Mathematical modelling of *Gyrodactylus* parasite dynamics: stochastic simulation-based approach and parameter estimation via modified sequential-type approximate Bayesian computation

Clement Twumasi [1, 2, 3*], Owen Jones [1], Joanne Cable [2], Andrey Pepelyshev [1]

[1] *School of Mathematics, Cardiff University, Abacws, Cardiff, CF24 4AG, UK*
[2] *School of Biosciences, Cardiff University, Cardiff, CF10 3AX, UK*
[3] *Department of Statistics, Oxford University, 29 St Giles', Oxford OX1 3LB, UK*

Abstract

Understanding host-parasite systems are challenging if biologists only adopt experimental approaches, whereas mathematical models can help uncover other in-depth knowledge about host infection dynamics. The gyrodactylid-fish system, like other host-parasite systems, is widely used to investigate ecological, evolutionary, and epidemiological problems. They are particularly amenable to experimental manipulation, and there is an individual-based model (IBM) to reproduce their within-host infection dynamics. However, spatial information on species-specific microhabitat preference and other relevant biological information for different *Gyrodactylus* strains across different fish populations are yet to be fully captured in the existing IBM. Therefore, this study aims to develop a novel individual-based stochastic simulation model via a hybrid τ-leaping to include these data to add to our understanding of the gyrodactylid-fish system's complexity. The infection dynamics of three gyrodactylid strains are compared across three host populations. A modified sequential-type approximate Bayesian computation (ABC) is developed for fitting this sophisticated model based on empirical data and an auxiliary stochastic model. A penalised local-linear regression for ABC post-processing analysis is proposed. A linear birth-death process with catastrophic extinction (B-D-C process) is considered the auxiliary model to refine the modified ABC's summary statistics, with other theoretical justifications and parameter estimation techniques of the B-D-C process provided. The B-D-C process simulation using τ-leaping also provided additional insights on accelerating the complex simulation model. The mathematical models can be adapted for other host-parasite systems, and the modified ABC methodologies can aid in efficiently calibrating other multi-parameter models with a high-dimensional set of correlating or independent summary statistics.

Keywords: *Gyrodactylus turnbulli*, *Gyrodactylus bullatarudis*, Individual-based Model, Approximate Bayesian Computation, Continuous-time Markov chains, Tau-leaping

1. Introduction

1.1. Background of the study

Mathematical modelling and simulation have become increasingly important with the advancement of theoretical and applied ecology [5]. Our understanding of biological processes and their interactions in biological systems has increased the development of modern computers and mathematical models, paralleling the need for more quantitative insight into complex biological problems [16]. Broadly, mathematical models can be classified as population-based models (PBMs), individual-based models (IBMs) or hybrid models (which leverage the relative advantages of PBMs and IBMs). PBMs dominate epidemiological and ecological studies compared to IBMs. The former frequently produce basic models, such as systems of ordinary differential or difference equations, that can be analysed mathematically and numerically [18]. In population ecology, the spatial-temporal variations in abundance and distribution of species (including plants, animals or humans) are often explored; whereas changes in the number of individuals of a given species or interacting species over time are usually described using dynamical system modelling (based on differential or difference equations). PBMs thus have been proposed to predict and understand these population dynamics. Broadly, continuous-time, discrete-time, and stochastic models are the three basic types of PBMs applied to problems in population ecological modelling [7]. Even though these PBMs often ignore some underlying biological or ecological realism of the system due to mathematical generalities, they provide a framework through which the occurrence of fundamental or complex relationships and processes of the system can be formulated and explored [13].

IBMs on the other hand, are a population and community modelling approach that allows for a high level of individual and interaction complexity [11]. IBM of population dynamics is a popular approach in current theoretical ecology [5], although it has only been used in a few parasitological studies so far [see 10, 19]. IBMs are typically more complicated than PBMs and are better suited to simulation experiments or *in silico* modelling than statistical analysis [18]. This class of models can simulate either single-species or multiple-species populations of biological systems with both motile and sessile individuals or organisms [5]. The usage of spatially explicit IBM simulation for biological systems necessitates a description of the environment (e.g., host) and each individual (parasite) living in it as well as interactions between individuals and the environment. Adapting computer or *in silico* models to simulate sophisticated biological processes has led to the generation of hypotheses and directions for experimentation [21]. Simulations cannot be considered a substitute for biological experiments in validating a biological hypothesis, but they can help biologists by enhancing experimental design and providing a broader understanding of experimental findings [16]. The *in silico* modelling of biological systems can thus help better understand how infections spread within or between host populations and provide a simple summary of epidemiological data [4]. Apart from making a substantial contribution to spatial pattern formation, IBMs also

give a means of determining the population-level outcomes of specific individual-level behaviour [5]. Computer simulations must be run repeatedly to offer information about average or typical population responses for IBMs to be reliable. These simulation experiments are better suited to answer specific questions about the model or the biological system rather than uncovering whole model behaviour due to the high dimensionality of model parameter space in some instances.

IBMs have however been criticised for lacking the formal framework and analytical procedures accessible in mathematical models made up of differential equations and Markov chains [11]. This is partly due to individual heterogeneity or complex interaction structures that can cause impacts on system dynamics that are difficult to account for using population-based framework [3]. To address this problem, understanding the transition from the most informative individual level to the levels at which system behaviour is typically observed is vital. Thus, a Markov chain approach can help derive and evaluate models on specific levels as well as understand the temporal and spatial patterns that may emerge in that transition [3]. A rigorous investigation of a family of IBMs that specify the dynamics of a complex system at the individual level using the Markov chain approach has been proposed by Banisch [3]. It adopts the lumpability and information theory to link the individual and population levels of observation, providing a basic framework for aggregation in individual-based and related computational models. The starting point is a microscopic Markov chain description of the dynamical process that is completely consistent with the dynamical behaviour of the IBM, which is derived by treating the state space of a large Markov chain as the set of all possible agent configurations [3]. This is known as a micro chain, and using the random mapping representation of a Markov process (defined in [15]), an explicit formal representation incorporating microscopic transition rates may be obtained for a class of models. The circumstances where the macro model is still Markov may be identified using well-known lumpability constraints, and in this case, a complete picture of the dynamics is obtained, including the transient stage, which is the most informative phase in applications. The sort of probability distribution used to construct the stochastic element of the model, which determines the updating process and drives the dynamics, plays a critical role in this regard. The problem of aggregation in IBMs, particularly the lumpability constraints, can be incorporated into a broader framework that employs information theory [defined in 23] to identify different levels and relevant scales in complex dynamical systems.

Izquierdo et al. [14] introduced the possibilities of using time-homogeneous Markov chains in the investigation of IBMs based on computer models (with a well-defined mathematical function written in a programming language, where pseudo-random number generators are used to simulate random variables in the computer models). They argued that when a computer model is analysed as a time-homogeneous Markov chain, many model features not explicit prior to the analysis become apparent. The key concept is to incorporate all possible agent system configurations as the state space of a huge Markov chain. While Izquierdo et al. (2009) relied on numerical computations to estimate the stochastic transition matrices of the models, Banisch [3] showed how to derive

the transition probabilities ($\hat{P}$) explicitly in terms of the update function (u) and a probability distribution (ω) accounting for the stochastic parts of the model. Thus, realisations of IBMs with a sequential update strategy can be thought of as random walks on regular graphs. In population ecology, agent-based models (ABMs) are a class of spatially explicit IBMs , where populations or systems of populations are simulated as being composed of discrete individual organisms [11]. For instance, consider a simple agent-based system defined by a set of N of agents (e.g., hosts or otherwise), where each one is characterised by individual attributes (e.g., physiological features, spatial location and behaviour) from a finite list of possibilities (denoted by S). Suppose $\Sigma = S^N$ is the configuration space representing the set all possible combinations of attributes of agents, and let $\mathbf{x} = (x_1, x_2, \cdots, x_N)$ with $x_i \in S$ for $i = 1, 2, \cdots, N$ denote an agent configuration (where $\mathbf{x}$ is a vector of discrete numbers). At each time step of the time-homogeneous Markov simulation model, the agents' attributes' updating procedure typically consists of two phases. First, a subset of agents is chosen at random using some probability distribution ω, and then, the attributes of the agents are updated according to a rule defined by u, which depends on the subset of agents selected at this time. IBMs can be represented using this specification by the so-called random map representation of Markov chains, under the existence and regularity conditions [17]. A more detailed mathematical theory and applications of Markov chain aggregation for ABMs are presented by Banisch [3], and Izquierdo et al. [14] has demonstrated an in-depth approach of *in silico* modelling using time-homogeneous Markov chains. A wide range of IBMs and techniques, as well as their applications in ecological modelling, have been explored by DeAngelis [8].

The initial goal of modelling host-parasite interactions of any biological system could be to summarise existing knowledge and build a formal representation of the system to make it easier to grasp the underlying complex processes and establish general qualitative assumptions with the help of existing empirical data [9]. Possible model representations of general biological systems include but are not limited to analytical formulations (e.g., deterministic dynamical systems, and stochastic processes) [20, 24, 27], computer-based or simulation-based models (e.g., Monte Carlo simulation algorithms) [1, 6], and graphical models (e.g., social network models, and directed or undirected graphical models) [2, 22]. After model identification and understanding the biological system being modelled, we then determine the relative importance of each of the numerous mechanisms involved in system dynamics (important, secondary, or irrelevant) [9]. A detailed description of the system is then required with explicitly stated assumptions and biologically relevant model parameters. The model can then test biological hypotheses (explaining the model's structure, parameter values, or underlying transition function) by comparing distinct sub-models dynamical behaviour that includes or excludes the hypotheses. Finally, the mathematical model is fitted or calibrated and validated based on the observed empirical data obtained from the underlying biological system [9]. Consequently, we can further forecast future system states based on observed previous states and assumptions about future mechanisms. Although quantitative forecasts are still subject to some

uncertainty following model validation, qualitative forecasts can be offered for several situations (only if past data are unavailable).

Consequently, the current study focuses on modelling the infection dynamics of a gyrodactylid-fish system within at least a standard 17-day experimental period by developing a novel individual-based (hybrid) stochastic simulation model to investigate the infrapopulation dynamics of this biological system. This hybrid simulation model takes advantage of the respective strengths of IBMs and PBMs and is designed to enhance the simulation of this biological system. The infection dynamics of three different parasite strains (with two strains of *Gyrodactylus turnbulli* and one strain of *G. bullatarudis*) are compared across three different fish populations (Ornamental stock, Lower Aripo River, and Upper Aripo River fish) over time. Our sophisticated stochastic model is constructed based on a multi-dimensional continuous-time Markov chain (CTMC) model via a hybrid τ-leaping simulation algorithm. For an individual fish host, the model simulates the movement of parasites for two age groups (young and older parasites) over the external surfaces of the fish within a standard 17-day infection period with population carrying capacity (dependent on host size and area of body regions). The model is parameterised by the birth, death and movement rates of young and older parasites in the presence or absence of the host's immune response. Host death is assumed to occur at a rate proportional to the total number of parasites on the fish. The unique caudal-rostral microhabitat preferences of the two distinct *Gyrodactylus* species, which were mathematically confirmed in our recent related study [25, in press]), are included in the stochastic model. The novel simulation model developed for the *Gyrodactylus*-fish system would provide a relatively realistic imitation of this biological system. Thus, it would facilitate experimental data collection and aid in investigating specific research questions and the system's complexity that may be difficult to control and implement experimentally.

Before developing our novel individual-based simulation model in the current study with the help of existing empirical data, we also relied on mathematical and biological insights of the gyrodactylid-fish system from our recent related study [25, in press]. The former study explored the spatial and temporal parasite dynamics of the two distinct co-infecting gyrodactylids (*Gyrodactylus turnbulli* and *G. bullatarudis*) by adopting a time-inhomogeneous multi-state Markov model (MSM) and other advanced statistical analyses. We answered three open biological questions concerning parasite microhabitat preferences, host survival, and parasite virulence over time based on the empirical data. The preference for parasites to move forward or backwards and the effective population carrying capacity are additional parameters that were estimated. Thus, the simulation model can provide relevant information about parasite numbers at different body locations of fish over time for a fish group based on specific demographic characteristics (such as the parasite strain, fish type, fish sex and fish size) and the underlying model parameters. The fish survival status and exact time to fish mortality are other essential outputs of the simulation model. To fit our novel stochastic simulation model in the current study, we propose a modified likelihood-

free parameter estimation methodology for complex model calibration via approximate Bayesian computation (ABC). In addition, a robust approach for ABC post-processing analysis with $L2$ regularisation is developed to correct for the imperfect mismatch between simulated and observed data. The modified ABC algorithm (with its pseudo-codes outlined in section 4, and dubbed in the current study as weighted-iterative ABC) is based on sequential Monte Carlo (SMC), sequential importance sampling (SIS), and ABC summary statistics weighting (adaptively computed based on the harmonic mean of previous and current summary statistics weights). A continuous-time Markov process dubbed the linear birth-death process with catastrophic extinction (B-D-C process) is further investigated and considered an auxiliary stochastic model (for the complex simulation model) to refine the ABC's summary statistics based on the B-D-C model parameter estimates, with other theoretical justifications and parameter estimation techniques of the B-D-C process provided (section 3). Finally, by adopting a new Bayesian hypothesis testing framework (whose decision rule integrates estimated credible intervals and a region of practical equivalence), we provide answers to other specific research questions of biological relevance (based on the adjusted posterior samples).

1.2. Paper structure

This study is structured in six main sections. The paper is organised as follows.

1.3. Contributions of the study

This interdisciplinary research work has made novel contributions to the gyrodactylid-fish system and the mathematical or ecological modelling community. Additionally, it has resulted in new findings and provided many directions for future studies of the gyrodactylid-fish system. It has also motivated future research regarding between-host parasitic transmission and mixed gyrodactylid infection for this host-parasite system, amongst others. Specifically, the current study has made the following contributions. The novel stochastic simulation developed in this study depends on model assumptions with some biological realism (informed by empirical data) about the gyrodactylid-fish system compared to the very useful existing IBM [26]. The existing IBM can provide a valuable tool for forecasting the development of gyrodactylid infections on single hosts and predicting optimal life history strategies of parasites [26]. However, within a realistic setting, the time to host immune response can occur at any time after infection, and localised immune response may depend on host and parasite genotype, the surface area of the body locations and host sex.

When examining the infrapopulation dynamics of gyrodactylids on their fish host, the existing IBM is yet to distinguish between fish's major body locations (tail fin, lower body, upper body, anal fin, dorsal fin, pelvic fins, pectoral fins, and head) including their respective surface areas, as well as to differentiate between young and old parasites and to loosen the restrictions on the maximum linear distance that parasites can move over time. For example, as individual host infections with *G. turnbulli* progress, parasites migrate from the caudal fin and body to the pectoral, pelvic, dorsal

and anal fins; a migration to potentially facilitate transmission [12]. There exist unique host-parasite strain-specific microhabitat preferences for different gyrodactylid strains across different host populations over time [as confirmed in 25, in press]. Thus, this spatial information must be incorporated when simulating the species-specific infrapopulation dynamics. Hence, the need for a more robust simulation model in the current study to include such relevant spatial information and specific demographic characteristics of the infecting parasite strain and the host (such as parasite strain, fish type, sex, size and survival status).

Based on the new epidemiological insights into the gyrodactylid-fish system by analysing empirical data in our recent related study [25, in press], the sophisticated (individual-based) stochastic simulation model is able to include data concerning parasite fecundity, age group (young or older parasites), parasite mortality, parasite mobility, and host immune response to understand the gyrodactylid-fish system's complexity better. The development of the modified ABC sampler (with SMC and SIS), dubbed weighted-iterative ABC, coupled with the proposed penalised ABC post-processing methodology (with $L2$ regularisation) can aid in a robust parameter estimation of both sophisticated and simple likelihood-free models sequentially across a whole population. Our robust ABC post-processing method (for adjusting the resulting ABC posterior and estimating its mean) is an extension of the standard ABC local-linear regression (to include an $L2$ regularisation term). It is considered an independent ABC final step after executing the weighted-iterative ABC to fit the novel simulation model. Unlike the standard ABC post-analysis methods, our proposed ABC posterior correction method is implementable even if the set of ABC summary statistics (possibly high-dimensional) is highly correlated in the neighbourhood of the observed summaries. In addition, it improves the standard ABC local-linear regression (with heteroscedastic errors) in the presence of multicollinearity, supercollinearity, outliers, or non-normal regression residuals.

The study has also demonstrated the use of an auxiliary model (for the developed complex individual-based stochastic simulation model), dubbed the linear birth-death process with catastrophic extinction (B-D-C process), to aid in refining our ABC summary statistics based on its parameter estimates. For the first time in this study, the exact analytical results of the B-D-C transition function and its theoretical moments are derived and numerically validated, as in the setting of discretely observed processes. Before ABC fitting of the novel stochastic simulation model, three different parameter estimation methods: maximum likelihood estimation (MLE), generalised method of moments (GMM), and embedded Galton-Watson (GW) estimation methods for the B-D-C model were developed and compared by exploring the trade-off between estimation accuracy (quantified by the estimation bias, variance, and mean square error) and computational speed based on different *in silico* simulation experiments (where parasite population size is large, moderate, or low). The two zero states of the B-D-C process (due to either the natural death of parasites or parasite population extinction after host mortality) were distinguished or separately set up in the aforementioned B-D-C parameter estimators (MLE, GMM, and GW).

Furthermore, we adapted and compared two hybrid τ-leaping algorithms to simulate the B-D-C process and identified which method is cost-effective (based on their respective simulation speed and fidelity). The simulation of the B-D-C process using a hybrid τ-leaping algorithm also provided additional insights on accelerating the complex stochastic simulation model (based on its derived leap size estimator) by proposing a reasonable error threshold based on the trade-off between simulation accuracy and computational speed. Prior to fitting our novel stochastic simulation model, the fidelity of our proposed ABC methodologies was numerically assessed at different proposal draws based on a simple modelling problem (with multivariate normal likelihood and known analytical posterior distribution).

Here, we investigated whether the resulting ABC approximation is mutually compatible with any Monte Carlo sample size ($N \geq 500$) or independent of N by determining the minimal number of proposal draws to achieve a good posterior estimation. The high computational cost of simulating data from the sophisticated individual-based simulation model, computing some components of the set of multidimensional ABC summary statistics (such as the B-D-C model parameters across a whole host population), and the quadratic cost of implementing ABC-SMC methods motivated this further exploration. Our ABC methodologies could also be modified and utilised in other studies to fit complex mathematical and simulation models

The findings from the current study can inform management decisions about the control of not only the gyrodactylid-fish system, but other problematic aquaculture diseases. The mathematical models developed in this study can be adapted for future predictions within and beyond the standard 17-day infection period for a particular *Gyrodactylus* strain across the different fish populations. In addition, the individual-based stochastic model can also be further modified to conduct biological experiments for mixed gyrodactylid parasite populations and can be expanded further for broader host-parasite systems.

All formulated mathematical theorems under **??????** of the current study are derived and proved for the first time (and thus not previously proposed in any other study). All the proposed algorithms, Algorithms 1-6 (with their respective pseudo-codes) under **??????**, are also developed for the first time in this study. All lemmas and Algorithms L1-L4 are presented in this current study based on previously published studies (with appropriate references provided accordingly). The first main work in this interdisciplinary research (corresponding to **??**) has been submitted to a high-impact peer-reviewed biological journal. Results from **??????** will be published in mathematics peer-reviewed journals. All R codes developed for statistical analyses and mathematical modelling (including their Jupyter Notebook HTML and source files) as well as the empirical data (for this study) can be found via the GitHub URL link (for reproducibility of results): github.com/twumasiclement/In-Silico-Modelling-of-Parasite-Dynamics.

Availability of data and materials

The datasets used in the current study are available from the corresponding author on reasonable request. The R codes used for all statistical analyses have been made publicly available for reproducibility of results via the URL link: https://github.com/twumasiclement/In-Silico-Modelling-of-Parasite-Dynamics.

Author Contribution

The authors conceived and designed the study and approved the final manuscript. OJ, JC, and AP supervised and reviewed the work prior to submission. CT and OJ developed the mathematical methodologies. JC provided the empirical data used in the study. CT performed all statistical analyses and wrote the article (for all accompanying R codes, see the URL link: https://github.com/twumasiclement/In-Silico-Modelling-of-Parasite-Dynamics).

Conflicts of Interest

All authors declare that there are no conflicts of interest.

Financial support

This study was funded by a Cardiff University Vice Chancellor's international scholarship scheme for research excellence to CT.

Ethical Standards

Not applicable.

References

[1] Asmussen, S. and Glynn, P. W. (2007). *Stochastic simulation: algorithms and analysis*, volume 57. Springer Science & Business Media.

[2] Ball, F., Britton, T., and Sirl, D. (2013). A network with tunable clustering, degree correlation and degree distribution, and an epidemic thereon. *Journal of Mathematical Biology*, 66(4):979–1019.

[3] Banisch, S. (2015). *Markov chain aggregation for agent-based models*. Springer.

[4] Becker, N. (1979). The uses of epidemic models. *Biometrics*, pages 295–305.

[5] Berec, L. (2002). Techniques of spatially explicit individual-based models: construction, simulation, and mean-field analysis. *Ecological modelling*, 150(1-2):55–81.

[6] Black, A. J. and McKane, A. J. (2012). Stochastic formulation of ecological models and their applications. *Trends in ecology & evolution*, 27(6):337–345.

[7] Bonsall, M. B. (2009). Population models. *Ecology-Volume II*, page 235.

[8] DeAngelis, D. L. (2018). *Individual-based models and approaches in ecology: populations, communities and ecosystems*. CRC Press.

[9] Ezanno, P., Vergu, E., Langlais, M., and Gilot-Fromont, E. (2012). Modelling the dynamics of host-parasite interactions: basic principles. In *New Frontiers of Molecular Epidemiology of Infectious Diseases*, pages 79–101. Springer.

[10] Gaba, S., Cabaret, J., Ginot, V., and Silvestre, A. (2006). The early drug selection of nematodes to anthelmintics: stochastic transmission and population in refuge. *Parasitology*, 133(3):345–356.

[11] Grimm, V. and Railsback, S. F. (2013). *Individual-based modeling and ecology*. Princeton university press.

[12] Harris, P. D. (1988). Changes in the site specificity of *Gyrodactylus turnbulli* Harris, 1986 (Monogenea) during infections of individual guppies (*Poecilia reticulata* Peters, 1859). *Canadian Journal of Zoology*, 66(12):2854–2857.

[13] Harvey, J. A., Malcicka, M., and Ellers, J. (2015). Integrating more biological and ecological realism into studies of multitrophic interactions. *Ecological Entomology*, 40(4):349–352.

[14] Izquierdo, L. R., Izquierdo, S. S., Galan, J. M., and Santos, J. I. (2009). Techniques to understand computer simulations: Markov chain analysis. *Journal of Artificial Societies and Social Simulation*, 12(1):6.

[15] Jost, J., Kell, M., and Rodrigues, C. S. (2015). Representation of Markov chains by random maps: existence and regularity conditions. *Calculus of Variations and Partial Differential Equations*, 54(3):2637–2655.

[16] Kaazempur-Mofrad, M., Bathe, M., Karcher, H., Younis, H., Seong, H., Shim, E., Chan, R., Hinton, D., Isasi, A., Upadhyaya, A., et al. (2003). Role of simulation in understanding biological systems. *Computers & structures*, 81(8-11):715–726.

[17] Levin, D. A. and Peres, Y. (2017). *Markov chains and mixing times*, volume 107. American Mathematical Society.

[18] Lloyd, A. (2007). Introduction to epidemiological modeling: basic models and their properties. *Networks*, pages 1–166.

[19] Louie, K., Vlassoff, A., and Mackay, A. (2007). Gastrointestinal nematode parasites of sheep: a dynamic model for their effect on liveweight gain. *International journal for Parasitology*, 37(2):233–241.

[20] Morris, A., Börger, L., and Crooks, E. (2019). Individual variability in dispersal and invasion speed. *Mathematics*, 7(9):795.

[21] Motta, S. and Pappalardo, F. (2013). Mathematical modeling of biological systems. *Briefings in Bioinformatics*, 14(4):411–422.

[22] Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V., and Trapman, P. (2015). Eight challenges for network epidemic models. *Epidemics*, 10:58–62.

[23] Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review*, 5(1):3–55.

[24] Twumasi, C., Asiedu, L., and Nortey, E. N. (2019). Markov chain modeling of HIV, tuberculosis, and Hepatitis B Transmission in Ghana. *Interdisciplinary perspectives on infectious diseases*, 2019.

[25] Twumasi, C., Cable, J., and Jones, O. (in press 2022). Spatial and temporal parasite dynamics: microhabitat preferences and infection progression of two co-infecting gyrodactylids. Preprint available at: *https://doi.org/10.13140/RG.2.2.33220.71044*.

[26] van Oosterhout, C., Potter, R., Wright, H., and Cable, J. (2008). Gyro-scope: An individual-based computer model to forecast gyrodactylid infections on fish hosts. *International Journal for Parasitology*, 38(5):541–548.

[27] Zipkin, E. F., Jennelle, C. S., and Cooch, E. G. (2010). A primer on the application of Markov chains to the study of wildlife disease dynamics. *Methods in Ecology and Evolution*, 1(2):192–198.

Appendices

Appendix 1: Number of surviving fish for the nine different host-parasite groups over time from days 1 to 17.