6 Jul. 2020.



A full explanation of the aesthetics and geoms available is available in the documentation of ggplot2, the R package<sup>4</sup> that provides the ground plotting functionality, and esquisse<sup>5</sup>, the GUI that wraps ggplot2.

<sup>4</sup> "ggplot2 - Tidyverse." <https://ggplot2.tidyverse.org/>. Accessed 6 Jul. 2020.

<sup>5</sup> "dreamRs/esquisse: RStudio add-in to make plots with ... - GitHub." <https://github.com/dreamRs/esquisse>. Accessed 6 Jul. 2020.



## Download

The analyzed time series as well as other data is available under **Export & Code > .csv**.

- .png: download the plot in png format.
- .pptx: download the plot as a PowerPoint document.
- .csv: .csv file containing the plotted time series.
- raw.csv: .csv file containing the scored time series before binning.
- .rds: a serialized object containing the plot and both binned and raw time series. These objects can be loaded into R via `load()`.
- The code area displays the R code that produces the shown plot from the binned dataset.



## TODO

- Support bout analysis (soon)
- Support cyclic analysis
- Convenient metadata generator
- Enable scripting interface