

A Computational Comparison of Alternatives to Including Uncertainty in Structured Population Models*

H.T. Banks, Jimena L. Davis and Shuhua Hu

Center for Research in Scientific Computation
Center for Quantitative Sciences in Biomedicine
North Carolina State University
Raleigh, NC 27695-8212

June 15, 2009

*On the occasion of the 2009 Festschrift in honor of Chris Byrnes and Anders Lindquist

Abstract

Two conceptually different approaches to incorporate growth uncertainty into size-structured population models have recently been investigated. One entails imposing a probabilistic structure on all the possible growth rates across the entire population, which results in a growth rate distribution model. The other involves formulating growth as a Markov stochastic diffusion process, which leads to a Fokker-Planck model. Numerical computations verify that a Fokker-Planck model and a growth rate distribution model can, with properly chosen parameters, yield quite similar time dependent population densities. The relationship between the two models is based on the theoretical analysis in [7].

1 Introduction

Class and size-structured population models, which have been extensively investigated for some time, have proved useful in modeling the dynamics of a wide variety of populations. Applications are diverse and include populations ranging from cells to whole organisms in animal, plant and marine species [1, 3, 5, 7, 8, 9, 12, 14, 17, 18, 19, 20, 21, 22, 24]. One of the intrinsic assumptions in standard size-structured population models is that all individuals of the same size have the same size-dependent growth rate. This does not allow for differences due to inherent genetic differences, chronic disease or disability, underlying

local environmental variability, etc. This means that if there is no reproduction involved then the variability in size at any time is totally determined by the variability in initial size. Such models are termed *cryptodeterministic* [16] and embody the fundamental feature that uncertainty or stochastic variability enters the population only through that in the initial data. However, the experimental data in [7] for the early growth of shrimp reveals that shrimp exhibit a great deal of variability in size as time evolves even though all the shrimp begin with similar size. It was also reported in [5, 9] that experimental size-structured field data on mosquitofish population (no reproduction involved) exhibits both dispersion and bimodality in size as time progresses even though the initial population density is unimodal. Hence, standard size-structured population models such as that first proposed by Sinko and Streifer [24] are inadequate to describe the dynamics of these populations. For these situations we need to incorporate some type of uncertainty or variability into the growth process so that the variability in size is not only determined by the variability in initial size but also by the variability in individual growth.

We consider here two conceptually different approaches to incorporating the growth uncertainty into a size-structured population model. One entails imposing a probabilistic structure on the set of possible growth rates permissible in the entire population while the other involves formulating growth as a stochastic diffusion process. In [7] these are referred to as probabilistic formulations and stochastic formulations, respectively. Because we are only interested in modeling growth uncertainty in this paper, for simplicity, we will not consider either reproduction and mortality rates in our formulations.

1.1 Probabilistic Formulation

The probabilistic formulation is motivated by the observation that genetic differences or non-lethal infections of some chronic disease can have an effect on individual growth. For example, in many marine species such as mosquitofish, females grow faster than males, which means that individuals with the same size may have different growth rates. The probabilistic formulation is constructed based on the assumption that each individual does grow according to a deterministic growth model $\frac{dx}{dt} = g(x, t)$ as posited in the Sinko-Streifer formulation, but that different individuals may have different size-dependent growth rates. Based on this underlying assumption, one partitions the entire population into (possibly a continuum of) subpopulations where individuals in each subpopulation have the same size-dependent growth rate, and then assigns a probability distribution to this partition of possible growth rates in the population.

The growth process for individuals in a subpopulation with growth rate g is assumed to be

described by the dynamics

$$\frac{dx(t; g)}{dt} = g(x(t; g), t), \quad g \in \mathcal{G}, \quad (1.1)$$

where \mathcal{G} is a collection of admissible growth rates. Model (1.1) combined with the probability distribution imposed on \mathcal{G} will be called the *probabilistic growth model* in this paper. Hence, we can see that for the probabilistic formulation, the growth uncertainty is introduced into the entire population by the variability of growth rates among subpopulations. In the literature, it is common to assume that growth rate is a nonnegative function, that is, no loss in size occurs. However, individuals may experience loss in size due to disease or some other involuntary factors. Hence, we will permit these situations in this formulation, but for simplicity we assume that growth rate in each subpopulation is either a nonnegative function or a negative function, that is, the size of each individual is either nondecreasing or decreasing continuously in its growth period.

With this assumption of a family of admissible growth rates and an associated probability distribution, one thus obtains a generalization of Sinko-Streifer model, called the growth rate distribution (GRD) model, which has been formulated and studied in [2, 4, 5, 6, 9, 10]. The model consists of solving

$$\begin{aligned} v_t(x, t; g) + (g(x, t)v(x, t; g))_x &= 0, \quad x \in (0, L), \quad t > 0, \\ g(0, t)v(0, t; g) &= 0 \text{ if } g \geq 0 \text{ or } g(L, t)v(L, t; g) = 0 \text{ if } g < 0, \\ v(x, 0; g) &= v_0(x; g), \end{aligned} \quad (1.2)$$

for a given $g \in \mathcal{G}$ and then “summing” (with respect to the probability) the corresponding solutions over all $g \in \mathcal{G}$. Thus if $v(x, t; g)$ is the population density of individuals with size x at time t having growth rate g , the expectation of the total population density for size x at time t is given by

$$u(x, t) = \int_{g \in \mathcal{G}} v(x, t; g) d\mathcal{P}(g), \quad (1.3)$$

where \mathcal{P} is a probability measure on \mathcal{G} . Thus, this probabilistic formulation involves a *stationary probabilistic structure* on a family of *deterministic dynamical systems*, and \mathcal{P} is the fundamental “parameter” that is to be estimated by either parametric or nonparametric methods (which depends on the prior information known about the form for \mathcal{P}). As detailed in [5, 10], the growth rate distribution model is sufficiently rich to exhibit a number of phenomena of interest, for example, dispersion and development of two modes from one.

Observe that if all the subpopulations have nonnegative growth rates, then we need to set $g(L, t)v(L, t; g) = 0$ for each $g \in \mathcal{G}$ in order to provide a conservation law for the GRD model. Specifically if L denotes the maximum attainable size of individuals in a life time, then it is reasonable to set $g(L, t) = 0$ (as commonly done in the literature). However, if we just consider the model in a short time period, then we may choose L sufficiently

large so that $u(L, t)$ is negligible or zero if possible. We observe that if there exist some subpopulations whose growth rates are negative, then we can not provide a conservation law for these subpopulations as $g(0, t) < 0$. Hence, in this case, once the size of an individual is decreased to below the minimum size, then that individual will be removed from the system. In other words, we exclude those individuals whose size go below the minimum size. This effectively serves as a sink for these subpopulations.

1.2 Stochastic Formulation

A stochastic formulation may be motivated by the acknowledgement that environmental or emotional fluctuations can have a significant influence on the individual growth. For example, the growth rate of shrimp are affected by several environmental factors [3] such as temperature, dissolved oxygen level and salinity. The stochastic formulation is constructed under the assumption that movement from one size class to another can be described by a stochastic diffusion process [1, 13, 16, 22]. Let $\{X(t) : t \geq 0\}$ be a Markov diffusion process with $X(t)$ representing size at time t (i.e., each process realization corresponds to the size trajectory of an individual). Then $X(t)$ is described by the Ito stochastic differential equation (we refer to this equation as the *stochastic growth model*)

$$dX(t) = g(X(t), t)dt + \sigma(X(t), t)dW(t), \quad (1.4)$$

where $W(t)$ is the standard Wiener process [1, 16]. Here $g(x, t)$ denotes the average growth rate (the first moment of rate of change in size) of individuals with size x at time t , and is given by

$$g(x, t) = \lim_{\Delta t \rightarrow 0^+} \frac{1}{\Delta t} \mathbb{E} \{ \Delta X(t) | X(t) = x \}. \quad (1.5)$$

For application purposes, we assume that g is a nonnegative function here. The function $\sigma(x, t)$ represents the variability in the growth rate of individuals (the second moment of rate of change in size) and is given by

$$\sigma^2(x, t) = \lim_{\Delta t \rightarrow 0^+} \frac{1}{\Delta t} \mathbb{E} \{ [\Delta X(t)]^2 | X(t) = x \}. \quad (1.6)$$

Hence, the growth process of each individual is stochastic, and each individual grows according to stochastic growth model (1.4). Thus, for this formulation the growth uncertainty is introduced into the entire population by the stochastic growth of each individual. In addition, individuals with the same size at the same time have the same uncertainty in growth, and individuals also have the possibility of reducing their size during a growth period.

With this assumption on the growth process, we obtain the Fokker-Planck (FP) or forward Kolmogorov model for the population density u , which was carefully derived in [22] among

numerous other places and subsequently studied in many references (e.g., [1, 13, 16]). The equation and appropriate boundary conditions are given by

$$\begin{aligned}
u_t(x, t) + (g(x, t)u(x, t))_x &= \frac{1}{2}(\sigma^2(x, t)u(x, t))_{xx}, & x \in (0, L), & t > 0, \\
g(0, t)u(0, t) - \frac{1}{2}(\sigma^2(x, t)u(x, t))_x|_{x=0} &= 0, \\
g(L, t)u(L, t) - \frac{1}{2}(\sigma^2(x, t)u(x, t))_x|_{x=L} &= 0, \\
u(x, 0) &= u_0(x).
\end{aligned} \tag{1.7}$$

Here L is the maximum size that individuals may attain in any given time period. Observe that the boundary conditions in (1.7) provide a conservation law for the FP model. Because both mortality and reproduction rates are assumed zero, the total number of individuals in the population is a constant given by $\int_0^L u_0(x)dx$. In addition, we observe that with the zero-flux boundary condition at zero (minimum size) one can equivalently set $X(t) = 0$ if $X(t) \leq 0$ for the stochastic growth model (1.4) in the sense that both are used to keep individuals in the system. This means that if the size of an individual is decreased to the minimum size, it remains in the system with the possibility to once again increase its size.

The discussions in Sections 1.1 and 1.2 indicate that these probabilistic and stochastic formulations are *conceptually* quite different. However, the analysis in [7] reveals that in some cases the size distribution (the probability density function of $X(t)$) obtained from the stochastic growth model is exactly the same as that obtained from the probabilistic growth model. For example, if we consider the two models

$$\begin{aligned}
\text{stochastic formulation: } dX(t) &= b_0(X(t) + c_0)dt + \sqrt{2t}\sigma_0(X(t) + c_0)dW(t) \\
\text{probabilistic formulation: } \frac{dx(t; b)}{dt} &= (b - \sigma_0^2 t)(x(t; b) + c_0), \\
& b \in \mathbb{R} \text{ with } B \sim \mathcal{N}(b_0, \sigma_0^2),
\end{aligned} \tag{1.8}$$

and assume their initial size distributions are the same, then we obtain at each time t the same size distribution from these two distinct formulations. Here b_0 , σ_0 and c_0 are positive constants (for application purposes), and B is a normal random variable with b a realization of B . Moreover, by using the same analysis as in [7] we can show that if we compare

$$\begin{aligned}
\text{stochastic formulation: } dX(t) &= (b_0 + \sigma_0^2 t)(X(t) + c_0)dt + \sqrt{2t}\sigma_0(X(t) + c_0)dW(t) \\
\text{probabilistic formulation: } \frac{dx(t; b)}{dt} &= b(x(t; b) + c_0), & b \in \mathbb{R} \text{ with } B \sim \mathcal{N}(b_0, \sigma_0^2),
\end{aligned} \tag{1.9}$$

with the same initial size distributions, then we can also obtain at each time t the same size distribution for these two formulations. In addition, we see that both the stochastic

growth models and the probabilistic growth models in (1.8) and (1.9) reduce to the same deterministic growth model $\dot{x} = b_0(x + c_0)$ when there is no uncertainty or variability in growth (i.e., $\sigma_0 = 0$) even though both models in (1.9) *do not* satisfy the mean growth dynamics

$$\frac{d\mathbb{E}(X(t))}{dt} = b_0(\mathbb{E}(X(t)) + c_0) \quad (1.10)$$

while both models in (1.8) *do*.

As remarked in [7], if in the probabilistic formulation we impose a normal distribution $\mathcal{N}(b_0, \sigma_0^2)$ for B , this is not completely reasonable in applications because the intrinsic growth rate b can be negative which results in the size having non-negligible probability of being negative in a finite time period when σ_0 is sufficiently large relative to b_0 . A standard approach in practice to remedy this problem is to impose a truncated normal distribution $\mathcal{N}_{[\underline{b}, \bar{b}]}(b_0, \sigma_0^2)$ instead of a normal distribution; that is, we restrict B in some reasonable range $[\underline{b}, \bar{b}]$. We observe that the stochastic formulation also can lead to the size having non-negligible probability of being negative when σ_0 is sufficiently large relative to b_0 . This is because $W(t) \sim \mathcal{N}(0, t)$ for any fixed t and hence decreases in size are possible. One way to remedy this situation is to set $X(t) = 0$ if $X(t) \leq 0$. Thus, if σ_0 is sufficiently large relative to b_0 , then we may obtain different size distributions for these two formulations *after* we have made these different modifications to each. The same anomalies hold for the solutions of the FP models and the GRD models themselves because we impose zero-flux boundary conditions in the FP model and put constraints on B in the GRD model. In this paper, we present some computational examples using the models in (1.8) and (1.9) to investigate how the solutions to the modified FP models and the modified GRD models change as we vary the values of σ_0 and \underline{b} .

The remainder of this paper is organized as follows. In Section 2 we outline the numerical scheme we use to numerically solve the Fokker-Planck model. In Section 3 we present computational examples using (1.8) and (1.9) to investigate the influence of the values of σ_0 and \underline{b} on the solutions to the FP model and the GRD model. Finally, we conclude the paper in Section 4 with some conclusions and further remarks.

2 Numerical Scheme to Solve the FP Model

For the computational results presented here, we used the finite difference scheme developed by Chang and Cooper in [15] to numerically solve the FP model (1.7). This scheme provides numerical solutions which preserve some of the more important intrinsic properties of the FP model. In particular, the solution is non-negative, is particle conserving in the absence of sources or sinks, and gives exact representations of the analytic solution upon equilibration.

In the following exposition, we assume that all the model parameters are sufficiently smooth to allow implementation of this scheme. For convenience, the following notation will be used in this section:

$$d(x, t) = \sigma^2(x, t), \quad F(x, t) = g(x, t)u(x, t) - \frac{1}{2}(d(x, t)u(x, t))_x, \quad h(x, t) = g(x, t) - \frac{1}{2}d_x(x, t).$$

Hence, we can rewrite F as

$$F(x, t) = h(x, t)u(x, t) - \frac{1}{2}d(x, t)u_x(x, t).$$

Let $\Delta x = L/n$ and $\Delta t = T/l$ be the spatial and time mesh sizes, respectively, where T is the maximum time considered in the simulations. The mesh points are given by $x_j = j\Delta x$, $j = 0, 1, 2, \dots, n$, and $t_k = k\Delta t$, $k = 0, 1, 2, \dots, l$. We denote by u_j^k the finite difference approximation of $u(x_j, t_k)$, and we let $u_j^0 = u_0(x_j)$, $j = 0, 1, 2, \dots, n$. The mid point between two space mesh points is given by $x_{j+\frac{1}{2}} = \frac{x_j + x_{j+1}}{2}$, and $h_{j+\frac{1}{2}}^k = g(x_{j+\frac{1}{2}}, t_k) - \frac{1}{2}d_x(x_{j+\frac{1}{2}}, t_k)$. The scheme to solve the FP model (1.7) is given by

$$\frac{u_j^{k+1} - u_j^k}{\Delta t} + \frac{F_{j+\frac{1}{2}}^{k+1} - F_{j-\frac{1}{2}}^{k+1}}{\Delta x} = 0, \quad j = 0, 1, 2, \dots, n, \quad k = 0, 1, 2, \dots, l-1. \quad (2.1)$$

Here $F_{j+\frac{1}{2}}^{k+1}$, $j = 0, 1, 2, \dots, n-1$ are defined by

$$\begin{aligned} F_{j+\frac{1}{2}}^{k+1} &= h_{j+\frac{1}{2}}^{k+1}u_{j+\frac{1}{2}}^{k+1} - \frac{1}{2}d_{j+\frac{1}{2}}^{k+1}\frac{u_{j+1}^{k+1} - u_j^{k+1}}{\Delta x} \\ &= h_{j+\frac{1}{2}}^{k+1}[\delta_j^{k+1}u_{j+1}^{k+1} + (1 - \delta_j^{k+1})u_j^{k+1}] - \frac{1}{2}d_{j+\frac{1}{2}}^{k+1}\frac{u_{j+1}^{k+1} - u_j^{k+1}}{\Delta x} \\ &= \left[\delta_j^{k+1}h_{j+\frac{1}{2}}^{k+1} - \frac{1}{2\Delta x}d_{j+\frac{1}{2}}^{k+1} \right] u_{j+1}^{k+1} + \left[(1 - \delta_j^{k+1})h_{j+\frac{1}{2}}^{k+1} + \frac{1}{2\Delta x}d_{j+\frac{1}{2}}^{k+1} \right] u_j^{k+1}, \end{aligned} \quad (2.2)$$

where $\delta_j^{k+1} = \frac{1}{\tau_j^{k+1}} - \frac{1}{\exp(\tau_j^{k+1}) - 1}$ with $\tau_j^{k+1} = \frac{2h_{j+\frac{1}{2}}^{k+1}\Delta x}{d_{j+\frac{1}{2}}^{k+1}}$. Note that if $h_{j+\frac{1}{2}}^{k+1} = 0$, then we do not need to figure out the value of $u_{j+\frac{1}{2}}^{k+1}$. Hence, we do not need to worry about δ_j^{k+1} in this case.

Define $f(\tau) = \frac{1}{\tau} - \frac{1}{\exp(\tau) - 1}$. By a Taylor series expansion, we know that $\exp(\tau) + \exp(-\tau) > 2 + \tau^2$. Hence, $f'(\tau) < 0$. Thus, f is monotonically decreasing. Note that $\lim_{\tau \rightarrow -\infty} f(\tau) = 1$ and $\lim_{\tau \rightarrow \infty} f(\tau) = 0$. Hence, $0 \leq \delta_j^{k+1} \leq 1$ for $j = 0, 1, 2, \dots, n-1$, $k = 0, 1, 2, \dots, l-1$. Thus, we can see that when this choice for $u_{j+\frac{1}{2}}^{k+1}$ is used in a first derivative,

the scheme continuously shifts from a backward difference ($\delta_j^{k+1} = 0$) to a centered difference ($\delta_j^{k+1} = \frac{1}{2}$) to a forward difference ($\delta_j^{k+1} = 1$).

To preserve the conservation law, we use $F_{-\frac{1}{2}}^{k+1} = 0$ and $F_{n+\frac{1}{2}}^{k+1} = 0$ to approximate boundary conditions $F(0, t_{k+1}) = 0$ and $F(L, t_{k+1}) = 0$ in the FP model, respectively. To the order of accuracy of the difference scheme, these numerical boundary conditions are consistent with the boundary conditions in the FP model. Note that scheme (2.1) can also be written as the following tridiagonal system

$$-a_{1,j}^{k+1}u_{j+1}^{k+1} + a_{0,j}^{k+1}u_j^{k+1} - a_{-1,j}^{k+1}u_{j-1}^{k+1} = u_j^k, \quad j = 0, 1, 2, \dots, n, \quad k = 0, 1, 2, \dots, l-1.$$

By (2.1), we have for $j = 1, 2, \dots, n-1$,

$$\begin{aligned} a_{1,j}^{k+1} &= \frac{\Delta t}{\Delta x} \left[\frac{1}{2\Delta x} d_{j+\frac{1}{2}}^{k+1} - \delta_j^{k+1} h_{j+\frac{1}{2}}^{k+1} \right] = \frac{\Delta t}{\Delta x} \left[\frac{h_{j+\frac{1}{2}}^{k+1}}{\exp(\tau_j^{k+1}) - 1} \right], \\ a_{0,j}^{k+1} &= 1 + \frac{\Delta t}{\Delta x} \left[(1 - \delta_j^{k+1}) h_{j+\frac{1}{2}}^{k+1} - \delta_{j-1}^{k+1} h_{j-\frac{1}{2}}^{k+1} \right] + \frac{\Delta t}{2\Delta x^2} \left(d_{j+\frac{1}{2}}^{k+1} + d_{j-\frac{1}{2}}^{k+1} \right) \\ &= 1 + \frac{\Delta t}{\Delta x} \left[\frac{\exp(\tau_j^{k+1})}{\exp(\tau_j^{k+1}) - 1} h_{j+\frac{1}{2}}^{k+1} + \frac{1}{\exp(\tau_{j-1}^{k+1}) - 1} h_{j-\frac{1}{2}}^{k+1} \right], \\ a_{-1,j}^{k+1} &= \frac{\Delta t}{\Delta x} \left[(1 - \delta_{j-1}^{k+1}) h_{j-\frac{1}{2}}^{k+1} + \frac{1}{2\Delta x} d_{j-\frac{1}{2}}^{k+1} \right] = \frac{\Delta t}{\Delta x} \left[\frac{\exp(\tau_{j-1}^{k+1})}{\exp(\tau_{j-1}^{k+1}) - 1} h_{j-\frac{1}{2}}^{k+1} \right]. \end{aligned}$$

By (2.1) with $j = 0$ and boundary condition $F_{-\frac{1}{2}}^{k+1} = 0$, we find that

$$\begin{aligned} a_{1,0}^{k+1} &= \frac{\Delta t}{\Delta x} \left[\frac{1}{2\Delta x} d_{\frac{1}{2}}^{k+1} - \delta_0^{k+1} h_{\frac{1}{2}}^{k+1} \right] = \frac{\Delta t}{\Delta x} \left[\frac{h_{\frac{1}{2}}^{k+1}}{\exp(\tau_0^{k+1}) - 1} \right], \\ a_{0,0}^{k+1} &= 1 + \frac{\Delta t}{\Delta x} \left[(1 - \delta_0^{k+1}) h_{\frac{1}{2}}^{k+1} + \frac{1}{2\Delta x} d_{\frac{1}{2}}^{k+1} \right] = 1 + \frac{\Delta t}{\Delta x} \left[\frac{\exp(\tau_0^{k+1})}{\exp(\tau_0^{k+1}) - 1} h_{\frac{1}{2}}^{k+1} \right], \\ a_{-1,0}^{k+1} &= 0. \end{aligned}$$

By (2.1) with $j = n$ and boundary condition $F_{n+\frac{1}{2}}^{k+1} = 0$, we find that

$$\begin{aligned} a_{1,n}^{k+1} &= 0, \\ a_{0,n}^{k+1} &= 1 + \frac{\Delta t}{\Delta x} \left[\frac{1}{2\Delta x} d_{n-\frac{1}{2}}^{k+1} - \delta_{n-1}^{k+1} h_{n-\frac{1}{2}}^{k+1} \right] = 1 + \frac{\Delta t}{\Delta x} \left[\frac{h_{n-\frac{1}{2}}^{k+1}}{\exp(\tau_{n-1}^{k+1}) - 1} \right], \\ a_{-1,n}^{k+1} &= \frac{\Delta t}{\Delta x} \left[(1 - \delta_{n-1}^{k+1}) h_{n-\frac{1}{2}}^{k+1} + \frac{1}{2\Delta x} d_{n-\frac{1}{2}}^{k+1} \right] = \frac{\Delta t}{\Delta x} \left[\frac{\exp(\tau_{n-1}^{k+1})}{\exp(\tau_{n-1}^{k+1}) - 1} h_{n-\frac{1}{2}}^{k+1} \right]. \end{aligned}$$

It is obvious that if we set $\Delta t < \frac{1}{\|h_x\|_\infty}$ and $\frac{\Delta t}{\Delta x} < \frac{1}{2\|h\|_\infty}$, then $a_{-1,j}^{k+1}$, $a_{0,j}^{k+1}$ and $a_{1,j}^{k+1}$ satisfy the following conditions

$$\begin{cases} a_{-1,j}^{k+1}, a_{0,j}^{k+1}, a_{1,j}^{k+1}, u_j^0 \geq 0, \\ a_{0,j}^{k+1} \geq a_{-1,j}^{k+1} + a_{1,j}^{k+1}. \end{cases} \quad j = 0, 1, 2, \dots, n, \quad k = 0, 1, 2, \dots, l-1, \quad (2.3)$$

which guarantee that $u_j^{k+1} \geq 0$, $j = 0, 1, 2, \dots, n$, $k = 0, 1, 2, \dots, l-1$ (see [15, 23]).

3 Numerical Results

For all the examples given in this section, the maximum time is set at $T = 10$. The initial condition in the FP model is given by $u_0(x) = 100 \exp(-100(x-0.4)^2)$, and initial conditions in the GRD model are given by $v_0(x; b) = 100 \exp(-100(x-0.4)^2)$ for $b \in [\underline{b}, \bar{b}]$. We set $c_0 = 0.1$, $b_0 = 0.045$, and $\sigma_0 = rb_0$, where r is a positive constant. We use $\Delta x = 10^{-3}$ and $\Delta t = 10^{-3}$ in the finite difference scheme to numerically solve the FP model.

Section 3.1 details results for an example where model parameters in the FP and the GRD models are chosen based on (1.8), and Section 3.2 contains results comparing the FP and the GRD models in (1.9). In these two examples, we vary the values of r and \underline{b} to illustrate their effect on the solutions to the FP and the GRD models.

3.1 Example 1

Model parameters in the FP and the GRD models in this example are chosen based on (1.8) and are given by

$$\begin{aligned} \text{FP model: } & g(x) = b_0(x + c_0), \quad \sigma(x, t) = \sqrt{2t}\sigma_0(x + c_0) \\ \text{GRD model: } & g(x, t; b) = (b - \sigma_0^2 t)(x + c_0), \quad \text{where } b \in [\underline{b}, \bar{b}] \text{ with } B \sim \mathcal{N}_{[\underline{b}, \bar{b}]}(b_0, \sigma_0^2). \end{aligned} \quad (3.1)$$

We choose $\underline{b} = b_0 - 3\sigma_0$ and $\bar{b} = b_0 + 3\sigma_0$. Let $r_0 = \frac{-3 + \sqrt{4b_0T + 9}}{2b_0T}$ (≈ 0.3182). It is easy to show that if $r < r_0$, then $g(x, t; b) = (b - \sigma_0^2 t)(x + c) > 0$ in $\{(x, t) | (x, t) \in [0, L] \times [0, T]\}$ for all $b \in [\underline{b}, \bar{b}]$. Here we just consider the case for $r < r_0$, i.e., the growth rate of each subpopulation is positive. To conserve the total number of the population in the system, we must choose L sufficiently large so that $v(L, t; b)$ is negligible for any $t \in [0, T]$ and $b \in [\underline{b}, \bar{b}]$. For this example we chose $L = 6$.

We observe that with this choice of $g(x, t) = (b - \sigma_0^2 t)(x + c_0)$ in the GRD model, we can analytically solve (1.2) by the method of characteristics, and the solution is given by

$$v(x, t; b) = \begin{cases} v_0(\omega(x, t); b) \exp(-bt + \frac{1}{2}\sigma_0^2 t^2) & \text{if } \omega(x, t) \geq 0 \\ 0 & \text{if } \omega(x, t) < 0, \end{cases} \quad (3.2)$$

where $\omega(x, t) = -c_0 + (x + c_0) \exp(-bt + \frac{1}{2}\sigma_0^2 t^2)$. Hence, by (1.3) we have

$$u(x, t) = \int_{\underline{b}}^{\bar{b}} v(x, t; b) \frac{\frac{1}{\sigma_0} \phi\left(\frac{b-b_0}{\sigma_0}\right)}{\Phi\left(\frac{\bar{b}-b_0}{\sigma_0}\right) - \Phi\left(\frac{\underline{b}-b_0}{\sigma_0}\right)} db, \quad (3.3)$$

where ϕ is the probability density function of the standard normal distribution, and Φ is its corresponding cumulative distribution function. In the simulations, the trapezoidal rule with $\Delta b = (\bar{b} - \underline{b})/128$ was used to calculate the integral in (3.3).

Snapshots of the numerical solution of the Fokker-Planck equation and the solution of the GRD model at $t = T$ with $r = 0.1$ (left) and $r = 0.3$ (right) are graphed in Figure 1. These

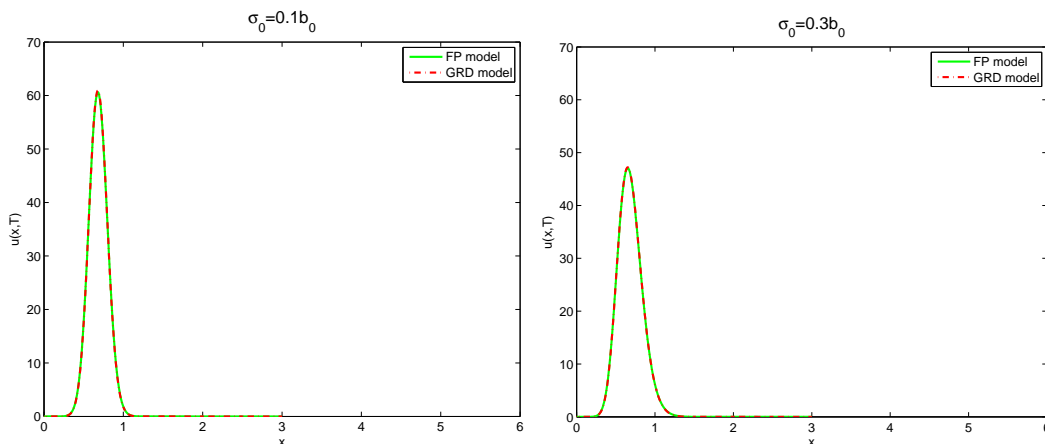


Figure 1: Numerical solutions $u(x, T)$ to the FP model and the GRD model with model parameters chosen as (3.1), where $\underline{b} = b_0 - 3\sigma_0$ and $\bar{b} = b_0 + 3\sigma_0$.

results, along with other snapshots (not depicted here) demonstrate that we do indeed obtain quite similar (in fact indistinguishable in these graphs) population densities for these two models and parameter values. This is because $\mathcal{N}_{[\underline{b}, \bar{b}]}(b_0, \sigma_0^2)$ is a good approximation of $\mathcal{N}(b_0, \sigma_0^2)$ (for this setup of \underline{b} and \bar{b}) and σ_0 is chosen sufficiently small so that the size distributions obtained in (1.8) are good approximations of size distributions obtained computationally with the GRD models and the FP models. Note that the population density $u(x, t)$ is just the product of the total number of the population and the probability density function.

3.2 Example 2

We consider model parameters in the FP and GRD models of (1.9). That is, we compare models with

$$\begin{aligned} \text{FP model: } & g(x, t) = (b_0 + \sigma_0^2 t)(x + c_0), \quad \sigma(x, t) = \sqrt{2t}\sigma_0(x + c_0) \\ \text{GRD model: } & g(x; b) = b(x + c_0), \text{ where } b \in [\underline{b}, \bar{b}] \text{ with } B \sim \mathcal{N}_{[\underline{b}, \bar{b}]}(b_0, \sigma_0^2). \end{aligned} \quad (3.4)$$

Because the growth rate g in the GRD model is a positive function if $b > 0$, we need to choose L sufficiently large so that $v(L, t; b)$ is negligible for any $t \in [0, T]$ in any subpopulation with positive intrinsic growth rate b . Doing so will conserve the total number in the population. Here we again chose $L = 6$.

With this choice of $g(x) = b(x + c_0)$ in the GRD model, we can again analytically solve (1.2) by the method of characteristics, and the solutions for subpopulations with nonnegative b (the boundary condition in (1.2) is $v(0, t; b) = 0$ in this case) is given by

$$v(x, t; b) = \begin{cases} v_0(\omega(x, t); b) \exp(-bt) & \text{if } \omega(x, t) \geq 0 \\ 0 & \text{if } \omega(x, t) < 0. \end{cases} \quad (3.5)$$

The solution for subpopulations with negative b (the boundary condition in (1.2) is $v(L, t; b) = 0$ in this case) is given by

$$v(x, t; b) = \begin{cases} v_0(\omega(x, t); b) \exp(-bt) & \text{if } \omega(x, t) \leq L \\ 0 & \text{if } \omega(x, t) > L, \end{cases} \quad (3.6)$$

where $\omega(x, t) = -c_0 + (x + c_0) \exp(-bt)$. We use these with (3.3) to calculate $u(x, t)$.

The numerical solutions of the Fokker-Planck equation and the corresponding solutions of the GRD model at $t = T$ with $r = 0.1, 0.3, 0.7, 0.9, 1.3$ and 1.5 are depicted in Figure 2, where $\underline{b} = \max\{b_0 - 3\sigma_0, 10^{-6}\}$ and $\bar{b} = b_0 + 3\sigma_0$. Let $r_0 = \frac{b_0 - 10^{-6}}{3b_0}$ (≈ 0.3333). It is easy to see that if $r \leq r_0$, then $\mathcal{N}_{[\underline{b}, \bar{b}]}(b_0, \sigma_0^2)$ is a good approximation of $\mathcal{N}(b_0, \sigma_0^2)$ as $\underline{b} = b_0 - 3\sigma_0$ in these cases. Figure 2 reveals that we obtained quite similar population densities for these two models for $r = 0.1$ and 0.3 , again because for these cases the size distributions obtained with (1.9) are good approximations of size distributions obtained by both the FP and GRD models. However, when $r > r_0$, the two solutions begin to diverge further as r increases. The reason is that $\mathcal{N}_{[\underline{b}, \bar{b}]}(b_0, \sigma_0^2)$ is no longer a good approximation of $\mathcal{N}(b_0, \sigma_0^2)$ because $\underline{b} = 10^{-6}$. This is greater than $b_0 - 3\sigma_0$ in these cases, which means the size distributions obtained with (1.9) are no longer a good approximation of size distributions obtained by the GRD model. Indeed, for the FP model with the case $r > r_0$, there exist some non-negligible fraction of individuals whose size is decreased, while in the GRD model the size of each individual always increases as b is always positive.

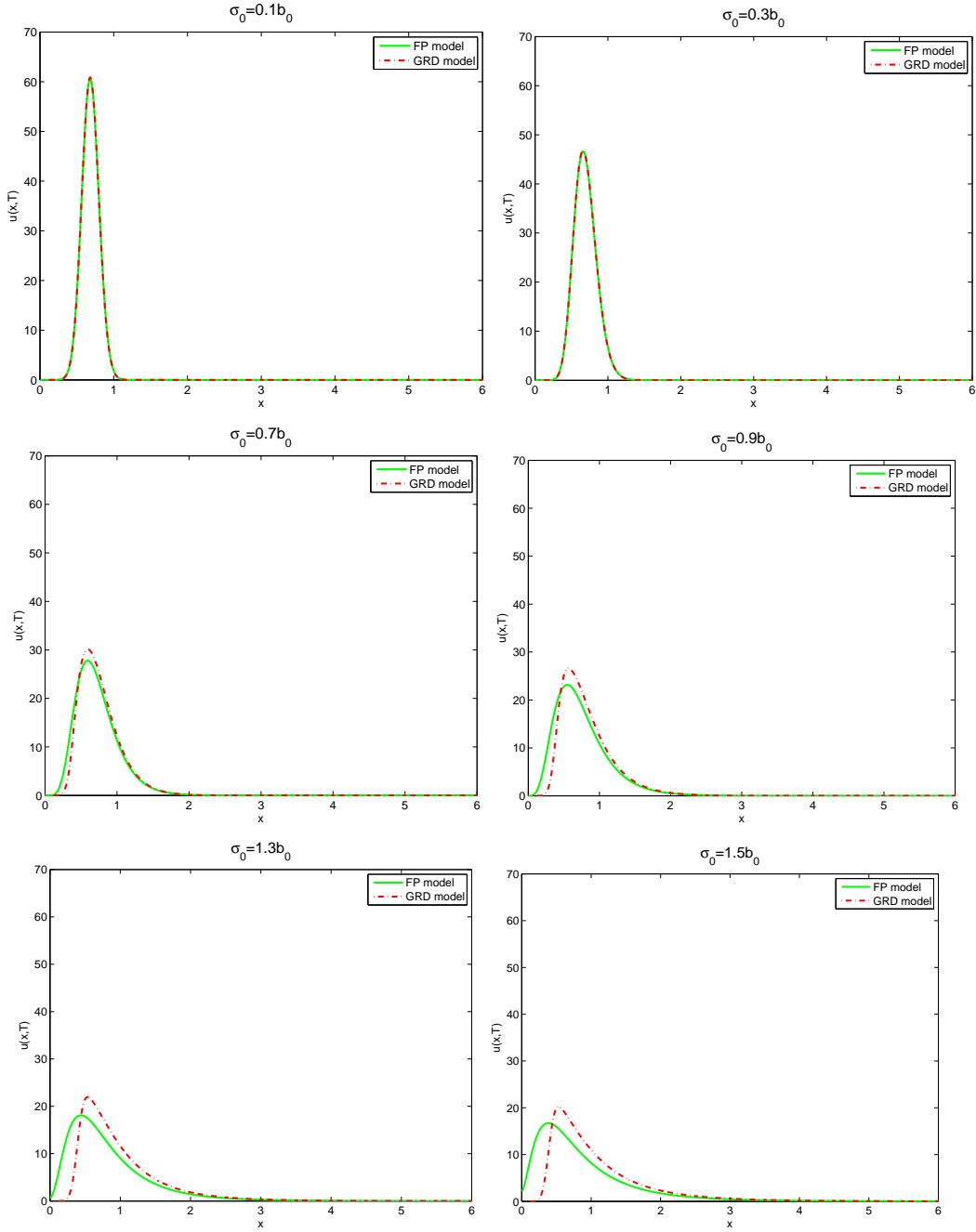


Figure 2: Numerical solutions $u(x, T)$ to the FP model and the GRD model with model parameters chosen as in (3.4), with $\underline{b} = \max\{b_0 - 3\sigma_0, 10^{-6}\}$ and $\bar{b} = b_0 + 3\sigma_0$.

Figure 3 illustrates the numerical solutions of the FP model and the solutions of the GRD model at $t = T$ with $r = 0.7, 0.9, 1.3$ and 1.5 , where $\underline{b} = b_0 - 3\sigma_0$ and $\bar{b} = b_0 + 3\sigma_0$. With

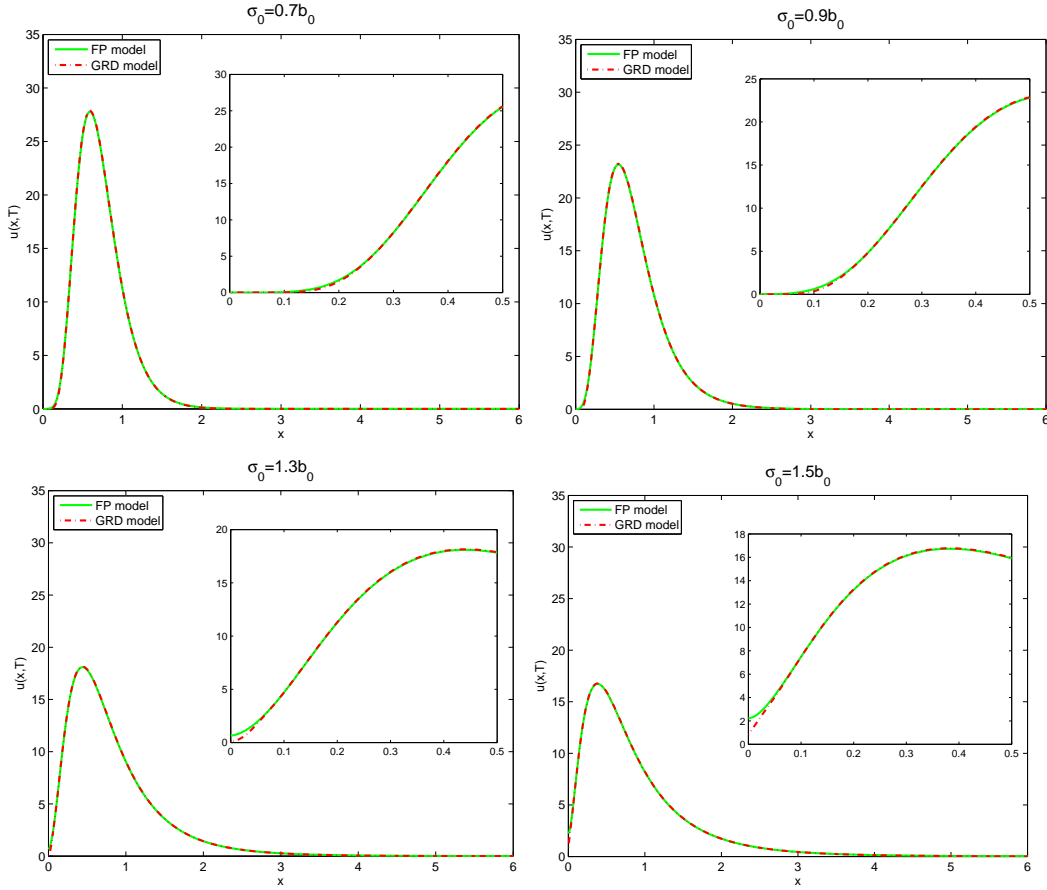


Figure 3: Numerical solutions $u(x, T)$ to the FP model and the GRD model with model parameters chosen as in (3.4), where $\underline{b} = b_0 - 3\sigma_0$ and $\bar{b} = b_0 + 3\sigma_0$. The embedded plots are enlarged snapshots of the plot in the region $[0, 0.5]$.

this choice of \underline{b} , we see that if $r > 1/3$, then there also exist some subpopulations in the GRD model with negative growth rates. Thus individuals in these subpopulations continue to lose weight and they will be removed from the population once their size is less than zero (the minimum size). If this situation occurs, then the total number in the population is no longer conserved, and this difficulty becomes worse as r becomes larger. However, for the FP model the total number of population is always conserved because of the zero-flux boundary conditions. In the FP model, once the size of individuals is decreased to the minimum size, they either stay there or they may increase their size in future time increments. From Figure 3 we can see that these two models yield pretty much similar solutions for $r = 0.7$ and 0.9 . This is because in these cases r is not sufficiently large, which results in the size having negligible probability being negative in the given time period. Thus most of individuals in the GRD model remain in the system. However, we can also see that for the cases $r = 1.3$ and $r = 1.5$, the solutions to the FP models and the GRD models diverge (at the left part of

the lower figures). This is because the size has non-negligible probability of being negative in these cases and these individuals with negative size in the GRD models are removed from the system.

4 Concluding Remarks

The computational results in this paper illustrate that, as predicted based on the analysis in [7], the Fokker-Planck model and the growth rate distribution model can, with properly chosen parameters in the individual growth dynamics, yield quite similar population densities. This implies that if one formulation is much more computationally difficult than the other, then we can use the easier one to compute solutions if we can find the corresponding equivalent forms. For example, the computational time needed to solve the Fokker-Planck model is usually much longer than that for growth rate distribution model for both examples given in Section 3. This is especially true when the initial population density is a sharp pulse, because then we need to employ a very fine mesh size to have a reasonably accurate solution to the FP model. In this case we can equivalently use the growth rate distribution model to compute the solution for the Fokker-Planck model when σ_0 is relatively small compared to b_0 .

In closing we note that the arguments of [7, 11] guarantee equivalent size distributions at any time t for the two formulations discussed in this paper. Moreover, while the GRD formulation is not defined in terms of a stochastic process, one can argue that there does exist an equivalent underlying stochastic process satisfying a random differential equation (but not a stochastic differential equation for a Markov process). It can be argued that while the corresponding stochastic processes have the same size distribution at any time t , they are *not* the same stochastic process. This can be seen, for example, by computing the covariances of the respective processes which are different [11].

Acknowledgment

This research was supported in part (HTB and SH) by grant number R01AI071915-07 from the National Institute of Allergy and Infectious Diseases, in part (HTB and SH) by the Air Force Office of Scientific Research under grant number FA9550-09-1-0226 and in part (JLD) by the US Department of Energy Computational Science Graduate Fellowship under grant DE-FG02-97ER25308.

References

- [1] L.J.S. Allen, *An Introduction to Stochastic Processes with Applications to Biology*, Prentice Hall, New Jersey, 2003.
- [2] H.T. Banks and K.L. Bihari, Modelling and estimating uncertainty in parameter estimation, *Inverse Problems*, **17** (2001), 95–111.
- [3] H.T. Banks, V.A. Bokil, S. Hu, A.K. Dhar, R.A. Bullis, C.L. Browdy and F.C.T. Allnutt, Modeling shrimp biomass and viral infection for production of biological countermeasures, CRSC-TR05-45, NCSU, December, 2005; *Mathematical Biosciences and Engineering*, **3** (2006), 635–660.
- [4] H.T. Banks, D.M. Bortz, G.A. Pinter and L.K. Potter, Modeling and imaging techniques with potential for application in bioterrorism, CRSC-TR03-02, NCSU, January, 2003; Chapter 6 in *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, (H.T. Banks and C. Castillo-Chavez, eds.) Frontiers in Applied Math, **FR28**, SIAM (2003), Philadelphia, PA, 129–154.
- [5] H.T. Banks, L.W. Botsford, F. Kappel and C. Wang, Modeling and estimation in size structured population models, LCDS-CCS Report 87-13, Brown University; *Proceedings 2nd Course on Mathematical Ecology*, (Trieste, December 8-12, 1986) World Press, Singapore, 1988, 521–541.
- [6] H.T. Banks and J.L. Davis, Quantifying uncertainty in the estimation of probability distributions, CRSC-TR07-21, December, 2007; *Math. Biosci. Engr.*, **5** (2008), 647–667.
- [7] H.T. Banks, J.L. Davis, S.L. Ernstberger, S. Hu, E. Artimovich, A.K. Dhar and C.L. Browdy, A comparison of probabilistic and stochastic formulations in modeling growth uncertainty and variability, CRSC-TR08-03, NCSU, February, 2008; *Journal of Biological Dynamics*, **3** (2009) 130–148.
- [8] H.T. Banks, J.L. Davis, S.L. Ernstberger, S. Hu, E. Artimovich and A.K. Dhar, Experimental design and estimation of growth rate distributions in size-structured shrimp populations, CRSC-TR08-20, NCSU, November, 2008; *Inverse Problems*, to appear.
- [9] H.T. Banks, B.G. Fitzpatrick, L.K. Potter, and Y. Zhang, Estimation of probability distributions for individual parameters using aggregate population data, CRSC-TR98-6, NCSU, January, 1998; in *Stochastic Analysis, Control, Optimization and Applications*, (Edited by W. McEneaney, G. Yin and Q. Zhang), Birkhauser, Boston, 1989, 353–371.
- [10] H.T. Banks and B.G. Fitzpatrick, Estimation of growth rate distributions in size structured population models, *Quart. Appl. Math.*, **49** (1991), 215–235.

- [11] H.T. Banks and S. Hu, An equivalence between nonlinear stochastic Markov processes and probabilistic structures on deterministic systems, in preparation.
- [12] H.T. Banks and H.T. Tran, *Mathematical and Experimental Modeling of Physical and Biological Processes*, CRC Press, Boca Raton, 2009.
- [13] H.T. Banks, H.T. Tran, and D.E. Woodward, Estimation of variable coefficients in the Fokker-Planck equations using moving node finite elements, *SIAM J. Numer. Anal.*, **30** (1993), 1574–1602.
- [14] G. Bell and E. Anderson, Cell growth and division I. A mathematical model with applications to cell volume distributions in mammalian suspension cultures, *Biophysical Journal*, **7** (1967), 329–351.
- [15] J.S. Chang and G. Cooper, A practical difference scheme for Fokker-Planck equations, *J. Comp. Phys.*, **6** (1970), 1-16.
- [16] T.C. Gard, *Introduction to Stochastic Differential Equations*, Marcel Dekker, New York, 1988.
- [17] M. Gyllenberg and G.F. Webb, A nonlinear structured population model of tumor growth with quiescence, *J. Math. Biol.*, **28** (1990), 671–694.
- [18] M. Kot, *Elements of Mathematical Ecology*, Cambridge University Press, Cambridge, 2001.
- [19] T. Luzyanina, D. Roose and G. Bocharov, Distributed parameter identification for a label-structured cell population dynamics model using CFSE histogram time-series data, *J. Math. Biol.*, to appear.
- [20] T. Luzyanina, D. Roose, T. Schenkel, M. Sester, S. Ehl, A. Meyerhans and G. Bocharov, Numerical modelling of label-structured cell population growth using CFSE distribution data, *Theoretical Biology and Medical Modelling*, **4** (2007), 1–26.
- [21] J.A.J. Metz and O. Diekmann, eds., *The Dynamics of Physiologically Structured Populations*, Lecture Notes in Biomathematics, Springer-Verlag, Berlin, 1986
- [22] A. Okubo, *Diffusion and Ecological Problems: Mathematical Models*, Lecture Notes in Biomathematics, Vol. 10, Springer-Verlag, Berlin, 1980.
- [23] R.D. Richtmyer and K.W. Morton, *Difference Methods for Initial-value Problems*, Wiley, New York, 1967.
- [24] J. Sinko and W. Streifer, A new model for age-size structure of a population, *Ecology*, **48** (1967), 910–918.