# Development of tools for processing and analysis of observational data on the activity of laboratory animals

Dmitriy Borisov
*Samara National Research University*
Samara, Russia
dimaborisov290699@yandex.ru

Aleksandr Blagov
*Samara National Research University*
Samara, Russia
alexander.blagov@gmail.com

Aleksey Inyushkin
*Samara National Research University*
Samara, Russia
ainyushkin@mail.ru

*Abstract*—**The work is devoted to the analysis of observational data on the activity of laboratory animals. The activity cycles obtained by measuring sensors installed in the laboratory are considered. The software tool has been developed that provides processing, visualization and analysis of the obtained data.**

*Keywords—data processing data analysis, python, circadian rhythms, locomotor activity*

## I. INTRODUCTION

The article considers the need to develop tools for processing, analysis and visualization of data obtained by observing laboratory rats and reading their activity on a running wheel using a sensor. According to [1-5], the development of such a toolkit is an important task for researchers in the field of data analysis. The results of the analysis are necessary to study the motor activity of laboratory animals.

At the Department of Animal and Human Physiology of the Samara National Research University named after Academician S.P. Korolev many different experiments and observations are carried out to study the behavior of laboratory animals. However, for high-quality and convenient conclusion about the experiment there are not enough tools that would provide detailed analysis in addition to processing and visualization of the obtained data.

The software package developed by the authors was implemented in the Python language. This high-level programming language is designed to analyze various data. In addition, the Python contains a huge number of libraries required for simple and convenient implementation of software solutions that include visualization of various graphs.

## II. THE LOCOMOTOR ACTIVITY RESEARCHING

Circadian rhythms are one of the most important biological rhythms of living beings. They are manifested in rhythmic changes in many behavioral, biochemical and physiological parameters for a period close to twenty-four hours [6]. Circadian rhythms, or as they are also called "daily rhythms" or "internal clock" can have a strong influence on the metabolic energy balance during the day [7]. This means that this rhythm can be used to control and analyze the appetite of various living beings, depending on their "internal clock". This discovery can be extremely useful in animal husbandry.

The mammals circadian rhythm is controlled by the suprachiasmatic nucleus of the hypothalamus, which contains the main oscillator of the organism [8]. Researches have shown that many different influencing factors cause changes in circadian rhythms [6, 8]. Various environmental events influence the recovery and formation of the rhythm: cyclic afferentation about the light level, eating regimen, and metabolic signals [6].

Research of the circadian biological rhythms is actual problem of modern biology and medicine. The increased interest to this problem is related to the facts that the functional characteristics of organs and systems of the body is impossible without their circadian physiological and biochemical rhythms. Disruption of the rhythms leads to the pathological conditions [9]. Among behavioural rhythms, sleep/wake rhythm is the most representative, and in the laboratory conditions, rodents are the most convenient object for the investigation of this rhythm. When the cage where laboratory animal is contained, is equipped with the running wheel, monitoring of the wheel revolutions gives a rich information on daily profile of the output drive from the suprachiasmatic nucleus circadian oscillator. Thereby the developing of software for the data obtained from the running wheel sensor processing is of practical importance for modern experimental chronobiology.

Observations of rats are made using a special sensor installed on each running wheel. Each one-eighth of a wheel's turn represents one conventional unit, called a bin. For every five seconds of observation, the computer records the number of bins that a particular rat makes during this period of time. Thus, the locomotor activity of a laboratory rat can be represented as a set of numbers.

It may notice that the rat has periods of decreasing activity when the rat is asleep, and periods of increasing when the rat is awake. These events are clearly described using the quantitative calculation of the partial revolutions of the running wheel. Based on the data obtained, it is possible to determine the period of locomotor activity of the rat. Thus, changes in the circadian rhythm are monitored over a long period of time, including before and after the influence of various factors.

## III. SENSOR DATA PROCESSING

The initial data received from the sensors is presented in the form of a table with six columns in a text document of the ".TXT" format (fig. 1).

Each column contains the following information: time since the start of observation, the number of bins made by five rats with a time interval of five seconds. For informational purposes, the name of each file is mainly the period of the experiment on rats indicating the beginning and end of observation with an accuracy of up to a minute.

Obviously, for a better analysis it is necessary to conduct long-term observations. However, this affects the number of rows in the resulting data document. This may result in a very long table. Therefore, for convenience, using the

developed software tool, all rows in the table are grouped with a time interval (step) that the user sets himself (rows are initially grouped in 5 second increments). In this case, the bins of each rat in the table will display the number of partial revolutions of the running wheel in the required period of time. The data processor asks for the necessary interval at which the grouping will be performed. It also warns about errors and whether ungrouped lines remained due to the multiplicity of the total time to the entered step. Finally the user gets the resulting set of bins for each rat, as in the example in Figure 2.

```
FileID_text
0       0       0       1       0       11
5       4       0       0       22      31
10      21      0       0       35      9
15      38      0       0       8       8
20      36      0       0       0       0
25      24      0       1       0       0
30      15      0       1       0       0
35      8       0       0       0       0
40      2       0       0       0       0
45      1       0       0       0       0
50      0       0       0       0       0
```

Fig. 1. The contents of the source file with data on rat sightings.

```
Enter the name of the data file BEFORE the drug injection:
24.11.09.  15-00 - 3.12.09. 14-56 Insulin.TXT
Enter the name of the data file AFTER the drug injection:
27.02.10 15-00 - 8.03.10 Insulin.txt
Enter a step for the handler
that is a multiple of five seconds (<value>): 1
Enter the step unit (<s/m/h>): m
Enter the start time of the experiment,
separated by a space (to the minute): 15 00
Attention! Data is not a multiple of a step.
 The remainder was discarded: 10 sec
Attention! Data is not a multiple of a step.
 The remainder was discarded: 10 sec
Rat data before injection:
Rat № 0: 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
Rat № 1: 1 0 2 2 6 0 86 89 5 34 0 0 0 0 0 2 0 0 1 6 7 0
Rat № 2: 0 8 15 3 9 2 0 0 0 0 0 0 0 1 0 1 1 1 1 3 2 5
Rat № 3: 0 5 7 218 194 90 4 2 0 0 0 0 4 4 6 3 1 5 2 1 22
Rat № 4: 1 0 0 1 3 6 1 1 2 0 3 0 0 1 0 0 0 0 1 1 5 1 1 1
Rat data after injection:
Rat № 0: 79 1 0 2 2 0 0 0 0 0 1 2 0 87 147 147 81 175 10
Rat № 1: 0 0 0 2 2 0 0 0 0 1 0 0 0 0 0 0 0 0 2 0 0 0 0
Rat № 2: 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 2 0
Rat № 3: 2 3 2 0 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
Rat № 4: 3 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Fig. 2. Output data after processing (in 1 minute increments).

In addition to the output of the processed data to the console, it is possible to enter them into separate text documents individually for each rat with a breakdown by the number of days passed (Figures 3-4). The resulting files are entered in a specially created directory with a unique name, depending on the processing step and the name of the file being processed. For this action to create a directory, the OS library is used, which is an interface for portable use of the operating system functionality [10].

```
Enter the received data in files? (<y/Y> - yes, otherwise - no): y
A new path for output data has been created:
 result/Rats_lm_24.11.09.  15-00 - 3.12.09. 14-56 Insulin.TXT/
A new path for output data has been created:
 result/Rats_lm_27.02.10 15-00 - 8.03.10 Insulin.txt/
```

Fig. 3. An example of creating a text file in the new directory.

```
Day 1:  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
306 255 383 282 292 317 259 331 160 317 403 316 310 38
Day 2:  249 344 256 144 338 184 214 201 273 441 381 34
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
101 0 1 2 3 1 1 5 0 0 1 1 4 0 3 72 3 195 85 1 3 0 0 0
```

Fig. 4. The file with data on rat activity in 1-minute increments (part shown).

An important feature of the developed software is the visualization of the processed data. To implement this functionality, the matplotlib package is used. Matplotlib is a library for building various kinds of two-dimensional graphs, histograms, diagrams in the Python with just a few lines of code [11].

In the developed application, a pop-up window is implemented for viewing, editing and scaling of five graphs on the behavior of each rat based on the processed data with a user-selected step as shown in Figure 5.
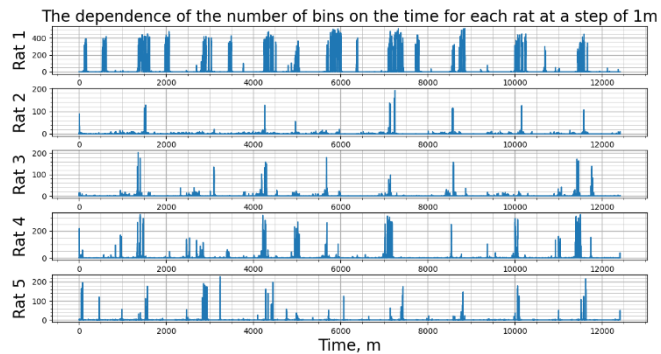


Fig. 5. The graphs of the dependence of the number of bins on time for each rat at a step of 1 minute (the X-axis shows the time in minutes, the Y-axis shows the number of bins).

## IV. DATA ANALYSIS AND DETERMINATION OF KEY CHARACTERISTIC

A comparative analysis of the behavior of a laboratory rat before and after exposure to any external facts consists in determining changes in the following indicators: the period of its circadian rhythm, acrophase and mesor, which describe the motor activity of the rat [12].

Finding a period of circadian rhythm is a rather non-trivial task. The difficulty is as follows. The period is calculated from the usual sequence of numbers (bins for a certain period of time), the distribution of which is rather approximately described by a sinusoid. The authors offer a fairly simple way to determine the period. It is based on the logic of the distribution of numbers and on one of the methods for finding the period of a sinusoid by a graphical method.

This method describes the time intervals between the maximum achieved values of the number of bins (amplitudes) over time periods equal to 24 hours, that is, equal to the approximate value of the circadian rhythm in the laboratory animal. The feature of the software implementation of the approach is as follows. The user can enter the start time of the day for each laboratory animal (for the subsequent discarding of excess data obtained on the first day before a given point in time).

This procedure is necessary in order for the general locomotor activity of the test subject to be described by a

cosine sine. This is necessary because the peak of activity in the rat is achieved at night. The algorithm for finding the biorhythm period is as follows.

The data is divided into equal time periods of 24 hours. At each interval, the amplitude is calculated. Then the time intervals between the nearest amplitudes are calculated, which are approximate values of the period of activity of the laboratory animal.

For a more accurate determination of the period, the arithmetic average of the obtained preliminary periods is calculated. According to the proposed method, the first value from the processed data can be taken as the first amplitude, since this is where the peak of activity is theoretically achieved. In addition, the first time period considered is usually 12 hours. A schematic representation of the algorithm is shown in Figure 6.
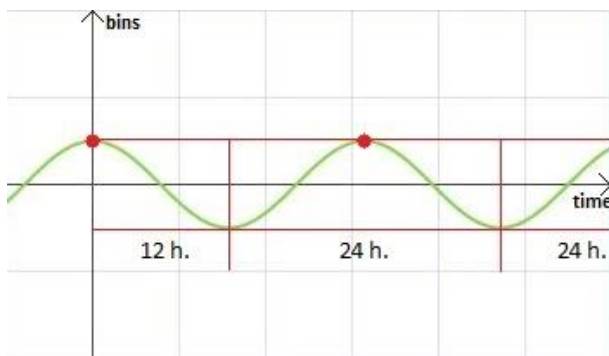


Fig. 6. Splitting data into segments (the X-axis shows the time, the Y-axis shows the number of bins).

It should be noted that during the operation of the algorithm abnormal preliminary periods can be thrown.

For example, this happens when one amplitude was found close to the end of the considered segment, and the other amplitude was calculated at the beginning of the next segment. This means that the junction of the segments is at the peak of the locomotor activity of the rat. As a result, the algorithm receives a critically small period value due to the two resulting maxima.

As a rule, such an anomaly in the algorithm provokes the opposite anomaly. With this anomaly, the resulting sample of periods will contain a critically large gap between the amplitudes. This is because the first amplitude is calculated at the beginning of the interval, and the next at the end of the other. As a result, a huge period of circadian rhythm for the laboratory rat enters the algorithm. Two opposite anomalies when calculating the average period value from the sample cancel each other out. In this case, a value close to true is obtained.

The next characteristic – an acrophase is determined by the moments of time when the rat's activity amplitude is reached. The method of finding the acrophase of circadian rhythm was chosen to find the arithmetic mean value among the moments of time when the amplitudes were reached. In other words, for each amplitude, its deviation from midnight in real time was calculated, then the average value among the received deviations was determined, which was taken as the acrophase of the circadian rhythm of the rat in question.

The mesor is the value of the average level of a sinusoid [13] or the average value of bins relative to each rat, so to

calculate it, just use the mean() function of the NumPy library.

Next, cosinor analysis is carried out - the method for approximating daily curves by harmonics (sinusoids) with a certain period, based on the least squares method [12].

To construct a sinusoid (Figure 7), in addition to the period and acrophase, a mesor and amplitude are required.
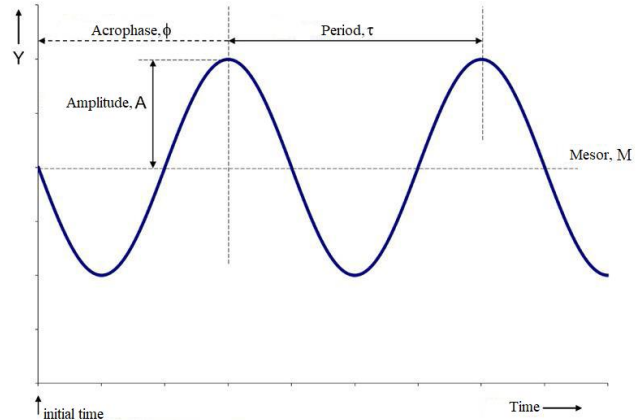


Fig. 7. The main characteristics of the circadian rhythm.

The calculated function for the sinusoid is represented by the formula (1), where M is the mesor, A – amplitude, "φ" – acrophase, "τ" – period, e(t) – random component describing other factors not included in the function.

$$Y(t) = M + A * \cos\left(\frac{2\pi t}{\tau} + \varphi\right) + e(t) \qquad (1)$$

Each of the obtained indicators is compared with the corresponding indicator before and after the influence of various influencing factors. Based on the comparison of indicators, the results and patterns of the influence of a factor on the behavior of a laboratory rat are formed.
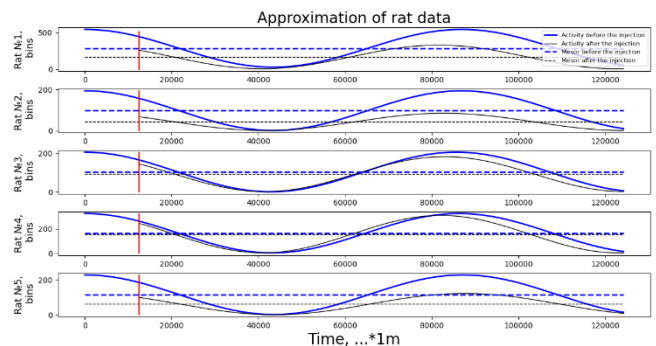


Fig. 8. Approximation of rat data (thick line - data before the injection of the insulin; thin line - data after injection of the insulin).

## V. Results

To conduct a comprehensive analysis using the developed software, we used data from observations of laboratory rats obtained for the periods from November 24, 2009 to December 3, 2009 and from February 27, 2010 to March 8, 2010 at the Department of human and animal physiology at Samara University. Based on these data, the Department conducted a research of the effect of insulin on the circadian rhythm of arbitrary locomotor activity in rats. The result of the reseacrh was the most pronounced shift of locomotor activity in the direction of phase advance at the time of injection of the medicine equal to ZT = 13 (analysis

of the acrophase shift revealed a significant advance of 5.48 ± 1.98 hours), a slight change in the total daily activity [6].

The results of data visualization by the software package are shown in Figure 8. The data obtained as a result of the analysis, shown in Table 1, generally correspond to the previously obtained results of the research on the effect of insulin on the locomotor activity of rats. In the future, the authors plan to refine the algorithms and methods for finding the main characteristics.

TABLE I.        MAIN CHARACTERISTIC

| CHARACTERISTIC | BEFORE | | | | | AFTER | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | RAT 1 | RAT 2 | RAT 3 | RAT 4 | RAT 5 | RAT 1 | RAT 2 | RAT 3 | RAT 4 | RAT 5 |
| MESOR, BINS | 28.8 | 0.7 | 1.18 | 4.56 | 1.45 | 6.8 | 0.75 | 2.08 | 4.39 | 0.27 |
| AMPLITUDE, BINS | 258 | 96.5 | 101.5 | 160.5 | 113.5 | 162 | 42 | 89.5 | 153 | 61 |
| PERIOD, SEC | 86880 | 86850 | 85665 | 86272.5 | 87157.5 | 81615 | 82830 | 82920 | 82035 | 87562.5 |
| ACROPHASE, HOUR:MINUTE | 19:59 | 22:46 | 0:24 | 22:10 | 1:49 | 20:43 | 6:31 | 7:44 | 8:33 | 21:35 |

## VI. CONCLUSION

The software tool designed to conduct a comprehensive analysis of the effect of various drugs on locomotor activity allows you to process rat observation data, present them in a visual form, and also determine the main characteristics of the circadian rhythm: activity period, acrophase, mesor. In addition, a function is constructed that describes the activity of laboratory animals in time. The software tool allows to analyze changes in the relevant characteristics before and after the influence of various influencing factors. The authors plan to expand the functionality of the developed software. In its final form, the software tool will allow more efficient and faster research of circadian biological rhythms and contribute to the emergence of new scientific discoveries.

## REFERENCES

[1] E.A. Shalimova, E.V. Shalnov and A.S. Konushin, "Camera parameters estimation from pose detections," Computer Optics, vol. 44, no. 3, pp. 385-392, 2020. DOI: 10.18287/2412-6179-CO-600.

[2] R.A. Shatalin, V.R. Fidelman and P.E. Ovchinnikov, "Incremental learning for abnormal behaviour detection algorithm based on principal components," Computer Optics, vol. 44, no. 3, pp. 476-481, 2020. DOI: 10.18287/2412-6179-CO-624.

[3] R.P. Bohush and I.Y. Zakharava, "Person tracking algorithm based on convolutional neural network for indoor video surveillance," Computer Optics, vol. 40, no. 1, pp. 109-116, 2020. DOI: 10.18287/ 2412-6179-CO-565.

[4] J. Cheng and W. Dai, "A framework of reading timestamps for surveillance video," Computer Optics, vol. 43, no 1, pp. 72-77, 2019. DOI: 10.18287/2412-6179-2019-43-1-72-77.

[5] N.A. Smelkina, R.N. Kosarev, A.V. Nikonorov, I.M. Bairikov, K.N. Ryabov, A.V. Avdeev and N.L. Kazanskiy, "Reconstruction of anatomical structures using statistical shape modeling," Computer Optics, vol. 41, no. 6, pp. 897-904, 2017. DOI: 10.18287/2412-6179-2017-41-6-897-904.

[6] A.N. Inyushkin, K.A. Mistryugov, D.S. Gromova, E.I. Pugachev, V.I. Belyakov and E.M. Inyushkina, "Influence of insulin on circadian rhythm of arbitrary locomotor activity of rats," Russian journal of physiology, vol. 97, no. 7, pp. 678-689, 2011.

[7] I. Boumans, E. Bokkers, G.J. Hofstede and I. Boer, "Explaining daily feeding patterns in pigs: modelling interaction between metabolic processes and circadian rhythm," Proceedings of the 50th Congress of the International Society for Applied Ethology, pp. 344-344, 2016.

[8] M.H. Hastings, M. Brancaccio and E.S. Maywood, "Circadian pacemaking in cells and circuits of the suprachiasmatic nucleus," Journal of neuroendocrinology, vol. 26, no. 1, 2014. DOI: 2-10. 10.1111/jne.12125.

[9] J.A. Evans and A.J. Davidson, "Health consequences of circadian disruption in humans and animal models," Progress in molecular biology and translational science," Academic Press, vol. 119, no. 1, pp. 283-323, 2013. DOI: 10.1016/B978-0-12-396971-2.00010-5.

[10] W.H. Walker, "Circadian rhythm disruption and mental health," Translational Psychiatry, vol. 10, no. 1, pp. 1-13, 2020. DOI: 10.1038/s41398-020-0694-0.

[11] Matplotlib: Python plotting - Matplotlib 3.1.3. documentation [Online]. URL: https://matplotlib.org/.

[12] D.E. Pastukhov, "Application of the kosinor-analysis method in chronobiological studies of microbial properties," Tyumen state University, 2016.

[13] L.A. Omelyanchik and N.B. Kolesnik, "Kosinor analysis of daily activity of pentose phosphate pathway dehydrogenases in the liver of mongrel rats," Topical issues of biology, ecology and chemistry. Zaporozhye national University, vol. 3, no. 2, pp. 28-34, 2011.