Supplement to "Mechanistic models for panel data: analysis of ecological experiments with four interacting species"

Bo Yang
Department of Biostatistics, University of Michigan
Jesse Wheeler
Department of Statistics, University of Michigan
Aaron A. King
Department of Ecology and Evolutionary Biology &
Center for the Study of Complex Systems, University of Michigan
Edward L. Ionides
Department of Statistics, University of Michigan

April 3, 2024

Contents

1 Model Diagram

Dynamic models addressed in this manuscript are comprehensively delineated through the descriptions provided within the main text, augmented by the supplementary details furnished in this section. Notwithstanding their complete mathematical specification, visual diagrams of dynamic systems frequently facilitate comprehension of the underlying equations. Hence, this section presents a model diagram, which corresponds to the SF model. It is imperative to recognize that, given the models are articulated through their mathematical formulations and numerical implementations, these diagrams do not constitute unique visual depictions. There exist alternative visualizations that might further aid in elucidating the models discussed herein.

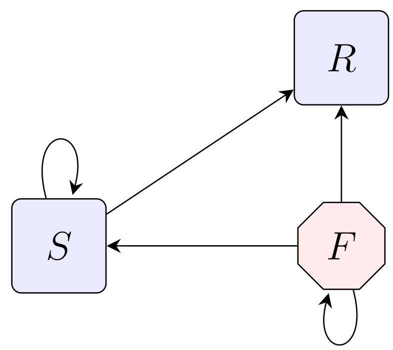

Figure 1: A flow diagram for the SF model. In this system, either species of *Daphnia* and the food resource contribute to reproduction, depicted as self-recycling loops for S, and F. As time progresses, individuals in F, S transition to the R state, signifying mortality.

2 Markov chain and differential equation interpretations in the dynamics

2.1 SF Model

The dynamics of the system are governed by a set of differential equations that describe the interactions between different species and their environment. The model consists of equations that represent the growth and interactions of *algae* (denoted as F) and two types of *Daphnia* populations, S_n and S_i, under certain conditions.

$$dS(t) = rS(t)dt - \theta_S S(t)dt - \delta S(t)dt + S(t)d\zeta_S \tag{1}$$

$$dF(t) = \alpha F(t)\left(1 - \frac{F(t)}{K_F}\right) \cdot dt - \gamma_n F(t)S(t) \cdot dt - \delta F(t) \cdot dt + \mu \cdot dt + F(t)d\zeta_F \tag{2}$$

$$\zeta_F \sim N\left(0, \sigma_F^2 dt\right) \tag{3}$$

$$\zeta_S \sim N\left(0, \sigma_S^2 dt\right) \tag{4}$$

This model includes parameters such as γ, α, σ_F, σ_S, θ, β, δ, and k_f, which are integral to understanding the dynamics between *algae* and *Daphnia* populations. The parameters α and K_F represent the growth rate and carrying capacity of *algae* , respectively, while β and θ describe the consumption rate of *algae* by Daphnia and the growth rate of *Daphnia*, respectively. The death rate of *Daphnia* is denoted by γ. The terms ζ_F and ζ_S introduce randomness into the growth processes of *algae* and *Daphnia*, modeling the stochastic nature of the ecosystem. In this model, we set the initial condition to be $M = 0, S = 3$ and $F = 16.667$ by Searles et al. For ease of understanding the model equations, we separate the differential equations (12)-(16) into distinct parts that each correspond to meaningful ecological phenomena:

Alga The change of the density of alga can be viewed as

$$dF_{birth}(t) = \alpha F(t)\left(1 - \frac{F(t)}{k_f}\right)dt + F(t)d\zeta_F \tag{5}$$

$$dF_{cons}(t) = \beta F(t)S(t)dt \tag{6}$$

$$dF_{spl}(t) = \delta F(t)dt \tag{7}$$

$$dF_{refill}(t) = \mu dt \tag{8}$$

$$dF(t) = dF_{birth}(t) - dF_{cons}(t) - dF_{spl}(t) + dF_{refill}(t) \tag{9}$$

In these equations, α represents the growth rate; k_f is the carrying capacity; ζ_F is the noise term; and β represents the rate at which alga are consumed. All of these are treated as unknown parameters. δ and μ are fixed parameters with $\delta = 0.013$ and $\mu = 0.37$, which model the food supply rate. We split the change of alga density (dF) into four parts: equation (17) represents the change of population due to the birth of alga and Brownian noise term; equation (18) is the amount of alga that is consumed by *Daphnia*; equations (19) and (20) indicate that change of alga density due to the process of alga refill and sampling. Together, these four equations explain the processes and causes of the changes in alga population density.

Daphnia The change of the density of *Daphnia* can be viewed as

$$dS_{grow}(t) = \theta S(t)F(t)dt + S(t)d\zeta_S \tag{10}$$
$$dS_{death}(t) = \gamma S(t)dt \tag{11}$$
$$dS_{spl}(t) = \delta S(t)dt \tag{12}$$
$$dS(t) = S_{grow}(t) - S_{death}(t) - S_{spl}(t) \tag{13}$$

In these equations, θ is the growth rate, ζ_S is the noise term, and γ is the death rate, all of which are unknown parameters. δ is a fixed parameter with $\delta = 0.013$. We split the change of *Daphnia* density (dS) into three parts. The equation (22) represents the change of population due to the birth of *Daphnia* and Brownian noise terms. While the equation (23) shows the amount of alga that is consumed by *Daphnia*, which will make the density decrease. The equation (24) indicates that change of *Daphnia* density due to the process of sampling. These three equations explain the processes and causes of the changes in *Daphnia* population density.

This model includes 9 parameters: $\gamma, \alpha, \sigma_F, \sigma_S, \theta, \beta, \delta, \mu$ and k_f. Among those parameters, α represents the rate of growth of algae. k_f shows the carrying capacity of algae in each bucket. β is the consumption rate of algae by certain species of *Daphnia*, θ shows the rate of growth of *Daphnia* with certain types of food and γ is the rate of death of certain species of *Daphnia*. Similar to the previous Lotka-Volterra model, we treat δ as a constant equal to 0.013 and $\mu = 0.37$ as a constant that implies the food refilling rate according to the experimental settings. ζ_F and ζ_S are Brownian noise terms, which show the randomness of the dynamic system. They are induced by a normal distribution with 0 mean,and $\sigma_F^2 \cdot dt, \sigma_S^2 \cdot dt$ to be variance respectively. Unlike the previous model, this model restricts the total food supply, more closely mimicking the experimental setting. One explanation for the observed precipitous decline in *Daphnia* population density is that there was a meaningful decline to the food supply throughout the experiment. Therefore the model equations permit both a natural decline in food levels and an increase of *Daphnia* food corresponding to the food supply rate of the experiment. Based on the experimental manipulations, we treated the food supply rate as a constant $\mu = 0.37 \cdot 10^5$ cells per day.

Using this model, we can investigate how the population dynamics of *Daphnia* are influenced by the density of algae with the condition that food levels without any food management may not be sufficient to describe *Daphnia* population dynamics. If the data are better explained using this model than the alternative hypotheses, it can be concluded that the density of algae plays a significant role in *Daphnia* population dynamics and the lack of food is potentially the reason of the precipitous decline in *Daphnia* population, as other factors may not obviously affect this treatment. Alternatively, if the model does not fit the data well, it's possible that the alga may not be the main reason that led to the change of population of *Daphnia*, which means other possible factors that may affect the population density of *Daphnia* should be considered.

Incorporating stochastic elements (ζ) into the model adheres to stochastic differential equation (SDE) frameworks, offering a refined perspective on population dynamics that includes the effect of random environmental variations. These components, depicted as

Gaussian white noise, add complexity to the model by portraying unpredictable external and internal influences on the populations. The variances σ_S^2 and σ_F^2 serve to measure the intensity of stochastic disturbances on both susceptible and infected population groups, alongside their alimentary resources. This methodology surpasses traditional deterministic models, providing a comprehensive and authentic analysis of ecological system dynamics.

2.2 SF2 Model

The following SDEs further explores the dynamics between two types of *Daphnia* populations (S_n and S_i) and their food source, *algae* (F), incorporating stochastic elements to reflect the randomness in natural processes:

$$dS_n(t) = r_n S_n(t)dt - \theta_{S_n} S_n(t)dt - \delta S_n(t)dt + S_n(t)d\zeta_{S_n} \tag{14}$$

$$dS_i(t) = r_i S_i(t)dt - \theta_{S_i} S_i(t)dt - \delta S_i(t)dt + S_i(t)d\zeta_{S_i} \tag{15}$$

$$\begin{aligned} dF(t) = \alpha F(t)\left(1 - \frac{F(t)}{K_F}\right) \cdot dt - \gamma_n F(t) S_n(t) \cdot dt - \gamma_i F(t) S_i(t) \cdot dt \\ - \delta F(t) \cdot dt + \mu \cdot dt + F(t) d\zeta_F \end{aligned} \tag{16}$$

$$\zeta_F \sim N\left(0, \sigma_F^2 dt\right) \tag{17}$$

$$\zeta_{S_n} \sim N\left(0, \sigma_{S_n}^2 dt\right) \tag{18}$$

$$\zeta_{S_i} \sim N\left(0, \sigma_{S_i}^2 dt\right) \tag{19}$$

These equations delineate the dynamics of a mixed *Daphnia* population, incorporating growth rates (r_n, r_i), the impact of *algae* consumption (θ_{S_n}, θ_{S_i}), and mortality rates (δ). The stochastic components (ζ_{S_n}, ζ_{S_i}, and ζ_F) introduce environmental variability into the model, rendering a more nuanced depiction of ecosystem dynamics. Such analysis illuminates the intricate relationships between species and their habitats, enhancing our comprehension of ecological complexities.

The incorporation of stochastic terms (ζ) aligns with the principles of stochastic differential equations (SDEs), presenting a nuanced view of population dynamics that accommodates for random environmental fluctuations. These elements, represented as Gaussian white noise, enrich the model by simulating unpredictable environmental and internal processes that impact the populations. The specified variances—$\sigma_{S_n}^2, \sigma_{S_n}^2$ and σ_F^2—quantify the magnitude of noise affecting both the susceptible and infected populations, as well as their food sources. This approach transcends the limitations of deterministic models, offering a layered and realistic exploration of the forces shaping ecological systems.

2.3 SIRPF Model

The SIRPF model presented encapsulates the dynamics between susceptible (S) and infected (I) populations, their food source (F), and a parasite population (P). Each differential equation accounts for various biological and ecological processes, augmented with stochastic terms to capture the inherent randomness of these systems.

$$dS(t) = rF(t)\left(S(t) + \xi I(t)\right)dt - \theta_S S(t)dt - pf_S S(t)Pdt - \delta S(t)dt + S(t)d\zeta_S \tag{20}$$
$$dI(t) = pf_S S(t)Pdt - \theta_I I(t)dt - \delta I(t)dt + I(t)d\zeta_I \tag{21}$$
$$dF(t) = \alpha F(t)\left(1 - \frac{F(t)}{K_F}\right)dt - \gamma F(t)\left(S(t) + \xi I(t)\right)dt - \delta F(t)dt + \mu dt + F(t)d\zeta_F \tag{22}$$
$$dP(t) = \beta\theta_I I(t)dt - f_S S(t)P(t)dt - f_I I(t)P(t)dt - \theta_{\mathrm{p}} P(t)dt - \delta P(t)dt + P(t)d\zeta_P \tag{23}$$
$$d\zeta_S \sim N\left(0, \sigma_S^2 dt\right) \tag{24}$$
$$d\zeta_I \sim N\left(0, \sigma_I^2 dt\right) \tag{25}$$
$$d\zeta_F \sim N\left(0, \sigma_F^2 dt\right) \tag{26}$$
$$d\zeta_P \sim N\left(0, \sigma_P^2 dt\right) \tag{27}$$

This model intricately details the interactions and feedback mechanisms between the populations under study. The equations describe the growth or decline of each population, modulated by both deterministic factors, such as reproduction and predation rates, and stochastic factors, captured through the $d\zeta$ terms representing environmental variability and other unpredictable influences. The inclusion of stochastic terms (ζ) adheres to the principles of stochastic differential equations (SDEs), offering a realistic portrayal of population dynamics by accounting for random fluctuations. These terms, modeled as Gaussian white noise, add depth to the analysis by allowing for the simulation of random processes affecting the populations, thereby extending beyond the deterministic skeleton of classical models. The variances $\sigma_S^2, \sigma_I^2, \sigma_F^2$, and σ_P^2 encapsulate the intensity of environmental and internal process noise affecting susceptible individuals, infected individuals, food resources, and parasite, respectively.

This model's capacity to intertwine biological realism with statistical rigor makes it a potent tool for understanding complex ecological interactions. By examining the system's sensitivity to parameter variations and the stochastic components' impact. Furthermore, statistical analysis of model outcomes, such as likelihood-based methods for parameter estimation or Bayesian inference for uncertainty quantification, can provide deeper understanding and predictive power, essential for conservation efforts and the management of biological resources.

2.4 SIRPF2 Model

The SIRPF2 model outlined below incorporates a detailed analysis of the interactions between two *Daphnia* populations (denoted as S_n and S_i), their infected counterparts (I_n and I_i), algae (F), and parasites (P). This model aims to capture the nuanced dynamics of ecological and epidemiological processes within a shared environment.

$$dS_n(t) = r_n F(t) \left(S_n(t) + \xi I_n(t)\right) dt - \theta_{S_n} S_n(t)dt - p_n f_{S_n} S_n(t)P(t)dt - \delta S_n(t)dt + S_n(t)d\zeta_{S_n} \tag{28}$$

$$dI_n(t) = p_n f_{S_n} S_n(t)Pdt - \theta_{I_n} I_n(t)dt - \delta I_n(t)dt + I_n d\zeta_{I_n} \tag{29}$$

$$dS_i(t) = r_i F(t) \left(S_i(t) + \xi I_i(t)\right) dt - \theta_{S_i} S_i(t)dt - p_i f_{S_i} S_i(t)P(t)dt - \delta S_i(t)dt + S_i(t)d\zeta_{S_i} \tag{30}$$

$$dI_i(t) = p_i f_{S_i} S_i(t)P(t)dt - \theta_{I_i} I_i(t)dt - \delta I_i(t)dt + I_i(t)d\zeta_{I_i}(t) \tag{31}$$

$$\begin{aligned} dF(t) = {} & \alpha F(t) \left(1 - \frac{F(t)}{K_F}\right) dt - \gamma_n F(t) \left(S_n(t) + \xi I_n(t)\right) dt - \gamma_i F(t) \left(S_i(t) + \xi I_i(t)\right) dt \\ & - \delta F(t)dt + \mu dt + F(t)d\zeta_F \end{aligned} \tag{32}$$

$$\begin{aligned} dP(t) = {} & \beta_n \theta_{I_n} I_n(t)dt + \beta_i \theta_{I_i} I_i(t)dt - f_{S_n} S_n(t)P(t)dt - f_{I_n} I_n(t)P(t)dt - f_{S_i} S_i(t)P(t)dt \\ & - \mathrm{f}_{\mathrm{I}_i} I_i(t)P(t)dt - \theta_{\mathrm{p}} P(t)dt - \delta P(t)dt + P(t)d\zeta_P \end{aligned} \tag{33}$$

$$d\zeta_{S_n} \sim N\left(0, \sigma_{S_n}^2 dt\right) \tag{34}$$

$$d\zeta_{I_n} \sim N\left(0, \sigma_{I_n}^2 dt\right) \tag{35}$$

$$d\zeta_{S_i} \sim N\left(0, \sigma_{S_i}^2 dt\right) \tag{36}$$

$$d\zeta_{I_i} \sim N\left(0, \sigma_{I_i}^2 dt\right) \tag{37}$$

$$d\zeta_F \sim N\left(0, \sigma_F^2 dt\right) \tag{38}$$

$$d\zeta_P \sim N\left(0, \sigma_P^2 dt\right) \tag{39}$$

This model is characterized by a series of stochastic differential equations that represent the temporal dynamics of each population and their interactions. The terms r_n and r_i denote the intrinsic growth rates of the non-infected *Daphnia* populations, while θ_{S_n}, θ_{S_i}, θ_{I_n}, and θ_{I_i} represent the respective mortality rates due to natural causes other than disease. The interaction between *Daphnia* populations and the algae is modeled by the consumption rates γ_n and γ_i, which also affect the *algae*'s growth dynamics along with its intrinsic growth rate α and carrying capacity K_F.

In the presented SDE model, the dynamics of susceptible and infected *Daphnia* populations, alongside their *algae* food source and parasites, are intricately modeled to incorporate both deterministic biological interactions and stochastic environmental variability. The inclusion of stochastic terms, denoted by $d\zeta$, adheres to a normal distribution with zero mean and variance proportional to dt, effectively embedding white noise into the system. This statistical approach allows the model to capture the inherent unpredictability of ecological processes, offering a nuanced representation that transcends traditional deterministic models. The SDE framework facilitates the exploration of population dynamics under random environmental influences, emphasizing the significance of stochasticity in ecological interactions.

3 Simulations

For the model and parameters swarm that result in the best AIC, we did 20 simulations on each unit of the panel dataViewing the figure, almost all simulations capture the data well. And the estimation of parameters swarm used for this simulation is shown in section Parameter.

3.1 SF Models

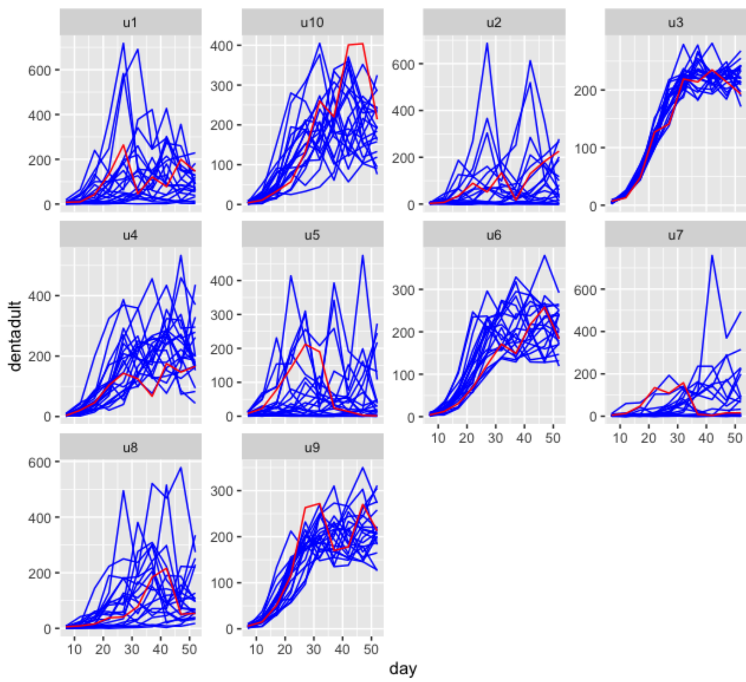

Figure 2: Simulation of Daphnia dentifera of SF dynamics

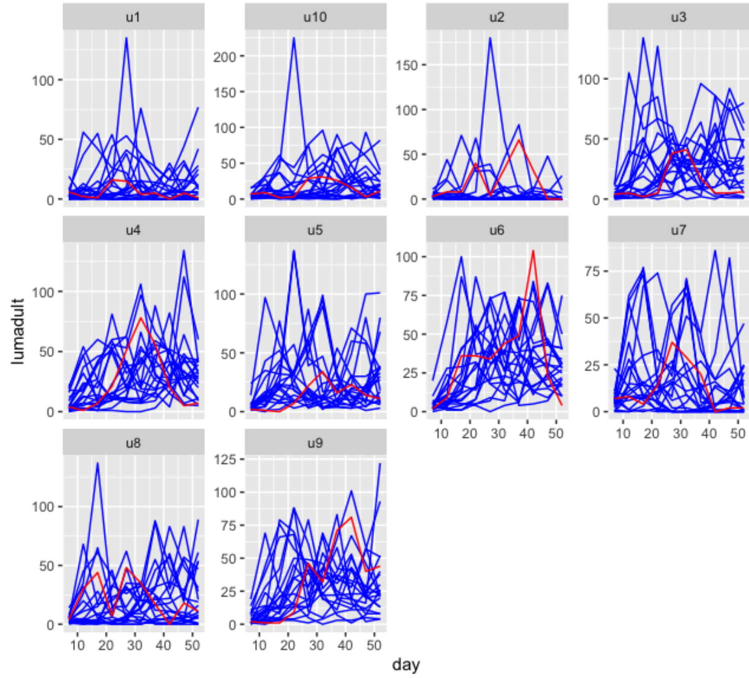

Figure 3: Simulation of Daphnia lumholtzi of SF dynamics

3.2 SF2 Models

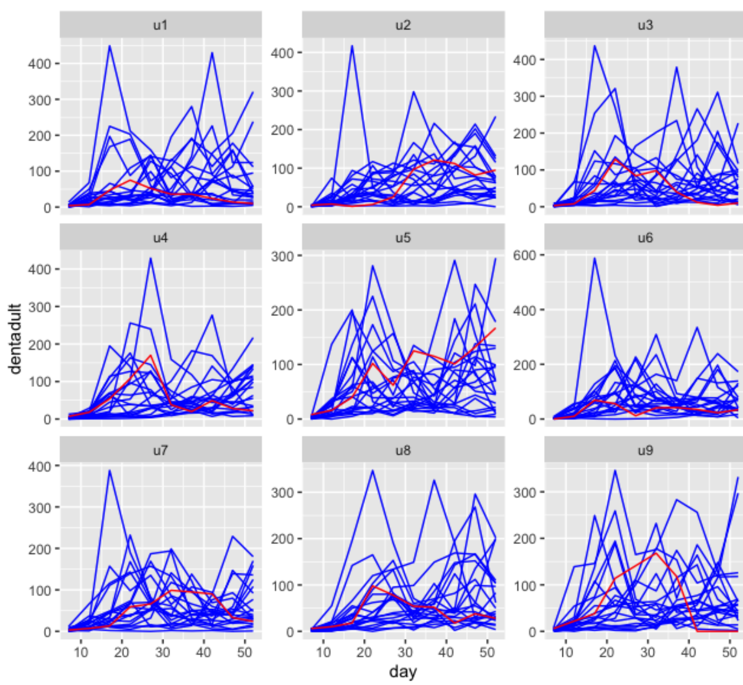

Figure 4: Simulation of Daphnia dentifera of SF2 dynamics

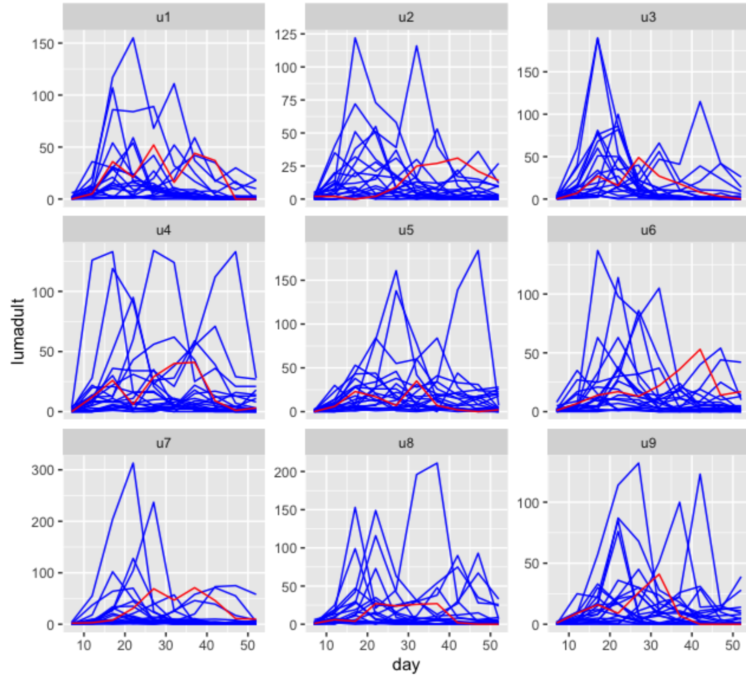

Figure 5: Simulation of Daphnia lumholtzi of SF2 dynamics

3.3 SIRPF Models

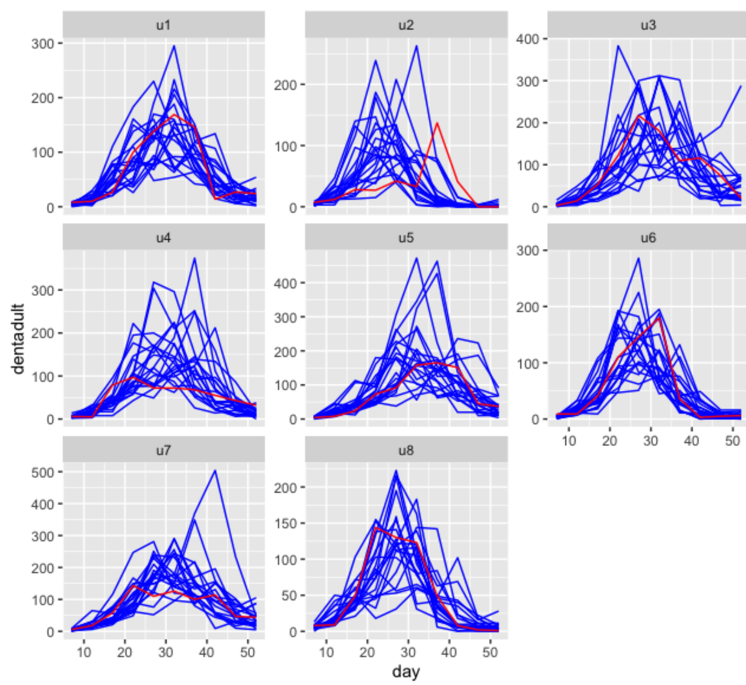

Figure 6: Simulation of Susceptible Daphnia dentifera of SIRPF dynamics

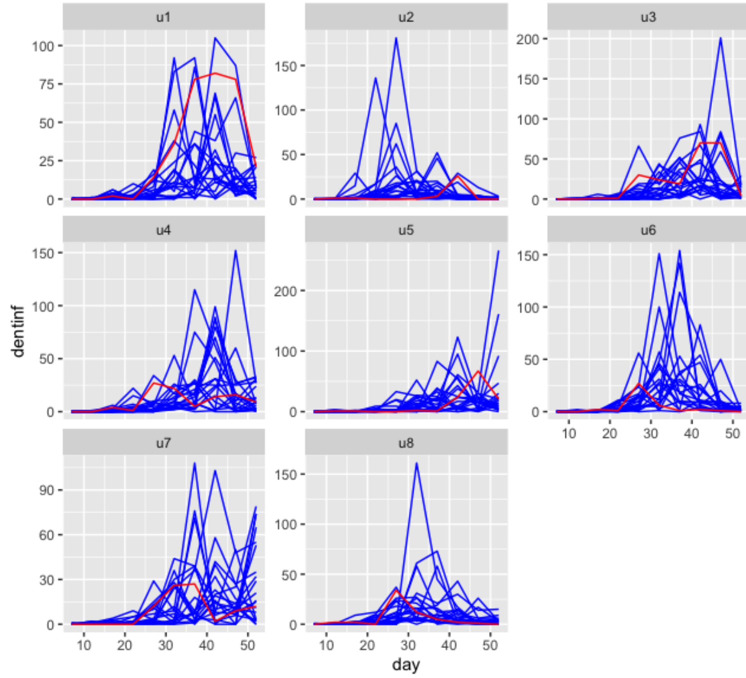

Figure 7: Simulation of Infected Daphnia dentifera of SIRPF dynamics

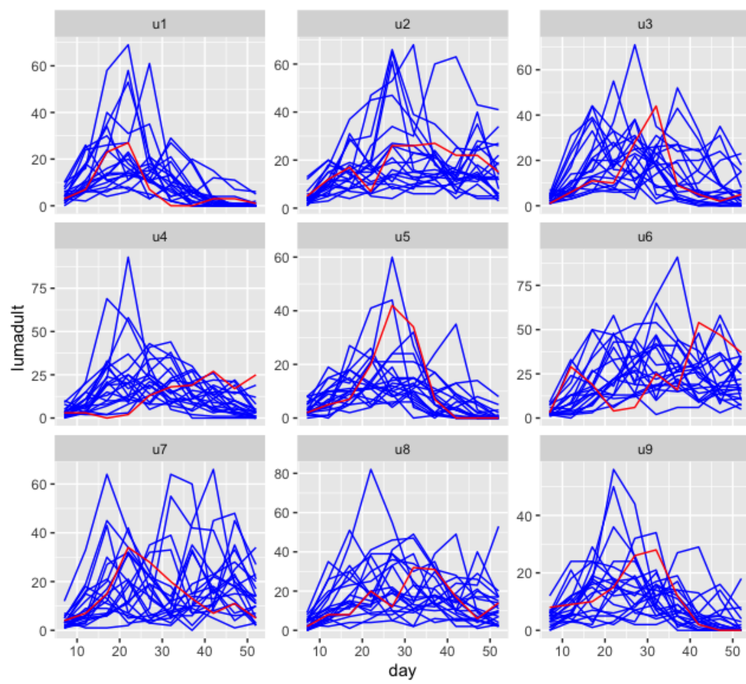

Figure 8: Simulation of Susceptible Daphnia lumholtzi of SIRPF dynamics

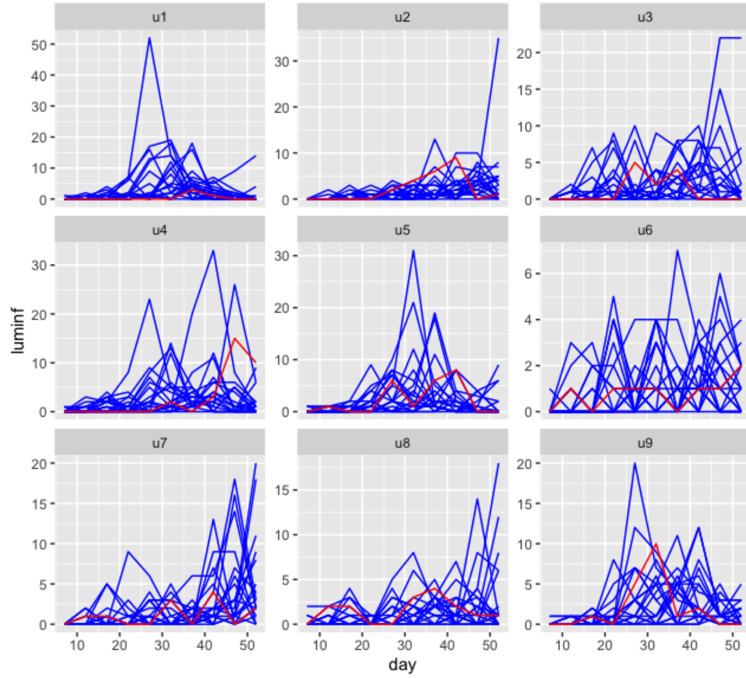

Figure 9: Simulation of Infected Daphnia lumholtzi of SIRPF dynamics

3.4 SIRPF2 Models

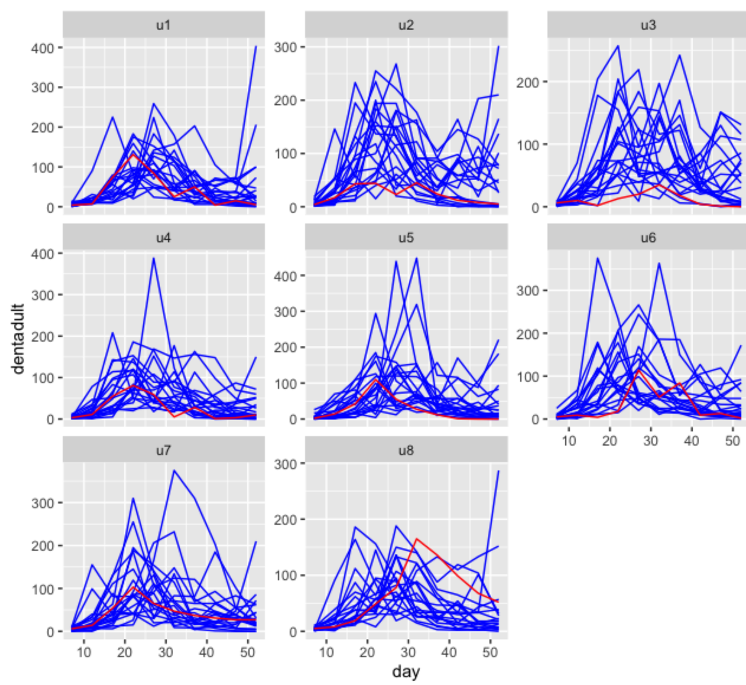

Figure 10: Simulation of Susceptible Daphnia dentifera of SIRPF2 dynamics

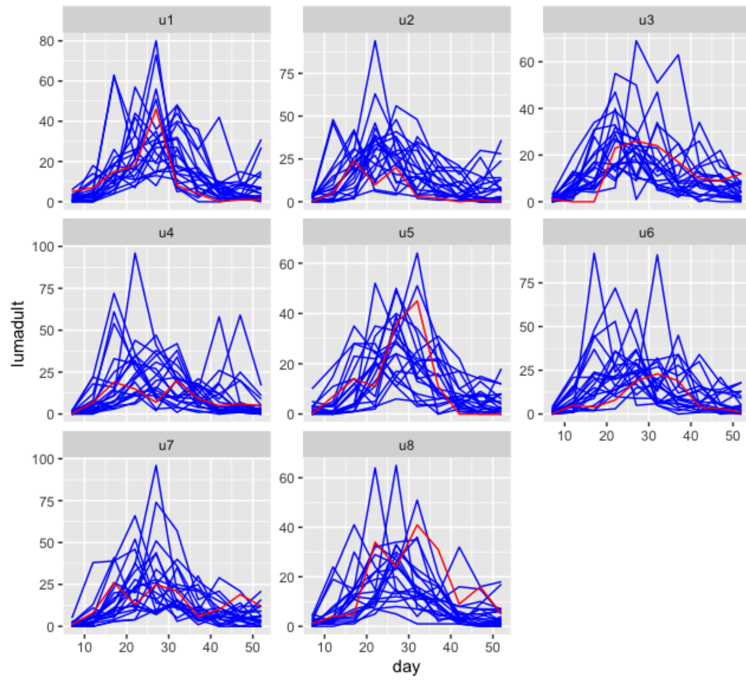

Figure 11: Simulation of Susceptible Daphnia lumholtzi of SIRPF2 dynamics

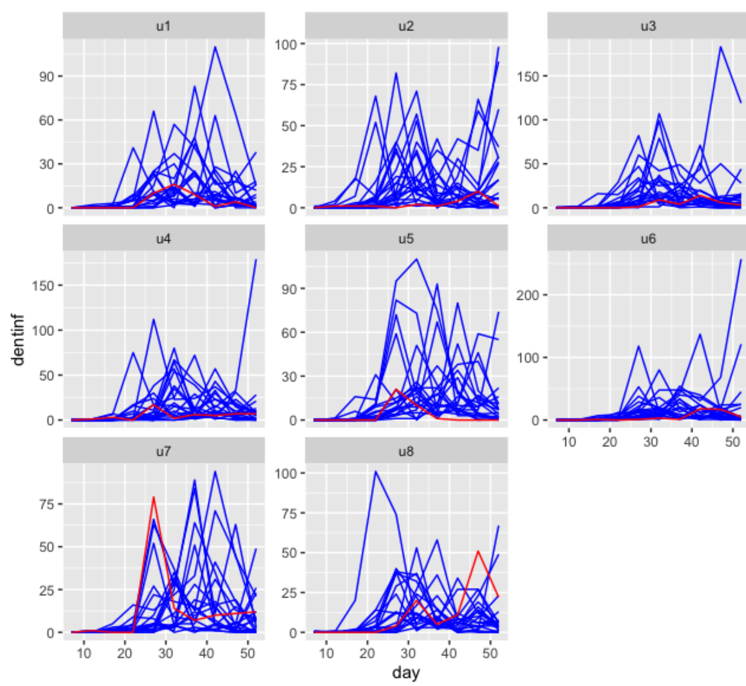

Figure 12: Simulation of Infected Daphnia dentifera of SIRPF2 dynamics

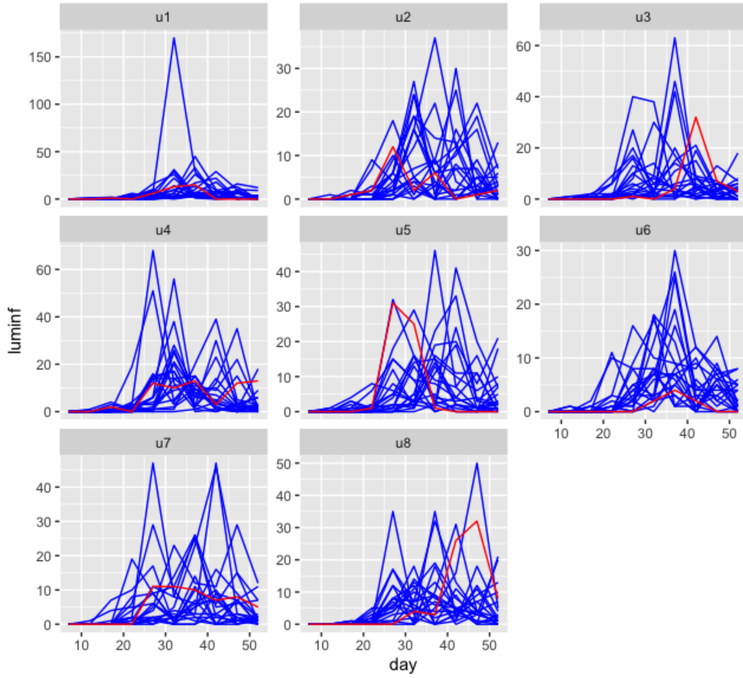

Figure 13: Simulation of Infected Daphnia lumholtzi of SIRPF2 dynamics

4 Results and Tables

Specific parameters	dimension	max log-likelihood	AIC
All shared	18	-566.82	1169.64
σ_S, σ_P	32	-553.03	1170.06
β	25	-560.04	1170.08
σ_S	25	-560.17	1170.34
σ_S, σ_I	32	-553.39	1170.78
p	25	-560.95	1171.92
β, θ_P	32	-554.24	1172.48
p, θ_P	32	-555.28	1174.56
σ_I	25	-564.21	1178.42
θ_P	25	-565.58	1181.16
β, p	32	-558.84	1181.16
σ_P	25	-565.93	1181.86
σ_I, σ_P	32	-563.33	1190.88

Table 1: This table includes the parameter comparison for SIR model with food management, which was newly introduced fitted to *Daphnia dentifera* with different choices of parameters to be either unit specific or shared

Specific parameters	dimension	max log-likelihood	AIC
All shared	18	-418.26	872.52
p	26	-411.02	874.04
β	26	-411.94	875.88
β, θ_P	34	-406.07	880.14
σ_S	26	-414.69	881.38
β, p	34	-406.81	881.62
p, θ_P	34	-408.07	884.14
$\sigma_I p$	34	-417.55	887.11
θ_P	26	-420.44	892.88
σ_P	26	-420.45	892.91
σ_S, σ_I	32	-412.94	893.88
σ_S, σ_P	32	-414.84	897.68
σ_I, σ_P	32	-417.19	902.38

Table 2: This table includes the parameter comparison for SIR model with food management, which was newly introduced fitted to *Daphnia lumholtzi* with different choices of parameters to be either unit specific or shared