Biomedical Signal Processing and Control 81 (2023) 104409

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Understanding the complex interplay of persistent and antipersistent regimes in animal movement trajectories as a prominent characteristic of their behavioral pattern profiles: Towards an automated and robust model based quantification of anxiety test data

Mikhail I. Bogachev [a,b,*], Asya I. Lyanova [c], Aleksandr M. Sinitca [c], Svetlana A. Pyko [a], Nikita S. Pyko [a], Alexander V. Kuzmenko [a], Sergey A. Romanov [c], Olga I. Brikova [c], Margarita Tsygankova [a], Dmitry Y. Ivkin [d], Sergey V. Okovityi [d], Veronika A. Prikhodko [d], Dmitrii I. Kaplun [c], Yuri I. Sysoev [d,e,f], Airat R. Kayumov [b]

[a] Centre for Digital Telecommunication Technologies, St. Petersburg Electrotechnical University "LETI", 5-F Professor Popov street, St. Petersburg, 197022, Russia
[b] Institute for Fundamental Medicine and Biology, Kazan Federal University, 18 Kremlevskaya street, Kazan, 420008, Tatarstan, Russia
[c] Department for Automation and Control Processes, St. Petersburg Electrotechnical University "LETI", 5-F Professor Popov street, St. Petersburg, 197022, Russia
[d] Department of Pharmacology and Clinical Pharmacology, St. Petersburg State Chemical Pharmaceutical University, 14 Professor Popov street, St. Petersburg, 197022, Russia
[e] Institute of Translational Biomedicine, St. Petersburg State University, 7-9 Universitetskaya embankment, St. Petersburg, 199034, Russia
[f] Pavlov Institute of Physiology, Russian Academy of Sciences, 3 Tiflisskaya street, St. Petersburg, 199034, Russia

ARTICLE INFO

Keywords:
Animal model
Zebrafish
Novel tank test
Detrended fluctuation analysis
Fractional Brownian motion
Random walk
Persistence
Antipersistence

ABSTRACT

Rapid advancement in computer vision technologies provides increasing opportunities for the quantitative characterization of animal behavior, although reduction of their analysis to several scalar metrics appears a common limitation for the representation of complex behavioral patterns. Here we suggest an alternative approach to the quantitative assessment of animal behavioral patterns by parameterization of a generalized scalable model based on fractional Brownian motion using detrended fluctuation analysis of the observational movement trajectories and validate it using novel tank test data. In a zebrafish model representative movement patterns are characterized by two asymptotic regimes, with persistent increments at small scales and antipersistent increments at large scales. A single crossover between these asymptotic regimes that appears a single free parameter of the animal movement model acts as a complementary behavioral indicator leading to a more explicit characterization of both stimulative and sedative effects. We show explicitly that the model can be also used for a robust estimation of interpretable scalar metrics commonly used in behavioral analysis leading to the emphasized differences between experimental groups. We believe that this approach, due to its universality, robustness and clear physical interpretation, is a perspective tool for the analysis of animal behavior complexity under various experimental and natural conditions.

1. Introduction

Analysis and interpretation of animal movement patterns in various experimental and natural environments is an important challenge in biomedical and other natural sciences, ranging from fundamental animal biology and ecological monitoring to applied biomedical and pharmacological studies, respectively [1,2], see also [3–6]. Recent advancements in computer vision technologies lead to increasing availability of video analysis based information including animal movement patterns in various biomedical research and ecological monitoring applications. In turn, better understanding of these movement patterns requires simple, interpretable, and scalable approaches to their quantitative characterization, leading to the development of fast, accurate, and cost-effective algorithms based on computer vision.

Experimental drug screening is a prominent application where analysis and interpretation of model organism movement patterns are widely used to overcome certain inevitable limitations of target-based

* Corresponding author at: Centre for Digital Telecommunication Technologies, St. Petersburg Electrotechnical University "LETI", 5-F Professor Popov street, St. Petersburg, 197022, Russia.
E-mail address: rogex@yandex.com (M.I. Bogachev).

https://doi.org/10.1016/j.bspc.2022.104409
Received 17 June 2022; Received in revised form 10 September 2022; Accepted 7 November 2022
Available online 1 December 2022
1746-8094/© 2022 Elsevier Ltd. All rights reserved.

in vitro assays that are particularly relevant for psychoactive drugs causing behavioral changes in animals that in turn reflect the complexity of their vertebrate nervous system response to the pharmacological stimuli [7].

For many years, the open field test remains the most widely used test for the quantitative characterization of behavioral patterns in experimental animal rodent models [8–12]. Over more than half a decade the open field test paradigm has developed into a powerful tool for the evaluation of animal locomotion and exploratory activity, as well as risk assessment and anxiety behavior, with a number of established quantitative markers for each of the behavioral characteristic studied.

Despite an impressive list of experimental models and behavior tests for rodents, zebrafish *Danio rerio* has an increasing popularity in neuropharmacological studies. Zebrafish as a model organism is highly sensitive to pharmacological manipulations that, due to substantial genetic, physiological and pharmacological homology with mammals [13], low cost, and reproducibility, represent an increasingly common model for high-throughput drug screening [14]. Despite the seemingly relative simplicity of the central nervous system, *Danio rerio* demonstrate an extensive behavior repertoire [15].

In recent years a significant success has been achieved in the understanding of drug-sensitive phenotypes of anxiety, such as geotaxis and scototaxis, which were pharmacologically and behaviorally validated [16–18]. To assess the anxiety-related behavior in zebrafish, several behavior paradigms such as the novel tank test, light-dark box, shoaling and several others have been proposed. The novel tank test is the most commonly used behavioral paradigm in adult zebrafish, conceptually similar to the rodent open field test and exploiting the natural behavior of zebrafish to seek protection in an unfamiliar environment by diving to the bottom and remaining there until they feel safe enough to explore [16], see also [19–24].

Interpretation of the results obtained in the novel tank test is commonly based on the quantitative assessment of a series of movement pattern characteristics, for example, the time spent in the bottom zone of the tank, as well as the number and total time of freezing bouts. Manual registration of behavior is often time-consuming and depends of the investigator's subjective point of view potentially reducing the validity of the results. Moreover, several parameters can hardly be scored manually (e.g. traveled distance or mean and maximum velocity). To partially overcome the above limitations, computer vision-based solutions are essential for the automated and unbiased analysis of large amounts of video data samples, and video-tracking systems are widely applied to zebrafish behavior research [25–27].

Among available solutions, Ethovision XT is the leading commercial software widely used for video tracking of zebrafish behavior, which can be successfully applied for social interaction tracking, anxiety-like behavior, exploration, swimming patterns, or just general activity in both adults and larvae [28–31]. Despite the fact that Ethovision XT has wide functionality and is applicable not only for zebrafish, but for almost any model organism, this software is not amenable to custom-made modifications.

In the common behavioral test analysis procedure, once the animal movement trajectories have been extracted, their further analysis is commonly reduced to a number of predefined characteristics, in most cases represented by a set of scalar metrics. For example, in the open field rodent test locomotion characteristics include the total distance traveled and the total zone entries, vertical activity is characterized by the rear frequency, rear duration and grooming, while the risk assessment is justified from the total stretch attend posture and the total sniffing events count, and the decision making patterns are typically interpreted based on the properties of the periphery zone return and corner zone return events. Altogether, the total anxiety level is justified from the decreased total locomotor activity, lower distance traveled in central zone, lower of time spent in central zone, higher fraction of time spent in the periphery zone, as well as in the corners [32,33]. Very similar characteristics are evaluated in the novel tank test, including the total distance, the average and the maximum velocity, fading frequency and cumulative duration, as well as durations and transitions between the top, central and bottom zones, respectively (see, e.g., [34–36] and references therein).

Although the above characteristics typically have a clear underlying physical interpretation, being analyzed as single variables, they are often insufficient to characterize the whole complexity of the animal movement patterns. In turn, more sophisticated multiparametric models are required to extract further significant information partially hidden in the interactions between these characteristics. One common way to combine multiple scalar parameters is based on multiple regression analysis such as binary logistic regression or maximum covariance analysis when it comes to distinguishing between two different movement scenarios (such as test vs. control comparison). Although these methods are rather simple, their disadvantages include specific model parameters for each particular binary discrimination scenario that are obtained during an inevitable learning procedure, as well as limited interpretability of the combined parameter used in the discrimination problem.

As a possible alternative, a model-based paradigm of animal movement pattern representation is often considered. Several models of animal movement trajectories can be found in the recent literature with application to various experimental and natural environments [37–40]. In particular, models based on the (fractional) Brownian motion (fBm) and its modifications has attracted attention in recent years [3–6]. One of the keynote features of fBm models is long-term persistence implying asymptotically unlimited correlation times that have been identified in multiple processes characterizing various natural complex systems in the last decades. The fBm based mathematical models are also scalable that makes them attractive for modeling time series with finite and often variable observation times. Finally, in marked contrast to models artificially combined from multiple scalar movement metrics, fBm based models have only few free parameters that are easily physically interpretable (for further details, we refer to [1] and references therein).

As we show below, in the context of this model, typical movement patterns during neuropharmacological experiments exhibit different crossover positions between persistent and antipersistent regimes depending on the applied pharmacological stimuli. Moreover, we also show how the model parameters can be associated with the conventional metrics such as the level crossing statistics characterizing zone transitions events, thus making the corresponding results reproducible from the model based perspective and interpretable in the context of commonly used animal movement metrics. We believe that the proposed approach, although validated here just for the novel tank test study, could be also applicable to other similar tests such as the rodent open field test.

2. Methods

2.1. Experimental protocol

Wild-type short-fin adult zebrafish (Danio rerio) (n = 96, 50:50 male:female ratio, 4–6 months old) were obtained from a local commercial supplier (Aksolotl Ltd., St. Petersburg, Russia) and housed in groups (15–20 fish per tank) in 40-L glass tanks filled with filtered facility water for at least 20 days prior to the novel tank testing (room and water temperature was maintained at 25–27 °C, and water pH at 7.0–7.5) [18]. Fluorescent ceiling-mounted light tubes provided room illumination on a 14/10 light/dark photoperiod cycle (lights on at 7:00 am and off at 9:00 pm). The animals were fed three times daily with commercial flake food Tetramin-Pro (Terta GMBH, Osnabrück, Germany). Animal experiments were approved by the Institutional Animal Care and Use Committee Reference: F-MORFO-SA-21, 20.01.2021, and fully adhered to the National and Institutional guidelines and regulations on animal experimentation in adherence to the principles

Table 1
Characterization of experimental groups, effectors, doses and notations.

Group notation	Animal count	Dose	Boxplot notation
9j control	n = 16	0 mg	0
9j 1 mg	n = 16	1 mg	+
9j 100 mg	n = 16	100 mg	++
Caffeine control	n = 16	0 mg	0
Caffeine 50 mg	n = 16	50 mg	+
Caffeine 100 mg	n = 16	100 mg	++

of the Declaration of Helsinki. All animals tested were included in final analyses, while for some indices individual values have been excluded when considering certain quantities sensitive to the presence of outliers (see below).

In the following, all animals were divided in six groups, with n=16 animals in each group, as outlined in Table 1.

Two groups of animals were treated with an investigational allylmorpholine derivative 9j ((E)-4-[1-(8-bromo-6-methyl-4-oxo-4H-chromen-3-yl)-4-methylpent-1-en-3-yl]morpholin-4-ium chloride) (1) that has been identified as a potent acetycholinesterase and butyrylcholinesterase inhibitor, as well as an N-methyl-D-aspartate glutamatergic receptor antagonist in vitro [41] at the concentrations of 1 mg L^{-1} and 100 mg L^{-1}, respectively, while the third group used as a relevant control. Of note, in a recent screening experiment in zebrafish, 9j was found to produce a dose-dependent sedative effect within a $1 \ldots 100$ mg L^{-1} concentration range in the Novel tank test [42].

From three remaining animal groups, two groups were treated with caffeine (1,3,7-trimethylpurine-2,6-dione) (2), a methylxanthine alkaloid, is known for its psychostimulant-like action, possibly due to the central A1 adenosine receptor antagonism. Acute exposure to 100 mg L^{-1} caffeine was shown previously to increase alertness and induce anxiety-like behavior in zebrafish in the novel tank test [43]. The remaining group served as a relevant control. Of note, to eliminate possible discrepancies due to the time lag between series of experiments with 9j and caffeine, two independent control groups have been used.

Behavioral testing was performed between 12.00 and 14.00 h. The novel tank apparatus was represented by a plexiglass container ($20 \times 20 \times 5$ cm, which restricted movement along the third dimension, thus enforcing "two-dimensional" zebrafish movement patterns), virtually divided by a marker line into the upper and lower halves [Egan, 2009]. The trials were recorded with a CNE-CWC web camera (CANYON Co., Gorinchem, The Netherlands) for further analyses. Further processing was performed offline to calculate the total distance (cm), average and maximum speed (cm/s), the stop duration (s) (velocity < 2 cm/s), time spent in the top zone (s), crosses count and the first ascent latency (s). For that, we applied an in-house developed algorithm and software written in Python using computer vision methods, indicated as module A in Fig. 1. For the initial validation of the developed algorithm and software tool, for the 9j test groups and the respective control group, we also applied Ethovision XT 11.5 software (Noldus IT, Wageningen, The Netherlands) as a reference.

2.2. The overall study design

The overall study design is summarized in Fig. 1 where the first pipeline (indicated by blue arrows) reproduces the conventional approach to the analysis of animal movement trajectories based on the estimation of seven scalar metrics similar to those implemented in the Ethovision XT software tool.

The second pipeline (indicated by violet arrows) represents an alternative model-based solution proposed here that focuses on the fluctuation analysis of animal movement trajectories and the estimation of the model parameters based on their fluctuation functions. In this scenario, the free model parameters are directly used in the decision making procedure to distinguish between characteristics movement patterns observed in the studied control and test groups, respectively.

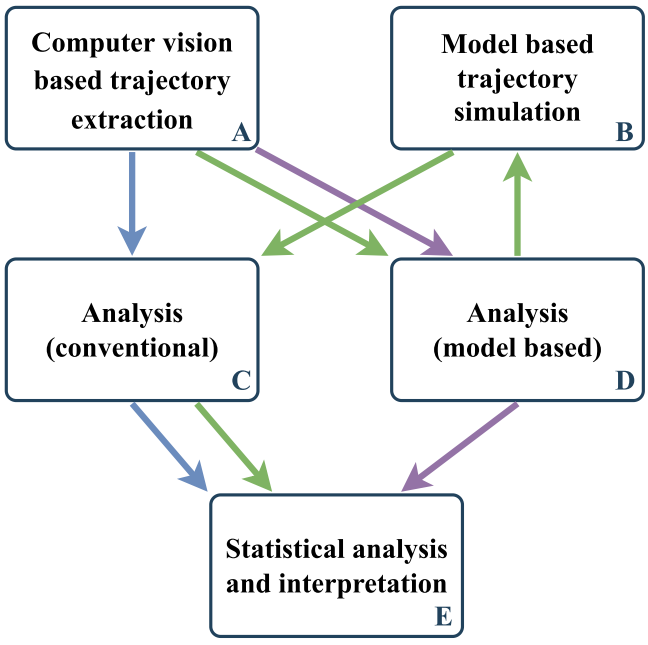

Fig. 1. The overall study design, including (A) computer vision based trajectory extraction, (B) a simulation algorithm that generates surrogate trajectories reminiscent to the observational ones, (C) conventional analysis based on the estimation of multiple scalar metrics, (D) an alternative approach based on the animal movement pattern model indentification by fluctuation analysis that we propose here, and (E) the final statistical analysis of the study results. Blue, violet and green arrows represent different analysis pipelines that we have tested in our study.

The third pipeline (indicated by green arrows) focuses on the reproducibility of the conventional metrics (originally obtained from the first pipeline) from the surrogate trajectories that are artificially simulated with the model parameters that best reproduce the statistical properties of the observational trajectories in terms of their fluctuation functions.

Finally, statistical analysis is considered both for the results of each single pipeline and for the combination of the results obtained using different methods, and thus represented by outcomes of different pipelines in order to combine potentially complementary features extracted by different approaches.

The modules and pipelines implemented in Python, as well as source data used for their validation, are provided in the Supplementary material.

2.3. Extraction of movement trajectories by computer vision method (module A)

In our in-house developed computer vision based algorithm, each individual frame extracted from the video sequence is subjected to the procedure as follows.

1. First, the color images are converted to grayscale as

$$GS_{ij} = 0.299R_{ij} + 0.587G_{ij} + 0.114B_{ij}, \quad (1)$$

where GS_{ij} is the grayscale pixel intensity, R_{ij}, G_{ij}, B_{ij} are the red, green, and blue channel intensities, respectively.
2. Next, the images are smoothed with Gaussian blur

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (2)$$

where x, y are the coordinates of the point, and σ is the standard deviation.
3. Finally, the image is binarized using the automatic Otsu thresholding algorithm.

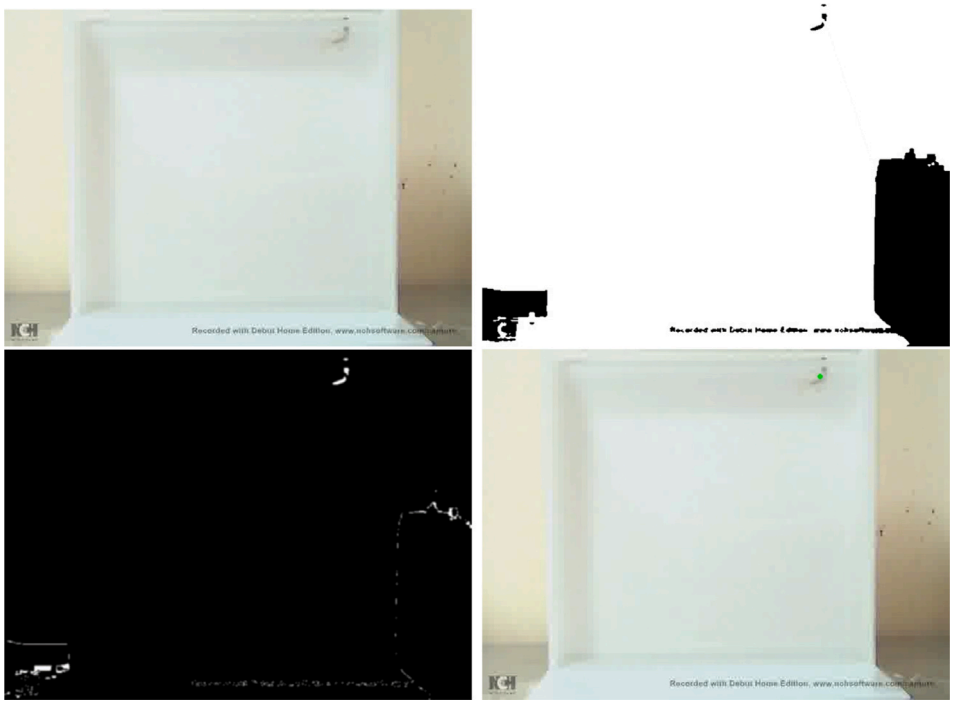

Fig. 2. Original frame from the video (upper left), frame after the Otsu thresholding algorithm (upper right), frame after background subtraction (lower left), visualization of the result (lower right).

In the following, we analyze the preprocessed frame sequence with algorithm 1. This algorithm is commonly used for different tasks of moving detection with a static background. The illustration of the algorithm is provided in Fig. 2.

Algorithm 1 Video processing

$Video$ = array[n, w, h, 1]
$BG = Video[0]$
for i in $(1, ..., len(Video))$ **do**
  $frame = Video[i]$
  $f = frame - BG$
  $cntrs, areas \leftarrow$ find contours and its areas in f
  reject $cntrs$ with $areas > threshold_{upper}$ or $areas < threshold_{lower}$
  $cntr_t \leftarrow$ select contour from $cntrs$ with largest area
  $M_{00}, M_{01}, M_{10}, M_{11} \leftarrow$ calculate Image moment for $cntr_t$
  $x_i = \frac{M_{10}}{M_{00}}$
  $y_i = \frac{M_{01}}{M_{00}}$
end for

The algorithm implemented in Python is provided together with the source data used for its validation in the Supplementary material.

2.4. Conventional metrics estimation (module C)

Seven conventional animal movement trajectory metrics have been estimated using the developed software module, indicated as module C in Fig. 1. In particular, the considered metrics include (location of respective metrics in the Ethovision XT 11.5 software is given in parentheses):

1. *Total distance* traveled by the fish during the experiment [cm] (Distance moved →center-point →Total).
2. *Average speed* defined as the total traveled distance over the duration of the experiment [cm/s] (Velocity →center-point →Mean).
3. *Stop duration* estimated as the total time the fish was in the fading state (Movement →Not Moving/center-point).
4. *Crosses count* indicating the total number of transition events between the top and the bottom zones during the experiment (count) (Bottom-top →Top/center-point →Frequency).
5. *Maximum speed* among local estimates in 5-point window (corresponding to 0.2 s duration) [cm/s] (Velocity →center-point →Maximum).
6. *Sum top time* estimated as the total time spent in the top zone of the tank (s) (In top zone →Top/center-point →Cumulative Duration).
7. *First ascent latency* indicating the time until the first ascent from the bottom to the top of the tank (Zone transition →center-point/Bottom/Top →Latency to First).

2.5. Estimation of model parameters (module D)

According to the module D in Fig. 1, for the identification of the animal movement trajectory model and the estimation of its parameters from empirical trajectories, we employed the detrended fluctuation analysis (DFA) originally introduced by Peng et al. [44] and thoroughly investigated since, for example, in [45]. In the DFA procedure, the cumulative sum of the data series is considered, $Y(i) \equiv \sum_{k=1}^{i} y_k$, $i = 1, \dots, L$, where L is the length of the data, which is sometimes also called "profile" or "landscape". Next, the entire profile is split into K_s windows of length s, and in each window the least mean squares polynomial fit $p_\nu(i)$ is calculated. Then, the variance

$$F_\nu^2(s) \equiv \frac{1}{s} \sum_{i=1}^{s} [Y((\nu - 1)s + i) - p_\nu(i)]^2 \tag{3}$$

is determined in each local window ν. DFA with polynomials p of order m is capable of eliminating background trends of order m in the profile, that corresponds to trends of order $m - 1$ in the raw data series. In general, the choice of polynomial order m used for the local trend fitting depends on the *a priori* assumptions regarding the complexity of the analyzed trajectory, although recent data indicate that for the majority of observational processes in natural sciences $m = 2$ appears sufficiently

accurate (see, e.g. [45]). In the last step one averages $F_\nu^2(s)$ over all windows and takes the square root to obtain the average fluctuation function $F(s)$.

$$F(s) \equiv \left[\frac{1}{K_s}\sum_{\nu=1}^{K_s} F_\nu^2(s)\right]^{1/2}. \tag{4}$$

It is known that for long-term correlated data the DFA fluctuation functions increase by a power law $F(s) \propto s^H$, where H is the Hurst exponent, irrespective of the order of the detrending polynomial. In the simple case of fully random ("white noise") data series $H = 1/2$. For short-term correlated records, $H \to 1/2$ for s above the correlation time $s_\times$. Compared to many widespread alternatives, the DFA method is comparatively robust to finite-size effects, providing accurate $F(s)$ estimates at scales $s \approx 8 \ldots L/4$.

An additional parameter for simulation is the mean square deviation D, which is estimated on the original trajectories. This parameter is not relevant for the analysis but requires a correct trajectory reconstruction.

2.6. Model-based trajectory generation (module B)

To simulate surrogate animal movement trajectories, we follow a standard procedure to generate data series with a given fluctuation function that is based on their basic properties [45]. The most common way to generate a stationary random series with a given Hurst exponent $0 < H < 1$ is by using the Fourier filter with transfer function $K(f) \propto f^{-\beta/2}$, where β is the exponent characterizing the decay of the spectral density $P(f) \propto f^{-\beta}$. According to the Wiener–Khinchin theorem, β can be obtained as $\beta = 1 - \gamma$, where γ is the exponent characterizing the decay of the autocorrelation function (ACF) $C(s) \propto s^{-\gamma}$, that is in one-by-one correspondence with the Hurst exponent $\gamma = 2 - 2H$. In the context of movement trajectories simulation, the above procedure could be used to generate movement increments.

Finally, the non-stationary trajectory itself can be obtained by further integration of the increments. It is well known that integration of the data series leads to the increment of the Hurst exponent H, while calculation of its moving average results in its integration over a limited scale span, leading to a characteristic crossover in the fluctuation function.

2.7. Statistical analysis (module E)

For the outliers removal, we employed the Tukey fence method based on the analysis of the interquartile range which is known for its resistance to the presence of extreme values and applicability to both normal and slightly skewed distributions. Next, to determine relative effects, we have re-normalized all empirically obtained values to the median of the same value obtained for the control group in the respective experimental series.

Distributions of movement metrics within the groups in many cases differed significantly from Gaussian distributions, as indicated by Kolmogorov–Smirnov and Shapiro–Wilk tests. Thus, for the assessment of the statistical significance of our results, we employed the Kruskal–Wallis statistical test for multiple comparisons and Mann–Whitney U-test for pairwise comparisons, respectively. However, since the one-factor analysis of variance (ANOVA), which is generally more powerful than the above non-parameteric tests, is known for its validity also for non-Gaussian distributed data with moderate distributional asymmetry, for a further confirmation, we also employed this test in combination with the Tamhane's T2 criterion for unequal variances for pairwise comparisons. To confirm statistically significant reduction of variances in the model based estimations of conventional scalar movement metrics compared to their direct estimates based on the video analysis procedure, we used Levene's statistical test. We consider the differences as statistically significant at the significance level $p < 0.05$.

To combine several potentially complementary metrics into a single decision making rule, we applied the logistic regression model approach. To test whether the prediction accuracy can be improved using a combination of metrics, we followed a stepwise procedure as described below. In the first step, for each experimental group we calculated all considered metrics. In the second step, we eliminated those metrics that did not indicate any statistically significant differences between the considered experimental groups according to the Mann–Whitney U-test. In the third step, possible combinations of the remaining metrics were included in the logistic regression model in a stepwise manner. The particular order of the metrics to be included in the model was determined by their Wald statistics, which are proportional to their significance levels as single pairwise classifiers. Once an additional classifier was included, we calculated the probability that the studied metrics belongs to one of the considered groups $p = 1/[1 + exp(-z)]$, where $z = B_0 + B_1 x_1 + B_2 x_2 + \cdots + B_m x_m$ is a linear combination of considered metrics x_i. The quality of the model fitting to the empirical data is quantified by the coefficient of determination R^2, which typically increases after the addition of classifiers, and converges to a plateau, when further addition of classifiers no longer improves the model fitting accuracy. In the fourth step, the effectiveness of the resulting model for pairwise classifications is validated by calculating p for each group and comparing it against a decision threshold $\Theta = 1/2$. Whether p is smaller or larger than Θ, the group is classified as belonging to the first or second group, respectively. In practical scenarios, the distributions of both individual metrics and their combinations in the logistic regression models for different experimental groups exhibit inevitable overlaps, leading to classification errors (for a recent example see, e.g., [46]). To quantify the classification accuracy, we next calculated the sensitivity or recall, also known as the true positive rate (TPR)

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}, \tag{5}$$

the specificity, also known as the true negative rate (TNR)

$$TNR = \frac{TN}{PN} = \frac{TN}{TN + FP}, \tag{6}$$

and the precision, also known as the positive predictive value (PPV)

$$PPV = \frac{TP}{TP + FP}, \tag{7}$$

where TP is the total number of true positive, TN of true negative, FP of false positive, and FN of false negative predictions, respectively.

3. Results

3.1. Conventional metrics estimation (the blue pipeline)

First, to validate the developed computer vision-based trajectories extraction procedure, for the animal groups treated with 9j and for the respective control group, we also estimated seven conventional metrics using both the Ethovision XT 11.5 software and using the original in-house developed algorithm. The results are summarized in Figs. 3 and 4, respectively.

3.2. Fluctuation analysis of observational trajectories and model identification (the violet pipeline)

Next, we analyze the same trajectories using DFA. Fig. 5 exemplifies the fluctuation functions for two test groups (a,c,d,f) with maximal effector doses and two respective control groups (b,e). The figure indicates that, at least on the average, to a very good approximation, the fluctuation functions exhibit two asymptotic regimes separated by a single characteristic crossover. At small scales $F(s) \propto s^2$ indicating pronounced persistence in the movement increments, while at large scales $F(s) \propto s$ indicating the emergence of antipersistence in the movement increments. Remarkably, these two regimes are reproduced quite well in all studied groups, especially for the y-axis projection of the movement trajectory, while for the x-axis projection, some

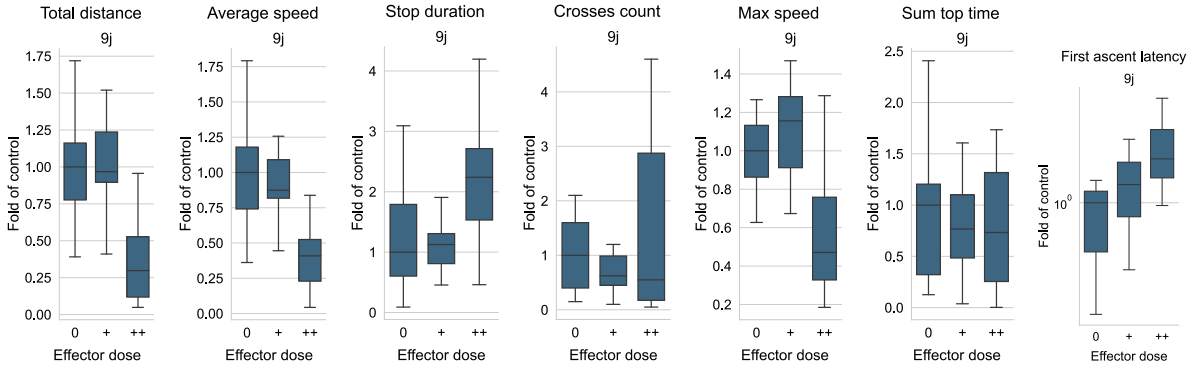

Fig. 3. Boxplots for seven different scalar metrics obtained by Ethovision XT 11.5 software.

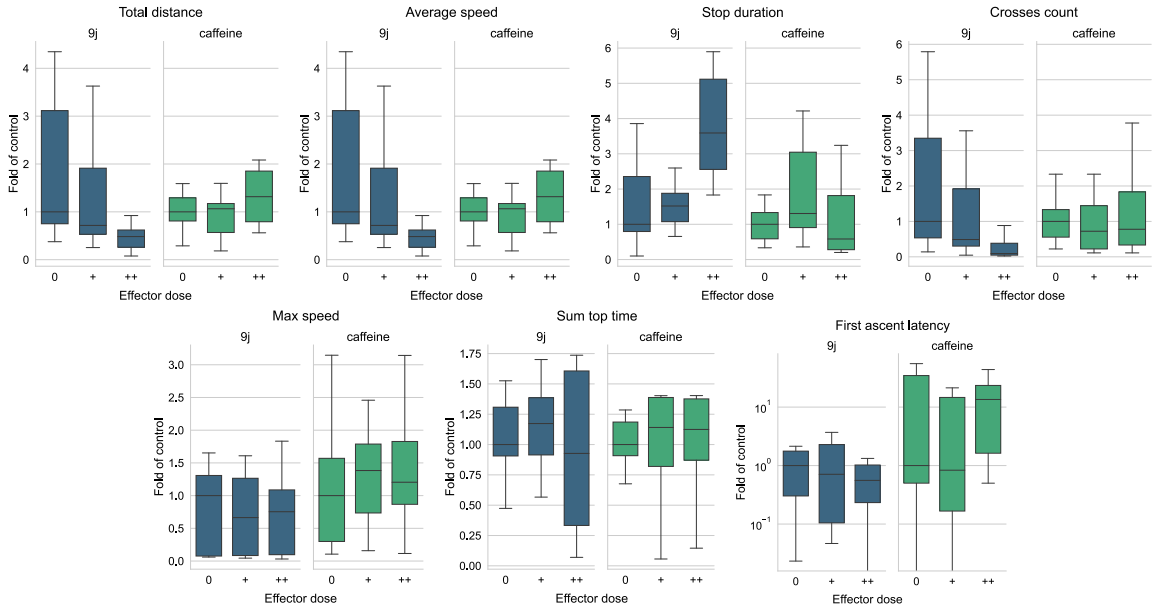

Fig. 4. Boxplots for the same seven metrics obtained by the original computer vision algorithm.

moderate deviations from this pattern can be observed at large scales only. When considering averaged fluctuation functions for each group, particular positions of the crossover between the two asymptotic scaling regimes likely represent a single free model parameter (for each projection of the movement trajectory) that explicitly reflects the alteration of the movement patterns under respective pharmacological stimuli. Accordingly, the respective crossover positions could act as an alternative to the conventional metrics, since they characterize the movement trajectory model, and thus the conventional metrics could be reproduced based on the model with respective parameters. Since two independent control groups have been used for the study with 9j and Caffeine, the central panels (b,e) indicating results for control groups shows two green bold curves in each. Similarly, two different dashed lines are used to illustrate rough asymptotic fits to the respective group averages. In other panels, additional curves (marked by different colors) indicate similar fluctuation functions for the simulated data, for typical model parameters. Moderate discrepancies could be observed between two control groups, indicating that further analysis and interpretation should rather focus on relative values to the respective controls.

To estimate the crossover position between the characteristic regimes $F(s) \propto s^2$ and $F(s) \propto s$, we next rotated the fluctuation functions by dividing $F(s)$ over $s^{3/2}$, where the exponent $3/2$ represents the average between the two asymptotic regimes $H = 2$ and $H = 1$, respectively. Accordingly, assuming that the asymptotic behavior holds, the "rotated" fluctuation function $F(s)/s^{3/2}$ would exhibit the same decay rates on both sides of the crossover, represented by a (smoothed) triangle a double logarithmic scale, and thus the position of the crossover can be obtained from its global maxima, as exemplified in Fig. 6 for the averaged fluctuation functions within each of the studied experimental groups. As shown in Figs. 7 and 8, the crossover positions reflect the specific response to the pharmacological stimuli indicated by explicit discrepancies between experimental groups. Of note, deviations of some fluctuation functions from the triangular shapes are an artifact of considerable zooming, in comparison with Fig. 6.

3.3. Model based trajectories simulation (the green pipeline)

In order to confirm the reproducibility of the conventional scalar movement metrics by the model, we next suggest a simulation algorithm that generates surrogate movement trajectories with the same statistical properties like in the empirical ones, based on just two model parameters s_x and s_y estimated for each individual animal movement trajectory. The algorithm is illustrated in Fig. 9. First, we generate white Gaussian noise (WGN) characterized by $H = 1/2$ at all scales. Second, we pass WGN through a $1/f$ filter leading to the enhancement of the Hurst exponent $H = 1$. Third, to obtain the resulting fluctuation function with two asymptotic regimes with $H = 2$ at small scales and $H = 1$ at large scales, separated by the desired crossover (see

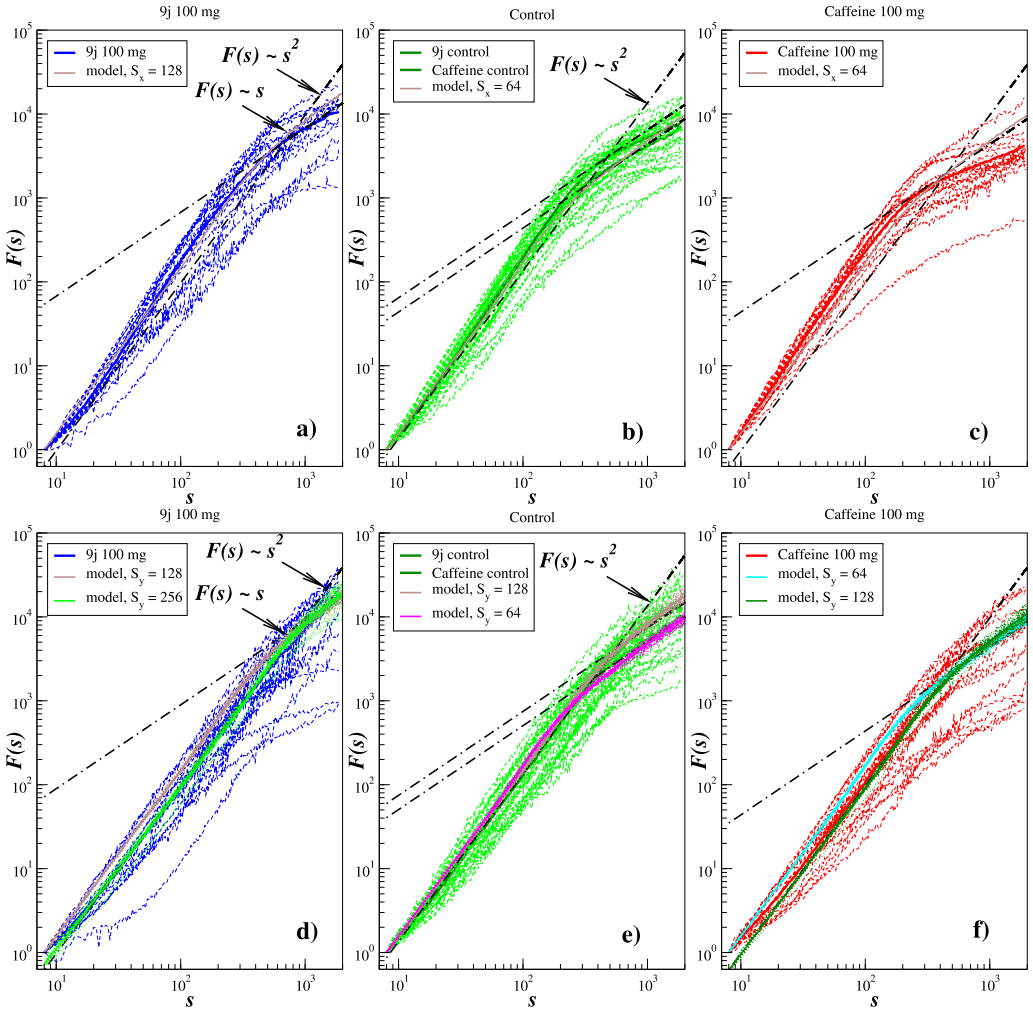

Fig. 5. Fluctuation functions for two test groups (a,c,d,f) with maximal effector doses and two respective control groups (b,e) for the movement projections on the x-axis (a–c) and y-axis (d–f), respectively, obtained by DFA. Thin dashed curves indicate fluctuation functions for each of the 16 individual trajectories in each group, while bold full curves represent averages over the thin curves. Black dashdot lines show approximations $F(s) \propto s$ and $F(s) \propto s^2$ as guides for the eye. For comparison, overlaying colored curves represent similar fluctuation functions obtained for some representative examples of the simulated trajectories characterized by crossover positions S_x and S_y, respectively, to illustrate similarities in their shapes and locations. Due to higher variability of fluctuation functions for the y-axis projections, fluctuation functions for two different model parameters are shown.

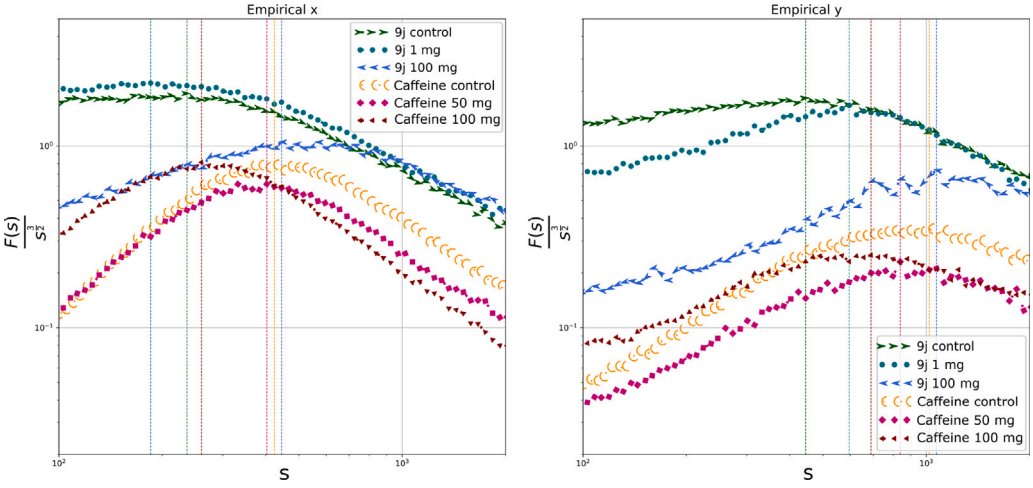

Fig. 6. The procedure of fluctuation function crossover position estimation based on finding the local maxima in $F(s)/s^{3/2}$ exemplified for the experimental groups with 9j and caffeine pharmacological stimuli, respectively.

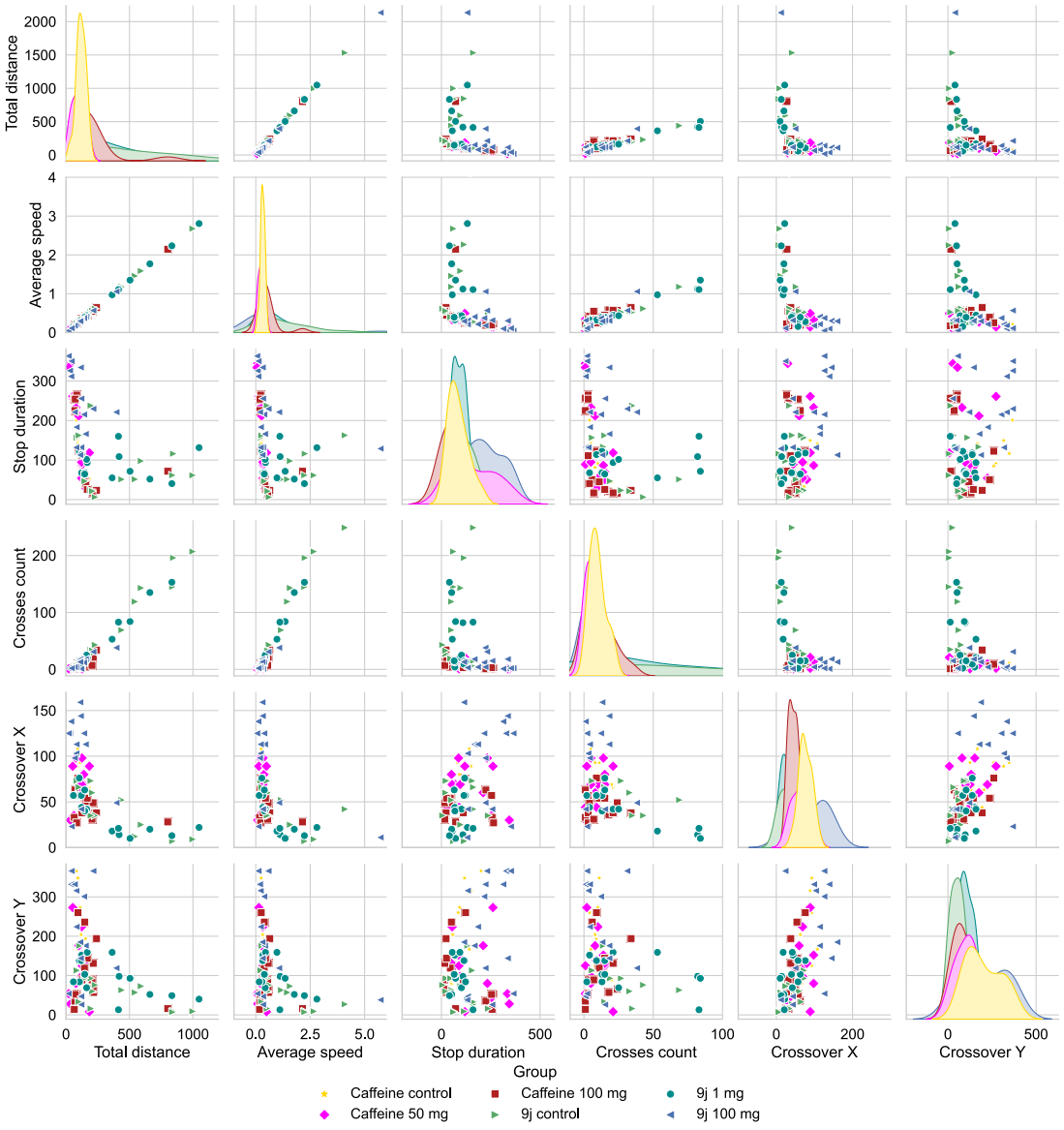

Fig. 7. A pairplot for the first four representative scalar metrics and two model parameter estimates, one for the movement along x-axis and the other along the y-axis, respectively.

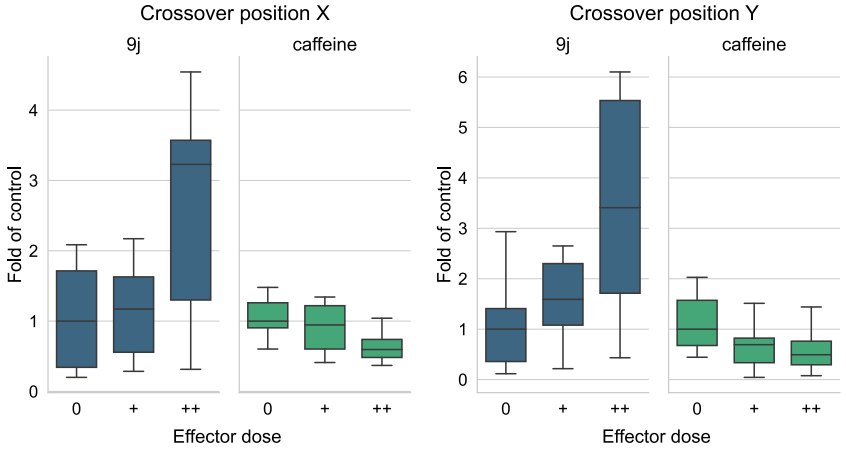

Fig. 8. Boxplots for the crossover positions s_x and s_y.

Fig. 9), we apply the moving average filter. To establish a one-by-one correspondence between the moving average filter size and the crossover position s_X, we have generated multiple surrogate series with variable moving average filter size and obtained the crossover positions

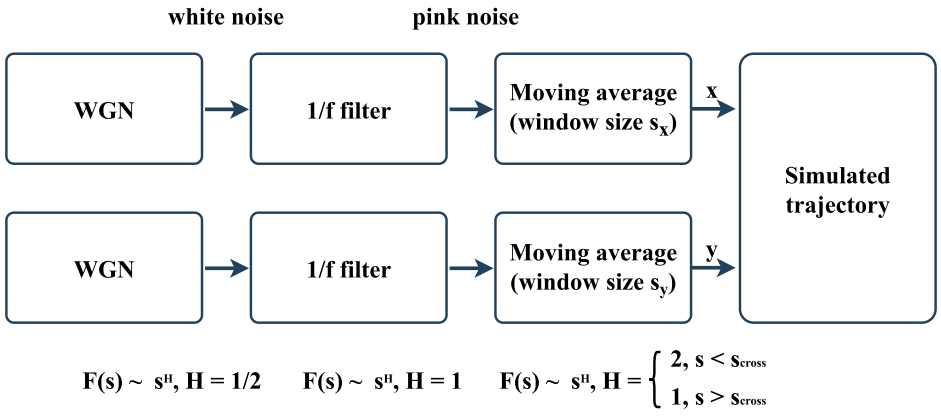

Fig. 9. Model based trajectories simulation algorithm.

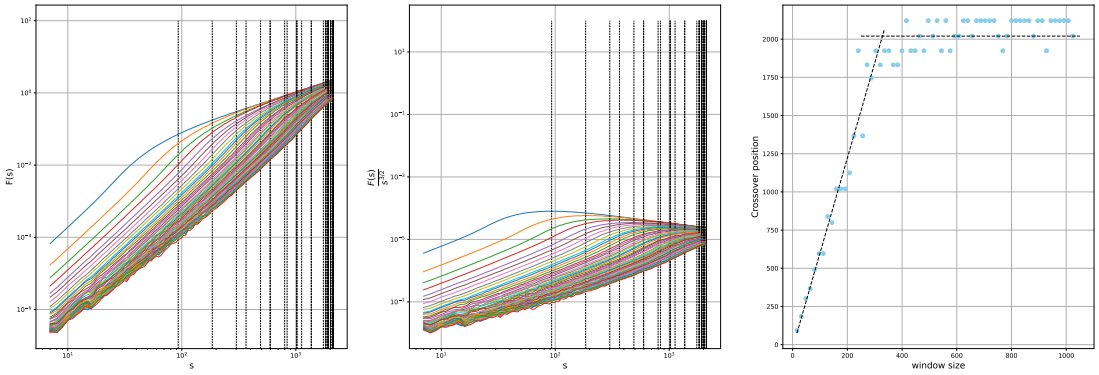

Fig. 10. (a) Fluctuation functions $F(s)$ for surrogate data series obtained by the simulation algorithm depicted in Fig. 9 with variable moving average filter size; (b) The same fluctuation functions but divided over $s^3/2$; (c) Linear regression of the crossover position as determined in (b) on the moving average filter size in the generation model. Convergence to an upper limit in (c) is the consequence of finite size effect, where particular saturation level and crossover position are dependent on the number of data points in the trajectory; for compatibility with observational data, the surrogate data series were of the same sizes as the empirical ones.

following the same procedure as for the empirical trajectories, as illustrated in Fig. 10. To confirm the accuracy of the simulation procedure, we show representative examples for typical model parameters in Fig. 5 leading to fluctuation functions that are closely reminiscent to the respective empirical group averages.

Finally, using the surrogate trajectories generated with characteristic model parameters s_x and s_y for each individual organism, we have estimated the conventional scalar movement metrics following the procedure identical to the one used previously the empirical trajectories. The respective boxplots are shown in Fig. 11.

The figure shows explicitly that the model based estimates of the conventional scalar movement metrics appear more stable, are characterized by smaller variances, and thus demonstrate more specific characteristics responses to the respective pharmacological stimuli, compared to their direct estimations from visual data. The above results clearly demonstrate that the model-based approach appears more robust in the presence of inevitable trajectory disturbances induced by noise and gaps, and thus represents a preferential approach to the analysis of empirical movement patterns. The numerical results are summarized in Table 2.

3.4. Statistical analysis

In the following, we focus on the comparisons of the conventional metrics and their indirect model based estimates. First, we consider the Caffeine test and compare the control group (1), the 50 mg test group (2) and the 100 mg test group (3) using both Kruskal–Wallis and ANOVA tests. The results summarized in Tables 3 and 4 indicate that, while none of the conventional metrics indicated statistically significant differences between the studied groups (except for Kruskal–Wallis for the first ascent latency, although not confirmed by ANOVA), due to high variations among estimated values, the model parameters for both x-axis and y-axis movement projections differed significantly with $p < 0.05$. Moreover, model based estimates for five out of seven conventional movement metrics, in marked contrast to their direct evaluations from the empirical trajectories, also indicated statistically significant differences between groups, according to both statistical tests employed. By contrast, the remaining two metrics, the sum stop time (the total fading duration) and the first ascent latency, indicated further enhancement of p-values in both Kruskal–Wallis and ANOVA tests. Speculatively, this could be interpreted as a more clear indication of the absence of significant discrepancies, contributing to the reduction of uncertainty (see Tables 5 and 6).

Next, we proceed with similar tests for the 9j experimental series. Again, both Kruskal–Wallis and ANOVA tests indicate significance differences between studied groups both in the model parameters s_x and s_y, as well as in five (out of seven) conventional scalar metrics when using their model based estimates, while their direct estimates from movement trajectories, as well as Ethovision based evaluations, resulted in significant discrepancies in just four metrics according to Kruskal–Wallis and two metrics according to ANOVA, respectively. Although for the crosses count metric with ANOVA and for the sum top time (the total duration in the upper tank zone) the significance threshold $p = 0.05$ has not been reached, at least in the latter case considerable improvements of the significance could be observed in both tests, indicated by the reduction of from $p \lessapprox 1$ to $p \gtrapprox 0.05$. The latter could be attributed to the reduction of within-group variances in many cases, as indicated explicitly by the results of the Levene's

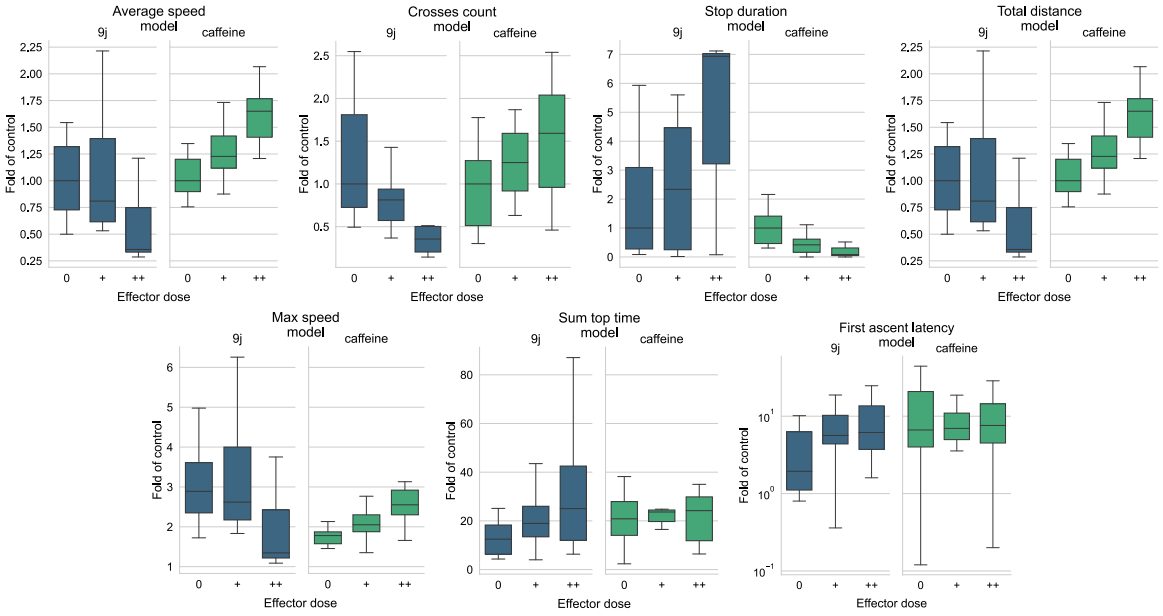

Fig. 11. Boxplots for the same seven conventional metrics as in Fig. 4, although this time obtained indirectly using the model based approach. The surrogate trajectories have been generated according to the procedure illustrated in Fig. 9 with moving average sizes as only free parameters obtained from fitted crossover positions as in Fig. 6 (although this time individually for each animal, separately for s_x and s_y) in accordance with the regression model obtained from Fig. 10. Next the surrogate trajectories have been subjected to the same analysis procedure as the empirical ones.

statistical test comparing the variances of the conventional and model based estimates of the same movement metrics in the same groups after their normalization to the medians of the respective control groups summarized in Table 7.

It is also interesting to note that direct estimates of conventional scalar metrics were not consistent with the Ethovision based evaluations in terms of particular metrics indicating significant discrepancies between studied groups, that raises questions regarding their reproducibility. In marked contrast, the crossover positions s_x and s_y remained significantly different between groups in all pairwise comparisons considered above also in the second iteration, when estimated from fluctuation functions obtained from surrogate data, indicating good reproducibility of the model parameters and further confirming the model self-consistency.

For pairwise comparison scenarios, the classification accuracy could be further improved by combining multiple criteria using the logistic regression model approach. In particular, for the experimental group treated with 100 mg Caffeine vs. respective control group, when considering only conventional metrics estimated directly from the trajectory (i.e. without using any kind of model based estimates yet), the combined model included the first accent latency and the total traveled distance, leading to the Nagelkerke's $R^2 = 0.567$, while sensitivity, specificity and precision estimates reached $TPR = 0.900$, $TNR = 0.917$, and $PPV = 0.900$, respectively. In contrast, when the crossover position s_x has been included in the combined model, in combination with the best complementary scalar metric to s_x that was the maximum speed observed during the test, $R^2 = 0.667$, $TPR = 1.000$, $TNR = 0.769$, and $PPV = 0.800$, respectively.

For the experimental group treated with 100 mg 9j vs. control group the combined model included the total distance traveled by the fish during the experiment and the crossover position s_x. Since crossover positions s_x and s_y are correlated (see Fig. 7, including one of the parameters (e.g., s_x) already provides with accurate classification, while further addition of the remaining parameter (in this case s_y) does not lead to significant improvement of the classification accuracy. In particular, when s_y has been chosen as a complementary parameter to the total traveled distance in the combined model, the Nagelkerke's $R^2 = 0.755$, while sensitivity, specificity and precision estimates reached $TPR = 0.929$, $TNR = 0.875$, and $PPV = 0.866$, respectively. Alternatively, inclusion of s_x in the combined model instead of s_y resulted in $R^2 = 0.867$, $TPR = 0.923$, $TNR = 0.933$, and $PPV = 0.923$, respectively. The above data indicate that, due to pronounced correlations between the model parameters s_x characterizing the movements along x- and y-axes, the respective crossover position parameters could generally substitute each other, reducing the total model complexity to a single free parameter even for the characterization of the two-dimensional movement pattern.

To further evaluate the robustness of the proposed approach to the presence of gaps in the data, we have repeated the above pairwise comparison for manually post-processed trajectories that mainly aimed at the interpolation of large gaps and removal of some artifacts that remained in the trajectories despite of the outlier elimination procedure during the automated analysis. Our results indicate that manual post-processing resulted only in a moderate improvement of the classification accuracy, leading to $R^2 = 0.916$, $TPR = 1.000$, $TNR = 0.909$, and $PPV = 1.000$, respectively.

4. Discussion

Statistical analysis and computer simulations of various movement patterns has a long history in applied mathematics and especially statistical physics. The most basic concept that dates nearly two centuries back is the simple Brownian motion, that also represents the limiting case for the random walk model. Under the assumption of statistically independent and identically distributed increments at each step, the root-mean square displacement of the random walker from the origin after s steps, according to the classic Fick's diffusion law, scales as $F(s) \propto s^{1/2}$. Possible generalizations of this concept for correlated increments include the fBm model with $F(s) \propto s^H$, a widely used approach for modeling data series with long-term persistence, including animal movement patterns in their natural environments. The above laws hold asymptotically, making these models completely scale-free, characterized by theoretically unlimited long-term correlations.

In contrast, experimental animal movement patterns, including the open field test and the novel tank test, premise that the motions are limited to a certain confined space, making asymptotically scale-free

Table 2
The summary of the first four representative conventional movement metrics and model parameters in all studied groups. Each group contained 16 animals. The fraction of gaps in the coordinate data obtained from the developed computer vision algorithms are denoted for each studied group.

Feature	Group	Dose	Empirical			Model	
			gaps, %	Median	IQR	Median	IQR
Total distance	9j	0	6.78	229.85	543.44	577.76	341.94
		1 mg	1.82	164.64	317.54	467.66	449.93
		100 mg	6.33	111.10	82.81	206.22	240.64
	Caffeine	0	4.67	115.31	55.80	272.48	82.56
		50 mg	14.62	122.76	69.79	334.48	81.81
		100 mg	4.36	151.86	122.09	449.63	98.18
Average speed	9j	0	6.78	0.62	1.46	1.55	0.92
		1 mg	1.82	0.44	0.85	1.25	1.21
		100 mg	6.33	0.3	0.22	0.55	0.64
	Caffeine	0	4.67	0.31	0.15	0.73	0.22
		50 mg	14.62	0.33	0.17	0.9	0.22
		100 mg	4.36	0.41	0.32	8.4	0.26
Stop duration	9j	0	6.78	61.6	96.1	52.2	146.5
		1 mg	1.82	93.6	49.6	121.5	219.25
		100 mg	6.33	221.2	157.65	360.6	198.1
	Caffeine	0	4.67	81.6	60.8	98.5	93.3
		50 mg	14.62	106.6	174.55	41.3	44.65
		100 mg	4.36	47.6	125.25	8.4	26.1
Crosses count	9j	0	6.78	43	121	166	180
		1 mg	1.82	21	69.5	135	61
		100 mg	6.33	4	14.5	59	49.75
	Caffeine	0	4.67	9	7	76	57.75
		50 mg	14.62	6.5	11	95	51.25
		100 mg	4.36	7	13.5	121	82
Crossover X	9j	0	6.78	35	48	9.5	2.75
		1 mg	1.82	41	37.5	8	0.75
		100 mg	6.33	113	79.5	18.5	12.5
	Caffeine	0	4.67	73	26	9	6.75
		50 mg	14.62	69	45	13	7
		100 mg	4.36	43.5	18.75	7	2
Crossover Y	9j	0	6.78	60	63	37	3.25
		1 mg	1.82	95.5	37.5	16	10
		100 mg	6.33	204.5	229.25	33	49
	Caffeine	0	4.67	180.5	161.75	37	23.5
		50 mg	14.62	125	88	18	18
		100 mg	4.36	89	84.5	13	9

Table 3
p-values for the Kruskal–Wallis test results for the Caffeine test, control group (1) vs. 50 mg group (2) vs. 100 mg group (3).

	Direct estimation	Model-based estimation
Total distance	0.099	**<0.001**
Average speed	0.099	**<0.001**
Stop duration	0.093	**<0.001**
Crosses count	0.411	**0.044**
Max speed	0.733	**<0.001**
Sum top time	0.500	0.930
First ascent latency	**0.037**	0.710
Crossover X	**<0.001**	**0.024**
Crossover Y	**0.025**	**0.015**

Table 4
p-values for the ANOVA test results for the Caffeine test, control group (1) vs. 50 mg group (2) vs. 100 mg group (3).

	Direct estimation	Model-based estimation
Total distance	0.080	**0.002**
Average speed	0.080	**0.002**
Stop duration	0.149	**0.001**
Crosses count	0.340	**0.047**
Max speed	0.802	**0.001**
Sum top time	0.940	0.882
First ascent latency	0.100	0.288
Crossover X	**0.032**	**0.020**
Crossover Y	**0.017**	**0.002**

Table 5
p-values for the Kruskal–Wallis test results for 9j control (group 4), 9j 1 mg (group 5) and 9j 100 mg (group 6).

	Ethovision results	Direct estimation	Model estimation
Total distance	**<0.001**	**0.006**	**0.006**
Average speed	**<0.001**	**0.006**	**0.006**
Stop duration	**0.010**	**<0.001**	**0.006**
Crosses count	0.793	**0.002**	**0.006**
Max speed	**<0.001**	0.980	**0.005**
Sum top time	0.923	0.541	0.058
First ascent latency	0.163	0.899	0.253
Crossover X	*N/A*	**0.006**	**0.001**
Crossover Y	*N/A*	**0.001**	**0.003**

Table 6
p-values for the ANOVA test results for 9j control (group 4), 9j 1 mg (group 5) and 9j 100 mg (group 6).

	Ethovision results	Direct estimation	Model estimation
Total distance	**<0.001**	0.316	**0.014**
Average speed	**<0.001**	0.316	**0.014**
Stop duration	**0.004**	**<0.001**	**0.002**
Crosses count	0.563	**0.050**	0.055
Max speed	0.231	0.996	**0.017**
Sum top time	0.871	0.749	**0.019**
First ascent latency	0.291	0.341	0.283
Crossover X	*N/A*	**0.004**	**<0.001**
Crossover Y	*N/A*	**<0.001**	**<0.001**

Table 7
p-values for the Levene's statistical test comparing the variances of the conventional and model based estimates of the same movement metrics in the same groups after their normalization to the medians of the respective control groups.

	Caffeine			9j		
	Control	50 mg	100 mg	Control	1 mg	100 mg
Total distance	0.316	**0.017**	**0.004**	**0.013**	0.203	0.231
Average speed	0.317	**0.017**	**0.004**	**0.013**	0.204	0.231
Stop duration	0.166	**<0.001**	**<0.001**	0.451	**<0.001**	**0.039**
Crosses count	0.344	0.3	0.536	**0.007**	**0.001**	**0.040**
Max speed	**0.001**	**0.009**	**0.001**	**0.001**	**0.003**	0.244
Sum top time	**0.01**	**0.007**	0.637	0.421	0.463	**0.024**
First ascent latency	0.482	0.967	0.137	**0.001**	**0.008**	**<0.001**

models hardly adequate, due to a breakdown of persistence above a certain scale. Moreover, the alternating movement patterns that are inevitable in confined space due to imminent reversals also assume the emergence of antipersistence at large scales. Accordingly, the resulting animal movement model is no longer expected to exhibit scale-free properties, but rather consist of at least two characteristic regimes, corresponding to a persistent random walk at small scales, that is substituted by an antipersistent pattern at large scales.

Our results indicate that these two asymptotic regimes $F(s) \propto s^2$ (which corresponds to persistent increments) at small scales and $F(s) \propto s$ (which corresponds to antipersistent increments) at large scales are separated by a single crossover at $s_\times$, which is a single characteristic scale parameter, that also appears a single free parameter in the proposed animal movement model. As we have shown above, the particular crossover position $s_\times$ explicitly reflects the characteristics of the animal movement pattern, while being sufficient to simulate the movement trajectories that reproduce the conventional animal movement metrics widely used to evaluate animal model test results.

Of note, since we have applied the DFA method with an inherent integration procedure to the trajectories, but not to the increments, the analysis procedure altogether is equivalent to the application of the DFA with double integration to the increments, that has been previously applied to characterize antipersistent data series, including physiological applications [47]. Instead, one could either apply DFA

to the increments, or apply DFA to the trajectories, but skip the initial integration step, this way estimating Hurst exponents of the increments, $H \approx 1$ at small and $H \approx 0$ at large scales, respectively. Although the above approach is fully legitimate and even more theoretically straightforward, to our opinion, in has two inherent drawbacks. First of all, since Hurst exponent is bounded by zero from below, asymptotic exponents close to zero cannot be distinguished from simple convergence to this bound, due to methodological limitations. Second, since coordinates are measured directly from video analysis, while increments are estimated, commonly by taking first differences, the relative noise levels in the increments are generally higher, leading to reduced robustness of the analysis algorithms.

Model-based estimates of movement parameters lead to a better classification effectiveness mainly because they are characterized by smaller variances. However, there is inevitable natural variance σ_N^2 in each experimental group, and thus it is important that this natural variability is retained despite of the chosen measurement and analysis procedure. To substantiate the above statement, let us assume that, at least in the first approximation, the natural variance σ_N^2 and the measurement variance σ_M^2 are statistically independent. Then, by definition, the total variance for each measurement procedure is given by $\sigma_T^2 = \sigma_M^2 + \sigma_N^2$, where σ_N^2 remains independent of the measurement procedure, and thus the minimum total variance σ_T^2 corresponds to the methodology leading to the minimal measurement variance σ_M^2 [48]. In our case, the model-based estimates of the same movement parameters are characterized by considerably smaller total σ_T^2 and measurement σ_M^2 variances, thus leading to a more specific characterization of the response to pharmacological stimuli.

As an outlook, we believe that similar approach could be applied to the analysis of other animal tests, such as the open field rodent test, as well as the analysis of animal movement patterns in their natural habitats, for example in ecological applications. In addition, further interpretation of the relationship between the movement model parameters and some of the conventional scalar movement metrics from the statistical physics perspective could be based on the statistics of level-crossing events generated by long-term correlated data series. Several metrics can be characterized using event based statistics, such as crossing events between the bottom and the top zones of the tank, fading episodes as stop and start events, and so on. From a theoretical perspective, emergence of long-range correlations in the data series lead to a considerable alteration of the distributions of return intervals between events determined by level crossings by the raw data series. While in the absence of correlations return interval distribution exhibits a simple exponential decay, in stationary data series with linear long-term correlations the respective distribution extends to the Weibull functional form that asymptotically follows a stretched exponential with the stretching parameter equals the correlation exponent γ, and thus also being in one-by-one correspondence with the Hurst exponent H, while in the presence of pronounced nonlinear correlations it converges to a power law decaying function [49–51]. For non-stationary data series, further generalization of the respective analytical treatments could be based on the superstatistical approach that is derived based on the law of total probability [52,53], like it has been done recently for various complex systems including various environmental indicators, as well as DNA structures, teletraffic intensity fluctuations and many others (for a few representative examples, we refer to [54–59] and references therein). We believe that similar approach could be also applied to generalized in higher-dimensional movement models, by using similar methods generalized to higher dimensions such as two-dimensional DFA [60] or even three-dimensional equivalents, in the relevant experimental setup, as well as to the analysis of animal movement patterns aiming at their more detailed analytical description, although this appears beyond the scope of our current study.

5. Conclusion

To summarize, in this work we analyzed movement patterns of zebrafish *D. rerio* in a novel tank test with an allylmorpholine derivative 9j exhibiting a dose-dependent sedative effect and, in contrast, caffeine exhibiting stimulative effect using detrended fluctuation analysis (DFA). Based on the results of this analysis, we have proposed a novel animal movement model based on the generalization of the fractional Brownian motion. The proposed model is capable of reproducing complex animal movement patterns this way overcoming typical limitations of the conventional analysis commonly reduced to a few scalar metrics that are often limited in their accuracy and reproducibility.

Our results indicate that characteristic animal movement patterns exhibit two asymptotic regimes, with persistent increments at small scales and antipersistent increments at large scales separated by a single crossover. The position of the crossover represents a single free parameter in the proposed animal movement model that explicitly reflects stimulative and sedative effects on the animal movement patterns leading to a better distinction between experimental groups. We also show that the model parameters, as well as the model based estimates of interpretable movement metrics commonly used in behavioral analysis appear more robust against measurement artifacts leading to a more explicit characterization of the movement patterns alteration in various experimental groups.

The proposed approach could further facilitate automated and objective characterization of animal model test results involving movement pattern analysis based on computer vision technologies. We also believe that the proposed approach, due to its universality, robustness and clear physical interpretation, could be a perspective tool for the analysis of behavioral complexity in various applications ranging from experimental drug screening to animal monitoring in their natural habitats.

CRediT authorship contribution statement

Mikhail I. Bogachev: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Funding acquisition. **Asya I. Lyanova:** Software, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Aleksandr M. Sinitca:** Software, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Svetlana A. Pyko:** Methodology, Validation. **Nikita S. Pyko:** Investigation. **Alexander V. Kuzmenko:** Software. **Sergey A. Romanov:** Software. **Olga I. Brikova:** Investigation. **Margarita Tsygankova:** Investigation. **Dmitry Y. Ivkin:** Resources. **Sergey V. Okovityi:** Resources. **Veronika A. Prikhodko:** Investigation. **Dmitrii I. Kaplun:** Project administration, Supervision, Writing – original draft. **Yuri I. Sysoev:** Conceptualization, Investigation, Methodology, Writing – original draft. **Airat R. Kayumov:** Methodology, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data are provided as Supplementary material.

Acknowledgments

This research was supported by the Ministry of Science and Higher Education under assignment FSEE-2020-0002. The experimental results were obtained using the equipment of the Center for Collective Use "Analytical Center of Saint Petersburg State Chemical and Pharmaceutical University" under the agreement No. 075-15-2021-685 dated July 26, 2021 with the financial support by the Ministry of Science and Higher Education. Y.I.S. would also like to acknowledge St. Petersburg State University for financial support (project ID: 93022798).

Appendix A. Supplementary data: program source codes and tracked animal movement data.

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.bspc.2022.104409.

References

[1] E.A. Codling, M.J. Plank, S. Benhamou, Random walk models in biology, J. R. Soc. Interface 5 (25) (2008) 813–834.

[2] M.B. Hooten, D.S. Johnson, B.T. McClintock, J.M. Morales, Animal Movement: Statistical Models for Telemetry Data, CRC Press, 2017.

[3] A.M. Reynolds, Scale-free animal movement patterns: Lévy walks outperform fractional brownian motions and fractional Lévy motions in random search scenarios, J. Phys. A 42 (43) (2009) 434006.

[4] D. Bearup, C.M. Benefer, S.V. Petrovskii, R.P. Blackshaw, Revisiting brownian motion as a description of animal movement: a comparison to experimental movement data, Methods Ecol. Evol. 7 (12) (2016) 1525–1537.

[5] M.B. Hooten, D.S. Johnson, Basis function models for animal movement, J. Amer. Statist. Assoc. 112 (518) (2017) 578–589.

[6] C.J. Torney, J.M. Morales, D. Husmeier, A hierarchical machine learning framework for the analysis of large scale animal movement data, Mov. Ecol. 9 (1) (2021) 1–11.

[7] Y. Agid, G. Buzsáki, D.M. Diamond, R. Frackowiak, J. Giedd, J.-A. Girault, A. Grace, J.J. Lambert, H. Manji, H. Mayberg, et al., How can drug discovery for psychiatric disorders be improved? Nat. Rev. Drug Discov. 6 (3) (2007) 189–201.

[8] S.C. Stanford, The open field test: reinventing the wheel, J. Psychopharmacol. 21 (2) (2007) 134–136.

[9] T.D. Gould, D.T. Dao, C.E. Kovacsics, The open field test, in: Mood and Anxiety Related Phenotypes in Mice, 2009, pp. 1–20.

[10] D. Perals, A.S. Griffin, I. Bartomeus, D. Sol, Revisiting the open-field test: what does it really tell us about animal personality? Anim. Behav. 123 (2017) 69–79.

[11] O. Sturman, P.-L. Germain, J. Bohacek, Exploratory rearing: a context-and stress-sensitive behavior recorded in the open-field test, Stress 21 (5) (2018) 443–452.

[12] A.-K. Kraeuter, P.C. Guest, Z. Sarnyai, The open field test for measuring locomotor activity and anxiety-like behavior, in: Pre-Clinical Models, Springer, 2019, pp. 99–103.

[13] A.M. Stewart, O. Braubach, J. Spitsbergen, R. Gerlai, A.V. Kalueff, Zebrafish models for translational neuroscience research: from tank to bedside, Trends Neurosci. 37 (5) (2014) 264–278.

[14] C.A. MacRae, R.T. Peterson, Zebrafish as tools for drug discovery, Nat. Rev. Drug Discov. 14 (10) (2015) 721–731.

[15] A.V. Kalueff, M. Gebhardt, A.M. Stewart, J.M. Cachat, M. Brimmer, J.S. Chawla, C. Craddock, E.J. Kyzar, A. Roth, S. Landsman, et al., Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond, Zebrafish 10 (1) (2013) 70–86.

[16] R.J. Egan, C.L. Bergner, P.C. Hart, J.M. Cachat, P.R. Canavello, M.F. Elegante, S.I. Elkhayat, B.K. Bartels, A.K. Tien, D.H. Tien, et al., Understanding behavioral and physiological phenotypes of stress and anxiety in zebrafish, Behav. Brain Res. 205 (1) (2009) 38–44.

[17] C. Maximino, T.M. De Brito, C.A.G. de Mattos Dias, A. Gouveia, S. Morato, Scototaxis as anxiety-like behavior in fish, Nat. Protoc. 5 (2) (2010) 209–216.

[18] A. Stewart, N. Wu, J. Cachat, P. Hart, S. Gaikwad, K. Wong, E. Utterback, T. Gilder, E. Kyzar, A. Newman, et al., Pharmacological modulation of anxiety-like phenotypes in adult zebrafish behavioral models, Prog. Neuro-Psychopharmacol. Biol. Psychiatry 35 (6) (2011) 1421–1431.

[19] R.E. Blaser, D.B. Rosemberg, Measures of anxiety in zebrafish (danio rerio): dissociation of black/white preference and novel tank test, PLoS One 7 (5) (2012) e36931.

[20] A.M. Stewart, J.F. Ullmann, W.H. Norton, M. Parker, C. Brennan, R. Gerlai, A.V. Kalueff, Molecular psychiatry of zebrafish, Mol. Psychiatry 20 (1) (2015) 2–17.

[21] B.D. Fontana, N.J. Mezzomo, A.V. Kalueff, D.B. Rosemberg, The developing utility of zebrafish models of neurological and neuropsychiatric disorders: A critical review, Exp. Neurol. 299 (2018) 157–171.

[22] S. Haghani, M. Karia, R.-K. Cheng, A.S. Mathuru, An automated assay system to study novel tank induced anxiety, Front. Behav. Neurosci. (2019) 180.

[23] D.G. Valcarce, J.M. Martínez-Vázquez, M.F. Riesco, V. Robles, Probiotics reduce anxiety-related behavior in zebrafish, Heliyon 6 (5) (2020) e03973.

[24] A. Golla, H. Østby, F. Kermen, Chronic unpredictable stress induces anxiety-like behaviors in young zebrafish, Sci. Rep. 10 (1) (2020) 1–10.

[25] S. Kato, T. Nakagawa, M. Ohkawa, K. Muramoto, O. Oyama, A. Watanabe, H. Nakashima, T. Nemoto, K. Sugitani, A computer image processing system for quantification of zebrafish behavior, J. Neurosci. Methods 134 (1) (2004) 1–7.

[26] A.S. Kane, J.D. Salierno, G.T. Gipson, T.C. Molteno, C. Hunter, A video-based movement analysis system to quantify behavioral stress responses of fish, Water Res. 38 (18) (2004) 3993–4001.

[27] J. Cachat, A. Stewart, E. Utterback, P. Hart, S. Gaikwad, K. Wong, E. Kyzar, N. Wu, A.V. Kalueff, Three-dimensional neurophenotyping of adult zebrafish behavior, PLoS One 6 (3) (2011) e17597.

[28] L. Grossman, E. Utterback, A. Stewart, S. Gaikwad, K.M. Chung, C. Suciu, K. Wong, M. Elegante, S. Elkhayat, J. Tan, et al., Characterization of behavioral and endocrine effects of lsd on zebrafish, Behav. Brain Res. 214 (2) (2010) 277–284.

[29] A. Stewart, R. Riehl, K. Wong, J. Green, J. Cosgrove, K. Vollmer, E. Kyzar, P. Hart, A. Allain, J. Cachat, et al., Behavioral effects of mdma (ecstasy) on adult zebrafish, Behav. Pharmacol. 22 (3) (2011) 275.

[30] T.O. Kolesnikova, S.L. Khatsko, O.S. Eltsov, V.A. Shevyrin, A.V. Kalueff, When fish take a bath: Psychopharmacological characterization of the effects of a synthetic cathinone bath salt 'flakka'on adult zebrafish, Neurotoxicol. Teratol. 73 (2019) 15–21.

[31] M.E. Wolter, K.R. Svoboda, Doing the locomotion: Insights and potential pitfalls associated with using locomotor activity as a readout of the circadian rhythm in larval zebrafish, J. Neurosci. Methods 330 (2020) 108465.

[32] G. Liebsch, A. Montkowski, F. Holsboer, R. Landgraf, Behavioural profiles of two wistar rat lines selectively bred for high or low anxiety-related behaviour, Behav. Brain Res. 94 (2) (1998) 301–310.

[33] N. Sestakova, A. Puzserova, M. Kluknavsky, I. Bernatova, Determination of motor activity and anxiety-related behaviour in rodents: methodological aspects and role of nitric oxide, Interdiscip. Toxicol. 6 (3) (2013) 126–135.

[34] P. Mathur, S. Guo, Differences of acute versus chronic ethanol exposure on anxiety-like behavioral responses in zebrafish, Behav. Brain Res. 219 (2) (2011) 234–239.

[35] A. Mustafa, E. Roman, S. Winberg, Boldness in male and female zebrafish (danio rerio) is dependent on strain and test, Front. Behav. Neurosci. (2019) 248.

[36] Y.I. Sysoev, D.A. Meshalkina, D.V. Petrov, S.V. Okovityi, P.E. Musienko, A.V. Kalueff, Pharmacological screening of a new alpha-2 adrenergic receptor agonist, mafedine, in zebrafish, Neurosci. Lett. 701 (2019) 234–239.

[37] R. Jeanson, S. Blanco, R. Fournier, J.-L. Deneubourg, V. Fourcassié, G. Theraulaz, A model of animal movements in a bounded space, J. Theoret. Biol. 225 (4) (2003) 443–451.

[38] T.A. Patterson, L. Thomas, C. Wilcox, O. Ovaskainen, J. Matthiopoulos, State–space models of individual animal movement, Trends Ecol. Evol. 23 (2) (2008) 87–94.

[39] P.E. Smouse, S. Focardi, P.R. Moorcroft, J.G. Kie, J.D. Forester, J.M. Morales, Stochastic modelling of animal movement, Philos. Trans. R. Soc. B 365 (1550) (2010) 2201–2211.

[40] R. Langrock, J.G.C. Hopcraft, P.G. Blackwell, V. Goodall, R. King, M. Niu, T.A. Patterson, M.W. Pedersen, A. Skarin, R.S. Schick, Modelling group dynamic animal movement, Methods Ecol. Evol. 5 (2) (2014) 190–199.

[41] N.M. Chernov, R.V. Shutov, O.I. Barygin, M.Y. Dron, G.L. Starova, N.N. Kuz'mich, I.P. Yakovlev, Synthesis of chromone-containing allylmorpholines through a Morita–Baylis–Hillman-type reaction, Eur. J. Organ. Chem. 2018 (45) (2018) 6304–6313.

[42] V. Prikhodko, Y. Sysoev, Sedative activity of a new allylmorpholine derivative in the novel tank diving test in zebrafish danio rerio, J. Biomed. 17 (3E) (2021) 161–164.

[43] A.O. Alia, M.L. Petrunich-Rutherford, Anxiety-like behavior and whole-body cortisol responses to components of energy drinks in zebrafish (danio rerio), PeerJ 7 (2019) e7546.

[44] C.-K. Peng, S. Buldyrev, S. Havlin, M. Simons, H. Stanley, A. Goldberger, Mosaic organization of dna nucleotides, Phys. Rev. E 49 (2) (1994) 1685.

[45] J. Kantelhardt, E. Koscielny-Bunde, H. Rego, S. Havlin, A. Bunde, Detecting long-range correlations with detrended fluctuation analysis, Phys. A 295 (3) (2001) 441–454.

[46] M.I. Bogachev, A.R. Kayumov, O.A. Markelov, A. Bunde, Statistical prediction of protein structural, localization and functional properties by the analysis of its fragment mass distributions after proteolytic cleavage, Sci. Rep. 6 (1) (2016) 1–19.

[47] J. Kantelhardt, Y. Ashkenazy, P.C. Ivanov, A. Bunde, S. Havlin, T. Penzel, J.-H. Peter, H.E. Stanley, Characterization of sleep stages by correlations in the magnitude and sign of heartbeat increments, Phys. Rev. E 65 (5) (2002) 051908.

[48] A. Sokolova, Y. Uljanitski, A.R. Kayumov, M.I. Bogachev, Improved online event detection and differentiation by a simple gradient-based nonlinear transformation: Implications for the biomedical signal and image analysis, Biomed. Signal Process. Control 66 (2021) 102470.

[49] M.I. Bogachev, J.F. Eichner, A. Bunde, Effect of nonlinear correlations on the statistics of return intervals in multifractal data sets, Phys. Rev. Lett. 99 (24) (2007) 240601.

[50] M.I. Bogachev, J.F. Eichner, A. Bunde, On the occurence of extreme events in long-term correlated and multifractal data sets, Pure Appl. Geophys. 165 (6) (2008) 1195–1207.

[51] M.I. Bogachev, I.S. Kireenkov, E.M. Nifontov, A. Bunde, Statistics of return intervals between long heartbeat intervals and their usability for online prediction of disorders, New J. Phys. 11 (6) (2009) 063036.

[52] C. Beck, E.G. Cohen, Superstatistics, Phys. A 322 (2003) 267–275.

[53] C. Beck, E.G. Cohen, H.L. Swinney, From time series to superstatistics, Phys. Rev. E 72 (5) (2005) 056133.

[54] A. Bunde, M.I. Bogachev, S. Lennartz, Precipitation and river flow: Long-term memory and predictability of extreme events, Extreme Events Nat. Hazards: Complex. Pers. 196 (2012) 139–152.

[55] A. Tamazian, V. Nguyen, O. Markelov, M. Bogachev, Universal model for collective access patterns in the internet traffic dynamics: A superstatistical approach, Europhys. Lett. 115 (1) (2016) 10008.
[56] O. Markelov, V.N. Duc, M. Bogachev, Statistical modeling of the internet traffic dynamics: To which extent do we need long-term correlations? Phys. A 485 (2017) 48–60.
[57] M.I. Bogachev, O.A. Markelov, A.R. Kayumov, A. Bunde, Superstatistical model of bacterial dna architecture, Sci. Rep. 7 (1) (2017) 1–12.
[58] B. Schäfer, C.M. Heppell, H. Rhys, C. Beck, Fluctuations of water quality time series in rivers follow superstatistics, Iscience 24 (8) (2021) 102881.
[59] Y. Itto, C. Beck, Superstatistical modelling of protein diffusion dynamics in bacteria, J. R. Soc. Interface 18 (176) (2020) 0927.
[60] G.-F. Gu, W.-X. Zhou, Detrended fluctuation analysis for fractals and multifractals in higher dimensions, Phys. Rev. E 74 (6) (2006) 061104.