

**NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS**

Faculty of Computer Science
Bachelor's Programme "Data Science and Business Analytics"

UDC _____

Research Project Report
on the topic **Analysis of a multi-species model of biological communities**

(interim, the first stage)

Fulfilled by the Student:

group #БПАД213

Mamedov Ali

Surname, First name, Patronymic, if any

Checked by the Project Supervisor:

Nikitin Alexey Antonovich

Surname, First name, Patronymic (if any), Academic title (if any)

Associate Professor of the Department of General

Mathematics of the VMK MSU

Job

MSU Faculty of Computational Mathematics and

Cybernetics

Place of Work (Company or HSE Department)

Date 12.02.2023

signature

Moscow 2023

Basic terms, definitions, and abbreviations	3
Introduction	4
1.1 Relevance	4
1.2 Goal	4
Review	5
2.1 Spatial model and its generations	5
2.2 Diverse models	5
2.2 Homogeneous and non-homogeneous models	7
Non-homogeneous and homogeneous spatial moment dynamics	9

Basic terms, definitions, and abbreviations

- A stochastic or random process is a mathematical object usually defined as a family of random variables. Stochastic processes are widely used as mathematical models of systems and phenomena that appear to vary in a random manner
- Microbial models are valuable tools in planning Hazard Analysis, Critical Control Point programs and making decisions in other fields of microbiology, as they provide the first estimates of expected changes in microbial populations when exposed to a specific set of conditions
- A biological community, in biology, is an interacting group of various species in a common location.

Introduction

1.1 Relevance

Simulations of stochastic, agent-based models are now widely used in the life sciences and social sciences, and are perceived as a key route to understanding complex processes where agents interact with neighbors. Although such simulations can give hints about the causes of emerging patterns, clear-cut answers usually entail going to the underlying mathematics. The use of spatial-moment dynamics is one of several ways of charting the ground between spatial, agent-based models and mathematical analysis.

The concept of spatial moments can be traced back to the early 20th century, when mathematicians and physicists began studying the moments of functions in one dimension. The concept was later extended to two and three dimensions, and today, spatial moments play a fundamental role in a wide range of applications.

At its core, spatial moments are used to calculate the moments of a function or image with respect to its centroid. The centroid is the weighted average of the coordinates of all the pixels or points in the image or object. Once the centroid is calculated, spatial moments can be used to describe the distribution of mass or intensity around it.

Spatial moments are a powerful mathematical tool used to describe the shape, size, and orientation of objects in two or three-dimensional space. They are widely used in image processing, computer vision, pattern recognition, microbiology, and machine learning applications to extract features and recognize objects.

1.2 Goal

And the main goal of this paper will be to find out how it happened that non-homogeneous spatial models prevailed over others and compare it with alternatives that people have.

Review

2.1 Spatial model and its generations

The definition of spatial moment was already given in the previous section, thus here we are going to look only at the features of models.

The table below shows advantages and disadvantages of considered model:

Spatial moment	
Pros	Cons
Spatial models can be useful in modeling the spread of microbial populations over space and time	Spatial models may oversimplify the behavior of individual agents in a system, focusing more on the macroscopic behavior of the population
Spatial models can incorporate environmental factors, such as nutrient availability or temperature gradients, that can affect microbial behavior	Spatial models may not be able to capture the complex interactions that occur in microbial systems
Spatial models can be used to test hypotheses about how spatial structure affects microbial populations	Spatial models may require large amounts of data to parameterize and validate, and may be sensitive to uncertainties in input data

The next table will show the summary of the features included in previous spatial moment-dynamic models:

	Movemen t	Birth	Death	Multi-type	NH
Bolker and Pacala (1997)	No	DI	DI+DD	No	No
Bolker and Pacala (1999)	No	DI	DI+DD	Yes	No
Lewis and Pacala (2000)	No	DI	No	No	Yes
Lewis (2000)	No	DI+DD	No	No	Yes
Dieckmann and Law (2000)	DI	DI+DD	DI+DD	Yes	No
Murrell and Law (2000)	DD	No	No	Yes	No
Bolker (2003)	No	DI	DI+DD	No	No
Murrell and Law (2003)	No	DI	DI+D	Yes	No
Murrell (2005)	DI	DI+DD	DI+DD	Yes	No
Plank and Law (2015) <- our model	DI+DD	DI+DD	DI+DD	Yes	Yes

Key: NH = non-homogeneous; DI = density-independent; DD = density-dependent

2.2 Diverse models

We managed to define the spatial model and mentioned that there are alternatives. Which for some reason are not used or are used, but not so frequently. Thus let's give definitions of the alternatives that are there:

- ❖ *Mathematical models*: Mathematical models are used to describe the behavior of microorganisms and the spread of infectious diseases over time, without considering the geographical aspect.
- ❖ *Statistical models*: Statistical models are used to analyze data and make predictions, often using regression or machine learning techniques, without considering the geographical aspect.
- ❖ *Agent-based models*: Agent-based models simulate the behavior of individuals in a population, including the spread of infectious diseases, without considering the geographical aspect.
- ❖ *Network models*: Network models are built on a graph-like structure, with nodes representing individual components of the system, such as genes, proteins, or individuals in a social network., without considering the geographical aspect.

These alternative models can be useful in microbiology, especially when the focus is on understanding the dynamics of disease spread or the behavior of microorganisms, but they may not capture the spatial heterogeneity that can be important in understanding the spread of infectious diseases for example.

Mathematical models	
Pros	Cons
Mathematical models can be used to make predictions about the behavior of microbial systems that are difficult or impossible to test experimentally	Mathematical models often make assumptions and simplify the complexity of biological systems, which can lead to inaccuracies or errors in predictions
Models allow researchers to test hypotheses and explore the effects of different variables on the behavior of microbial systems	Models can only provide predictions based on the input parameters and assumptions used, and it is important to validate model predictions through experimental studies
Models can be used to optimize experimental design and predict the most effective conditions for microbial growth, product formation, or other outcomes of interest	Models can be affected by uncertainty in the input parameters, which can affect the accuracy of their predictions

Mathematical models are often less expensive and time-consuming than experimental studies, making them an attractive option for exploring a wide range of scenarios and hypotheses	Mathematical models may not capture the spatial heterogeneity that can be important in understanding the spread of infectious diseases
--	--

Statistical models	
Pros	Cons
Provide insights into relationships between microbial populations and environmental factors	Spatial models can provide more detailed and accurate information about the distribution of microbial populations
Identify important variables that influence microbial growth and survival	Statistical models may oversimplify complex interactions between microbial populations and their environment
Help predict the effects of different treatments or interventions on microbial populations	Statistical models may not capture the spatial heterogeneity that can be important in understanding the spread of infectious diseases.

ABM	
Pros	Cons
ABMs allow for a high level of detail in modeling individual behaviors and interactions, which can be important in microbiology	ABMs can be computationally intensive and require a lot of computational power
ABMs can be used to explore emergent phenomena, such as the behavior of bacterial populations under different conditions	ABMs can be difficult to parameterize, and it can be challenging to validate them
ABMs can incorporate stochasticity and randomness, which is important in modeling microbial systems that often exhibit variability	ABMs may not always capture the macroscopic behavior of a system, as they focus on individual agents

Lattice models	
Pros	Cons

They can provide a holistic view of complex microbial systems and their interactions	They rely heavily on accurate and comprehensive data, which may be difficult to obtain
They can identify key microbial players and their functions within the system	They can be computationally intensive and require specialized expertise to build and analyze
Network models can have crowding effects built in through exclusion of more than one individual from a cell	
They can predict the behavior of microbial communities in response to environmental changes	

2.2 Homogeneous and non-homogeneous models

Another way we can divide models into groups that are useful in microbiology is by comparing homogeneous and non-homogeneous spatial models. The main differences between them are:

1. Homogeneous models assume that all individuals in the population are identical and behave in the same way. Non-homogeneous models, on the other hand, allow for individual variation within the population
2. Homogeneous models are typically simpler and easier to analyze than non-homogeneous models. Non-homogeneous models can be more complex and require more data to parameterize
3. Homogeneous models are often used to study the dynamics of large populations of microorganisms. Non-homogeneous models are better suited for studying small populations or populations with complex interactions

The combination of large- and small-scale spatial structures in non-homogeneous settings using spatial-moment dynamics is important and requires more attention. Ecologists prefer this approach because the second spatial moment, which is a fundamental measure of spatial structure in plant communities, can be expressed as a pair correlation function. However, from a mathematical perspective, other methods of incorporating spatial structure into dynamics are equally important. Using low-order spatial moments requires a closure of the hierarchy of moment equations, which can be limiting. Second-order closures can provide acceptable approximations for stochastic processes, but their understanding is limited and remains a subject of ongoing mathematical research.

Non-homogeneous and homogeneous spatial moment dynamics

3.1 The main difference in characteristics approach

Actually, the main differences were already mentioned in the previous section, however there is something to add

Homogeneous spatial moments assume that the distribution of points is homogeneous, which means that the density of points is the same across the entire spatial domain. Non-homogeneous spatial moments relax this assumption and allow for the density of points to vary across the domain

Here are some of the main differences between spatial homogeneous and non-homogeneous moments in each of these three sections:

- ❖ First moment. In homogeneous point processes, the first moment is the same across the entire domain, indicating that the mean location of the points is constant. In non-homogeneous point processes, the first moment can vary across the domain, indicating that the density of points varies across the domain
- ❖ Second moment. In homogeneous point processes, the second moment is a constant matrix that describes the scale and shape of the distribution, but does not vary across the domain. In non-homogeneous point processes, the second moment can vary across the domain, indicating that the distribution is not the same scale and shape in all areas of the domain.
- ❖ Third moment. In homogeneous point processes, the third moment is zero, indicating that the distribution is symmetric. In non-homogeneous point processes, the third moment can be non-zero, indicating that the distribution is asymmetric in some areas of the domain.

3.2 The main difference in formulas approach

Here we will consider formulas for the first time.

Non-homogeneous models:

Let's represent the number of organisms at location x and time t as $N(x, t)$. In such a case, the governing equations involving birth, death, and movement events could be represented as follows:

Birth: The birth of an individual at location x can be represented by a birth rate $\beta(N(x,t))$, which is generally a function of the local population size $N(x, t)$.

Death: Similarly, the death of an individual at location x can be represented by a death rate $\delta(N(x,t))$.

Movement: The movement of individuals can be represented by a dispersal kernel $m(x, y)$, which defines the probability that an individual moves from location y to x .

Incorporating these events, the moment equations can be defined as follows:

First moment (intensity function):

The change in the first moment over time can be represented as the difference between the birth and death rates, integrated over all locations:

$$d\lambda(x,t)/dt = \int [\beta(N(x,t)) - \delta(N(x,t))] dx$$

Second moment (pair correlation function):

The change in the second moment can be affected by births, deaths, as well as movement of individuals:

$$dE[N(x,t)N(y,t)]/dt = \int [\beta(N(x,t)) + \beta(N(y,t)) - \delta(N(x,t)) - \delta(N(y,t))] dx dy + \iint m(x, y) [E[N(x,t)N(z,t)] + E[N(y,t)N(z,t)]] dz dx dy$$

Third moment (triplet correlation function):

The change in the third moment includes similar terms for births, deaths, and movements:

$$dE[N(x,t)N(y,t)N(z,t)]/dt = \int [\beta(N(x,t)) + \beta(N(y,t)) + \beta(N(z,t)) - \delta(N(x,t)) - \delta(N(y,t)) - \delta(N(z,t))] dx dy dz + \iiint m(x, y) [E[N(w,t)N(x,t)N(z,t)] + E[N(w,t)N(y,t)N(z,t)]] dw dx dy dz$$

These equations give the rate of change of the moments in time due to births, deaths, and movement. The actual form of the birth and death rates $\beta(N)$ and $\delta(N)$, and the dispersal kernel

$m(x, y)$, would depend on the specific biological system and the assumptions made in the model. In practice, these equations are often quite complex and may need to be solved numerically.

Homogeneous models:

In homogeneous systems, the events of birth, death, and movement can also be considered, and their impact on the population can be quantified through similar equations as for non-homogeneous systems. However, because the intensity function is constant in a homogeneous system, these equations are typically simpler.

Let's consider a homogeneous system, where the intensity function $\lambda(t)$ is constant across the entire spatial region:

Birth: The birth rate of individuals is represented by $\beta(\lambda(t))$, typically assumed to be a constant times the population size, resulting in a term $\beta * \lambda(t)$.

Death: The death rate of individuals is represented by $\delta(\lambda(t))$, which also usually scales with the population size, resulting in a term $\delta * \lambda(t)$.

Movement: In a homogeneous system, the overall effect of movement is typically considered to be negligible on the spatial scale because of the assumption of homogeneity, which implies that the population is evenly distributed.

Incorporating these events, the moment equations can be defined as follows:

First moment (intensity):

The change in the first moment over time can be represented as the difference between the birth and death rates: $d\lambda(t)/dt = \beta * \lambda(t) - \delta * \lambda(t)$

Second moment (pair correlation function):

The change in the second moment could include contributions from the births and deaths: $dE[N(N-1)]/dt = 2 * \beta * E[N(N-1)] - 2 * \delta * E[N(N-1)]$

Third moment (triplet correlation function):

The change in the third moment includes similar terms for births and deaths:

$$dE[N(N-1)(N-2)]/dt = 3 * \beta * E[N(N-1)(N-2)] - 3 * \delta * E[N(N-1)(N-2)]$$

These equations describe how the moments change over time in response to births and deaths in a homogeneous population. The constant intensity assumption simplifies the equations, as the spatial distribution is not changing over time.

List of Resources

1. M. Plank and R. Law, Spatial point processes and moment dynamics in the life sciences: a parsimonious derivation and some extensions (2015)
2. M. A. Lewis and S. Pacala, Modeling and analysis of stochastic invasion processes (2000)
3. M. J. Keeling and L. Danon, Mathematical modelling of infectious diseases (2009)
4. T. Britton and F. Giardina, Introduction to statistical inference for infectious diseases
5. J. Illian, A. Penttinen, H. Stoyan and D. Stoyan, Statistical Analysis and Modelling of Spatial Point Patterns(2008)
6. Bolker and Pacala (1999)
7. Lewis and Pacala (2000)
8. Lewis (2000)
9. Dieckmann and Law (2000)
10. Murrell and Law (2000)
11. Bolker (2003)
12. Murrell and Law (2003)
13. Murrell (2005)
14. Johnson, M. T., & Munson, A. D. (2018). Using spatial moment equations to model dispersion in individual-based simulations
15. "Stochastic population processes and their applications" by Nisbet and Gurney (1976).
16. "Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions," by Tilman and Kareiva (1997).
17. "Moment equations for spatial ecological systems" by Bolker and Pacala (1997).
18. O'Sullivan, D., & Perry, G. L. (2013). Spatial simulation: exploring pattern and process. John Wiley & Sons.
19. "Patterns of patchy spread in deterministic and stochastic models of biological invasion and biological control" Sergei V. Petrovskii