

# README

---

## General description

This folder `SI_swim--trackcode` contains files that allow performing the statistical analysis of swimmer trajectories obtained from optical tracking experiments. These tracking experiments produce an xml files. An example of typical output files are included here in the folder `data/Swimmers in sucrose - changing viscosity`, containing 4 different observations of swimmer tracks in xml files:

```
GUV 5Biotin-95DOPC-Rho + S500-OG_magnet ON_1kg magnet_in
sucrose_001_Tracks.xml
GUV 5Biotin-95DOPC-Rho + S500-OG_magnet ON_1kg magnet_in
sucrose_001b_Tracks.xml
GUV 5Biotin-95DOPC-Rho + S500-OG_magnet ON_1kg magnet_in
sucrose_002b_Tracks.xml
GUV 5Biotin-95DOPC-Rho + S500-OG_magnet ON_1kg magnet_in
sucrose_003_Tracks.xml
```

This data is here analysed by the scripts written in Python programming language. One easy way to install Python is to download Anaconda distribution from <https://www.anaconda.com>. Then open `Terminal` (Mac, Linux), or command line tool `CMD` (Windows), and navigate from the command line (repeat `cd foldername`) to this folder `SI_swim_trackcode`. Then individual scripts can be run simply by typing the following on the command line, for example:

```
python swim_total_distance_histogram.py
```

Note that it is important that the folder structure is preserved, i.e. all the python scripts are inside the `swim-trackcode` folder, which must also include the `data` folder containing the folder with the actual data (such as `Swimmers in sucrose - changing viscosity` folder in our example). Note that additional data folders can be provided inside the folder `data`. An access to data can be defined through the `swim_config.py` file. A snapshot of this file shows:

```
path_in = './data/Swimmers in sucrose - changing viscosity'

datafile1 = 'GUV 5Biotin-95DOPC-Rho + S500-OG_magnet ON_1kg magnet_in
sucrose_001_Tracks.xml'
datafile2 = 'GUV 5Biotin-95DOPC-Rho + S500-OG_magnet ON_1kg magnet_in
sucrose_001b_Tracks.xml'
datafile3 = 'GUV 5Biotin-95DOPC-Rho + S500-OG_magnet ON_1kg magnet_in
sucrose_002b_Tracks.xml'
datafile4 = 'GUV 5Biotin-95DOPC-Rho + S500-OG_magnet ON_1kg magnet_in
sucrose_003_Tracks.xml'

files = [datafile1, datafile2, datafile3, datafile4]
```

The `path_in` variable allows to change the data directory. For example, if our new data folder is named `my_new_data` then we could specify:

```
path_in = './data/my_new_data'
```

The variables `datafile1`-`datafile4` hold the names of the actual xml data files. New file names can be given here, also a different number of datafiles can be used. All data file variables then have to be added in the list `files` as:

```
files = [datafile1, datafile2, datafile3, ..., datafileN]
```

## Python scripts

There are three Python scripts that can be used for the analysis once the `swim_config.py` file sets the data paths:

```
swim_total_distance_histogram.py  
swim_velocity_from_total_distance.py  
swim_velocity_instantaneous_fluctuations.py
```

They are used to produce the plots of distance histograms, velocity whisker plots, and the actual raw instantaneous velocities (red lines), cumulative distance (blue lines), and autocorrelation function (green line) over the swimmer trajectories, respectively. Again, the way these files are run from the command line is as:

```
python swim_total_distance_histogram.py  
python swim_velocity_from_total_distance.py  
python swim_velocity_instantaneous_fluctuations.py
```

The output of the statistical evaluation is placed in a folder `results` created by the scripts. This output contains self-descriptive `PDF` files with the actual publication quality plots, and text files (extension `txt`) containing the data that can be used to replot the plots in a different application such as excel or origin.