

Article

Mathematical Modeling and Simulation of Logistic Growth

Camilla Pelagalli [†] , Stefano Faccio [†]  and Paolo Casari ^{*} 

Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy

^{*} Correspondence: paolo.casari@unitn.it[†] These authors contributed equally to this work.

Abstract: We propose a reproducible pipeline of work consisting of the time-driven simulation of discrete logistic growth based on the corresponding master equation, focusing on demographic variation under a carrying capacity limit. The mathematical modeling that leads to the stochastic implementation is presented in a step-by-step fashion to statistically ground the designed simulation. The main parameters of the system, whose settings include extreme values, are varied to analyze the simulation behavior and explore the empirical limits of its applicability, minimizing the distance between the theoretical and observed carrying capacity through parameter tuning. After such tuning, a single simulation scenario is chosen and compared with the state-of-the-art Gillespie algorithm, which adopts a contrasting event-driven approach. The output analysis of these two strategies and the assessment of their statistical significance highlight the trade-off between adherence to the model and the computational effort of the proposed approach, while shedding light on multiple facets of logistic growth, including discrepancies between continuous and discrete models.

Keywords: logistic growth; population ecology; continuous vs. discrete time; deterministic vs. stochastic model; Gillespie algorithm; parameter tuning; analysis of residuals

1. Introduction

As defined in [1], *population ecology* studies the relationships between a group of individual organisms of the same species and their environment, i.e., a particular area, taking into consideration biotic (living) and abiotic (non-living) components of the ecosystem. In general, a population can be measured in terms of its *size*, or the total number of individuals. In particular, the *density*, or the number of individuals per unit volume or unit area, plays a key role when competition for resources, such as food, water, sunlight and living space, affects population growth [1].

In this paper, we consider the discrete-time event simulation of *density-dependent* population growth, limited by a specific habitat *carrying capacity*, which is outlined in [1] as the maximum sustainable population that preserves the environment's living conditions, without causing its degradation. Motivated by recent papers about the applicability of the logistic growth model to evaluate the mutation proportion in epidemics [2] and study the COVID-19 pandemic along with other infectious diseases [3,4], we were interested in the model's development and constraints and how they could be translated into tunable parameters in a simulated system, so far populated by synthetic replicable creatures, potentially representing an organism or pathogen.

Logistic growth is indeed an essential analytical model for ecologists and demographers that can be extended to stock markets and forecasting [5]; online deterministic simulators (<https://www.biointeractive.org>, accessed on 14 February 2025) are available



Academic Editor: Sergey Suslov

Received: 3 January 2025

Revised: 24 March 2025

Accepted: 4 April 2025

Published: 16 April 2025

Citation: Pelagalli, C.; Faccio, S.; Casari, P. Mathematical Modeling and Simulation of Logistic Growth. *Appl. Sci.* **2025**, *15*, 4409. <https://doi.org/10.3390/app15084409>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

to explore logistic growth from a theoretical point of view, and stochastic simulations are presented in the context of biological evolution (<https://www.youtube.com/@PrimerBlobs>, accessed on 14 February 2025). The Gillespie algorithm [6] has become the standard for the generation of time evolution trajectories of finite populations over continuous time [7], and it was employed as a reference to evaluate the performance of our simulation.

Our aim is to present a reproducible pipeline of work, from the mathematical modeling of the logistic growth master equation to its translation into a stochastic simulation, along with its subsequent parameter tuning and output analysis. Specifically, we shed light on the discrepancies between continuous and discrete growth, and between deterministic and stochastic growth, which tend to be implicit and scattered across the available literature and experiments. We adopt an interdisciplinary approach, bridging mathematics, statistics, computer science and, tangentially, biology, combined with a step-by-step imprint, to help a potential researcher to overcome the complexity and pitfalls of logistic growth modeling.

The specific contributions of our work are as follows:

- The implementation of a straightforward simulation of the logistic growth model, characterized by parameters that explicitly match the underlying mathematical variables and parameter ranges; these variables can be easily modified by the user, taking into account repeatability purposes and limited computational resources via the customizable setting of the number of simulations and iterations within a single simulation scenario;
- The definition of a basic simulation framework that can be tailored to investigate real-world empirical systems, making use of the extensive and reproducible testing performed on a non-specific synthetic population across extreme living conditions;
- The assessment of the limits of the proposed system through a detailed comparison with the state-of-the-art Gillespie algorithm in terms of statistical significance and residual analysis.

2. Background

The science of ecology refers to the *Malthusian law* as the first founding principle of population dynamics [8]. The Malthus model is applicable to an isolated population of individuals, not interacting with other populations, with infinite food and space resources; hence, growth is unlimited, and the rate at which the population changes is proportional to the population itself:

$$\frac{dN}{dt} = k N. \quad (1)$$

The parameter k is the *relative* growth rate (percent per unit time), to be distinguished from the *absolute* growth rate, $\frac{dN}{dt}$ (people per unit time) [5]. When the initial population is large enough and its evolution is observed for a long period of time, population growth can be mapped as a *continuous* variable and explained by the *exponential model*

$$N_t = N_0 e^{kt} \quad (2)$$

whose *discrete* version is the *geometric progression*

$$N_{t+1} = N_t (1 + r)^t \quad (3)$$

where $r = e^k - 1$ [5,9]. The difference between these two cases resides in the value of their base and the interpretation of the *continuous* and *discrete* growth rates, represented by the parameters k and r , which, respectively, account for *theoretical* and *empirical* populations. As shown in Table 1, if $r < 0.1$ and k is obtained accordingly with the formula $\ln(1 + r)$, the corresponding *percentage error (PE)* decreases, since the population size increases more slowly over time. In this way, empirical growth, characterized by integer-valued increments,

approximates continuous growth, consisting of real-valued fluctuations, while minimizing their discrepancy.

Table 1. Comparison of discrete and continuous growth by evaluation of corresponding r and k rates. Ranges: $0.01 \leq r \leq 1$, $0.01 \leq k \leq 0.693$, $0.499\% \leq PE \leq 44.27\%$, respectively. Decreasing r below 0.1 implies a significant PE decrease with respect to k , allowing discrete growth (Equation (3)) to catch up with continuous growth (Equation (2)) in accordance with the original Malthus model (Equation (1)).

Continuous Growth	Discrete Growth	Percentage Error
$N_t = N_0 e^{kt}$	$N_{t+1} = N_t (1 + r)^t$	$PE = 100 (r - k) / k$
$k = \ln(1 + 1.0) = 0.693$	$r = 1.00$	44.270%
$k = \ln(1 + 0.5) = 0.405$	$r = 0.50$	23.315%
$k = \ln(1 + 0.2) = 0.182$	$r = 0.20$	9.696%
$k = \ln(1 + 0.1) = 0.095$	$r = 0.10$	4.921%
$k = \ln(1 + 0.05) = 0.049$	$r = 0.05$	2.480%
$k = \ln(1 + 0.02) = 0.020$	$r = 0.02$	0.997%
$k = \ln(1 + 0.01) = 0.010$	$r = 0.01$	0.499%

In general, if $r = 0$ and $k = 0$, the population remains constant; if $r < 0$ and $k < 0$, the population tends to become extinct; and if $r > 0$ and $k > 0$, the population “explodes” as long-term behavior. However, growth without boundaries is unrealistic: a population is regulated by the availability of resources and damage caused by, e.g., natural disasters, disease, and war. Although the Malthusian law provided a good approximation of the world population and that of many regions for decades and even centuries, Malthus himself noted that the population would grow exponentially if unchecked, whereas the food supply would grow linearly, leading to the population exceeding the food supply [9].

The Belgian mathematician Pierre F. Verhulst, under the guidance of Adolphe Quetelet, proposed a differential equation (the original equation is proposed in the *Recerches mathématiques sur la loi d'accroissement de la population* [10], page 8), known as the *logistic equation*, that takes into account the upper limit to population growth [11]:

$$\frac{dN}{dt} = kN \left(\frac{L - N}{L} \right) \tag{4}$$

Here, L stands for the carrying capacity of the environment, representing the largest population that can be supported, while k is the continuous growth rate. The logistic equation can be obtained from the Malthus model, assuming that the relative growth rate of the population is a linearly decreasing function of N

$$\frac{1}{N} \frac{dN}{dt} = k - cN \tag{5}$$

where $0 < c \leq \frac{k}{N}$. The constant c acts as a *crowding coefficient*: the greater it becomes, the smaller $\frac{dN}{dt}$ becomes, to account for the limiting resources of an over-crowded environment. When

$$c = \frac{k}{N} \tag{6}$$

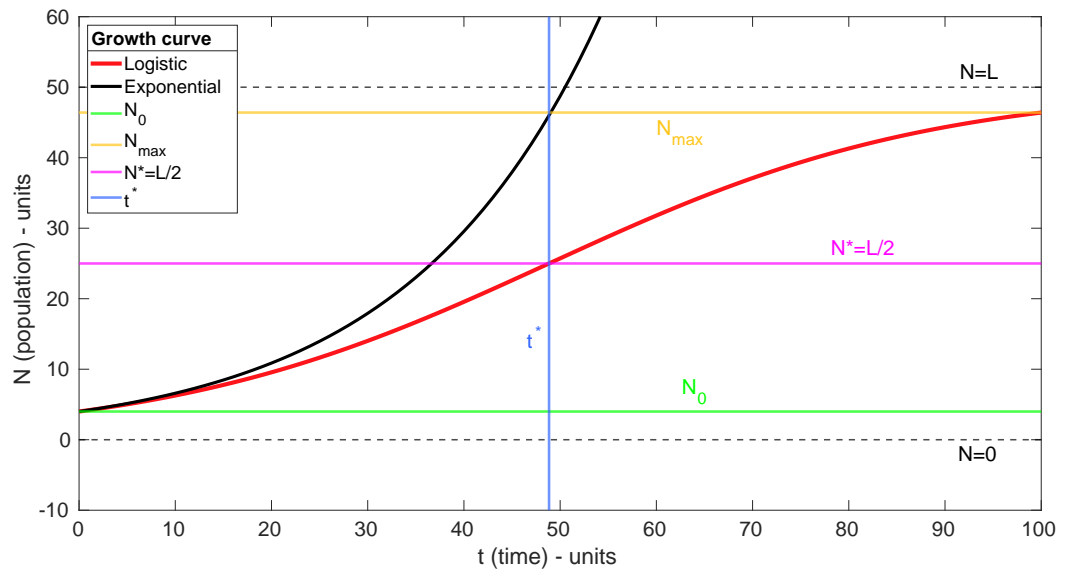
there is no total change in the population size. Solving for N , we obtain

$$N = \frac{k}{c} = L \tag{7}$$

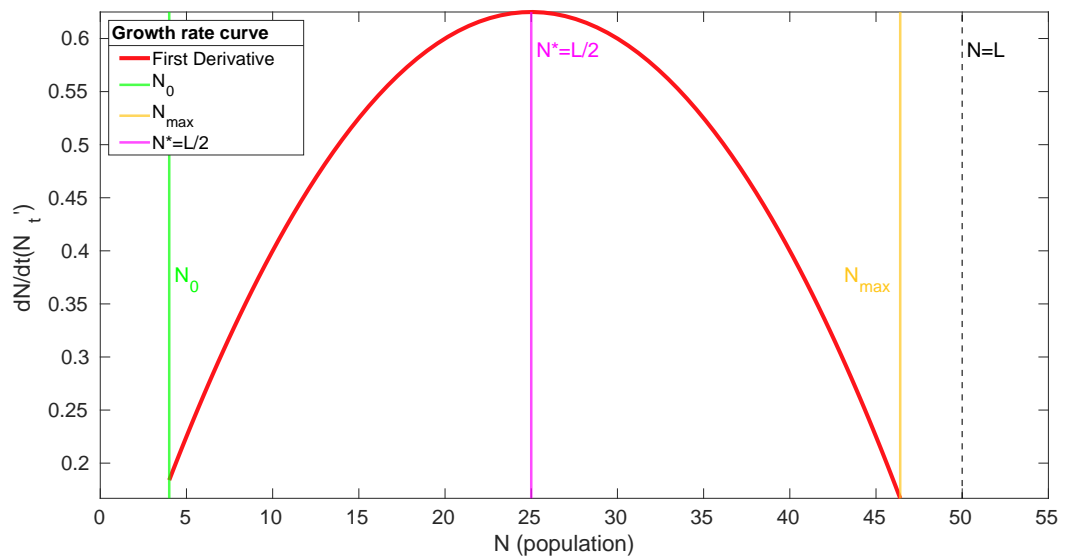
as the carrying capacity [5]. The analytical solution to the logistic equation (Figure 1a) is the *sigmoid function*

$$N_t = \frac{N_0 L}{(L - N_0)e^{-kt} + N_0} \tag{8}$$

whose first derivative N'_t (Figure 1b) is positive and increasing for $N_0 < N < L/2$, resulting in N_t being concave up. For $L/2 < N < N_{max}$, the slope $\frac{dN}{dt}$ is negative and N'_t is decreasing, making N_t concave down [5,12].



(a) Sigmoid function: population size as a function of time.



(b) Phase plot: growth rate as a function of population size.

Figure 1. Continuous model of logistic growth. Parameter settings: initial population $N_0 = 4$; maximum population (at time $t = 100$); $N_{max} = 46.4$; crowding coefficient $c = 0.001$; continuous growth rate $k = 0.05$; carrying capacity $L = 50$; inflection population $N^* = L/2 = 25$; inflection time $t^* = \frac{1}{r} \ln\left(\frac{L-N_0}{N_0}\right) = 48.85$. (a) The exponential growth (Malthusian) model is displayed as a reference. N^* and t^* are, respectively, the x and y coordinates of the inflection point, obtained by setting $N''_t = 0$, when the logistic function grows the fastest and then flips its concavity, growing slower until L is reached [13]. (b) At N^* , the acceleration changes from increasing to decreasing, i.e., the second derivative is zero at the top of the hump.

The *discrete* version of logistic growth

$$N_{t+1} = N_t \left[1 + r \left(1 - \frac{N_t}{L} \right) \right], \tag{9}$$

also called the *master* equation, has no general closed-form solution, and it was the reference method for our discrete-time event simulation. The ecologist Robert May studied this equation for populations and discovered that it could produce very complicated dynamics [14]. However, by keeping r small enough (Table 1), the population is prevented from changing in size rapidly over short periods of time, and rounding errors are minimized (Figure 2) [15].

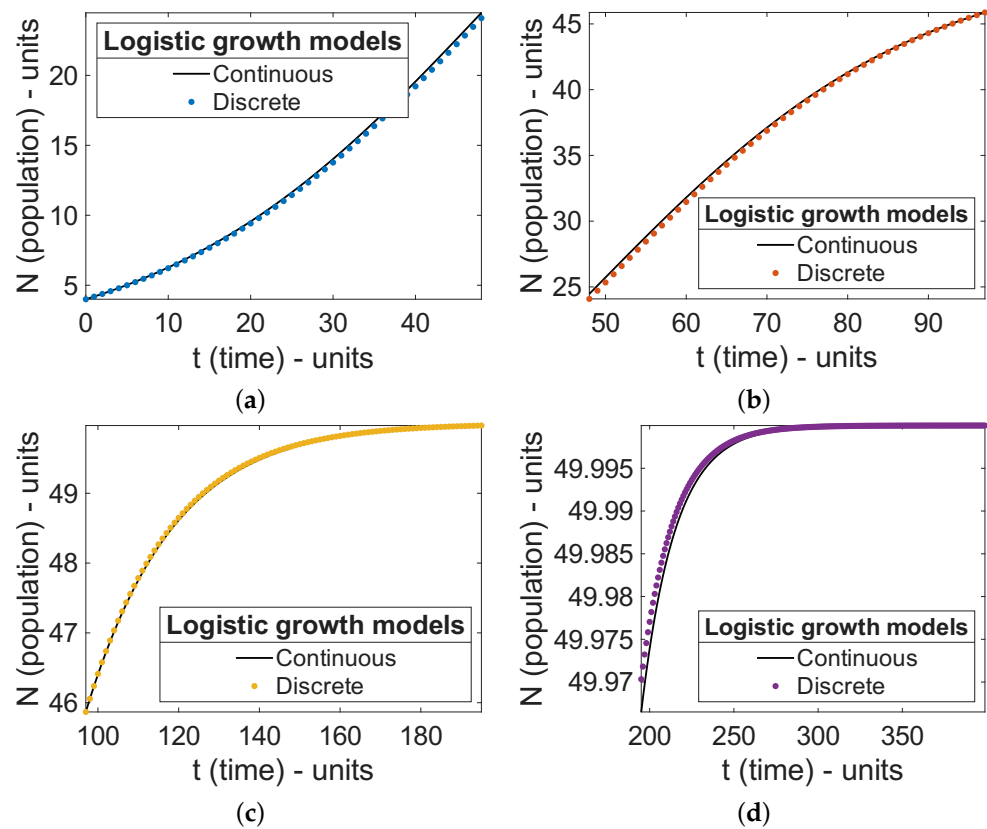


Figure 2. Approximation of continuous logistic growth within a discrete model in subsequent time frames. (a) Time range: $[0, 49]$; $t^* = 49$; (b) time range: $[49, 98]$; $2t^* = 98$; (c) time range: $[98, 196]$; $4t^* = 196$; (d) time range: $[196, 400]$; $8t^* = 392$. Discrete = continuous growth rate $r = k = 0.05$; initial population $N_0 = 4$; carrying capacity $L = 50$. Discrete population values, respectively, underestimate and overestimate continuous values in the time ranges (a,b) and (c,d); then, both model realizations converge to L .

3. Materials and Methods

3.1. Proposed Simulation of Logistic Growth

The simulation system is defined as a set of synthetic creatures placed on a remote island with limited resources, and the population’s growth is monitored over time. Two entities can be identified: the *environment*, which is *passive*, since it is represented by the constraint of the carrying capacity, and the *creature*, which is *active* and hence a *process*, being able to *reproduce*, causing the “birth” (or appearance) of another creature, and *die*, reducing the overall population. Reproduction and death represent the simulation *events*, and the *state variable* under observation is the current size N of the population as a density, with the unit area being an island, at each time step or iteration. Following the approach

of [16], we assume an *initial population* N_0 , and we use R and D to denote the per capita *reproduction* and *death rates* per unit time as a function of N , respectively. The simulation *activities* account for a member of the population (or creature) A reproducing, when they instantiate a copy of themselves, $A \rightarrow A + A$, and dying, when they disappear from the environment, $A \rightarrow \emptyset$. As defined in [16], the change in the population size (ΔN) during a time interval (Δt) is then given as

$$\Delta N = N_{t+\Delta t} - N_t = (R \Delta t N) - (D \Delta t N) \tag{10}$$

By letting $\Delta t \rightarrow 0$, we obtain the absolute growth rate, or the overall expected change of a population

$$\frac{dN}{dt} = (R - D) N \tag{11}$$

where $R - D$ is the net per capita *increase rate* per unit time [16], corresponding to the continuous growth rate k from Equation (1) of the Malthus model (see Section 2), which we approximate with the discrete growth rate r , as required by a discrete-time event simulation, deriving $R = D + r$.

In accordance with the criteria of limited growth, the population has to plateau: the expected change per time step has to become zero when the maximum number of creatures that the environment can support is reached. Equations (10) and (11) require an additional term to adjust the death rate, based on crowding:

$$\Delta N = N_{t+\Delta t} - N_t = (R \Delta t N) - (D \Delta t N) - (c \Delta t N^2) \tag{12}$$

$$\Delta t \rightarrow 0 : \frac{dN}{dt} = (R - D - cN) N \tag{13}$$

where the crowding coefficient c is equivalent to its theoretical counterpart, being a constant, and the term cN of Equation (13) is small in the presence of a limited number of creatures N and large when too many creatures are sharing the environment, matching Equation (5) of the logistic growth model (see Section 2). The increase rate updates to $R - (D + cN)$, where $D + cN$ represents the *death rate adjusted for crowding*, while the value of $R - D$ is kept fixed to r . In this way, the increase rate is zeroed when $c = \frac{R-D}{N}$ and $N=L$, in line with Equations (6) and (7) (Section 2), meeting the constraint of the carrying capacity L , redefined accordingly as $\frac{R-D}{c}$.

The proposed simulation is based on the mathematical model represented by Equations (12) and (13), with parameters R, D, c , and N_0 , treating R and $D + cN$ as *stochastic* components, by transforming them into continuous random variables that take on probability values. Time advances in steps, with a fixed increment of $\Delta t = 1$, and, at each timestep, for each creature that makes up the current population N , starting from N_0 , we draw R and $D + cN$ as samples from the uniform distribution. To account for the corresponding random processes of reproduction and death adjusted for crowding, we apply the *rejection sampling* method [17], which requires *threshold* probability values for comparison with the stochastic ones. Threshold values result from the exploration and tuning of the simulation parameters with respect to the mathematical model described in Section 3.3, and they were employed to test each creature with the following conditions.

- If the drawn probability for R is smaller than the corresponding threshold set for R , then the creature reproduces, introducing a new one in the system.
- If the drawn probability for $D + cN$ is smaller than its threshold, then the same creature dies. We note that the threshold probability value for $D + cN$ is dynamic, depending on the current value of N at each timestep.

- If both probability samples satisfy or do not satisfy their matching inequalities, then “nothing happens”: either the same creature reproduces, dies, and is replaced by its own copy, or it keeps living without causing any change.

After taking into account all current N creatures, the value of N is updated. As a consequence, the number of creatures that die is subtracted from the number of creatures that are born, leading to either an increase, decrease, or zero net change in the population for the next timestep.

The number of timesteps or iterations is an additional parameter of the system, since it is otherwise non-terminating: when the carrying capacity L is reached, the value of N becomes stationary. Depending on the research question and computational resources, users can arbitrarily interrupt the simulation. For example, they can focus on the short-term, logistic behavior of the system or on the long-term, steady-state behavior. Conforming to the method of *independent replications* [17], the number of simulations can also be set as a system parameter, complying with customized accuracy standards. To handle multiple simulation runs, we have set matching sequential prime numbers as seeds, employing the PRNG that implements the Mersenne Twister 1997 algorithm from the *Math.NET* Project (<https://numerics.mathdotnet.com/Random>, accessed on 14 February 2025).

To conduct a statistical analysis of the output, we ran the simulation 1000 times, with a specific parameter setting derived from the exploration and tuning steps of the pipeline, whose detailed description and reported results are, respectively, available in Sections 3.3 and 4.3. We mainly focused on three behaviors of the system (see Section 4.1):

- *short-term*, which represents the *sigmoid* phase, including lag, acceleration, and deceleration, also named *head*, with $t \in [t_0, 2t^*]$;
- *steady-state*, or *tail*, when the system hits the carrying capacity limit and becomes stationary, with $t \in [2t^*, t_{max}]$;
- *long-term*, which comprises both *head* and *tail*, accounting for the whole evolution of the system, with $t \in [t_0, t_{max}]$.

We captured the trend of the simulated growth through two main measures: the iteration mean $\mu(N_i)$ and the iteration variance $\sigma^2(N_i)$. The iteration (sample) mean $\mu(N_i)$ is computed, on iteration i , by summing over the set \mathcal{S} of simulations performed and dividing by $|\mathcal{S}|$:

$$\mu(N_i) = \frac{\sum_{j \in \mathcal{S}} N_{ij}}{|\mathcal{S}|} \quad (14)$$

where N_{ij} is the current population per simulation, or the value N_i of N at timestep i , corresponding to simulation j . Taken together, the iteration means are representative of population evolution, averaging N at a given time unit, across all the independent simulation runs, and they were used to compute raw residuals with respect to the two theoretical models used as mathematical references, continuous and discrete, to assess and compare the magnitude of the approximation errors caused by stochastic modeling and the employed time-driven approach. The iteration (sample) variance $\sigma^2(N_i)$ is equally computed on iteration i , summing over all simulations in \mathcal{S} , and dividing by $|\mathcal{S}| - 1$ to obtain an unbiased estimate:

$$\sigma^2(N_i) = \frac{\sum_{j \in \mathcal{S}} (N_{ij} - \mu(N_i))^2}{|\mathcal{S}| - 1} \quad (15)$$

Variability can thus be tracked at each timestep, allowing the discovery of spread patterns across the different phases that the system undergoes.

3.2. Comparison with the Gillespie Algorithm

In order to assess the statistical significance of the iteration mean $\mu(N_i)$ values obtained by our simulation, we evaluated how the same information, produced by the application of the Gillespie algorithm (exact method) [6] 1000 times, would compare to the continuous model and the corresponding raw residuals. One of the main differences between the proposed simulation and the state-of-the-art implementation of the Gillespie algorithm for logistic growth [18] resides in the interpretation of time: by taking finite unit steps in time (mirroring the discrete logistic master equation), we dictated that both reproduction and death may occur in the same step, updating the probability. However, these steps are not always synchronous. Birth (by reproduction) and death can be considered Poisson processes, meaning that the waiting time (τ) until either process occurs is exponentially distributed [19,20]. The underlying continuous-time Markov process has the following events and transition rates:

- Birth, $N \rightarrow N + 1$, with rate $\gamma = rN$ accounting for the left-hand side of the subtraction within the logistic equation $\frac{dN}{dt} = rN\left(1 - \frac{N}{L}\right)$.
- Death, $N \rightarrow N - 1$, with rate $\omega N = \frac{rN^2}{L}$, corresponding to the right-hand side, $\frac{dN}{dt} = rN\left(1 - \frac{N}{L}\right)$. This rate is dynamic, depending on the current value of N .

The corresponding transition probabilities are given as:

- $P(N \rightarrow N + 1) = \frac{\gamma}{\gamma + \omega N}$.
- $P(N \rightarrow N - 1) = \frac{\omega N}{\gamma + \omega N}$.

They are used for rejection sampling: if the sample drawn from the uniform distribution is smaller than $P(N \rightarrow N + 1)$, then the creature reproduces; otherwise, it dies. This makes the two events mutually exclusive [21], another major difference with respect to the proposed simulation, where they can happen at the same time, resulting in no net change. To determine how long (τ) it would take for a birth (or death) event to occur, a random number from an exponential distribution with a mean of $\frac{1}{\gamma + \omega N}$ is drawn (cf. theorem on Poisson arrivals). As a consequence, time points were not evenly spaced across the simulations, so we took the mean of the current N at each time point over the individual simulations and weighted each count by the length of time for which it existed [19]. In this way, we could compute the $\mu(N_i)$ values and analyze the distribution of raw residuals with respect to the continuous model for both the Gillespie simulation and the discrete master equation simulation that we chose to implement. In particular, we divided logistic growth into three time windows, relative to the short-term, steady-state, and long-term phases, to scrutinize the corresponding residuals' histograms and normal probability plots, adding valuable information for the comparison of the two simulation approaches: time-driven, like ours, and event-driven, as in the Gillespie algorithm.

3.3. Parameter Space Exploration and Tuning

We have performed the designed simulation 100 times, with different parameter settings, to explore the resulting scenarios with an unbiased approach and single out a configuration of parameters to conduct the output analysis upon and the comparison with Gillespie's method. We have kept r , or $R - D$, and the number of iterations fixed to 0.05 and 500, respectively: lowering the discrete growth rate to 0.02 or 0.01, to obtain a percentage error closer to 0%, implied increasing the simulation execution time, so we chose the first r greater than 0.02 from Table 1, while the selected number of iterations showed all the phases of logistic growth, including the long-term behavior.

We chose the ranges of the initial population (N_0), crowding coefficient (c), and reproduction (R) or death ($D = R - r$) probability to be $1 \leq N_0 \leq 51$, $10^{-5} \leq c \leq 10^{-4}$,

and $0.05 \leq R \leq 1$ ($0 \leq D \leq 0.95$), with matching steps of 2, 10^{-5} , and 0.05. In this way, the theoretical L shifts from 5000 to 500 creatures (by one order of magnitude), and the parameter space dimension is 26 (values of N_0) \cdot 20 (values of R) \cdot 10 (values of c) \cdot 500 (number of iterations), for a total of 5200 different combinations of parameter settings and $2.6 \cdot 10^6$ values of $\mu(N_i)$, or iteration means.

After restricting the parameter space dimension to the last 200 iterations, when growth is already stationary, we take the (grand) mean of the corresponding $\mu(N_i)$ values, for each parameter setting, to obtain the observed L values (95% confidence intervals (computed with the bootstrap method) of the 5200 observed L values are given in Supplementary Table S1). The scaled absolute difference ΔL between the observed and theoretical L was used as a measure of the distance of the stochastic model from the mathematical one, and the interplay between the simulation parameters c , N_0 , R , and D was analyzed to understand their impacts on the ability of the population to reach the theoretical L . In particular, tuning consisted of the observation of the pairwise relationships between ΔL as a percentage and each parameter evaluated regarding its range of values, seeking to choose the one that minimized the gap between the theoretical and observed L . The resulting R , D , and c values were used to set the corresponding R and $D + cN$ threshold probability values of the rejection sampling process, as described in Section 3.1. We note that the tuning of N_0 ensures that the simulation eventually reaches steady-state behavior, by setting the initial N of the $D + cN$ threshold probability value.

We have also performed a small-scale test of the PRNG by analyzing 100 different streams of data points, matching 100 single simulation runs, each containing 500 iterations, multiplied by the two probabilities applied for R and $D + cN$, respectively, and by the number of creatures alive per iteration, which we set to 100, for a total of 100,000 data points per stream. To check the absence of correlations between different streams, we reviewed their pairwise Spearman rank correlation coefficient matrices and r Pearson coefficients, while the one-sample Kolmogorov-Smirnov and chi-squared goodness-of-fit tests were performed on each stream to see whether it followed a uniform distribution.

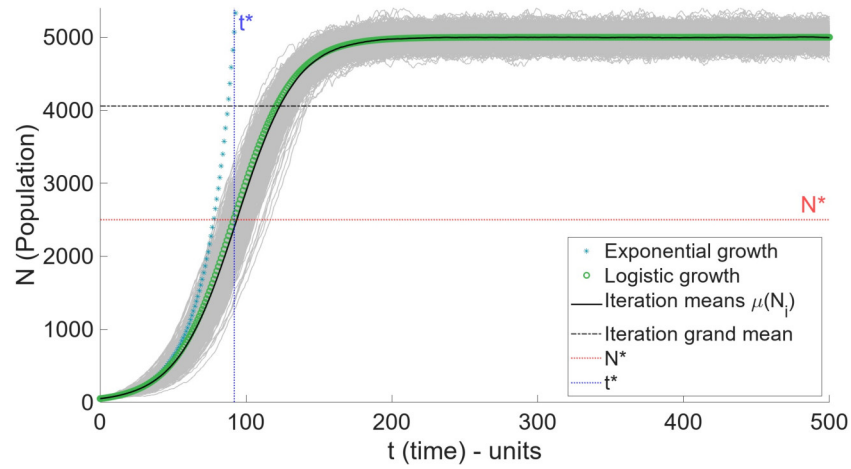
4. Results and Discussion

4.1. Proposed Simulation of Logistic Growth

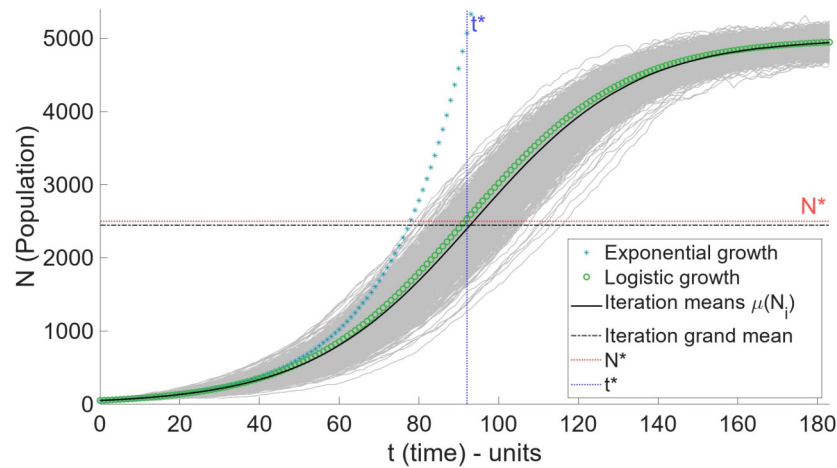
We ran the simulation system 1000 times, with the parameter settings that resulted from the exploration and tuning phase of the pipeline, as thoroughly explained in Section 4.3. We focused on both the long-term (Figure 3a) and short-term (Figure 3b) behaviors of the system, computing the $\mu(N_i)$ values to capture the trends of the simulated growth. The iteration grand mean of the long-term plot (~ 4059) moved closer to $L = 5000$ (the carrying capacity), while it was almost at the inflection population $N^* = 2500$ (~ 2447) in the short-term plot.

A closer look at the distribution of the $\mu(N_i)$ and $\sigma^2(N_i)$ values (95% confidence intervals for each $\mu(N_i)$ and $\sigma^2(N_i)$ were computed with the bootstrap method and are given in Supplementary Tables S2 and S3) in the two scenarios through the box plots in Figure 4a,b, respectively, reveals the impact of the tail on the iteration means: when the tail is cut, their median shifts from 4994 to 2366, matching the iteration grand mean decrease towards the inflection population. The number of outliers moves from 99 out of 501 data points (19.76%) to 0 in the head. This is due to the iteration means belonging to the exponential phase of the logistic growth, which appear as outliers with respect to the means calculated when the population approaches L . The variance has a similar pattern, with 151 outliers out of 501 data points (30.14%) in the long-term behavior and 0 in the short-term case. However, the median of the $\sigma^2(N_i)$ values increases from 9648 to 167.70

and the 75th percentile stretches from 102.20 to 593.80 (almost sixfold). This variability seems to be flattened by values approaching L from above or below at steady state.



(a) **Long-term behavior (head+tail).** The steady-state phase (*tail*) starts approximately after $2t^*$ and it is terminated by the simulation within 501 iterations. *Specific settings:* iterations = 501; iteration grand mean = 4058.85.



(b) **Short-term behavior (head).** The lag, acceleration, and deceleration phases, included in the time range $[0, 2t^*]$, are displayed, giving rise to the familiar symmetrical S-shaped curve. *Specific settings:* iterations = 184; iteration grand mean = 2446.72.

Figure 3. Logistic growth: simulation against theory. (a) Long-term behavior (*head+tail*); (b) short-term behavior (*head*). *Parameter settings:* simulations = 1000; theoretical inflection population $N^* = 2500$; theoretical inflection time $t^* = 92$; $R = 0.1$; $D = 0.05$; $c = 10^{-5}$; $N_0 = 51$.

Raw residuals, derived from the comparison of our simulation with the two grounding theoretical models, continuous (CM, Figure 5b) and discrete (DM, Figure 5c), were computed by the subtraction of the respective N_i values at the current time t from the corresponding value of $\mu(N_i)$ to look for heteroskedasticity. A similar pattern of variation emerged through a roughly symmetric V shape in the short term, which converged around $\Delta N = 0$ in the long term, with one main difference: the minima were approximately -138 in CM and almost -48 in DM, in line with the underlying simulation implementation. From the analysis of the 95% confidence intervals of the $\mu(N_i)$ values, we can globally state that a greater number of population values from DM falls into the $\mu(N_i)$ CIs than from CM (Figure 5d), and, locally, considering subsequent time frames (Figure 5e,f), the DM

population values fit better than the *CM* ones in both the short-term and long-term ranges, especially in the latter. We also noted that both the *CM* and *DM* population points ranging from t_0 to $2t^*$ had a similar number of true and false outcomes with respect to the iteration means, validating the symmetry arising from the raw residuals.

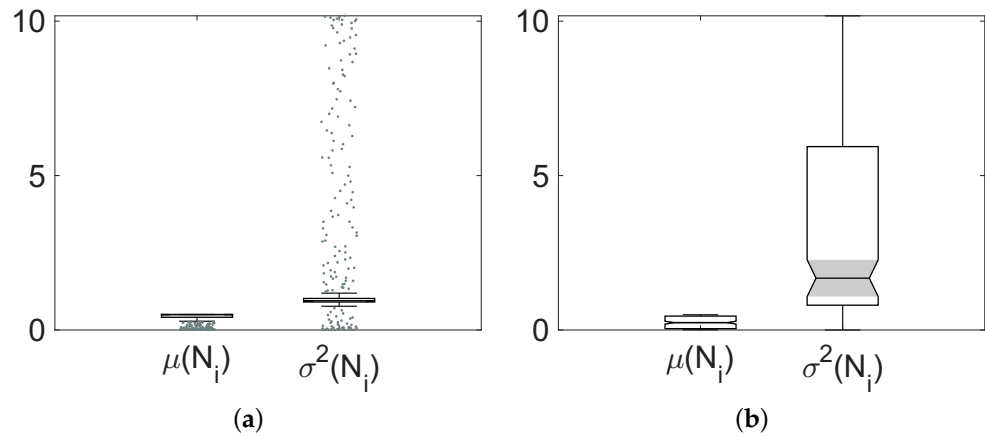


Figure 4. Box plot of the iteration means $\mu(N_i)$ and variances $\sigma^2(N_i)$. (a) Long-term behavior: $t \in [t_0, t_{max}]$. The steady-state phase of logistic growth ($t \in [2t^*, t_{max}]$) is included, turning the iteration means of the exponential phase into outliers, while the median of $\mu(N_i)$ approaches L . (b) Short-term behavior: $t \in [t_0, 2t^*]$. The steady-state phase is excluded, resulting in the absence of outliers, while the median of $\mu(N_i)$ decreases towards the inflection population. The median of $\sigma^2(N_i)$ increases from (a) to (b), where the sigmoid symmetry is restored. Values on the y axis need to be multiplied by 10^4 .

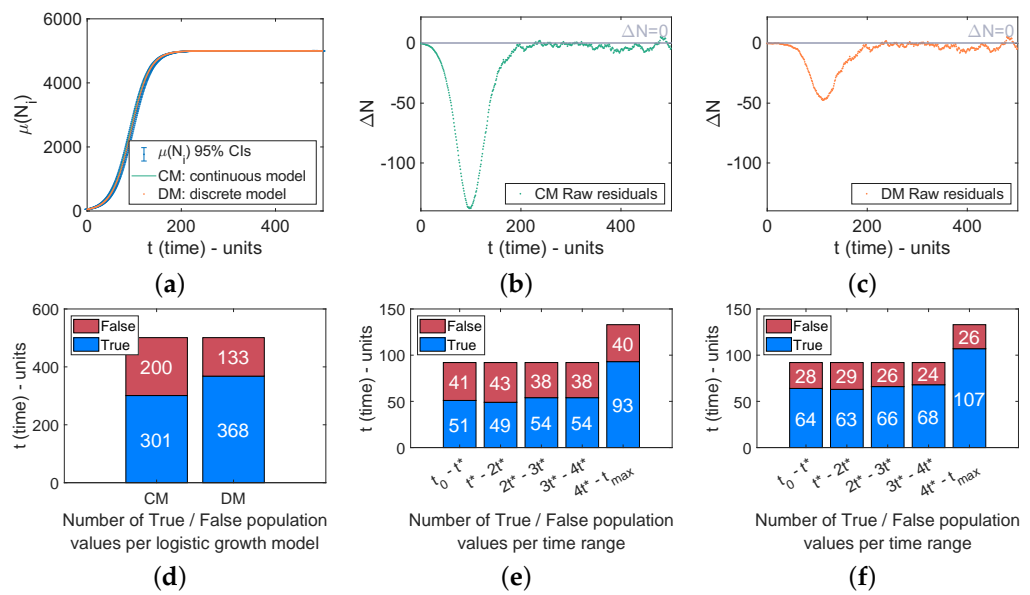
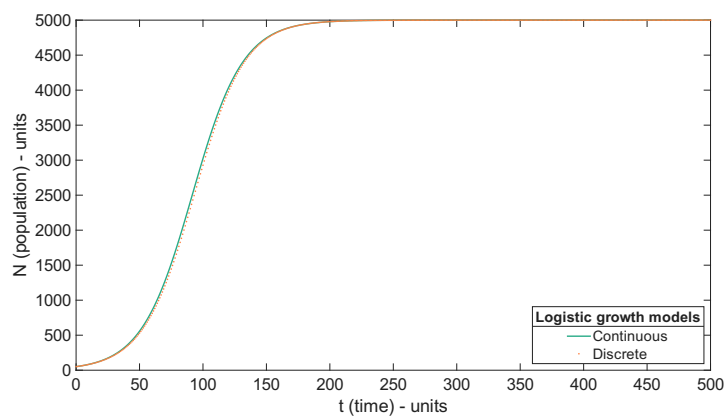
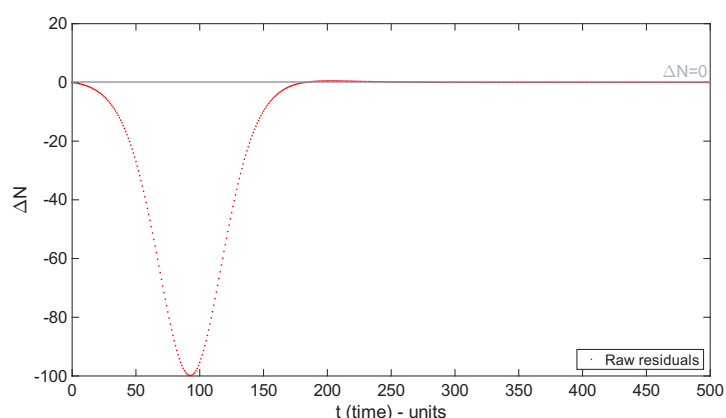


Figure 5. Comparison between continuous (*CM*) and discrete (*DM*) model fittings of the iteration means. (a) Overview of *CM* and *DM* fitting of the experimental $\mu(N_i)$ values; (b) difference in population (ΔN) between $\mu(N_i)$ and *CM* per time unit; (c) difference in population (ΔN) between $\mu(N_i)$ and *DM* per time unit; (d) comparison of *CM* and *DM* numbers of population (N) values belonging (true) or not (false) to $\mu(N_i)$ 95% CIs; (e) same comparison, within *CM*, per time range; (f) same comparison, within *DM*, per time range.

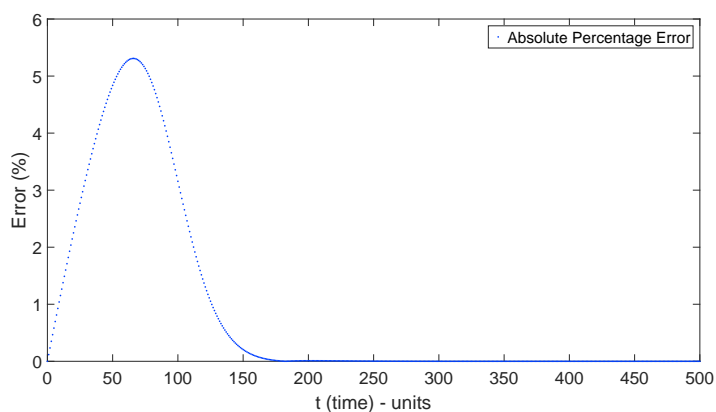
These results led us to directly compare the two theoretical models, so we plotted the ΔN values between the *CM* and *DM* values of N across time (Figure 6b), causing the symmetrical V shape, characteristic of the two models, to appear with a minimum of approximately -100 at iteration $i = 93$.



(a) **Overview of the two theoretical models.** It can already be seen that the orange dots representing the discrete model do not fall inside the corresponding green line of the continuous model, creating gaps.



(b) **Difference in population (ΔN) between DM and CM.** Raw residuals are computed for each timestep t , subtracting the CM current population N from the DM current N , obtaining an absolute difference between the two models.



(c) **Absolute percentage error between DM and CM.** Residuals are normalized for each timestep t through the use of the APE_i metric, obtaining a relative difference between the two models, shifting the focus to the error pattern.

Figure 6. Comparison of continuous (CM) and discrete (DM) deterministic models of logistic growth. Parameter settings: discrete = continuous growth rate $r = k = 0.05$; carrying capacity $L = 5000$; initial population $N_0 = 51$; time range $[0, 500]$; inflection time $t^* = 92$.

Up to t^* , DM increasingly underestimates CM, in harmony with an exponential phase of growth, which causes DM to round N down to the nearest integer; afterwards, the difference is gradually re-absorbed during the deceleration phase of logistic growth;

at steady state, $\Delta N = 0$ for both models. As a consequence, heteroskedasticity in the raw residuals of the iteration means does not imply a simulation implementation-related error but is a constitutive characteristic of our discrete-event simulation trying to keep up with continuous growth. Figure 6c represents the same information portrayed by Figure 6b, but using a different metric that interprets residuals as rounding error results, derived from DM approximating CM (as also shown in Figure 2), through the APE_i or *iteration absolute percentage error*, which reflects the absolute value of the difference between the current population N_i of the discrete model DM , named $N_{i,DM}$, and N_i of the continuous model $N_{i,CM}$, divided by $N_{i,CM}$. The error is then multiplied by 100 to obtain the matching percentage:

$$APE_i = 100 \frac{|N_{i,DM} - N_{i,CM}|}{N_{i,CM}} \tag{16}$$

In this way, we can compare DM with CM across time units, taking into account the current population N_i of the mathematical model underneath the proposed simulation, normalizing each deviation with respect to its continuous reference. We thus obtain a clearer representation of the error pattern, given that the residuals are now properly scaled, differently from Figure 6b, where the residuals have their raw values. The APE_i reaches its maximum value at iteration $i = 66$, where $N_{i,DM} = 1091.91$ and $N_{i,CM} = 1033.94$ (Figure 6a), with an error of 5.31%, while $\Delta N = -57.97$ (Figure 6b), which is not the greatest raw deviation from CM . This additional metric shows a slightly different V shape that loses its symmetry in favor of an initial linear deviation trend. Rounding errors are absorbed earlier in time and stabilization around the carrying capacity happens similarly, shedding more light on the constitutive discrepancy between the two theoretical models of logistic growth.

4.2. Comparison with the Gillespie Algorithm

An overall view of population growth according to the theoretical continuous model (CM) superimposed with the weighted iteration mean $\mu(N_i)$ values of 1000 Gillespie simulation runs is given in Figure 7a. The number of Gillespie iterations required to match the tail behavior of CM , stopping at 501 iterations as in our simulation, was around 200,000: as explained in Section 3.2, with Gillespie’s being an event-driven approach, its iterations actually represent the number of events that ensures that the stopping time mean across all the performed simulations is approximately 501, aligning with the CM and proposed simulation unit timesteps, while also allowing an easier graphical assessment. The distribution of raw residuals, shown in Figure 7b, is computed by subtracting the CM value of N_i at the current time t from the corresponding Gillespie $\mu(N_i)$ value (95% CIs are not available for Gillespie iteration means due to the time and computational costs of the bootstrap method applied to the required number of iterations). The maximum raw residual of Gillespie is around -27 (Figure 7a), a value lower than that in our simulation against the underlying discrete model and in the theoretical continuous model against its discrete version. This may be explained by Gillespie correctly accounting for the inherent fluctuations and correlations that are necessarily ignored in the deterministic formulation. In addition, this algorithm never approximates infinitesimal time increments dt by finite timesteps Δt , in harmony with population levels that change suddenly and sharply with time [6]. However, the similar V shape in the short-term residuals (Figure 7b) confirms the presence of an unavoidable bias, probably depending on the intrinsic symmetry of the sigmoid.

To further investigate this discrepancy in the residuals, we divided logistic growth into the three time windows of short-term or head $[t_0, 2t^*]$, long-term or head + tail $[t_0, t_{max}]$, and only steady-state or tail $[2t^*, t_{max}]$, respectively, given how the iteration means and variances of the proposed simulation behaved in the first and second windows (cf. Fig-

ure 4 of Section 4.1). The distribution of the raw residuals was observed for Gillespie through the histogram plots of Figure 7c–e. Figure 8a–c refer to our simulation. The same rationale was applied to the normal probability plots: Figure 7f–h are relative to Gillespie and Figure 8d–f match the proposed simulation.

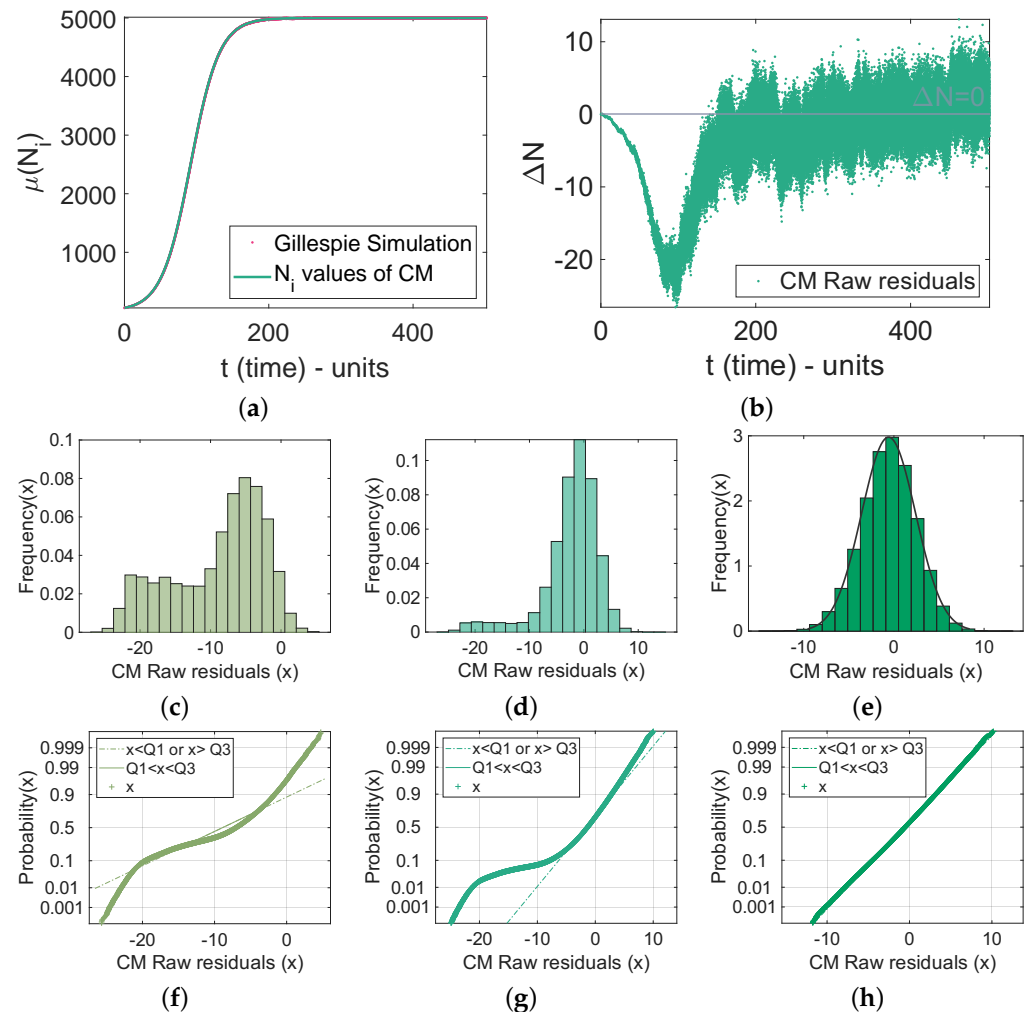


Figure 7. Gillespie simulation. (a) Overview of the CM fitting of the experimental Gillespie $\mu(N_i)$ values; (b) difference in population (ΔN) between Gillespie $\mu(N_i)$ values and CM per time unit; (c) CM histogram plot of the *head* (short-term) raw residuals, time range: $[t_0, 2t^* (\sim 184)]$; (d) CM histogram plot of the *head+tail* (long-term) raw residuals, time range: $[t_0, t_{max} (\sim 501)]$; (e) CM histogram plot of the *tail* (steady-state only) raw residuals, time range: $[2t^*, t_{max}]$; (f) CM normal probability plot of the *head*, NRSME indicator: 0.01; (g) CM normal probability plot of the *head+tail*, NRSME indicator: 0.009; (h) CM normal probability plot of the *tail*, NRSME indicator: 0.38.

Regarding our simulation of the discrete master equation, the short-term histogram (Figure 8a) shows greater multi modality than the long-term histogram (Figure 8b), reflecting the deletion of tail $\mu(N_i)$ values. The long-term histogram (Figure 8b) is indeed extremely left-skewed, and the smaller bars match the sigmoid growth outliers. The corresponding Figure 8d,e normal probability plots resemble a symmetric distribution of residuals, as the V shape of Figure 5b and the significant horizontal deviation from the reference line of Figure 8e are in agreement with the long histogram tail. The short-term and long-term residuals of our simulation do not exhibit a normal distribution. Only the last window (Figure 8c,f) seems to have a close-to-normal distribution.

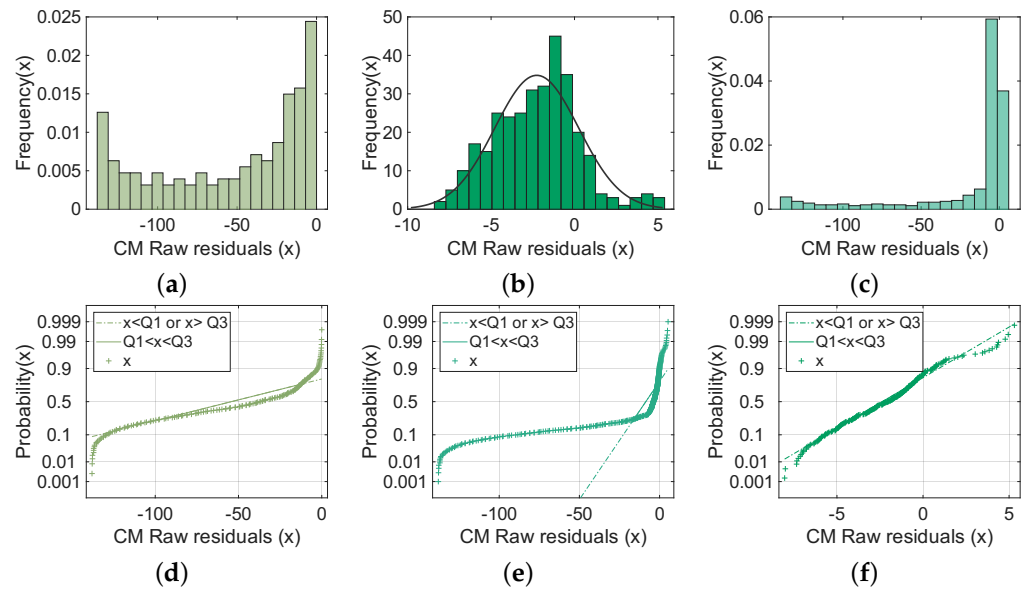


Figure 8. Discrete master equation simulation. (a) CM histogram plot of the *head* (short-term) raw residuals, time range: $[t_0, 2t^*(184)]$; (b) CM histogram plot of the *head+tail* (long-term) raw residuals, time range: $[t_0, t_{max}(501)]$; (c) CM histogram plot of the *tail* (steady-state only) raw residuals, time range: $[2t^*, t_{max}]$; (d) CM normal probability plot of the *head*, NRSME indicator: 0.04; (e) CM normal probability plot of the *head+tail*, NRSME indicator: 0.03; (f) CM normal probability plot of the *tail*, NRSME indicator: 0.39.

Regarding the Gillespie $\mu(N_i)$ raw residuals, noting that the number of points under scrutiny is much greater than in our simulation, their distribution in Figure 7e,h is consistent with normality. Moreover, the long-term histogram’s shape (Figure 7d), aside from the bottom-left tail clearly visible from the normal probability plot, tends toward normality. Instead, the short-term histogram (Figure 7c) is bimodal, probably in concordance with the acceleration/deceleration phases of logistic growth (a mixture of Gaussians could explain the bimodality of the frequency distribution). Not plotting the distribution of residuals and rather relying only on the NRMSE indicator would have been misleading, as the NRMSE compares the iteration means with the background continuous model population values. Indeed, a low NRMSE does not exclude the possibility that a model is biased [22]. We find that the continuous-time approach of the Gillespie algorithm yields more statistically significant $\mu(N_i)$ values, but does not adjust for the inherent bias of the logistic growth continuous model: being time-series data, the error terms are correlated in the sigmoid phase of growth.

Two main limitations of the proposed simulation emerged from the comparison with the Gillespie implementation. The first one is synchronous updating: by checking whether or not the event of reproduction or death happens in each timestep t of length Δt , we simultaneously increase or decrease the N_i counter and save the changes that take place at the end of the current iteration i , making the two events mutually inclusive [23] (see Section 3.1). For example, u creatures are born and v die at timestep t ; if $u > v$, then the counter is updated by adding $u - v$; otherwise, if $u < v$, then $v - u$ is subtracted from N_i . If $u = v$, the update is zeroed by the subtraction and addition of the same amount of creatures. As a consequence, during rejection sampling, for each creature, the drawn probabilities for the reproduction rate R and death rate adjusted for crowding, $D + cN$, against their threshold probabilities do not take into account the other creatures that were born or died in the meantime, including the possible reproduction of the creature under current testing. In this way, a bias is introduced, which Gillespie resolves by adopting an event-driven design choice, making the two events mutually exclusive, as previously described

in Section 3.2. The second limitation involves the time resolution: Δt is implicitly set to 1, so the system cannot record sudden or subtle changes, being unable to fulfill a potentially useful requirement of biological systems. In contrast with the Gillespie algorithm, mapped over a continuous-time Markov process, whose simulation clock jumps from event time to the next event time, the results of our rejection sampling are always approximate, given that time is discretized [23] and periods of inactivity cannot be skipped over [24], since each Δt can be compared to a “black box” where multiple overlapping events can happen. For example, if N_0 is not large enough (see Section 4.3), the population might become extinct very rapidly, with the short-term behavior of the system being absorbed into a constant zero line, without the system being able to record its decline, because it could happen in the same Δt . The opposite extreme condition, i.e., if N_0 is too large (see Section 4.3), might imply fast stabilization around the carrying capacity limit, causing the subsumption of the short-term by the long-term behavior, without the possibility to track it while in progress.

A notable advantage of the proposed simulation is a consequence of the time resolution limitation: the number of iterations required to reach the carrying capacity, leading to the steady-state phase of the system, is significantly smaller than the exact Gillespie iterations. Performing 501 steps versus 200,000 is less demanding overall from the computational resources and time perspectives: case studies necessitating the observation of large populations’ evolution over indefinitely long periods of time, which focus on trends and general patterns rather than fine-grained changes, could effectively benefit from the reduced weight placed on resource allocation, especially if the simulation is predicted to be run for a long time period. Furthermore, simulation adherence to the theoretical model, which is more accurate in Gillespie’s simulation, is traded off for statistical consistency in the output analysis, as performed by the proposed simulation: increasing the number of independent replications leads to a larger number of reproducible results to analyze and better stochasticity management. Finally, the procedure implemented to evenly space the Gillespie timesteps does not facilitate an average behavior assessment across simulations and with respect to the underlying *CM*, in contrast with our simulation, which does not require further computation after all simulation runs. The timesteps are predefined, conforming to the master equation (*DM*), and it can be straightforwardly compared with *CM* by feeding the sigmoid function with integer timesteps.

4.3. Parameter Space Exploration and Tuning

To visually determine how each simulation parameter affected the distance of the stochastic model from the mathematical one, we produced a four-dimensional plot, represented by Figure 9, where points represent the absolute value of the difference (ΔL) between the theoretical and observed carrying capacity L , scaled by the theoretical L , for each setting of N_0 , c , and $R(= D + 0.05)$. In this way, we obtained weighted differences, comparable with each other, given that the theoretical L ranged from 5000 to 500. We recall that the last 200 iterations out of 500 were used to compute the observed L , i.e., the grand mean of 200 $\mu(N_i)$ values. The range of ΔL values, depicted by color shading as the 4th dimension, is in logarithmic scale to accentuate the “yellowish” half-circle pattern that can be detected in each vertical section created by c , suggesting that lower values of N_0 lead to a higher ΔL , independently of c and R . This could be explained by the exponential growth phase: if not enough creatures exist, reproduce, and balance out the dead ones, the population could become extinct very quickly. Out of 100 simulations, the majority could reach 0 as the current N_0 before even reaching the inflection time. Extreme values of R (blue dots), or at least those symmetrically distanced from the center values (e.g., values far from $R = 0.5$ and $D = 0.45$), seemed to imply the shortening of ΔL towards 0, while no significant effect was seen on the outcome because of parameter c ’s variation.

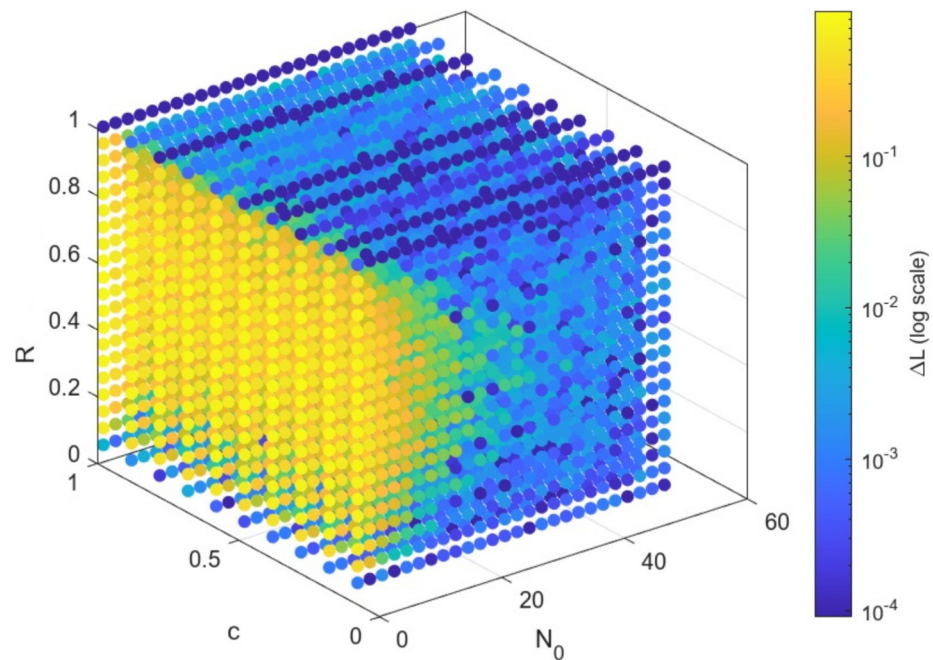


Figure 9. Parameter space hypercube. Each 4-dimensional point represents a single value of ΔL , the absolute value of the difference between the theoretical and observed carrying capacity L , scaled by the theoretical L , for each setting of N_0 , c , and $R(= D + 0.05)$, resulting in 5200 different combinations. Color shading matches the ΔL range: yellow and blue dots, respectively, account for low and high ΔL values. The use of a logarithmic scale eases pattern detection. *Parameter settings:* discrete growth rate $r = 0.05$; initial population $N_0 \in [1, 51]$, step = 2; reproduction probability $R \in [0.05, 1]$, step = 0.05; crowding coefficient $c \in [10^{-5}, 10^{-4}]$, step = 10^{-5} ; death probability $D \in [0, 0.95]$, step = 0.05; simulations = 100. Values on the c axis need to be multiplied by 10^{-5} .

The pairwise relationship of ΔL as a percentage with each parameter setting is illustrated by Figure 10. The scatter plot of ΔL against N_0 (Figure 10a) confirms that at least 30 creatures are necessary to obtain $\Delta L < 3\%$. Choosing $N_0 = 51$, the upper limit of the range, is a clear tuning outcome. For c and R , two box plots (Figure 10b,c) were produced to evaluate the distribution of ΔL within the respective ranges. Regarding c , the 75th percentile is approximately $\leq 3\%$ and the upper adjacent is $< 8\%$ for all ten settings, with a median that is very close to 0%, ranging between 0.13% ($c = 10^{-5}$; theoretical $L = 5000$) and 0.95% ($c = 10^{-4}$; theoretical $L = 500$). A greater L probably allows for smoother growth and a more stable steady state: the addition or subtraction of a single creature affects $L = 5000$ by $\pm 0.02\%$, compared with $\pm 0.2\%$ when $L = 500$, the lower limit of the range. Instead, the R variation shows a roughly symmetric pattern. The values of 0.5 and 0.55 record upper adjacent limits close to 20% and 75th percentiles $> 8\%$. Probabilities farther from the center yield shorter interquartile ranges. The extremes of $R = 0.05$ and $R = 1$ should be excluded because they, respectively, represent creatures not dying except for overcrowding and those always reproducing. The R value recording the lowest median (of 0.15%) is 0.95, followed by $R = 0.1$, with a median of 0.22%. We cannot choose R from an absolute biological perspective: we simulated the growth of an *in silico* population, so any R and D values are feasible. However, in a real world-scenario, $R > 50\%$ is not sustainable. Remembering that $R - D$ is kept constant and the growth rate never changes, this pattern may be due to the stochasticity of the model, so their setting cannot be tuned unambiguously or may not be tunable at all. We speculate that values of R closer to the

lower and upper extremes result in a lower probability of simulation failure (P_f), i.e., when the population becomes stuck at 0. For example, in a single-creature and single-iteration scenario, under a value of L so great that death by overcrowding becomes negligible (by approximation), $P_f = (1-R)D = (1-R)(R-0.05)$. If $R = 0.95$ or $R = 0.1$, then P_f is 0.045 for both; otherwise, if $R = 0.5$ or $R = 0.55$, P_f peaks at 0.225. For $R = 0.6$ or $R = 0.45$, $P_f = 0.22$, following the same symmetric pattern of $\Delta L(\%)$ from Figure 10c.

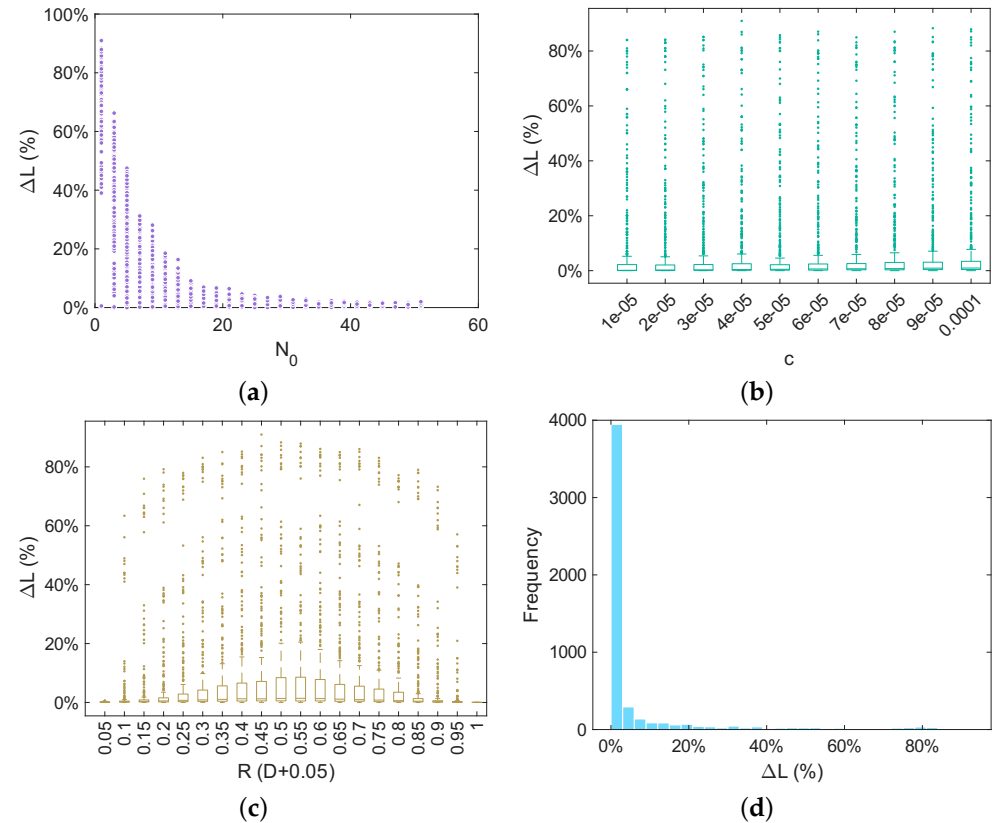


Figure 10. Evaluation of ΔL and the system variables. (a) Pairwise relationship between N_0 (initial population) and ΔL ; (b) box plot of ΔL for each c (crowding coefficient) value; (c) box plot of ΔL for each R (reproduction) and D (death) probability value; the ΔL gaps arise from the inclusion of all N_0 values in the range $[1, 51]$; (d) frequency distribution of ΔL values across all parameter settings. ΔL is shown as a percentage to facilitate visual comparison across plots. $\Delta L(\%) = \frac{\text{observed } L - \text{theoretical } L}{\text{theoretical } L} \cdot 100$; $R = D + 0.05$.

We thus decided to choose the setting of $R = 0.1$ to proceed with the pipeline of the output analysis. (The value of R could be replaced by any other setting to carry out the output analysis, probably yielding different results. Moreover, the same output analysis could be applied to all values in the R range in parallel for further comparison). This is a realistic option that also minimizes ΔL , while the choice of $N_0 = 51$ and $c = 10^{-5}$ is motivated by the tuning results.

The frequency distribution of ΔL across all simulations (Figure 10d) validated the overall consistency of the stochastic model. Additionally, the outcome of the small-scale PRNG testing on the 100 different streams of data points corresponding to 100 single simulation runs indicated that 8 out of 10,000 possible stream couples of the Spearman rank correlation coefficient matrix, excluding the main diagonal, resulted in $r_s \geq 0.01$ or $r_s \leq -0.01$; the same results were also obtained with the r Pearson coefficient. The one-sample Kolmogorov–Smirnov and chi-squared goodness-of-fit tests showed that 96 and 95 streams of random numbers, out of 100, followed a uniform distribution, corroborating the statistical soundness of the simulation system.

5. Conclusions

We have successfully designed a generalizable and reproducible pipeline of work for logistic growth modeling, simulation, tuning, and performance evaluation and highlighted the differences between the implementation of a time-driven simulation, such as ours, which was mapped from the discrete master equation, and the implementation of an event-driven simulation, such as Gillespie's, which yielded more statistically significant results, when compared to the continuous model of logistic growth. We were able to distinguish the pitfalls of our simulation from the intrinsic characteristics of the mathematical models thanks to an in-depth analysis that only an interdisciplinary approach, combined with a clear didactic purpose, could foster.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app15084409/s1>.

Author Contributions: Conceptualization, C.P. and S.F.; Methodology, C.P. and S.F.; Software, C.P. and S.F.; Formal analysis, C.P.; Resources, P.C.; Data curation, S.F.; Writing—original draft, C.P. and S.F.; Writing—review & editing, P.C.; Visualization, C.P. and S.F.; Supervision, P.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the presented results can be fully reproduced on both simulation and output analysis sides, since the code is available under GPL-3.0 license in the GitLab Manta UniTn repository named *Simulation-of-Logistic-Growth* (Repository web URL: <https://git.manta.disi.unitn.it/stefano.faccio/simulation-of-logistic-growth>, accessed on 3 April 2025), by executing the *C# project* and *Matlab* files. The files are organized as follows: the folder *Parameter space exploration* contains all the code used in Sections 3.3 and 4.3 while the folder *Focus on Single Simulation* refers to Sections 3.1 and 4.1. The PRNG testing can be found into folder *Pseudo Random Number Generator Testing*, also included at the end of Sections 3.3 and 4.3 and the Supplementary Tables S1–S3 can be found in the folder *Supplementary Tables*.

Conflicts of Interest: The authors declare no conflicts of interest.

Notation

The following table summarizes the notation used in this manuscript.

Notation	Definition	Description
N	Population value	Generic reference to the number of creatures in the designed environment
N_0	Initial population	Number of creatures at the first time-step t_0 of the model or simulation
N_{max}	Final population	Number of creatures at the maximum (or last) time-step t_{max} of the model or simulation
N^*	Inflection population	$N^* = L/2$ is the x coordinate of the inflection point, obtained by setting the second derivative of the sigmoid function to zero: $N_t'' = 0$
N_i	Current population	Number of creatures at the current time-step t or iteration i of the model or simulation

N_{ij}	Current population per simulation	N_i corresponding to simulation j
N_t	Continuous model of growth	N_t is used to map both the exponential and logistic equation (analytical solution), to highlight the transition from theoretical, unlimited to limited growth
N_{t+1}	Discrete model of growth	N_{t+1} is used to map both discrete versions of the exponential and logistic equation (master equation), to highlight the transition from theoretical, unlimited to limited growth
t	Current time-step or time-unit (time)	Current iteration i of the model or simulation
t_0	First time-step	First iteration i of the model or simulation
t_{max}	Maximum time-step	Last iteration i of the model or simulation
t^*	Inflection time	$t^* = \frac{1}{r} \ln\left(\frac{L-N_0}{N_0}\right)$ is the y coordinate of the inflection point, obtained by setting the second derivative of the sigmoid function to zero: $N_t'' = 0$
$\frac{dN}{dt}$	Absolute growth rate	Population per unit of time-step t
Δt	Delta time	Unit time interval or length of each time-step t : $\Delta t = 1$
ΔN	Delta population	This abbreviation accounts for a difference between two population N values, depending on the paper context: <ul style="list-style-type: none"> • change of population size, as numerator of the absolute growth rate, or the overall expected change of a population $\frac{\Delta N}{\Delta t}$ (Sections 2 and 3.1) • difference between our simulation and DM values of N_i for each time-step t (Section 4.1) • difference between our simulation and CM values of N_i for each time-step t (Section 4.1) • difference between DM and CM values of N_i for each time-step t, or $N_{i,DM} - N_{i,CM}$ (Section 4.1) • difference between Gillespie and CM values of N_i for each time-step t (Section 4.2)
k	Continuous relative growth rate	Percent of population per time-step t of the continuous model
r	Discrete relative growth rate	Percent of population per time-step t of the discrete model
R	Reproduction rate or threshold probability parameter	Per-capita reproduction rate per time-step t as a function of N (simulation model), turned into a threshold probability value for comparison with the corresponding stochastic component of the simulation implementation (rejection sampling)
D	Death rate or probability parameter	Per-capita death rate per time-step t as a function of N (simulation model), turned into a probability value that makes up the threshold probability of $D + cN$

$D + cN$	Death rate adjusted for crowding or threshold probability parameter	Per-capita death rate per time-step t as a function of cN to account for over-crowding (simulation model), turned into a threshold probability value for comparison with the corresponding stochastic component of the simulation implementation (rejection sampling)
L	Carrying capacity	The largest population that can be supported by the environment. We distinguish between <i>theoretical</i> L and <i>observed</i> L to respectively account for the model and simulation, given that the latter fluctuates above and below N_{max}
c	Crowding coefficient	Constant value accounting for the limiting resources of an over-crowded environment: <ul style="list-style-type: none"> • $c = \frac{k}{L}$ (continuous model) • $c = \frac{r}{L}$ (discrete model) • $c = \frac{R-D}{L}$ (simulation model)
$R - D$	Increase rate	Net per-capita increase rate per time-step t under unlimited (exponential) population growth: <ul style="list-style-type: none"> • $R - D = k$ (continuous model) • $R - D = r$ (discrete model and simulation model)
$R - (D + cN)$	Increase rate adjusted for crowding	Net per-capita increase rate per time-step t under limited (logistic) population growth
\mathcal{S}	Set of simulations	\mathcal{S} indicates the set of all simulation runs
$ \mathcal{S} $	Number of simulations	$ \mathcal{S} $ represents the cardinality of the simulation set, hence the total number of simulation runs
CM	Continuous Model of Logistic Growth	Abbreviation used to facilitate visualization on plots and readability on text, due to its recurrent use as the theoretical continuous reference of the proposed simulation
DM	Discrete Model of Logistic Growth	Abbreviation used to facilitate visualization on plots and readability on text, due to its recurrent use, being the master equation at the base of the proposed simulation
γ	Birth transition rate	Along with the event of birth ($N \rightarrow N + 1$), the transition rate is a property of the underlying continuous-Markov process. In particular, γ stands for the left hand-side of the subtraction within the logistic equation. $\gamma = rN$
ωN	Death transition rate	Along with the event of death ($N \rightarrow N - 1$), the transition rate is a property of the underlying continuous-Markov process. ωN stands for the left hand-side of the subtraction within the logistic equation. $\omega N = \frac{rN^2}{L}$
τ	Waiting time	How long it takes for a birth (or death) event to occur, according to an exponential distribution with mean $\frac{1}{\gamma + \omega N}$
$\mu(N_i)$	Iteration mean	The iteration (sample) mean averages N at a given time-step t , across all the independent simulation runs in set \mathcal{S} : $\mu(N_i) = \frac{\sum_{j \in \mathcal{S}} N_{ij}}{ \mathcal{S} }$

$\sigma^2(N_i)$	Iteration variance	The iteration (sample) variance tracks N variability at each time-step t , across all the independent S simulation runs, correcting for bias: $\sigma^2(N_i) = \frac{\sum_{j \in S} (N_{ij} - \mu(N_i))^2}{ S - 1}$
PE	Percentage Error	Difference between the discrete rate r , representing empirical growth, and the continuous growth rate k , standing for theoretical growth, as a percentage of k : $PE = 100 \frac{r - k}{k}$
APE_i	Iteration Absolute Percentage Error	For each iteration i , this metric computes the absolute value of the difference between the DM current population N_i ($N_{i,DM}$) and the CM current population N_i ($N_{i,CM}$) divided by $N_{i,CM}$. The resulting value is multiplied by 100 to get the corresponding percentage: $APE_i = 100 \frac{ N_{i,DM} - N_{i,CM} }{N_{i,CM}}$
$NRMSE$	Normalized Root Mean Square Error	This metric compares the iteration mean $\mu(N_i)$ values with the background continuous model (CM) population values. $NRMSE = \sqrt{\frac{\sum_{t=1}^{max} (\mu(N_i) - N_{i,CM})^2}{t_{max}}}$
PRNG	Pseudo Random Number Generator	Employment of the Mersenne Twister 1997 algorithm for the inclusion of the rejection sampling method into the development of the proposed Monte Carlo discrete-time event simulation

References

1. Miller, G.; Spoolman, S. *Essentials of Ecology*; Brooks/Cole: Pacific Grove, CA, USA, 2009.
2. Zhao, S.; Hu, I.; Lou, J.; Chong, M.K.; Cao, L.; He, D.; Zee, B.C.; Wang, M.H. The mechanism shaping the logistic growth of mutation proportion in epidemics at population scale. *Infect. Dis. Model.* **2023**, *8*, 107–121. [[CrossRef](#)] [[PubMed](#)]
3. Shen, C.Y. Logistic growth modelling of COVID-19 proliferation in China and its international implications. *Int. J. Infect. Dis.* **2020**, *96*, 582–589. [[CrossRef](#)] [[PubMed](#)]
4. Pelinovsky, E.; Kurkin, A.; Kurkina, O.; Kokoulina, M.; Epifanova, A. Logistic equation and COVID-19. *Chaos, Solitons 'I&' Fractals* **2020**, *140*, 110241. [[CrossRef](#)]
5. Hughes-Hallett, D.; Gleason, A.; McCallum, W. *Calculus: Single and Multivariable*, 6th ed.; Wiley: New York, NY, USA, 2012.
6. Gillespie, D.T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **1977**, *81*, 2340–2361. [[CrossRef](#)]
7. Pineda-Krch, M. GillespieSSA: Implementing the Gillespie Stochastic Simulation Algorithm in R. *J. Stat. Softw.* **2008**, *25*, 1–18. [[CrossRef](#)]
8. Berryman, A.A. On principles, laws and theory in population ecology. *Oikos* **2003**, *103*, 695–701. [[CrossRef](#)]
9. Malthus, T.R. *An Essay on the Principle of Population*; Number Malthus1798 in History of Economic Thought Books; McMaster University Archive for the History of Economic Thought: Hamilton, ON, Canada, 1798.
10. Académie royale des sciences, d.l.e.d.b.a.d.B. *Nouveaux mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*; L'Académie Royale de Bruxelles et de l'Université Louvain: Louvain, Belgium, 1845.
11. Verhulst, P.F. A Note on the Law of Population Growth. In *Mathematical Demography: Selected Papers*; Springer Berlin Heidelberg: Berlin/Heidelberg, Germany, 1977; pp. 333–339. [[CrossRef](#)]
12. Lipkin, L.; Smith, D. Logistic growth model. In *Convergence*; MAA: Washington, DC, USA, 2004.
13. Cavallini, F. Fitting a Logistic Curve to Data. *Coll. Math. J.* **1993**, *24*, 247–253. [[CrossRef](#)]
14. May, R. Simple Mathematical Models With Very Complicated Dynamics. *Nature* **1976**, *26*, 457. [[CrossRef](#)] [[PubMed](#)]
15. Gordon, S.P. Comparing the Discrete and Continuous Logistic Models. *PRIMUM* **2008**, *18*, 449–455. [[CrossRef](#)]

16. Takasu, F. Lecture 10: Logistic Growth Models. Available online: https://gi.ics.nara-wu.ac.jp/~takasu/lecture/old/archives/L10_logistic-growth.pdf (accessed on 28 June 2006).
17. Boudec, J. *Performance Evaluation of Computer and Communication Systems*; Computer and Communication Sciences; EFPL Press: Lausanne, Switzerland, 2010.
18. Pineda-Krch, M.; Cannoodt, R. *GillespieSSA: Gillespie's Stochastic Simulation Algorithm (SSA)*; R Package Version 0.6.2; 2022. Available online: <https://cran.r-project.org/web/packages/GillespieSSA/index.html> (accessed on 3 April 2025).
19. Chure, G. Tutorial 6: Stochastic Simulation of Constitutive Expression. 2018. Available online: https://www.rpgroup.caltech.edu/ncbs_pboc/code/t06_gillespie_algorithm.html (accessed on 2 July 2024).
20. Sánchez-Taltavull, D. Stochastic Modelling in Mathematical Biology. 2013. Available online: <https://www.ub.edu/simba/slides/simba130304-DSanchez.pdf> (accessed on 2 July 2024).
21. Policriti, A. Stochastic Simulation (and Gillespie's Algorithm). Available online: <https://users.dimi.uniud.it/~alberto.policriti/home/sites/default/files/bioinformatica-supe/Stochastic-Simulation-advanced.pdf> (accessed on 2 July 2024).
22. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*; Springer: Berlin/Heidelberg, Germany, 2013.
23. Masuda, N.; Vestergaard, C.L. *Gillespie Algorithms for Stochastic Multiagent Dynamics in Populations and Networks*; Cambridge University Press: Cambridge, UK, 2022. [[CrossRef](#)]
24. Sloot, P. 1 Introduction to Simulation and Modeling. 1.4 Model Execution: Event Driven vs. Time Driven. 2013. Available online: https://artemis.wszib.edu.pl/~sloot/1_4.html (accessed on 12 July 2012).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.