

# SCIENTIFIC REPORTS



OPEN

## *In-silico* experiments of zebrafish behaviour: modeling swimming in three dimensions

Violet Mwaffo<sup>1</sup>, Sachit Butail<sup>2</sup> & Maurizio Porfiri<sup>1</sup>

Received: 02 August 2016  
Accepted: 28 November 2016  
Published: 10 January 2017

Zebrafish is fast becoming a species of choice in biomedical research for the investigation of functional and dysfunctional processes coupled with their genetic and pharmacological modulation. As with mammals, experimentation with zebrafish constitutes a complicated ethical issue that calls for the exploration of alternative testing methods to reduce the number of subjects, refine experimental designs, and replace live animals. Inspired by the demonstrated advantages of computational studies in other life science domains, we establish an authentic data-driven modelling framework to simulate zebrafish swimming in three dimensions. The model encapsulates burst-and-coast swimming style, speed modulation, and wall interaction, laying the foundations for *in-silico* experiments of zebrafish behaviour. Through computational studies, we demonstrate the ability of the model to replicate common ethological observables such as speed and spatial preference, and anticipate experimental observations on the correlation between tank dimensions on zebrafish behaviour. Reaching to other experimental paradigms, our framework is expected to contribute to a reduction in animal use and suffering.

Experimental studies with animal subjects face ethical issues that demand the exploration of alternative approaches to reduce the number of subjects, refine experimental designs, and replace the use of live animals (3Rs)<sup>1</sup>. While animal studies are the ultimate cornerstone against which new hypotheses should be tested, computer simulations, or *in-silico* experiments, constitute a powerful tool in preclinical research. *In-silico* experiments could be used to: (i) gather preliminary evidence in favour or against a specific hypothesis through systematic pilot trials on the computer<sup>2</sup>; (ii) achieve statistical power in an animal study via careful selection of the number of trials, independent and dependent variables, and frequency of measurement<sup>3</sup>; and (iii) reinforce the scientific merit of existing observations through comparison against synthetic data and statistical analysis informed by modelling insight<sup>4–7</sup>.

In recent years, we have witnessed a surge in the number of *in-silico* experiments, including molecular biology<sup>8</sup>, genetic assays<sup>9</sup>, tumour growth<sup>10</sup>, dermatology<sup>11</sup>, bone remodelling<sup>12</sup>, organ failure<sup>13</sup>, and toxicology<sup>14</sup>. These studies have helped build thrust in the potential of *in-silico* experiments and develop a strong link between simulations and experiments, which are more often coupled in new preclinical studies. However, the use of computer simulations in behavioural neuroscience has been limited, partly due to the complexity of animal behaviour and partly because of the unavailability of refined three-dimensional (3D) data needed to calibrate and validate mathematical models underlying such simulations<sup>15–20</sup>.

Although animal research has mainly focused on rodents<sup>21</sup> and primates<sup>22</sup>, zebrafish (*Danio rerio*) are attaining popularity for both technical and economic reasons. Zebrafish are genetically similar to humans, especially with respect to diseases<sup>23,24</sup>, where a strong genetic similarity is found<sup>25</sup>. Zebrafish embryos and larvae are transparent, which afford the possibility of monitoring inflammatory processes and indicators of disease without surgery<sup>24</sup>. Further, unlike mammalian animal models, zebrafish can be bred in large numbers<sup>26</sup>, have a short inter-generation time<sup>27</sup>, and can be stocked at high densities<sup>24</sup>. These factors are fuelling intensive research in zebrafish behavioural phenotyping, whereby locomotory patterns are continuously being identified and categorized to create a comprehensive catalogue to aid in measuring zebrafish behaviour<sup>28</sup>, for use in neurobehavioural research.

<sup>1</sup>Department of Mechanical and Aerospace Engineering, New York University Tandon School of Engineering, Brooklyn, NY, USA. <sup>2</sup>Department of Mechanical Engineering, Northern Illinois University, DeKalb, IL, USA. Correspondence and requests for materials should be addressed to M.P. (email: mporfiri@nyu.edu)

A large number of experimental protocols have been proposed to investigate zebrafish behaviour, often borrowing from murine models. For example, classical choice<sup>29</sup>, place preference<sup>30</sup>, open field<sup>31</sup>, and light/dark preference tests<sup>32</sup> have been successfully used to investigate the neurobiological basis of social behaviour, study fear and anxiety, and explore the effect of psychoactive compounds and genetic manipulations<sup>33–36</sup>. In these experiments, zebrafish behaviour is measured in terms of locomotory patterns that are typically scored from labour intensive manual observations and seldom using automated computer tracking. Only recently, technical breakthroughs in automated tracking have enabled the complete reconstruction of zebrafish swimming in 3D<sup>17,37–40</sup>. *In-silico* experiments may significantly contribute to the design of new test batteries, by enabling the systematic analysis of dependent and independent variables while isolating key experimental factors that may act as confounds during testing. Further, *in-silico* experiments may be used to identify false positives and negatives, by providing an empirically-based ground truth on which to contextualize new experimental observations<sup>41,42</sup>.

The first computational model of zebrafish swimming has been proposed by our group in refs 43, 44, 45 and 46. In this series of studies, we have proposed a data-driven modelling framework to predict the turn rate dynamics and speed modulation of zebrafish swimming in a shallow water tank. To capture the burst-and-coast swimming style of zebrafish, where sporadic bursts of the tail are followed by coasting phases without tail beating<sup>47,48</sup>, we have proposed the use of jump processes, often employed in financial engineering to model extreme events in the time evolution of assets prices<sup>49,50</sup>. Specifically, in refs 43, 44, the turn rate dynamics of zebrafish is described through a mean reverting stochastic process that incorporates jumps. The jump persistent turning walker (JPTW) accounts for the sudden and fast turns observed in the yaw rate of zebrafish swimming. In ref. 46, we have contemplated the possibility of speed modulation during in-plane manoeuvring, where fish slow down their motion when turning. This modelling framework has been recently extended to study social behaviour in ref. 45 by demonstrating the role of speed modulation in regulating collective dynamics of a shoal of zebrafish. Using a similar approach based on stochastic processes, a recent study<sup>51</sup> has simulated zebrafish orientation within a kinematic model, by sampling from a probability density function that accounts for a spherical visual perception field. These models, built and calibrated on 2D datasets, are expected to suffer from the same limitations that are attributed to 2D behavioural analysis, including unrealistic representation of swimming activity and novelty seeking behaviour<sup>17</sup>; incomplete characterization of stress-related behaviours, such as thigmotaxis and geotaxis<sup>28</sup>; and limited comparability of different experimental assays due to unmeasured variations in depth<sup>17</sup>.

Here, we present a new 3D model for zebrafish swimming, extending our previous work to three dimensions and encapsulating wall interaction and speed modulation. We model both yaw and pitch rates as JPTW, and we treat speed as a mean reverting square process<sup>52</sup>. Wall avoidance is modelled as a decaying function of time to collision<sup>53,54</sup>, and is further adapted to include collision avoidance at high speeds. Model parameters are calibrated with 3D trajectory data of adult zebrafish swimming in a parallelepipedal tank with a rectangular base for 10 minutes<sup>38</sup>. Model validation is performed by conducting a simulation study and comparing speed and indirect behavioural measures, such as positional preference along the horizontal and vertical directions, with experimental observations. To further evaluate if our model can be used to contribute to replacement and reduction, we conduct a second simulation study, investigating the effect of tank dimensions on zebrafish behaviour. We systematically vary the tank dimensions for 60 combinations of length, width, and height, and compare our results with data collated from the literature. We demonstrate the use of *in-silico* experiments in a prospective study, which would require thousands of subjects and months of experiments to explore the parameter space utilized in our numerical simulations.

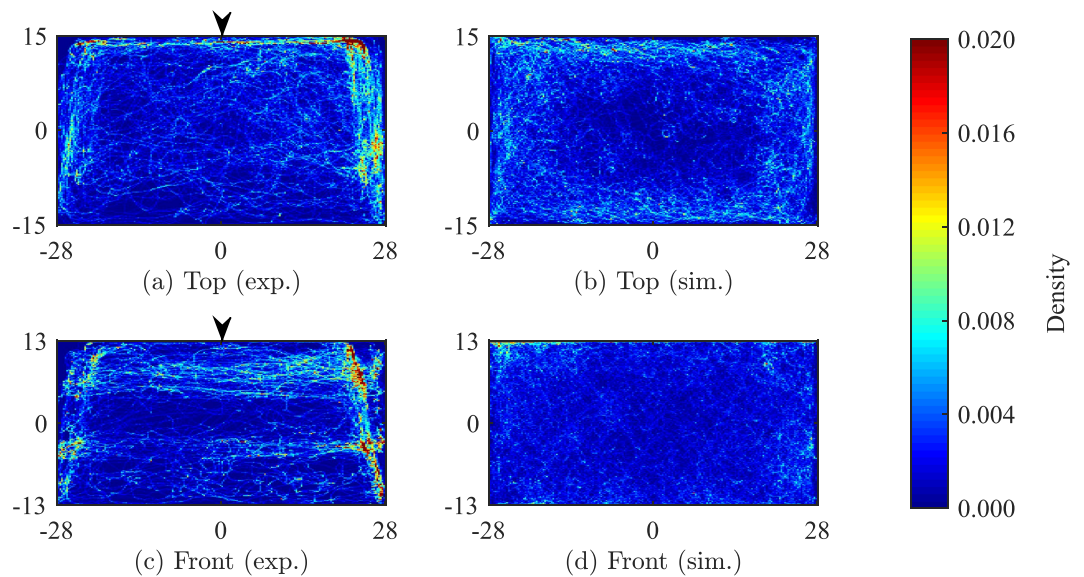
## Results

**First *in-silico* experiment for model calibration.** Position distribution from the top view over all of 10 trials show that fish explore all sections of the tank with a marked preference for the corners and side walls of the tank (Fig. 1(a)). The lack of symmetry in the density along the tank width is related to the fact that the front portion of the tank is not covered with contact paper to enable tracking using the front camera. The front view indicates a similar exploratory behaviour along the depth of the tank (Fig. 1(c)) with more time spent near the side walls. Model predictions yield a response that is qualitatively similar to experimental observations, with larger density toward the walls (Fig. 1(b)).

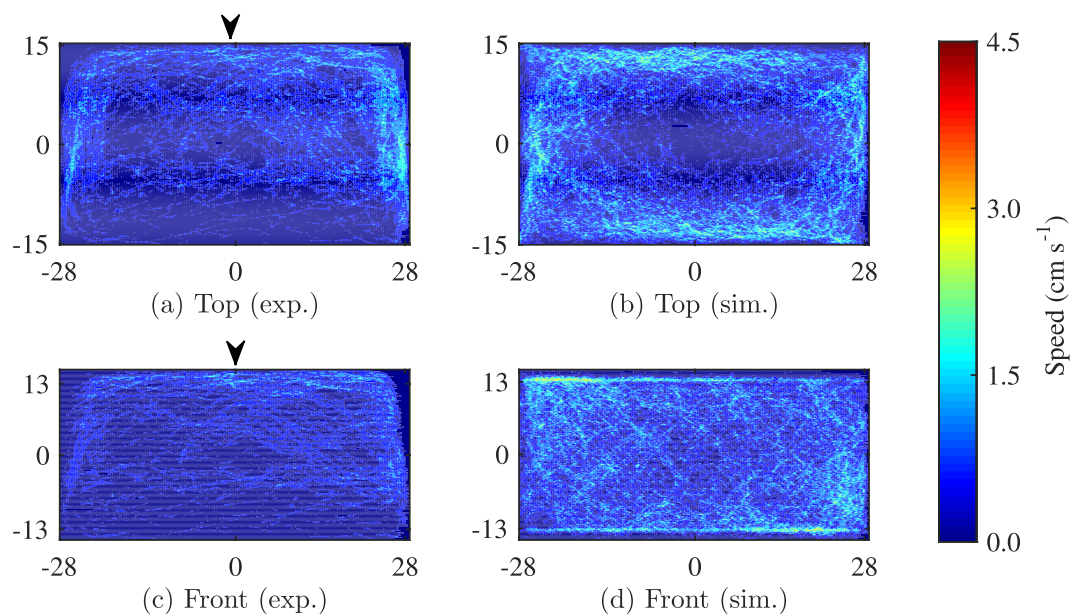
However, the model does not anticipate the wall preference observed in the experimental data. Different from experiments, where fish tend to prefer the clear side of the tank to the side covered with contact paper, simulation results show no such preference for a particular side (Fig. 1(d)). This is because the model implements the same interaction rule on each side. Along the front view simulation results do not reproduce preference to the smaller sides or to the water surface which is different from tank bottom (Fig. 1(d)).

Experimental results show that speed distribution within the tank is asymmetric along the tank plane (Fig. 2(a)) with fish swimming at a higher speed near the side walls and the front wall, which is open for the camera to record. An asymmetry in experimental data is also visible along tank depth with higher speeds near the water surface (Fig. 2(c)). Simulation results (Fig. 2(b) and (d)) are in agreement with experimental observations, whereby they predict larger speeds in the vicinity of the walls. Different from experiments, simulation results do not differ along the front wall, as the model implements the same interaction rule on each wall. In terms of individual motion characteristics, Fig. 3 compares a sample trajectory of a live fish with a simulated fish, offering evidence for the model ability to replicate the sharp turn associated with avoidance manoeuvres and the fragmented tracks within the bulk of the tank.

Figures 4, 5 and 6 compare the behavioural response between live and simulated fish using a non-parametric Kruskal-Wallis test to examine the null hypothesis that the values are sampled from different distributions with different median. Figure 4 displays experimental observations and model predictions on fish activity, measured



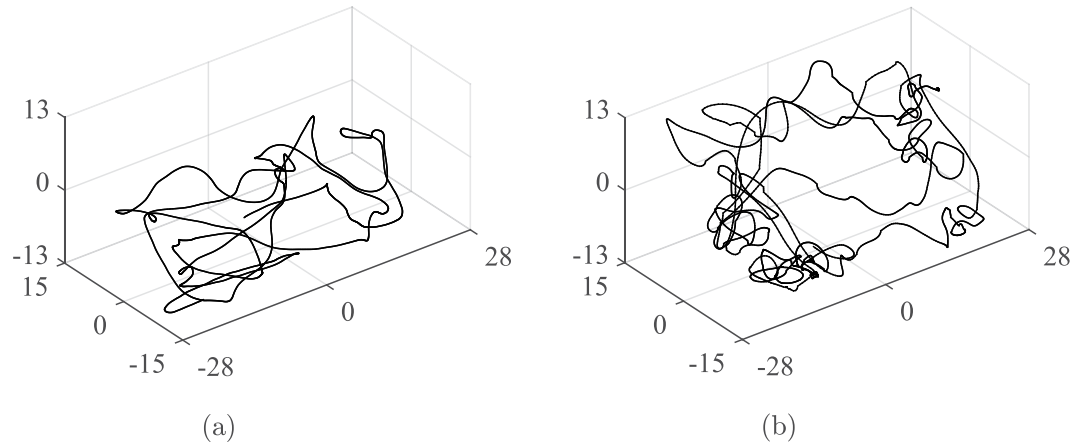
**Figure 1.** Position distribution of nine experimental (a and c) and 10 simulated (b and d) fish, along the length (a and b) and the water column (c and d). The bin size is  $0.28 \times 0.15$  cm for the top view and  $0.28 \times 0.13$  cm for the front view. The density is saturated at 0.02 to ease the visualization of the data, although the maximum value is 0.07. Arrows indicate the tank sides not covered with white contact paper.



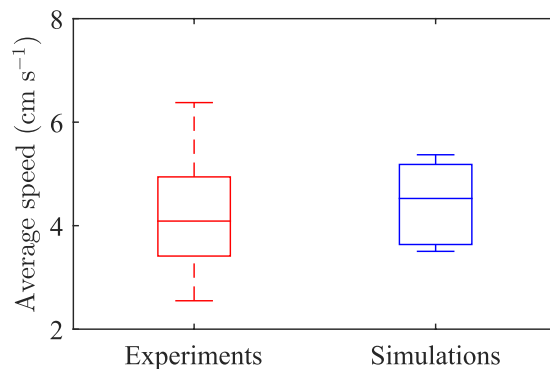
**Figure 2.** Speed distribution of nine experimental (a and c) and a 10 simulated (b and d) fish, along the length (a and b) and the water column (c and d). The bin size is  $0.28 \times 0.15$  cm for the top view and  $0.28 \times 0.13$  cm for the front view. The speed is saturated at  $4.5 \text{ cm s}^{-1}$  corresponding also to the maximum averaged speed value. Arrows indicate the tank sides not covered with white contact paper.

through the average speed. Specifically, we find that model predictions are not significantly different from experimental data ( $\chi^2(1) = 0.8229$ ,  $p = 0.3643$ ).

Figure 5 presents a detailed comparison between experimental data and model predictions on fish spatial preference along the water tank. The model is successful in anticipating the time budgeting along the water column, whereby simulation data and experiments are similar (time spent in the top third:  $\chi^2(1) = 0.0229$ ,  $p = 0.8798$  and time spent in the bottom third:  $\chi^2(1) = 0.0229$ ,  $p = 0.8798$ ). The number of entries predicted by the model is also statistically equivalent to experimental observations (number of entries in the top third:  $\chi^2(1) = 0.1169$ ,  $p = 0.7324$  and number of entries in the bottom third:  $\chi^2(1) = 0.1450$ ,  $p = 0.7033$ ).



**Figure 3.** Sample trajectories of experimental (a) and simulated fish (b) in 3D for two minutes.



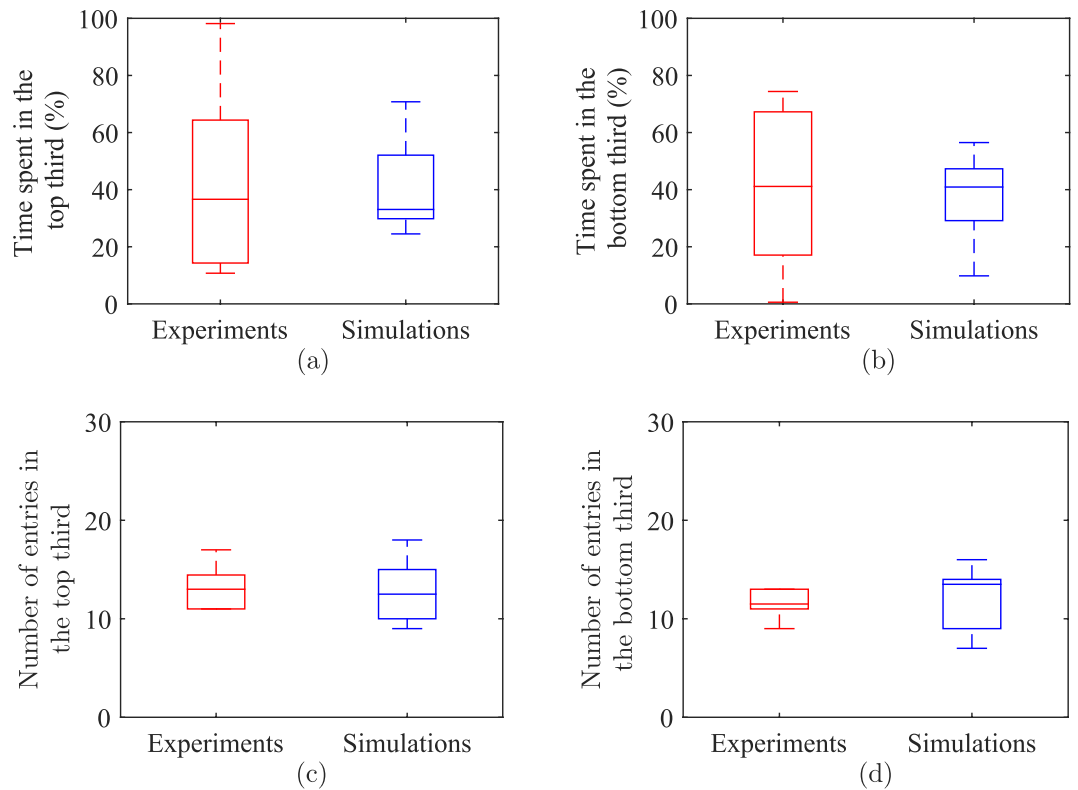
**Figure 4.** Comparison of average speed between live and simulated fish. Experimental data refer to Table 1 in ref. 38. The bottom and top of each box are the first and third quartiles, the band inside the box is the median, and whiskers identify one standard deviation above and below the mean.

Finally, Fig. 6 offers validation for the model capability to anticipate fish positional preference along the length of the tank. In line with all the previous comparisons, model predictions are found to be statistically indistinguishable from experimental data (time spent in the lateral sides:  $\chi^2(1) = 1.12$ ,  $p = 0.2899$  and number of entries in the lateral sides:  $\chi^2(1) = 0.2808$ ,  $p = 0.5961$ ).

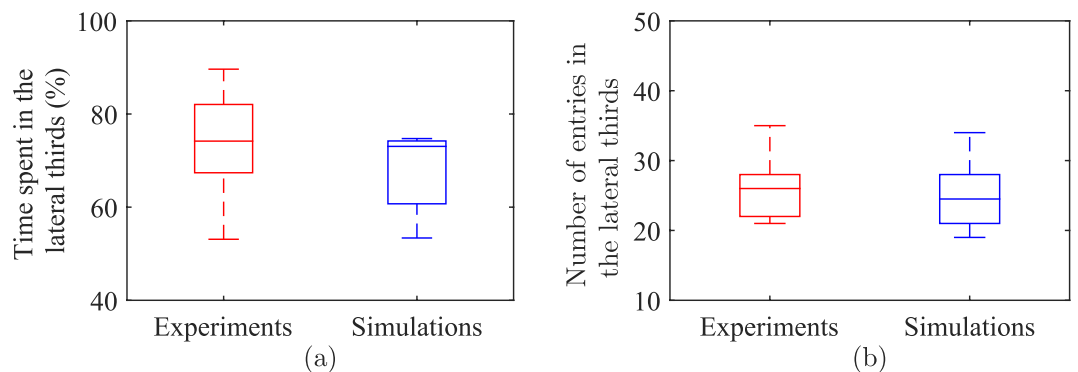
**Second *in-silico* experiment for model evaluation.** Our simulation results presented in Fig. 7 suggest that fish speed is an increasing function of the tank planar dimensions across different depths. Specifically, using a GLM analysis, fish speed is correlated to tank width for all the tested aspect ratios, namely, 0.25 (Dispersion = 0.160,  $F_{1,13} = 15.8$ ,  $p < 0.01$ ), 0.5 (Dispersion = 0.247,  $F_{1,13} = 6.49$ ,  $p = 0.0243$ ), 0.75 (Dispersion = 0.222,  $F_{1,13} = 19.8$ ,  $p < 0.01$ ), and 1 (Dispersion = 0.145,  $F_{1,13} = 41.6$ ,  $p < 0.01$ ). By aggregating the simulation data with respect to the total surface area (Fig. 8), we find a strong correlation between the logarithm of the total surface area and the average speed (Dispersion = 0.168,  $F_{1,58} = 94.4$ ,  $p < 0.01$ ). By performing the same analysis on experimental data from Table 1, we uncover a similar, albeit weaker, trend between the average speed and the total surface area (Dispersion = 2.01,  $F_{1,30} = 27.4$ ,  $p < 0.01$ ). Although the GLM analysis shows a higher dispersion in experimental data compared to simulation results, the slope is also steeper. Thus, experimental data indicate the presence of a stronger effect than computational predictions.

## Discussion

In cellular and molecular biology, mathematical models of regulatory networks, supported by high-throughput data collection, have enabled the seamless integration of computer simulations and experimental research<sup>8,11–14</sup>. Benefits of an *in-silico* approach include: the possibility of setting up advanced starting points for testing novel hypotheses; speeding up the validation of challenging propositions; and putting forward a direct, tangible opportunity to reduce the extent of animal use. A necessary step toward bringing such benefits to the field of animal behaviour is the formulation of mathematical models that can faithfully reconstruct animal locomotion of individuals<sup>53</sup>, which in turn better informs collective motion models<sup>55–57</sup>. In the case of zebrafish, a model organism for preclinical studies, a comprehensive catalogue of observed behaviours<sup>28</sup> and high-volume datasets from the literature<sup>17,37–40</sup> make a compelling case for the development of a framework to support *in-silico* experiments.



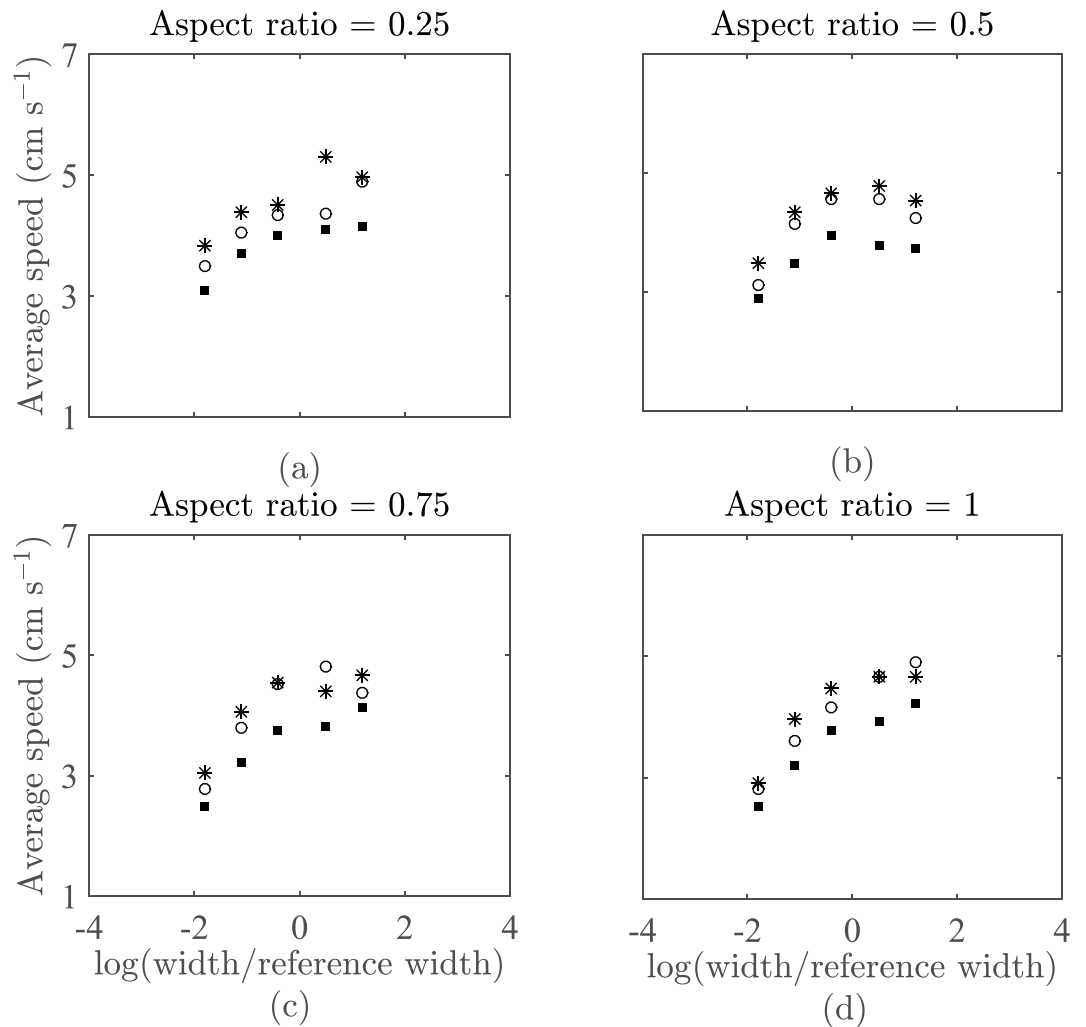
**Figure 5.** Comparison of fish positional preference along the water column between live and simulated fish: time spent in the top third (a); time spent in the bottom third (b); number of entries in the top third (c); and number of entries in the bottom third. Experimental data refer to Fig. 7 of ref. 38. The bottom and top of each box are the first and third quartiles, the band inside the box is the median, and whiskers identify one standard deviation above and below the mean.



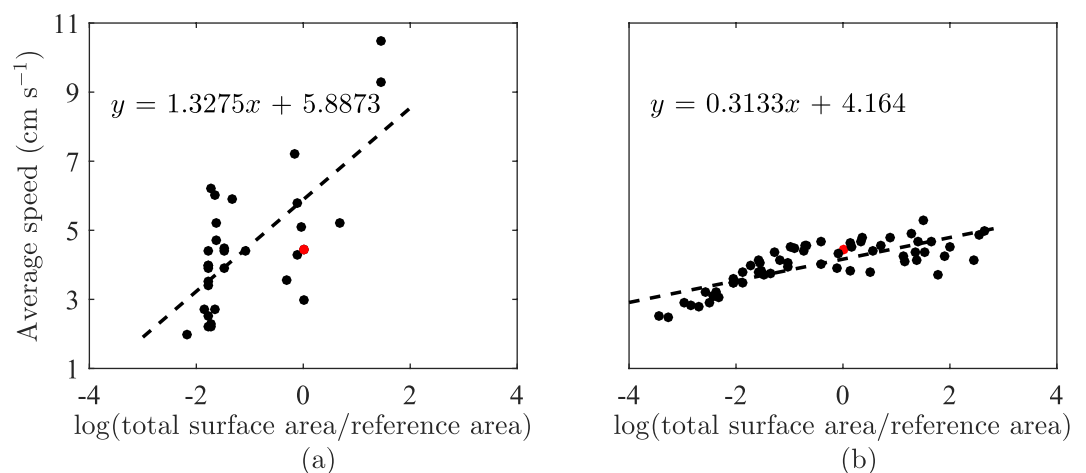
**Figure 6.** Comparison of fish positional preference along the length of the tank between live and simulated fish: time spent in the lateral sides of the tank (a) and number of entries in the lateral sides (b). Experimental data refer to Figs 6 and 7 of ref. 38. The bottom and top of the each are the first and third quartiles, the band inside the box is the median, and whiskers identify one standard deviation above and below the mean.

An accurate model that is able to robustly reproduce the full catalogue of zebrafish behaviours is far from accomplished. Instead, we describe here a model that conforms to basic zebrafish locomotory patterns in 3D, laying the foundations for *in-silico* zebrafish experiments. We extend our previous approach<sup>43,44</sup> of using jump diffusion stochastic processes to reproduce segments of zebrafish turning motion in 2D to a more realistic scenario, where fish swim in 3D adjusting their speed and interacting with the tank walls throughout an entire experiment. Comparison with experimental data demonstrates that our model is successful in qualitatively reproducing zebrafish swimming behaviour, and quantitatively predicting activity and positional preference, in terms of average speed, time budgeting, and number of entries in longitudinal and transversal sections of the tank. Visual differences in positional preference and activity are likely due to complex interactions with clear tank walls, water surface, and tank bottom, as well as difference in interactions due to wall size. The model consists





**Figure 7.** Simulated fish speed by varying tank depth and width with aspect ratio 0.25 (a), 0.5 (b), 0.75 (c), and 1. The symbols star, dot, and square represent tank depths of 10, 20 and 30 cm, respectively. The function  $\log(\bullet)$  refer to the natural logarithm and the width is divided by the width of the experimental tank in ref. 38, which is taken to be the reference width.



**Figure 8.** Effect of the total tank surface on fish speed obtained from (a) data compilation from literature, and (b) *in-silico* experiments of zebrafish swimming in an open field tank. The function  $\log(\bullet)$  refer to the natural logarithm, the total surface area is divided by the surface area of the experimental tank in ref. 38, which is taken to be the reference area. For reference, we display in red the point corresponding to the experimental data in ref. 38 used to calibrate the model.

of 16 parameters, which are calibrated from complete experimental trajectories of zebrafish swimming in 3D. Although the parameter space is large, the identification procedure is relatively robust, as evidenced by the limited variability in model parameters across the nine subjects considered in the experiment. Also, by automatically isolating instances when fish interact with the wall, the procedure is effectively partitioned in two consequent steps, reducing the number of parameters to be determined at once. Model parameters indicate that yaw and pitch rate parameters, such as the relaxation rate and variability have similar values. This is because the fish orientation is defined in the tank inertial frame, and turning up towards the water surface sideways may be interpreted as pitching in the inertial frame. On the other hand, jump intensity in the pitching direction is higher than that in the yaw direction, which is likely caused by stress-related behaviour in experimental subjects<sup>17,35,58,59</sup>. In a novel environment, zebrafish have been shown to repeatedly change their swimming depth<sup>40,60,61</sup>, which in our model is associated with the jump process in the pitch direction.

The potential of the model to complement empirical research is demonstrated through *in-silico* experiments that seek to understand the effect of tank dimensions on zebrafish activity—a study which would otherwise require thousands of animals and multiple experimental setups. These computational experiments reveal that the average speed of zebrafish should increase with all of the tank dimensions. By aggregating data with respect to the overall tank surface, we find a strong correlation with the average speed. A correlation between the tank surface and average speed is also evidenced from experimental data collated from the literature, even though the slope of the correlation is different from the simulation data. While it is likely that fish activity is influenced by a number of factors that are not constant across the collated dataset, such as illumination<sup>62</sup> and habituation time<sup>63,64</sup>, we could explain the observed correlation between speed and surface area through two different reasons. First, fish speed may increase due to fewer instances of boundary avoidance as the tank size increases. Second, speed may increase due to thigmotaxis<sup>61</sup>, produced by thrashing or escape behaviours along the tank walls and corners<sup>28</sup>. This stress-related behaviour may be accompanied by high accelerations<sup>28</sup>, which would, in turn, result into a larger average speed.

In its present incarnation, our model appears to weight the second possibility more than the first. This is also evident from the speed distribution plots of simulated data, which suggest increasing speeds in the vicinity of the side walls, bottom wall, and free surface. Indeed, our model explicitly captures fish positional preference for walls and corners in the wall interaction function. However, our model does not directly encode for the first possibility, namely increased speed due to fewer encounters. This behaviour could be directly related to increased exploratory tendency in real fish, leading to higher average speeds<sup>59</sup> for larger tanks size, and thus the steeper slope observed in the experimental data. Encoding such behaviours within our model may likely be accomplished by setting model parameters such as relaxation rate or wall avoidance as a function of tank dimensions. The nature of this dependence, however, will require detailed experiments that allow an accurate and reliable calibration of the model. Future work will focus on further validating this claim through new experiments with a wider range of tank sizes, and investigating the dependency of the model on other experimental factors, such as varying illumination<sup>62</sup> and habituation time<sup>63,64</sup>, which could influence the observed correlation.

An approach similar to modelling wall interaction could be utilized to describe response to external stimuli, such as drawing towards novel objects or light, avoidance of predators, alarm reactions to sudden threats, and shoaling toward conspecifics. For example, with respect to shoaling behavior, the present model would be used to represent the individual dynamics of each zebrafish, interacting as a networked dynamical system. The network of interactions would then be based on relative orientation and position, with weights associated with each cue that can be calibrated using experimental data building on existing models for fish interactions in 2D<sup>53,65</sup>. The model could also be expanded to encapsulate innate behaviours, such as home-base behaviour and habituation<sup>64</sup>. Home-base behaviour may be integrated into the model by including preference for the initial region of the environment that decays as a function of time, with the rate of decay calibrated on the basis of actual experiments. Habituation, which manifests as a time effect in fish behaviour as it enters a new environment<sup>63,64</sup>, could entail creating a virtual, tighter boundary within the environment that expands to the full tank size as time progresses.

The model may also be expanded to describe the effect of psychoactive compounds, through a controlled dependence of model parameters on pharmacological treatment. For example, in the context of the 2D JPTW, we have demonstrated in ref. 44 that alcohol treatment controls both relaxation rate and jump intensity. In the proposed 3D model, the effect of ethanol should also manifest in the speed modulation and wall interaction. As ethanol concentration is varied, zebrafish should modify their swimming style, since both the tail-beat amplitude and the tail-beat frequency have been shown to robustly decrease for high ethanol concentrations<sup>66</sup>. As a result, the parameters describing the instantaneous change in fish speed and the coupling with the turn rate dynamics should be adjusted as a function of ethanol concentration. The parameters related to the interaction with the wall should depend on ethanol concentration, whereby exposure to intermediate concentrations have been shown to reduce wall interaction, while higher concentration could influence the extent of the interaction<sup>67</sup>.

Creating the scientific and technical knowledge for affording comprehensive *in-silico* experiments on zebrafish will require a systems-level approach<sup>68</sup>. This entails not only identifying and modelling individual behaviours, but also understanding how one behaviour may lead to another. For example, habituation may lead to reduced activity and exploration, and interaction may lead to increased sociality. The proposed model should be considered as a first necessary step toward *in-silico* research. Among the several directions that will be pursued to improve model capability, zebrafish response to conspecifics is perhaps most promising, as it will require augmentation of data-driven response functions<sup>45,53</sup> in 3D, to immediately test social behaviour.

Recent years have witnessed sustained movements that strenuously advocate the respectful treatment and welfare of animals in laboratory research. With approximately twenty million animals used in research and testing annually, policy-makers, citizens, and scientific authorities alike yearn for alternative approaches, while retaining the fundamental instrument of scientific investigation and the improvement of human health. Here, we have proposed the systematic integration of computer modelling as a new, unprecedented means to complement

Reference	Tank shape	Dimensions (cm)	Time (s)	Water volume (l)	Total surface (cm <sup>2</sup> )	Avg. speed (cm s <sup>-1</sup> )	Specific source
88	Rectangular parallelepiped	23.5 × 13.5 × 13.0	300	1.0	1408.8	6.2	Fig. 1d: 2D data available
Exp. 1 <sup>89</sup>	Trapezoidal prism	27.9 × 6.4 × 15.2	300	1.4	1408.8	2.3	Fig. 1: speed estimated from 2D travelled distance
Exp. 2 <sup>89</sup>	Trapezoidal prism	27.9 × 6.4 × 15.2	300	1.4	1408.8	2.2	Fig. 4: speed estimated from 2D travelled distance
37	Trapezoidal prism	27.9 × 7.1 × 15.2	360	1.5	1339.4	4.0	Figs 1 and 2: 3D data available
Exp. 1 <sup>58</sup>	Trapezoidal prism	27.9 × 7.1 × 15.2	360	1.5	1339.4	2.2	Fig. 1a: speed estimated from 2D travelled distance
Exp. 1 <sup>59</sup>	Trapezoidal prism	27.9 × 7.1 × 15.2	360	1.5	1483.7	2.7	Fig. 5b: speed estimated from 2D travelled distance
Exp. 2 <sup>59</sup>	Trapezoidal prism	27.9 × 7.1 × 15.2	360	1.5	1339.4	3.9	Fig. 6b: speed estimated from 2D travelled distance
Exp. 1 <sup>81</sup>	Trapezoidal prism	28.0 × 7.0 × 15.0	360	1.5	1337.6	2.5	Fig. 1c: 3D data available
Exp. 2 <sup>81</sup>	Trapezoidal prism	28.0 × 7.0 × 15.0	1800	1.5	1337.6	4.4	Fig. 1d: 2D data available
Exp. 1 <sup>90</sup>	Trapezoidal prism	28.0 × 7.0 × 15.0	360	1.5	1337.6	3.5	Fig. 1a: speed estimated from 2D travelled distance
60	Trapezoidal prism	28.9 × 7.4 × 15.1	900	1.5	1527.7	5.2	Fig. 5a: 2D data available
91	Trapezoidal prism	28.9 × 15.1 × 7.1	360	1.5	1527.7	4.7	Fig. 3: speed estimated from 3D travelled distance
40	Trapezoidal prism	27.0 × 7.0 × 15.0	360	1.5	1237.5	2.7	Fig. 5a: 3D data available
92	Trapezoidal prism	28.0 × 7.0 × 15.0	300	1.5	1337.6	4.0	Table 1: 2D data available
Exp. 3 <sup>81</sup>	Cylinder	12.0 × 19.0	1200	2.0	893.0	2.0	Fig. 4: 2D data available
Exp. 3 <sup>90</sup>	Cylinder	21.0 × 24.0	360	4.2	1483.7	6.0	Fig. 2a: 2D data available
66	Rectangular parallelepiped	30.0 × 10.0 × 15.0	60	4.5	1800.0	4.4	Fig. 1b: 2D data available
Exp. 1 <sup>93</sup>	Rectangular parallelepiped	30.0 × 10.0 × 15.0	300	4.5	1800.0	3.9	Fig. 1c: 2D data available
Exp. 2 <sup>93</sup>	Rectangular parallelepiped	30.0 × 10.0 × 15.0	300	4.5	1800.0	4.5	Fig. 2c: 2D data available
94	Rectangular parallelepiped	20.0 × 12.0 × 25.0	600	6.0	2080.0	5.9	Fig. 5: speed estimated from 2D travelled distance
Exp. 2 <sup>95</sup>	Cylinder	21.0 × 24.0	360	8.3	1339.4	3.4	Fig. 2b: 2D data available
39	Rectangular parallelepiped	25.0 × 25.0 × 18.0	1200	8.8	2650.0	4.4	Fig. 1f: 3D data available
96	Parallelepiped	40.0 × 30 × 30	1800	36.0	6600.0	7.2	Table 2: 2D data available
97	Parallelepiped	54 × 30 × 30.0	600	24.3	5760.0	3.6	Fig. 8: 3D data available
98	Parallelepiped	50.0 × 25.0 × 30.0	1800	37.0	6940.0	4.3	Fig. 5: speed estimated from 2D travelled distance
99	Parallelepiped	50.0 × 30.0 × 25.0	3600	37.0	7000.0	5.8	Fig. 1: speed estimated from 2D travelled distance
100	Parallelepiped	56.0 × 30.0 × 30.0	900	40.6	7832.0	3.0	Fig. 6: 2D data available
101	Parallelepiped	56.0 × 30.0 × 30.0	600	43.7	7488.0	5.1	Fig. 4: 2D data available
102	Cylinder	90.0 × 10.0	300	63.6	15543.0	5.2	Fig. 4: 2D data available
38	Parallelepiped	56.0 × 30.0 × 30.0	600	43.7	7832.0	4.5	Table 1: 3D data available
103	Parallelepiped	120.0 × 120.0 × 20.0	300	144.0	33600.0	10.5	Fig. S1 (Supporting info): 2D data available
104	Parallelepiped	120.0 × 120.0 × 20.0	300	144.0	33600.0	9.3	Fig. 6: 2D data available

**Table 1. Tank shape, dimensions, duration of the experiment, water volume, total surface area, and average speed from 32 different experiments reported in the literature.** Values from control conditions were used to ensure that fish were not exposed to experiment-specific stimuli. Water volume may not match the tank volume due to lower water levels than tank height. Total surface area is taken as the sum of the area of all water surfaces except when water depth is not available, in which case the tank is assumed to be full (the same rule is used for the computation of the water volume). The dimensions for the rectangular parallelepiped and trapezoid are given in term of (length(cm) × width (cm) × depth (cm)) whereas for the trapezoidal tank shape, only the maximum length and the width at the top surface are reported. The cylinder shape is given in term of (diameter (cm) × height (cm)).

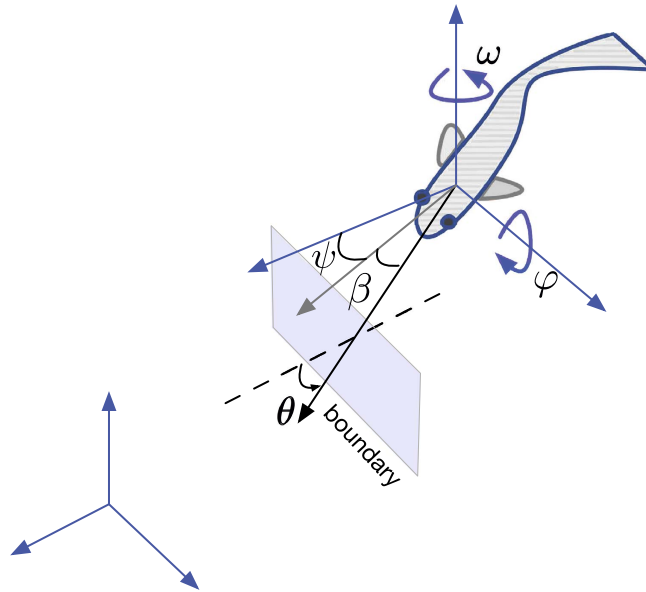
preclinical research on animal behaviour, focusing on the zebrafish-animal-model. While our simulations cannot replace *in-vivo* experiments, they can constitute an economical (time and budget-wise) complement to inform new experimental design and engineering approaches<sup>69–72</sup>, along with the definition of mathematically grounded measurable parameters<sup>73–76</sup>, leading to a consistent reduction of experimental subjects.

## Materials and Methods

We begin by formulating the model for the time evolution of the yaw rate, pitch rate, and speed, all of which are assumed to be coupled. Environmental interactions with the boundaries—tank walls, tank bottom, and water surface—are incorporated into the yaw and pitch rates in the form of avoidance reactions. We then detail the experimental data used to calibrate the 3D model, followed by model calibration where individual parameters are estimated using maximum likelihood estimation (MLE).

**Ethics statement.** The data used in this paper are obtained from the control condition of trajectory dataset published in ref. 38. The experimental methods in ref. 38 were performed in accordance with the relevant guidelines and regulations, and were approved by the University Animal Welfare Committee of New York University under protocol number 131424, and the Animal Welfare Oversight Committee of the Polytechnic Institute of New York University (now New York University Tandon School of Engineering which was independent from New York University until 2013 with respect to animal care) under protocol number AWOC-2013-103.





**Figure 9. Fish orientation in terms of the azimuth ( $\psi$ ) and elevation ( $\beta$ ) in the tank inertial frame (blue).** Fish yaw rate, pitch rate, and speed are modelled as stochastic processes. Fish interaction with the boundary is modelled as a function of the predicted time to collision and the angle made by the projected fish heading to the normal at the impact point  $\theta$ .

**Modelling zebrafish motion in 3D.** Fish yaw rate in the tank frame (Fig. 9) is modelled as a jump persistent turning walker (JPTW), following our previous work on zebrafish 2D swimming in shallow water tanks<sup>43</sup>. Briefly, the instantaneous variation at time  $t$  of the yaw rate,  $d\omega_t$  (rad  $s^{-1}$ ), is described as the sum of a deterministic term, which quantifies the deviation from a desired value, and two stochastic terms, which represent variability in the yaw rate in the form of increments of a Wiener process and intermittent large turns in the yaw direction in the form of a jump process. Specifically, the JPTW for the yaw rate reads

$$d\omega_t = \eta^\omega (f^\omega - \omega_t) dt + \sigma^\omega dW_t^\omega + dJ_t^\omega, \quad (1)$$

where  $\eta^\omega$  ( $s^{-1}$ ) is the relaxation rate at which the yaw tends to reach the desired value  $f^\omega$ , which encapsulates wall interaction;  $W_t^\omega$  is the standard Wiener process, whose time increment is a Gaussian with variance equal to  $dt$ ;  $\sigma^\omega$  (rad  $s^{-3/2}$ ) is a weighting parameter used to scale the Wiener process; and  $J_t^\omega$  is the yaw rate jump process.

We model the wall interaction  $f^\omega$  as a decaying function following the approach in ref. 54, and parametrize this function with the time to collision<sup>53</sup>. Denoting by  $\theta^\omega$  (rad) the angle between the normal to the projected wall point of impact and the fish heading vector projected on the horizontal plane, the wall interaction function is

$$f^\omega(t^\omega, \theta^\omega) = k_W \left( \frac{\text{sgn}(\theta^\omega)}{t^\omega} - w_a \cos \theta^\omega \right), \quad (2)$$

where  $\text{sgn}(\bullet)$  is the sign function;  $k_W > 0$  characterizes the intensity of the wall interaction;  $w_a$  ( $s^{-1}$ ) is a non-negative parameter that captures the tendency to follow the wall; and  $t^\omega$  (s) is the time to collision, estimated as the time it takes for the fish to reach the wall at its current velocity. Note that the wall interaction function depends on the time  $t$  through both  $\theta^\omega$  and  $t^\omega$ . As such, the simulated fish turns away from the wall with higher intensity as the time-to-collision approaches zero, and for large values, the time-to-collision does not affect fish swimming. The wall interaction function (2) differs from the one proposed in ref. 53 in terms of the second summand  $-k_W w_a \cos \theta^\omega$ , which is included to describe wall following or thigmotaxis, a commonly observed behaviour in zebrafish<sup>28</sup>.

Similar to ref. 43, the yaw rate jump process,  $J_t^\omega$ , is modelled as

$$J_t^\omega = \sum_{j=1}^{\nu_t^\omega} Z_j^\omega, \quad (3)$$

Here,  $Z_j^\omega$ , for  $j = 1, 2, \dots$ , are independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and standard deviation equal to  $\gamma^\omega$  ( $Z_j^\omega \sim \mathcal{N}(0, \gamma^\omega)$ ) and  $\nu_t^\omega$  is a counting process defined such that, for time  $r \leq t$ ,  $\nu_t^\omega - \nu_r^\omega$  is a Poisson random variable with parameter  $\lambda^\omega(t - r)$ . The parameters  $\gamma^\omega$  (rad  $s^{-1}$ ) and  $\lambda^\omega$  ( $s^{-1}$ ) measure the severity and frequency of the jumps in the yaw rate.

Fish pitch rate in the tank frame (Fig. 9) is also modelled as a JPTW, since zebrafish pitch rate can reach values beyond three times the standard deviation with arrival time following an exponential distribution (Fig. 10). The instantaneous change in pitch rate  $d\varphi_t$  is thus modelled as

$$d\varphi_t = \eta^\varphi (f^\varphi - \varphi_t)dt + \sigma^\varphi dW_t^\varphi + dJ_t^\varphi. \tag{4}$$

Here,  $\eta^\varphi(\text{s}^{-1})$  is the relaxation rate in reaching the desired value  $f^\varphi$ , which captures the tendency to avoid boundaries along the tank depth, that is, tank bottom surface and water surface;  $W_t^\varphi$  is the standard Wiener process, which is weighted by the pitch rate variability  $\sigma^\varphi(\text{rad s}^{-3/2})$ ; and  $J_t^\varphi$  is the pitch rate jump process.

Similar to boundary avoidance in the yaw rate, boundary avoidance along tank depth,  $f^\varphi$  is parameterized by the time to collision  $t^\varphi$  and the angle between the normal to the projected boundary point of impact and the fish heading vector projected on the vertical plane oriented along the tank depth  $\theta^\varphi$

$$f^\varphi(t^\varphi, \theta^\varphi) = k_W \left( \frac{\text{sgn}(\theta^\varphi)}{t^\varphi} - w_a \cos \theta^\varphi \right), \tag{5}$$

where  $k_W > 0$  and  $w_a$  are defined in Eq (2). To reduce the number of parameters in the proposed model, we describe the interaction with the bottom wall and the free surface through the same parameters used for the side walls interaction.

Similar to the yaw, the pitch jump process is also modified to avoid collisions beyond a critical time by forcing a jump as

$$J_t^\varphi = \sum_{j=1}^{\nu_t^\varphi} Z_j^\varphi, \tag{6}$$

where  $Z_j^\varphi$ , for  $j = 1, 2, \dots$ , are i.i.d. Gaussian random variables with zero mean and standard deviation  $\gamma^\varphi$ , and  $\nu_t^\varphi$  is a counting process defined such that, for time  $r \leq t$ ,  $\nu_t^\varphi - \nu_r^\varphi$  is a Poisson random variable with parameter  $\lambda^\varphi(t - r)$ .

Fish speed is found to vary about an average mean value while always staying positive (Fig. 11). We therefore model instantaneous change in fish speed,  $v_t(\text{cm s}^{-1})$  as a Cox-Ingersoll-Ross process<sup>52</sup>

$$dv_t = \eta^v (\mu^v f^v - v_t)dt + \sigma^v \sqrt{v_t} dW_t^v. \tag{7}$$

Here,  $\eta^v(\text{s}^{-1})$  is the adjustment rate controlling the rate at which fish speed reverts to its long term average value  $\mu^v(\text{cm s}^{-1})$ ;  $f^v = f^v(\omega_t, \varphi_t)$  is a scaling function of  $\mu^v$ ; and  $W_t^v$  is a standard Wiener process weighted by the speed volatility  $\sigma^v(\text{cm}^{1/2} \text{s}^{-1})$ . Different from Eqs (1) and (4), the Wiener process is weighted by the state variable, whereby we multiply the speed volatility by the square root of the speed. This allows for scaling the Wiener process proportionally to the instantaneous speed, ensuring that the speed is always nonnegative, provided that the remaining model parameters satisfy  $2\eta^v \mu^v \geq (\sigma^v)^2$ . The scaling function  $f^v$  is defined such that

$$f^v(\omega_t, \varphi_t) = \frac{A^v}{|\omega_t| + |\varphi_t|}, \tag{8}$$

where  $A^v(\text{s}^{-1})$  is a positive parameter to be determined. This functional dependence ensures that the speed decreases with the pitch and yaw rate, whereby the fish is expected to reduce its speed when undertaking a manoeuvre in either the vertical or horizontal plane. The stochastic differential equations Eqs (1), (4) and (7) define the proposed 3D model of zebrafish swimming.

Beyond the coupling through the function  $f^v$  in Eqs (7) and (8), we assume that speed, pitch rate, and yaw rate processes are pairwise linearly correlated via the corresponding stochastic terms<sup>77</sup>. The proposed correlation structure is

$$dW_t^v dW_t^\varphi = \rho_W^{v\varphi} dt, \tag{9}$$

$$dW_t^v dW_t^\omega = \rho_W^{v\omega} dt, \tag{10}$$

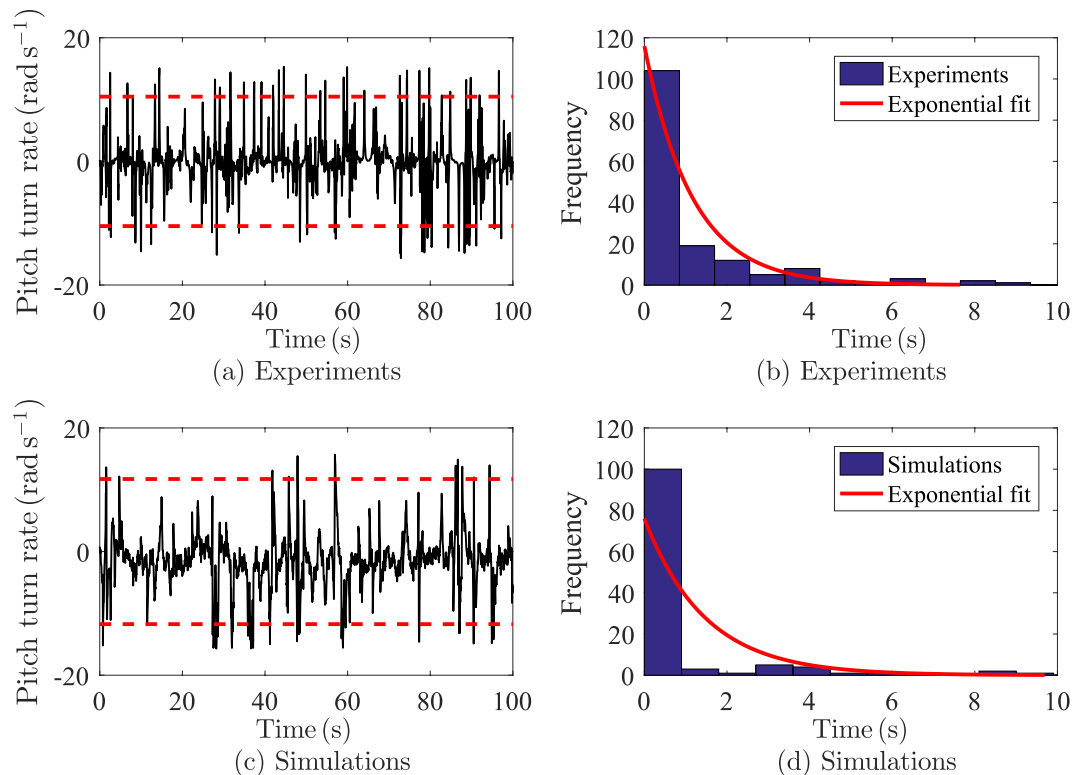
where  $\rho_W^{v\varphi} < 0$  is the correlation coefficient between the speed and the pitch rate Wiener processes;  $\rho_W^{v\omega} < 0$  is the correlation coefficient between the speed and the yaw rate Wiener processes; Based on experimental data where absence or low correlation is observed between yaw and pitch rate, we hypothesize no correlation between pitch and yaw rates, and between the speed Wiener process and the pitch and yaw rate jumps, that is,

$$dW_t^\varphi dW_t^\omega = 0, \tag{11}$$

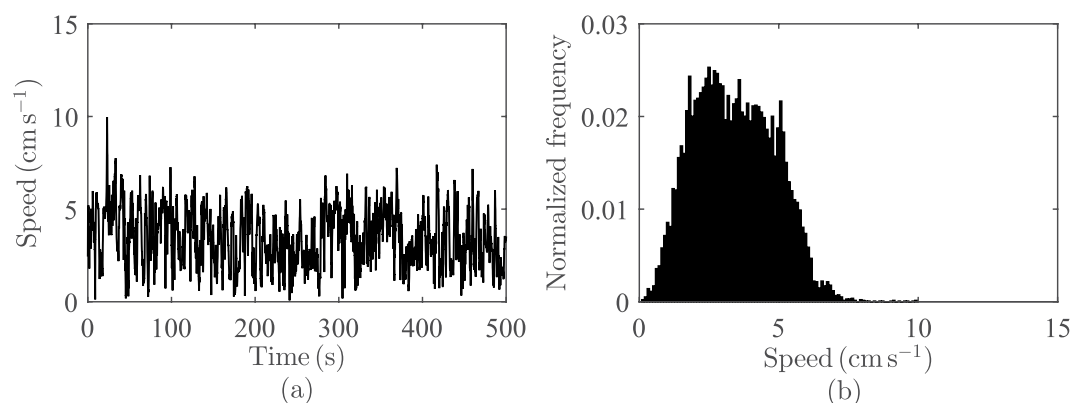
$$dW_t^v Z_t^\varphi = 0, \tag{12}$$

$$dW_t^v Z_t^\omega = 0. \tag{13}$$

**Model calibration.** The experimental dataset of 3D zebrafish swimming used to calibrate the model is described first, followed by our approach for parameter estimation, which is a combination of MLE and locally weighted regression. Next, we present the discrete form of the 3D fish swimming dynamics that is utilized to



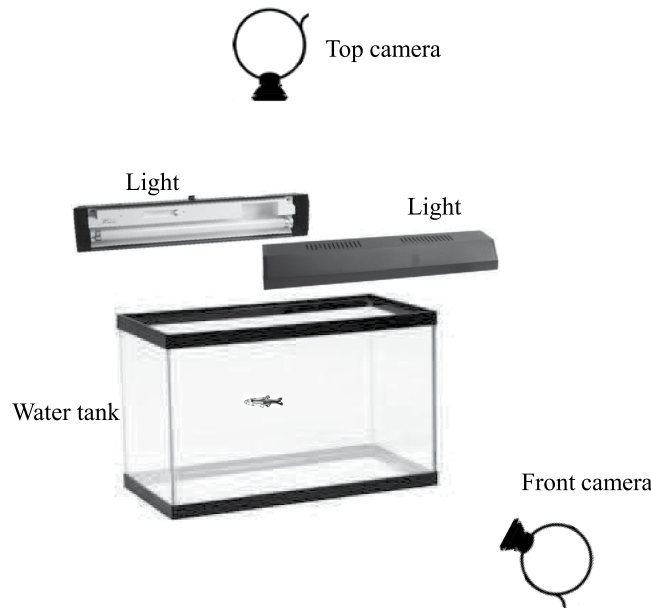
**Figure 10.** Sample experimental and simulated fish pitch rate time series, (a) and (c), and distribution of the jumps arrival time in seconds, (b) and (d). Jumps are identified with values of the pitch rate over three times the standard deviation in dashed lines (red). The arrival time is the time between two consecutive jumps, and can be fitted using an exponential distribution in solid line (red).



**Figure 11.** Sample fish speed time series (a) and distribution (b). Speed is modelled as a stochastic process that maintains a positive value.

simulate trajectories. Finally, we detail the *in-silico* experiments that are used to validate and evaluate the proposed model.

**Experimental dataset for model calibration.** We calibrate the model using a subset of the dataset collected as part of an earlier study<sup>38</sup>. Specifically, the data comprises of 3D trajectories of 10 experimentally naive zebrafish, filmed one at a time, in a  $56 \times 30 \times 30$  cm compartment within a  $76 \times 30 \times 30$  cm glass tank (Fig. 12) for 10 minutes each after 10 minutes of habituation. The 56 cm length compartment was separated from two side regions, 10 cm in length each, with a glass partition. The water depth in the tank was maintained at 26 cm. The zebrafish were filmed using two camera arranged to view the top and front side of the tank thus obtaining orthogonal views. The tank surface on four sides, where camera view was not required was covered with white contact paper for increased contrast during tracking. A custom tracking system was utilized to obtain



**Figure 12. Illustration of the experimental setup.** All sides, except the top and front, are covered with white contact paper.

horizontal and vertical trajectories of fish swimming which were then merged to form three-dimensional trajectory data. Additional details on the experimental procedure and the tracking algorithm can be found in ref. 38.

Fish 3D position data was filtered using a Daubechies wavelet filter<sup>78</sup> to remove body movement due to tail-beat motion<sup>54</sup>. Similar to refs 43, 54, the curvature of the fish centroid trajectory was utilized to estimate the time trace of fish speed, pitch rate, and yaw rate. Stress-related measures, including freezing and thrashing, were computed automatically from the trajectory data<sup>79</sup>. A single trial where fish was found to freeze more than 70% of the time was discarded from the analysis, reducing the total number of trials to nine.

**Parameter estimation.** The full 3D swimming model excluding boundary avoidance consists of 16 parameters denoted by  $\Theta = (\Omega, \Phi, \Upsilon, P)$ . These include: four independent parameters and two identical parameters each for yaw rate  $\Omega = (\eta^\omega, \sigma^\omega, \gamma^\omega, \lambda^\omega, k_w, w_a)$  and pitch rate  $\Phi = (\eta^\varphi, \sigma^\varphi, \gamma^\varphi, \lambda^\varphi, k_w, w_a)$ ; four independent parameters for speed  $\Upsilon = (\mu^v, \eta^v, \sigma^v, A^v)$ ; and two correlation coefficients  $P = (\rho_W^{v\varphi}, \rho_W^{v\omega})$ . Model parameters were estimated using MLE with a discrete time Euler approximation<sup>80</sup>. Denoting the time index by  $k = 1, 2, \dots$ , where  $t_k - t_{k-1} = \Delta t$  (s), the discrete forms of Eqs (1), (4) and (7) are

$$\begin{aligned} \omega(k) &= f^\omega (1 - e^{-\eta^\omega \Delta t}) + \omega(k-1)e^{-\eta^\omega \Delta t} + \sigma^\omega \sqrt{\frac{1}{2\eta^\omega}} (1 - e^{-2\eta^\omega \Delta t}) \\ &\times \varepsilon^\omega(k) + \gamma^\omega \Delta v^\omega(k\Delta t) \zeta^\omega(k), \end{aligned} \quad (14)$$

$$\begin{aligned} \varphi(k) &= f^\varphi (1 - e^{-\eta^\varphi \Delta t}) + \varphi(k-1)e^{-\eta^\varphi \Delta t} + \sigma^\varphi \sqrt{\frac{1}{2\eta^\varphi}} (1 - e^{-2\eta^\varphi \Delta t}) \\ &\times \varepsilon^\varphi(k) + \gamma^\varphi \Delta v^\varphi(k\Delta t) \zeta^\varphi(k), \end{aligned} \quad (15)$$

$$v(k) = \mu^v f^v \eta^v \Delta t + v(k-1)(1 - \eta^v \Delta t) + \sigma^v \sqrt{v(k-1)\Delta t} \varepsilon^v(k). \quad (16)$$

Similar to ref. 43, we hypothesize that a single jump occurs at most in a time step and accordingly, approximate the jump processes  $\Delta v^\omega(k\Delta t)$  and  $\Delta v^\varphi(k\Delta t)$  by Bernoulli random variables with parameters  $\lambda^\omega \Delta t$ , and  $\lambda^\varphi \Delta t$  respectively. In Eqs (14), (15) and (16),  $\varepsilon^v(k)$ ,  $\varepsilon^\varphi(k)$  and  $\varepsilon^\omega(k)$  are all correlated Gaussian random variables.

We estimate the model parameters separately for fish swimming near the boundary and far from it. Specifically, we partition the fish trajectory data based on the distance from the wall, where position data further from 2 body-lengths (BL)  $\approx 6$  cm of the wall is considered to be not affected by the wall interaction response. This threshold of 2BL represents a rounded up value for measuring zebrafish peripheral activity<sup>61,81</sup>. Accordingly, we first estimate the fish state at time index  $k$  denoted by  $X(k) = [\omega(k), \varphi(k), v(k)]^T$ , and at  $k-1$  without the boundary avoidance functions as  $\Pi(k-1) = [\omega(k-1)e^{-\eta^\omega \Delta t}, \varphi(k-1)e^{-\eta^\varphi \Delta t}, \mu^v \eta^v \Delta t + v(k-1)(1 - \eta^v \Delta t)]^T$ . Using the above notation, the fish discrete state  $X(k)$  admits the following distribution:

$$X(k) \sim \Pi(k - 1) + \varepsilon(k) + \zeta(k), \tag{17}$$

where  $\varepsilon(k) = \mathcal{N}(\mathbf{0}, \mathcal{S}_W)$  and  $\zeta(k) = \mathcal{N}(\mathbf{0}, \mathcal{S}_J)$ , with  $\mathbf{0} = [0, 0, 0]^T$  being the 3D null vector, and the covariance matrices for Wiener process,  $\mathcal{S}_W$ , and jump processes,  $\mathcal{S}_J$ , are

$$\mathcal{S}_W = \begin{bmatrix} \sigma_{W,11} & \sigma_{W,21} & \sigma_{W,31} \\ \sigma_{W,12} & \sigma_{W,22} & \sigma_{W,23} \\ \sigma_{W,13} & \sigma_{W,32} & \sigma_{W,33} \end{bmatrix} \text{ and } \mathcal{S}_J = \begin{bmatrix} \sigma_{J,11} & \sigma_{J,21} & \sigma_{J,31} \\ \sigma_{J,12} & \sigma_{J,22} & \sigma_{J,23} \\ \sigma_{J,13} & \sigma_{J,32} & \sigma_{J,33} \end{bmatrix},$$

with

$$\sigma_{W,11} = \frac{(\sigma^\omega)^2}{2\eta^\omega} (1 - e^{-2\eta^\omega \Delta t}), \tag{18}$$

$$\sigma_{W,12} = \sigma_{W,21} = 0, \tag{19}$$

$$\sigma_{W,13} = \sigma_{31}^W = \sigma^v \sigma^\omega \rho_W^{v\omega} \sqrt{v(k-1) \frac{\Delta t}{2\eta^\omega} (1 - e^{-2\eta^\omega \Delta t})}, \tag{20}$$

$$\sigma_{W,22} = \frac{(\sigma^\varphi)^2}{2\eta^\varphi} (1 - e^{-2\eta^\varphi \Delta t}), \tag{21}$$

$$\sigma_{W,23} = \sigma_{W,32} = \sigma^v \sigma^\varphi \rho_W^{v\varphi} \sqrt{v(k-1) \frac{\Delta t}{2\eta^\varphi} (1 - e^{-2\eta^\varphi \Delta t})}, \tag{22}$$

$$\sigma_{W,33} = (\sigma^v)^2 v(k-1) \Delta t, \tag{23}$$

and

$$\sigma_{J,11} = (\gamma^\omega)^2, \tag{24}$$

$$\sigma_{J,12} = \sigma_{J,21} = 0, \tag{25}$$

$$\sigma_{J,13} = \sigma_{J,31} = 0, \tag{26}$$

$$\sigma_{J,22} = (\gamma^\varphi)^2, \tag{27}$$

$$\sigma_{J,23} = \sigma_{J,32} = 0, \tag{28}$$

$$\sigma_{J,33} = 0. \tag{29}$$

We introduce the system augmented covariance matrix  $\mathcal{S}_{\text{aug}}$ , assumed positive definite, as

$$\mathcal{S}_{\text{aug}} = \mathcal{S}_W + \begin{bmatrix} \mathbf{0} & \delta^\varphi & \delta^\omega \\ \delta^\varphi & \delta^\varphi & \delta^\varphi \delta^\omega \\ \delta^\omega & \delta^\varphi \delta^\omega & \delta^\omega \end{bmatrix} \mathcal{S}_J,$$

where  $\delta^\varphi$  and  $\delta^\omega$  are equal to 1 when a jump occurs in the pitch or yaw rates respectively, and 0 otherwise. Using this augmented covariance matrix, the likelihood probability density function for the state of the fish at a given time index  $k$  is

$$f_\Theta^{(k)}(X(k)|X(k-1)) = \sum_{\delta^\varphi, \delta^\omega \in \{0,1\}} g(\lambda^\varphi, \lambda^\omega, \delta^\varphi, \delta^\omega) \sqrt{(2\pi)^3 \det(\mathcal{S}_{\text{aug}})} \\ \times \exp\left[-\frac{1}{2}(X(k) - \Pi(k-1))^T \mathcal{S}_{\text{aug}}^{-1} (X(k) - \Pi(k-1))\right],$$

where  $g(\lambda^\varphi, \lambda^\omega, \delta^\varphi, \delta^\omega) = (\lambda^\varphi \Delta t)^{\delta^\varphi} (\lambda^\omega \Delta t)^{\delta^\omega} (1 - \lambda^\varphi \Delta t)^{1-\delta^\varphi} (1 - \lambda^\omega \Delta t)^{1-\delta^\omega}$  is the probability of a jump to occur in either pitch or yaw turn rate and  $\det(\bullet)$  is the matrix determinant. Considering a given 3D time series of zebrafish speed, pitch, and yaw turn rates  $\{\hat{X}(k)\}_{k=1}^T$ , where  $T$  is the length of the dataset, model calibration is performed by maximizing the log-likelihood function<sup>50</sup>



$$\mathcal{L}_{3D\text{model}}(\Theta) = \sum_{k=2}^T \log f_{\Theta}^{(k)}(\hat{X}(k)|\hat{X}(k-1)). \quad (30)$$

The MLE is implemented using the non-linear constraint algorithm available in the optimization toolbox of MATLAB (R2015a; MathWorks, Natick, Massachusetts, USA)<sup>82</sup>. To initialize and reduce the size of the large parameter set, three separate likelihood functions, corresponding to the three stochastic differential equations associated with, yaw rate, pitch rate, and speed, are utilized to estimate the values of a subset of the parameter set, consisting of the 11 independent parameters, with no correlation. Next, these values are used as initial values to maximize the 3D likelihood function in Eq (30). Table 2 lists the average value of estimated parameters. The full set of calibrated model parameters for each fish are available in the supporting information.

For wall interaction functions  $f^v$  and  $f^p$ , regression fits are unable to capture two behaviours: wall interaction with low times-to-collision during which the fish performs a jump manoeuvre to prevent a collision, and wall following<sup>28</sup>. While simulating wall interaction with jumps require further complexity in the form of additional jump term, wall following is likely not captured due to the fit performed on trajectories. Specifically, trajectories will be rendered discontinuous due to specific threshold used to partition them into free swimming or wall interaction. In reality, this threshold may not be even a constant value. Accordingly, an instance of wall following may be captured in the form of two different collision avoidance manoeuvres if the fish swims beyond the assigned threshold even once.

During our simulations we found that we can achieve both these behaviours by increasing the value of  $k_w$ . Accordingly, we select the value of  $k_w = 2.33$  for  $w_a = 0.88 \text{ s}^{-1}$ . The parameter  $A^v$  for the scaling function  $f^v$  is estimated in two steps by utilizing a locally weighted regression to estimate the decay function of the absolute turn rate and the speed, and by applying a simple linear regression to approximate the decay rate which is averaged to 2.92 for the 9 fish datasets.

**Simulation of fish trajectories.** Fish are simulated using Eqs (14), (15) and (16), that is, by Euler integrating the stochastic differential equations to obtain fish heading velocity at each time step. In particular, identifying fish position at  $k$  as  $\mathbf{r}(k) \in \mathbb{R}^3$  and its orientation in 3D through the azimuth  $\psi(k) \in [-\pi, \pi)$  and elevation  $\beta(k) \in [-\frac{\pi}{2}, \frac{\pi}{2})$  in the tank frame (Fig. 9), the position and orientation at the next time index is

$$\begin{aligned} \mathbf{r}(k+1) &= \mathbf{r}(k) + v(k+1) \begin{bmatrix} \cos \psi(k+1) \cos \beta(k+1) \\ \sin \psi(k+1) \cos \beta(k+1) \\ \sin \beta(k+1) \end{bmatrix} \Delta t, \\ \psi(k+1) &= \psi(k) + \omega(k+1) \Delta t, \\ \beta(k+1) &= \beta(k) + \phi(k+1) \Delta t, \end{aligned} \quad (31)$$

where  $\omega(k)$ ,  $\phi(k)$ , and  $v(k)$  are obtained as in Eqs (14), (15) and (16). The correlated Gaussian random variables  $\varepsilon^v(k)$ ,  $\varepsilon^\omega(k)$ , and  $\varepsilon^\phi(k)$  are sampled on uncorrelated standard Gaussian random variables  $\xi^v(k)$ ,  $\xi^\omega(k)$ , and  $\xi^\phi(k)$  with zero mean and unit standard deviation through<sup>77</sup>

$$\varepsilon^\omega(k) = \xi^\omega(k), \quad (32)$$

$$\varepsilon^\phi(k) = \xi^\phi(k), \quad (33)$$

$$\varepsilon^v(k) = \rho_w^{v\omega} \xi^\omega(k) + \rho_w^{v\phi} \xi^\phi(k) + \sqrt{1 - (\rho_w^{v\omega})^2 - (\rho_w^{v\phi})^2} \xi^v(k). \quad (34)$$

**In-silico experiments.** Beyond demonstrating the accuracy of the proposed model on the experimental dataset in ref. 38, we consider an independent *in-silico* experiment addressing conditions which were not part of the dataset used for model calibration.

In the first computational experiment, a total number of 10 fish are simulated, each for 600 seconds with  $\Delta t = 0.04 \text{ s}$  within a  $56 \times 30 \times 30 \text{ cm}$  bounded environment, replicating the experimental setup in ref. 83. Fish position is initialized so that all fish start at right top corner of the tank and the azimuth and elevation are oriented at a uniformly random direction. Model parameters are selected randomly to uniformly lie within one standard deviation of the average values (Table 2) obtained after model calibration on the real experimental data. This selection for the model parameters attempts at facilitating the process of replicating model predictions, and mitigating the dependence of inter-individual variability of the subjects on the predictions of ethological observables.

In the second computational experiment, we explore the role of tank dimensions (length, width, and depth) on zebrafish response. For a rectangular parallelepiped, we consider three different values for the depth (10, 20 and 30 cm), five different values for the tank width (5, 10, 20, 50 and 100 cm), and four different values aspect ratios (ratio of width to length: 0.25, 0.5, 0.75 and 1) width. Overall, we, simulate a total of 60 different tanks, spanning typical sizes considered in the literature (Table 1). We perform a total of 10 simulations for each tank geometry, similar to the first computational experiment. Each simulation is initialized as the first *in-silico* experiment. We set  $\Delta t = 0.04 \text{ s}$  and we simulate each fish for 12 minutes. Model parameters are randomized as in the first experiment.

Parameter	Description	Mean $\pm$ std
$\eta_\omega$	Relaxation rate (yaw)	$3.32 \pm 0.46$ (s <sup>-1</sup> )
$\sigma_\omega$	Turn rate variability (yaw)	$3.00 \pm 0.68$ (rad s <sup>-3/2</sup> )
$\gamma_\omega$	Jump intensity (yaw)	$3.27 \pm 1.15$ (rad s <sup>-1</sup> )
$\lambda_\omega$	Jump frequency (yaw)	$0.42 \pm 0.08$ (s <sup>-1</sup> )
$\eta_\varphi$	Relaxation rate (pitch)	$3.72 \pm 0.50$ (s <sup>-1</sup> )
$\sigma_\varphi$	Turn rate variability (pitch)	$3.73 \pm 0.63$ (rad s <sup>-3/2</sup> )
$\gamma_\varphi$	Jump intensity (pitch)	$4.52 \pm 0.97$ (rad s <sup>-1</sup> )
$\lambda_\varphi$	Jump frequency (pitch)	$0.47 \pm 0.02$ (s <sup>-1</sup> )
$\mu_v$	Long term mean speed	$4.49 \pm 1.26$ (cm s <sup>-1</sup> )
$\eta_v$	Speed adjustment rate	$0.26 \pm 0.23$ (s <sup>-1</sup> )
$\sigma_v$	Speed volatility	$0.62 \pm 0.27$ (cm <sup>1/2</sup> s <sup>-1</sup> )
$\rho_{v\omega}^W$	Correlation speed and yaw rate Wiener processes	$-0.24 \pm 0.67$
$\rho_{v\varphi}^W$	Correlation speed and pitch rate Wiener processes	$-0.45 \pm 0.62$
$k_W$	Intensity of the wall interaction	2.33
$w_a$	Propensity to follow the wall	0.88 (s <sup>-1</sup> )
$A^v$	Speed decay coefficient	2.92 (s <sup>-1</sup> )

**Table 2. Estimated model parameters from experimental data.**

**Analysis.** The trajectories of the 10 simulated fish in the first experiment are utilized to generate a distribution of position and speed along the horizontal and vertical sections of the environment for visual comparison with the experimental data. All 10 simulated fish trajectories are then processed to compute ethological observables used for analysis in fish behaviour experiments<sup>31,36,84</sup>, such as: average speed; spatial preference along the water column scored from the time spent and number of entries in the top or bottom third of the tank; and the spatial preference along the tank length, measured from number of entries in either of the left or right third of the tank and the time spent therein. These observables are then compared with experimental data from<sup>38</sup> using a non-parametric Kruskal-Wallis test<sup>85</sup> which returns the chi-square statistics ( $\chi^2(\cdot)$ ) with degree of freedom ( $\cdot$ ), and the  $p$ -value.

For the simulations of the second experiment, fish behaviour is analysed in terms of the average speed, which is often reported in the technical literature as a measure of activity<sup>60,86</sup> (Table 1). The correlation between the fish speed and each tank dimension is assessed using a generalized linear regression model (GLM)<sup>87</sup>. The same model is used to aggregate the simulation data and demonstrate the dependence of the speed on the overall tank surface area. The latter dependence is verified through comparison with experimental observations, which are also aggregated in terms of the total surface area. The total surface area accounts for the entire surface of the water, including the tank sides and the free surface. The experimental data listed in Table 1 is collated from 32 different zebrafish experiments, across a range of tank geometries and dimensions. We utilize the command *glmfit* of Matlab which returns the dispersion, the  $F$ -test statistics, and the  $p$ -value on the regression model assuming Gaussian error distribution. All the statistical tests are performed with a significance level set at 0.05.

## References

- Russell, W. & Burch, R. *The Principles of Humane Experimental Technique*. (Universities Federation for Animal Welfare (UFAW), Methuen, London, UK, 1959).
- Ekins, S., Mestres, J. & Testa, B. *In silico* pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br. J. Pharmacol.* **152**, 9–20 (2007).
- Thomas, L. & Juanes, F. The importance of statistical power analysis: an example from animal behaviour. *Anim. Behav.* **52**, 856–859 (1996).
- Chuaqui, R. F. *et al.* Post-analysis follow-up and validation of microarray experiments. *Nat. Genet.* **32**, 509–514 (2002).
- Huston, M., DeAngelis, D. & Post, W. New computer models unify ecological theory. *BioScience* **38**, 682–691 (1988).
- Kumar, V. S. & Maranas, C. D. Growmatch: an automated method for reconciling *in silico* *in vivo* growth predictions. *PLoS Comput. Biol.* **5**, e1000308 (2009).
- Segura, D., Mahadevan, R., Juárez, K. & Lovley, D. R. Computational and experimental analysis of redundancy in the central metabolism of geobacter sulfurreducens. *PLoS Comput. Biol.* **4**, e36 (2008).
- Di Ventura, B., Lemerle, C., Michalodimitrakis, K. & Serrano, L. From *in vivo* to *in silico* biology and back. *Nature* **443**, 527–533 (2006).
- Hayashi, M. Three Rs in mutation research-From *in vivo* to *in silico*- evaluation. *Proc. 6th World Congress Alt.: Anim. Use Life Sci.* **14**, 9–13 (2007).
- Sanga, S. *et al.* Predictive oncology: a review of multidisciplinary, multiscale *in silico* modeling linking phenotype, morphology and growth. *Neuroimage* **37**, S120–S134 (2007).
- Natsch, A., Emter, R. & Ellis, G. Filling the concept with data: integrating data from different *in vitro* and *in silico* assays on skin sensitizers to explore the battery approach for animal-free skin sensitization testing. *Toxicol. Sci.* **107**, 106–121 (2009).
- Geris, L., Vander Sloten, J. & Van Oosterwyck, H. *In silico* biology of bone modelling and remodelling: regeneration. *Phil. Trans. R. Soc. A.: Math. Phys. Eng. Sci.* **367**, 2031–2053 (2009).
- An, G. *In silico* experiments of existing and hypothetical cytokine-directed clinical trials using agent-based modeling. *Crit. Care Med.* **32**, 2050–2060 (2004).
- Raunio, H. *In silico* toxicology-non-testing methods. *Front. Pharmacol.* **2** (2011).
- Rohrseitz, N. & Fry, S. N. Behavioural system identification of visual flight speed control in *Drosophila melanogaster*. *J. R. Soc. Interface* **8**, 171–185 (2011).

16. Lillacci, G. & Khammash, M. Parameter estimation and model selection in computational biology. *PLoS Comput. Biol.* **6**, e1000696 (2010).
17. Cachat, J. *et al.* Three-dimensional neurophenotyping of adult zebrafish behavior. *PLoS ONE* **6**, e17597 (2011).
18. Kalueff, A. V., Wheaton, M. & Murphy, D. L. What's wrong with my mouse model?: advances and strategies in animal modeling of anxiety and depression. *Behav. Brain Res.* **179**, 1–18 (2007).
19. Karakostas, T., Hsiang, S., Boger, H., Middaugh, L. & Granholm, A. C. Three-dimensional rodent motion analysis and neurodegenerative disorders. *J. Neurosci. M* **231**, 31–37 (2014).
20. Kitano, H. Computational systems biology. *Nature* **420**, 206–210 (2002).
21. Robinson, T. E. & Becker, J. B. Enduring changes in brain and behavior produced by chronic amphetamine administration: a review and evaluation of animal models of amphetamine psychosis. *Brain Res. Rev.* **11**, 157–198 (1986).
22. DeLong, M. R. Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.* **13**, 281–285 (1990).
23. Dooley, K. & Zon, L. I. Zebrafish: a model system for the study of human disease. *Curr. Opin. Genet. Dev.* **10**, 252–6 (2000).
24. Lieschke, G. J. & Currie, P. D. Animal models of human disease: zebrafish swim into view. *Nat. Rev. Genet.* **8**, 353–67 (2007).
25. Howe, K. *et al.* The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498–503 (2013).
26. Brand, M., Granato, M. & Nüsslein-Volhard, C. Keeping and raising zebrafish. *Zebrafish* **261**, 7–37 (2002).
27. Kishi, S. *et al.* The zebrafish as a vertebrate model of functional aging and very gradual senescence. *Exp. Geront.* **38**, 777–786 (2003).
28. Kalueff, A. V. *et al.* Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond. *Zebrafish* **10**, 70–86 (2013).
29. Spence, R., Gerlach, G., Lawrence, C. & Smith, C. The behaviour and ecology of the zebrafish, (*Danio rerio*). *Biol. Rev.* **83**, 13–34 (2008).
30. Wright, D. & Krause, J. Repeated measures of shoaling tendency in zebrafish (*Danio rerio*) and other small teleost fishes. *Nat. Protoc.* **1**, 1828–1831 (2006).
31. Maximino, C. *et al.* Measuring anxiety in zebrafish: a critical review. *Behav. Brain Res.* **214**, 157–171 (2010).
32. Stewart, A. *et al.* Neurophenotyping of adult zebrafish using the light/dark box paradigm. *Zebrafish Neurobehav. Protoc.* 157–167 (2011).
33. Gerlai, R. High-throughput behavioral screens: the first step towards finding genes involved in vertebrate brain function using zebrafish. *Molecules* **15**, 2609–2622 (2010).
34. Guo, S. Linking genes to brain, behavior and neurological diseases: what can we learn from zebrafish? *Genes Brain and Behav.* **3**, 63–74 (2004).
35. Kalueff, A. V., Stewart, A. M. & Gerlai, R. Zebrafish as an emerging model for studying complex brain disorders. *Trends Pharmacol. Sci.* **35**, 63–75 (2014).
36. Stewart, A. *et al.* Modeling anxiety using adult zebrafish: a conceptual review. *Neuropharmacology* **62**, 135–143 (2012).
37. Cachat, J. M. *et al.* Video-aided analysis of zebrafish locomotion and anxiety-related behavioral responses. In *Zebrafish Neurobehav. Protoc.* 1–14 (Springer, New York, USA, 2011).
38. Ladu, F. *et al.* Live predators, robots, and computer-animated images elicit differential avoidance responses in zebrafish. *Zebrafish* **12**, 205–214 (2015).
39. Maaswinkel, H., Zhu, L. & Weng, W. The immediate and the delayed effects of buspirone on zebrafish (*Danio rerio*) in an open field test: a 3D approach. *Behav. Brain Res.* **234**, 365–374 (2012).
40. Stewart, A. M. *et al.* A novel 3D method of locomotor analysis in adult zebrafish: Implications for automated detection of CNS drug-evoked phenotypes. *J. Neurosci. M* **255**, 66–74 (2015).
41. Macr, S., Mason, G. J. & Würbel, H. Dissociation in the effects of neonatal maternal separations on maternal care and the offspring's hpa and fear responses in rats. *Eur J. Neurosci.* **20**, 1017–1024 (2004).
42. Richter, S. H., Garner, J. P. & Würbel, H. Environmental standardization: cure or cause of poor reproducibility in animal experiments? *Nat. Methods* **6**, 257–261 (2009).
43. Mwaffo, V., Anderson, R. P., Butail, S. & Porfiri, M. A jump persistent turning walker to model zebrafish locomotion. *J. R. Soc. Interface* **12**, 20140884 (2015).
44. Mwaffo, V. & Porfiri, M. Turning rate dynamics of zebrafish exposed to ethanol. *Int. J. Bifurcat. Chaos* **25**, 1540006 (2015).
45. Zienkiewicz, A., Barton, D. A. W., Porfiri, M. & di Bernardo, M. Leadership emergence in a data-driven model of zebrafish shoals with speed modulation. *Eur. Phys. J. Spec. Top.* **224**, 3343–3360 (2015).
46. Zienkiewicz, A., Barton, D. A. W., Porfiri, M. & di Bernardo, M. Data-driven stochastic modelling of zebrafish locomotion. *J. Math. Biol.* **71**, 1081–1105 (2015).
47. Budick, S. A. & O'Malley, D. M. Locomotor repertoire of the larval zebrafish: swimming, turning and prey capture. *J. Exp. Biol.* **203**, 2565–2579 (2000).
48. Müller, U. K. & van Leeuwen, J. L. Swimming of larval zebrafish: ontogeny of body waves and implications for locomotory development. *J. Exp. Biol.* **207**, 853–868 (2004).
49. Jorion, P. On jump processes in the foreign exchange and stock markets. *Rev. Financ. St.* **1**, 427–445 (1988).
50. Tankov, P. *Financial Modelling with Jump Processes* (CRC Press, Boca Raton, FL, 2003).
51. Collignon, B., Séguret, A. & Halloy, J. A stochastic vision-based model inspired by zebrafish collective behaviour in heterogeneous environments. *R. Soc. Open Sci.* **3**, 150473 (2016).
52. Cox, J. C., Ingersoll, J. E. & Ross, S. A. A theory of the term structure of interest rates. *Econometrica* **53**, 385–407 (1985).
53. Gautrais, J. *et al.* Deciphering interactions in moving animal groups. *PLoS Comput. Biol.* **8**, e1002678 (2012).
54. Gautrais, J. *et al.* Analyzing fish movement as a persistent turning walker. *J. Math. Biol.* **58**, 429–445 (2009).
55. Romey, W. L. Individual differences make a difference in the trajectories of simulated schools of fish. *Ecol. Model.* **92**, 65–77 (1996).
56. Löber, J., Ziebert, F. & Aranson, I. S. Collisions of deformable cells lead to collective migration. *Sci. Rep.* **5**, 9172–9172 (2014).
57. Vicsek, T. & Zafeiris, A. Collective motion. *Phys. Rep.* **517**, 71–140 (2012).
58. Cachat, J. *et al.* Measuring behavioral and endocrine responses to novelty stress in adult zebrafish. *Nat. Protoc.* **5**, 1786–1799 (2010).
59. Egan, R. J. *et al.* Understanding behavioral and physiological phenotypes of stress and anxiety in zebrafish. *Behav. Brain Res.* **205**, 38–44 (2009).
60. Rosemberg, D. B. *et al.* Differences in spatio-temporal behavior of zebrafish in the open tank paradigm after a short-period confinement into dark and bright environments. *PLoS ONE* **6**, e19397 (2011).
61. Blaser, R. E. & Rosemberg, D. B. Measures of anxiety in zebrafish (*Danio rerio*): dissociation of black/white preference and novel tank test. *PLoS ONE* **7**, e36931 (2012).
62. Blaser, R. E. & Penalosa, Y. M. Stimuli affecting zebrafish (*Danio rerio*) behavior in the light/dark preference test. *Physiol. Behav.* **104**, 831–837 (2011).
63. Jia, J., Fernandes, Y. & Gerlai, R. Short-term memory in zebrafish (*Danio rerio*). *Behav. Brain Res.* **270**, 29–36 (2014).
64. Wong, K. *et al.* Analyzing habituation responses to novelty in zebrafish (*Danio rerio*). *Behav. Brain Res.* **208**, 450–457 (2010).
65. Butail, S. & Mwaffo, V. & Porfiri, M. Model-free information-theoretic approach to infer leadership in pairs of zebrafish. *Phys. Rev. E* **93**, 042411 (2016).
66. Bartolini, T., Mwaffo, V., Butail, S. & Porfiri, M. Effect of acute ethanol administration on zebrafish tail-beat motion. *Alcohol* **49**, 721–725 (2015).

67. Gerlai, R., Ahmad, F. & Prajapati, S. Differences in acute alcohol-induced behavioral responses among zebrafish populations. *Alcohol. Clin. Exp. Res.* **32**, 1763–1773 (2008).
68. Kitano, H. Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002).
69. Gerlai, R. Animated images in the analysis of zebrafish behavior. *Curr. Zool.* zow077 (2016).
70. Frohnwieser, A., Murray, J. C., Pike, T. W. & Wilkinson, A. Using robots to understand animal cognition. *J. Exp. Anal. Behav.* **105**, 14–22 (2016).
71. Krause, J., Winfield, A. F. T. & Deneubourg, J. Interactive robots in experimental biology. *Trends Ecol. Evol.* **26**, 369–75 (2011).
72. Du, R., Li, Z., Youcef-Toumi, K. & Alvarado, P. V. Robot fish. *Berlin, Heidelberg: Springer.* **10**, 978–3 (2015).
73. DeLellis, P. *et al.* Collective behaviour across animal species. *Sci. Rep.* **4**, 3723 (2014).
74. Kolpas, A., Moehlis, J. & Kevrekidis, I. G. Coarse-grained analysis of stochasticity-induced switching between collective motion states. *Proc. Natl. Acad. Sci.* **104**, 5931–5 (2007).
75. Wang, X. R., Miller, J. M., Lizier, J. T., Prokopenko, M. & Rossi, L. F. Quantifying and tracing information cascades in swarms. *PLoS ONE* **7** (2012).
76. Butail, S., Mwaffo, V. & Porfiri, M. Model-free information-theoretic approach to infer leadership in pairs of zebrafish. *Phys. Rev. E* **93**, 042411 (2016).
77. Van Emmerich, C. Modelling correlation as a stochastic process. *Preprint BUW-AMNA* **6** (2006).
78. Daubechies, I. & Sweldens, W. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.* **4**, 247–269 (1998).
79. Kopman, V., Laut, J., Polverino, G. & Porfiri, M. Closed-loop control of zebrafish response using a bioinspired robotic-fish in a preference test. *J. R. Soc. Interface* **10**, 20120540 (2013).
80. Pedersen, A. R. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scand. J. Statist.* **22**, 55–71 (1995).
81. Grossman, L. *et al.* Characterization of behavioral and endocrine effects of lsd on zebrafish. *Behav. Brain Res.* **214**, 277–284 (2010).
82. Moles, C. G., Mendes, P. & Banga, J. R. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res.* **13**, 2467–2474 (2003).
83. Stewart, A. M., Gaikwad, S., Kyzar, E. & Kalueff, A. V. Understanding spatio-temporal strategies of adult zebrafish exploration in the open field test. *Brain Res.* **1451**, 44–52 (2012).
84. Kalueff, A. V. & Cachat, J. M. *Zebrafish Models in Neurobehavioral Research* (Springer, New York, USA, 2011).
85. Kruskal, W. H. & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Assoc.* **47**, 583–621 (1952).
86. Blaser, R. & Gerlai, R. Behavioral phenotyping in zebrafish: comparison of three behavioral quantification methods. *Behav. Res. Methods* **38**, 456–469 (2006).
87. Neter, J., Kutner, M. H., Nachtsheim, C. J. & Wasserman, W. *Applied Linear Statistical Models* vol. 4 (Irwin, Chicago, USA, 1996).
88. Mathur, P. & Guo, S. Differences of acute versus chronic ethanol exposure on anxiety-like behavioral responses in zebrafish. *Behav. Brain Res.* **219**, 234–239 (2011).
89. Levin, E. D., Bencan, Z. & Cerutti, D. T. Anxiolytic effects of nicotine in zebrafish. *Physiol. Behav.* **90**, 54–58 (2007).
90. Riehl, R. *et al.* Behavioral and physiological effects of acute ketamine exposure in adult zebrafish. *Neurotoxicol. Teratol.* **33**, 658–667 (2011).
91. Rosemberg, D. B. *et al.* Behavioral effects of taurine pretreatment in zebrafish acutely exposed to ethanol. *Neuropharmacology* **63**, 613–623 (2012).
92. Stewart, A. M., Grossman, L., Collier, A. D., Echevarria, D. J. & Kalueff, A. V. Anxiogenic-like effects of chronic nicotine exposure in zebrafish. *Pharmacol. Biochem. Behav.* **139**, 112–120 (2015).
93. Seibt, K. J. *et al.* Antipsychotic drugs prevent the motor hyperactivity induced by psychotomimetic MK-801 in zebrafish (*Danio rerio*). *Behav. Brain Res.* **214**, 417–422 (2010).
94. Gerlai, R., Lee, V. & Blaser, R. Effects of acute and chronic ethanol exposure on the behavior of adult zebrafish (*Danio rerio*). *Pharmacol. Biochem. Behav.* **85**, 752–761 (2006).
95. Cachat, J. *et al.* Unique and potent effects of acute ibogaine on zebrafish: the developing utility of novel aquatic models for hallucinogenic drug research. *Behav. Brain Res.* **236**, 258–269 (2013).
96. Kato, S. *et al.* A computer image processing system for quantification of zebrafish behavior. *J. Neurosci. M* **134**, 1–7 (2004).
97. Ruberto, R., Mwaffo, V., Singh, S., Neri, D. & Porfiri, M. Zebrafish response to a robotic replica in three dimensions. *R. Soc. Open Sci.* **3**, 160505 (2016).
98. Fernandes, Y. & Gerlai, R. Long-term behavioral changes in response to early developmental exposure to ethanol in zebrafish. *Alcohol. Clin. Exp. Res.* **33**, 601–9 (2009).
99. Tran, S. & Gerlai, R. Time-course of behavioural changes induced by ethanol in zebrafish (*Danio rerio*). *Behav. Brain Res.* **252**, 204–213 (2013).
100. Bartolini, T. *et al.* Zebrafish response to 3D printed shoals of conspecifics: the effect of body size. *Bioinspir. Biom.* **11**, 026003 (2016).
101. Ladu, F., Mwaffo, V., Li, J., Macr, S. & Porfiri, M. Acute caffeine administration affects zebrafish response to a robotic stimulus. *Behav. Brain Res.* **289**, 48–54 (2015).
102. Ladu, F., Butail, S., Macr, S. & Porfiri, M. Sociality modulates the effects of ethanol in zebra fish. *Alcohol. Clin. Exp. Res.* **38**, 1–9 (2014).
103. Butail, S., Polverino, G., Phamduy, P., Del Sette, F. & Porfiri, M. Influence of robotic shoal size, configuration, and activity on zebrafish behavior in a free-swimming environment. *Behav. Brain Res.* **275**, 269–280 (2014).
104. Butail, S., Bartolini, T. & Porfiri, M. Collective response of zebrafish shoals to a free-swimming robotic fish. *PLoS ONE* **8**, e71623 (2013).

## Acknowledgements

This work was supported by the National Science Foundation under Grant # CMMI-1433670 and Grant # CMMI-1505832. The authors are thankful to Dr. Mario di Bernardo, Dr. Simone Macrì, and Mr. Tommaso Ruberto for valuable discussion during the preparation of this work.

## Author Contributions

V.M. conceived the model; designed *in-silico* experiments; performed the simulations; analysed the data; and wrote the paper. S.B. analysed the data and wrote the paper. M.P. conceived the model; designed *in-silico* experiments; analysed the data; and wrote the paper.

## Additional Information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Mwaffo, V. *et al.* *In-silico* experiments of zebrafish behaviour: modeling swimming in three dimensions. *Sci. Rep.* **7**, 39877; doi: 10.1038/srep39877 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017