

Expected and Predicted Curves of Markov Stochastic Processes and a New Interpretation of Deterministic Models

Miloslawa Sokol (✉ milka@biol.uw.edu.pl)

Faculty of Biology, Biological and Chemical Research Centre, University of Warsaw

Research Article

Keywords: Markov process, Expected curve, Predicted curve, Birth-and-death process, Logistic model, Lotka-Volterra model, Genetic drift, Hardy-Weinberg law

Posted Date: May 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-468716/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Expected and Predicted Curves of Markov Stochastic Processes and a New Interpretation of Deterministic Models

Miłosława Sokół

*^aFaculty of Biology, Biological and Chemical Research Centre, University of
Warsaw, Zwirki i Wigury 101, Warsaw, 02-089, , Poland
[*]milka@biol.uw.edu.pl*

Abstract

Two curves characterizing a set of realizations of Markov stochastic process were analysed: the expected curve and the predicted curve. Although the definition of the predicted curve is new, curves with the same formulas are known in science as so-called deterministic models. This can be seen in the examples of the most well-known models in biology: logistics, Lotka-Volterra and Hardy-Weinberg. Attention was paid to the efficiency of calculating these both curves and their interpretation as the central tendency characteristics of Markov stochastic processes.

Keywords: Markov process; Expected curve; Predicted curve; Birth-and-death process; Logistic model; Lotka-Volterra model; Genetic drift; Hardy-Weinberg law

1. Introduction

All numerical random variables can be characterized by a "central tendency" – the central or typical value for their probability distribution. The most popular of these is an expected value, despite the fact that it does not always exist ([8], [20]). Numerical stochastic processes can be characterized in a similar way. Expected values of stochastic process are usually calculated in application articles ([11], [18], [3]).

In this paper, the multidimensional stochastic processes are considered, in which the states are elements of $X = \mathbb{N}^n \times \mathbb{R}^m$. However, it is convenient to separately discuss about discrete-valued (states in \mathbb{N}^n) and continuous-valued

(states in \mathbb{R}^m) processes. Therefore, four types of stochastic processes are being considered: discrete-time discrete-valued (stochastic chains), discrete-time continuous-valued, continuous-time discrete-valued and continuous-time continuous-valued.

Homogeneous processes are most often considered in mathematics. Meanwhile, non-homogeneous processes have greater application in biology due to the diurnal and seasonal rhythms of many natural phenomena. In addition, theoretical considerations of Markov stochastic processes are much easier for non-homogeneous stochastic processes as opposed to the special case: homogeneous processes. This is due to the fact that the integrals in the formulas for the transition probabilities unfortunately have the antiderivatives of homogeneous processes, but with extremely complex formulas, as described in [22]).

All known stochastic models that play a major role in biology have the Markov property. This assumption is accomplished in models, which recreate the course of events that occur in reality and the probabilities of these events at time $[t, t + 1)$ depend on the state of the system at time t ([1], [15], [7]). This can be done in all object-oriented models ([10], [12], [5]) because everything that happens in time step $[t, t + 1)$ depends on the state of the program at time t in these models. It is impossible to program such models in a different way. All past events can be considered only by programming some carrier of information about these events which exists at time t . Sometimes, a numerical variable without the Markov property is calculated using object-oriented models, but then it is usually the sum of the variables that form a multidimensional Markovian model.

All stochastic processes $(\xi_t)_{t \in \Upsilon}$ with the Markov property yield a set of realizations, which can be characterized by the expected realization. It is a function of:

$$\Upsilon \ni t \rightarrow E\xi_t \in X$$

It will be referred to as the expected curve. The expected curve is difficult to calculate even for simple stochastic processes. However, the set of realizations of the stochastic processes has a second good curve, which forms a recursive sequence or an easily computable differential equation. It will be referred to as the predicted curve. Its formal definition is the purpose of this paper. The second aim is to compare it with the expected curve.

In this paper, it will be shown that for the most well-known models used in ecology and evolution, the formulas for the predicted curves turn out to be the same as for the so-called deterministic population models. This fact

completely changes their interpretation. In addition, the formulas derived in this paper will allow the correct determination of stochastic and deterministic models of the same phenomenon.

2. Results

2.1. Basic definitions

Stochastic processes formed by object-oriented models illustrate how desirable it is to generalize the theory of Markovian stochastic processes.

A stochastic process is defined as a collection of random variables: $(\xi_t)_{t \in \Upsilon}$ indexed by a subset Υ of $[0, \infty)$ and having values in X (a set of states). A sigma-algebra $\sigma(X)$ with a measure μ is defined for subsets of X . For $X = \mathbb{N}^d$ or $X = \mathbb{R}^d$ this is a Lebesgue measure. Moreover, there exists a probability space $(\Omega, \sigma(\Omega), P)$, which allows the calculation of the probability of some subsets of the set Ω (set of the process realizations).

All Markov processes can be defined by highly generalized counterparts of the stochastic or intensity (state-transition) matrix or kernel. It is a multidimensional object containing time-dependent functions whose values are probabilities or instantaneous probability rates. Depending on the type of the stochastic process, the set Ψ of such functions is defined as:

- $\Psi = \{f : \Upsilon \rightarrow [0, 1]\}$ for discrete-time, discrete-valued stochastic processes.
- $\Psi = \{f : \Upsilon \rightarrow [0, \infty)\}$ for discrete-time, continuous-valued stochastic processes.
- $\Psi = \{f : \Upsilon \rightarrow \mathbb{R}; f \text{ is integrable}\}$ for continuous-time, discrete-valued stochastic processes.

Let X be a space of states with numbers. Thus, $X = \mathbb{N}^n$ for discrete-valued processes and $X = \mathbb{R}^n$ for continuous-valued processes. Let $X_0 \subseteq X$. Most often, the set X_0 is assumed to be bounded.

The generalizations of a stochastic matrix, a stochastic kernel, an intensity (state-transition) matrix and an intensity kernel to a multidimensional Markov stochastic process have the following form:

Definition 2.1. *Let $(\xi_t)_{t \in \Upsilon}$ be a discrete-time, discrete-valued Markov stochastic process. A stochastic matrix is a set of numbers formed by the function $p : X_0 \times X \rightarrow \Psi$ and $p(x, y)(t, t + 1) = P\{\xi_{t+1} = y | \xi_t = x\}$.*

This function has the property: $\forall_{x \in X_0} \sum_{y \in X} p(x, y)(t, t + 1) = 1$. Each function of this property defines the discrete-time, discrete-valued Markov stochastic process.

Definition 2.2. Let $(\xi_t)_{t \in \Upsilon}$ be a discrete-time, continuous-valued Markov stochastic process. A stochastic kernel is a set of numbers formed by the function $p : X_0 \times X \rightarrow \Psi$ and $\int_U p(x, y)(t, t + 1) dy = P\{\xi_{t+1} \in U - \{x\} | \xi_t = x\}$ for all U measurable subsets of X .

This function has the property $\forall_{x \in X_0} \int_X p(x, y)(t, t + 1) \leq 1$. Each function of this property defines the discrete-time, continuous-valued Markov stochastic process.

Definition 2.3. Let $(\xi_t)_{t \in \Upsilon}$ be a continuous-time, discrete-valued Markov stochastic process. An intensity matrix is a set of numbers formed by the function $q : X_0 \times X \rightarrow \Psi$ and $q(x, y)(t, t + 1) = \lim_{\Delta \rightarrow 0} \frac{P\{\xi_{t+\Delta} = y | \xi_t = x\} - P\{\xi_t = y | \xi_t = x\}}{\Delta}$.

This function has the properties: $\forall_{x \in X_0} \forall_{y \in X, y \neq x} q(x, y)(t) \geq 0$ and $\forall_{x \in X_0} q(x, x)(t) \leq 0$, and $\forall_{x \in X_0} \sum_{y \in X} q(x, y)(t) = 0$. Each integrable function of these properties defines the continuous-time, discrete-valued Markov stochastic process.

Continuous-time continuous-valued stochastic processes are rarely modelled, and therefore are not termed as such in this terminology. They are created in textbooks using the Wiener process ([6], [2]). However, the natural generalization of the Markov process described earlier is quite different. It should still be a process in which a finite number of random events can occur per unit of time. Yet these events can be vectors of real numbers, and the moments of these events can be any real number. Its realizations are not continuous. They are step functions with arbitrarily long and high step size. These processes have the intensity kernel which is an integrable and bounded function $q : X_0 \times X \rightarrow \Psi$ and each integrable and bounded function $q : X_0 \times X \rightarrow \Psi$ defines a continuous-time continuous-valued Markov stochastic process. Because no one such process is as known as processes described in this paper, the theory of them will be omitted.

The definitions 2.1 - 2.3 show that stochastic/intensity matrices/kernels may not be square. Every good programmer who writes object-oriented programs restricts the set of events for which probabilities or probability rates can be calculated. Otherwise, his program may hang, or in the worst case, produce erroneous results. Most often, when a state outside of X_0

is reached during the simulation, the simulation stops before the assumed completion time T_{max} .

At the beginning, the above generalizations due to technical limitations seemed irrelevant. Meanwhile, the past study proved that for each Markovian stochastic process with a probability or intensity matrix or kernel (corresponding to the above definitions), there exists a probability space that is completely consistent with the Kolmogorov's system of axioms, which evaluates the frequency of occurrence of certain realizations in the simulated process ([21], [23]). Moreover, the formulas for these probabilities were derived and they are used in this paper.

The class of Markov stochastic processes, having a stochastic matrix or an intensity matrix, a stochastic kernel or an intensity kernel, is very broad. Since each finite quadratic stochastic matrix or kernel can be extended to non-quadratic by making $p(x, y)(t, t + 1) = 0$ for $y \in X - X_0$ and each quadratic intensity matrix or kernel can be extended to non-quadratic ones by making $q(x, y)(t) = 0$ for $y \in X - X_0$, such class of Markov processes includes all the well-known processes discussed in the textbooks.

Sometimes calculation a stochastic or intensity matrix or kernel for an object-oriented model is not easy, but it is worth doing. They allow calculation of the expected curve $E\xi : \Upsilon \ni t \rightarrow E\xi_t \in \mathbb{R}^n$ and a similar concept, which is referred to as the predicted curve. Both curves allow the prediction of the behaviour of the model during the simulation and the possibility to limit the values of the model parameters to suit one's own needs.

2.2. *Expected and predicted curves in discrete-time discrete-value processes*

In this section, the concepts of expected and predicted curves are illustrated for a birth-and-death process in discrete time. However, they are defined in a general way. + A two-dimensional stochastic matrix in independent-dependent orientation (put into a Cartesian coordinate system) often has a characteristic structure. The probabilities of values slightly greater than 0 in each column are concentrated in a relatively small area. If the points of this matrix are coloured with different shades of gray depending on the probability, then a dark stripe is formed in the area of such matrix. This allows predicting the shapes of the stochastic process realizations started from the same state. For the discrete-time birth-and-death process, the dark stripe courses above the diagonal for small values and below its for large values of the population size Fig.1).

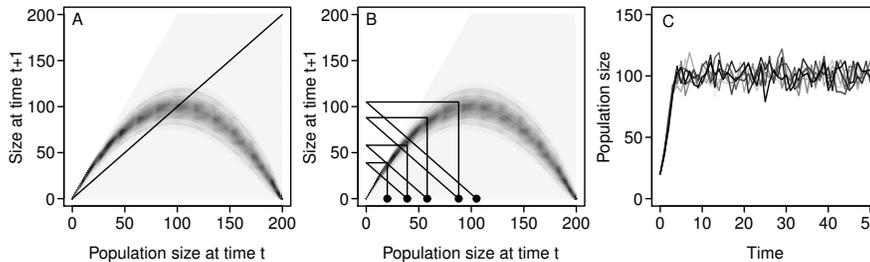


Figure 1: A. The stochastic matrix for the discrete-time birth-and-death process with the probability of birth of an individual depending on the population size according to the formula $p_r(k) = 1 - 0.005k$ and the probability of death of an individual depending on the population size according to the formula $p_s(k) = 0.005k$. Then, the probability that the population size changes from k to n during time $[t, t + 1)$ is equal to:

$$p(k, n)(t) = \sum_r \binom{k}{r} \binom{k}{k-n+r} (p_r(k))^r (1 - p_r(k))^{k-r} (p_s(k))^{k-n+r} (1 - p_s(k))^{n-r}.$$

The higher probabilities are drawn in darker colours.

B. The idea of the formation realization of the stochastic is process, as given by the stochastic matrix.

C. Seven sample realizations of this process.

This process has a non-quadratic stochastic matrix with positive probabilities in the area 200×400 , and the set $X_0 = 0, 1, \dots, 200$ cannot be extended.

The expected value at time t can only be calculated for realizations existing by the time t or more. At the beginning, a quadratic submatrix containing the probabilities of transition from state k to state n will be defined for $n \in X_0$.

Definition 2.4. A conditional quadratic submatrix of the stochastic matrix $p : X_0 \times X \rightarrow \{f : \Upsilon \rightarrow [0, 1]\}$ is a function $\check{p} : X_0 \times X_0 \rightarrow \{f : \Upsilon \rightarrow [0, 1]\}$ such that $\check{p}(x, y) = p(x, y) / (\sum_{y \in X_0} p(x, y))$.

For a one-dimensional stochastic process, it is a well-known quadratic stochastic matrix. The s-product of quadratic stochastic matrices is a probability matrix, where a state X transits to y in time $[t, t + s)$. It will be noted as: $p^s(t, t + s)$. This property allows calculating the expected value of a stochastic process $(\xi_t)_{t \in \Upsilon}$ of non-quadratic stochastic matrices.

$$E\xi_1 = \sum_{x_1 \in X} x_1 p(x_0, x_1)(0, 1) \tag{1}$$

$$E\xi_t = \sum_{x_{t-1} \in X_0} \sum_{x_t \in X} x_t \check{p}^{t-1}(x_0, x_{t-1})(0, t-1) p(x_{t-1}, x_t)(t-1, t).$$

The definition of the expected curve has a standard form:

Definition 2.5. *An expected curve of a discrete-time stochastic process $(\xi_t)_{t \in \Upsilon}$ is a function $F_{exp} : \Upsilon \ni t \rightarrow E\xi_t \in X$.*

Calculating the expected curve is quite laborious and sometimes the only way to predict the expected value for major t is by iteratively traversing the ergodic effects. However, there is another curve that characterizes the set of realizations of the stochastic process. There are many reasons to call this characteristic a predicted curve. The idea of calculating this curve is presented in Fig.2 .

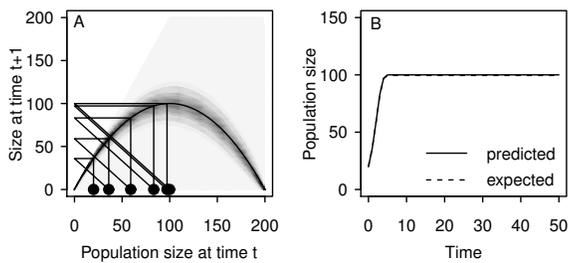


Figure 2: A. The stochastic matrix for the discrete-time birth-and-death process shown in Fig. 1, where each column has a conditional expected value (given by the formula $k + p_r(k) * k - p_s(k) * k$ for any k). The idea of forming a predicted curve is that the ends of the vertical segments lie exactly on the function $x \rightarrow x + p_r(x) * x - p_s(x) * x$. B. Predicted and expected curves for this process. The asymptotic value of the predicted curve equals 100. The expected curve tends to 0 (if $0 \in X_0$ but it takes an extremely long time to drop to 0) or 99.4974 (if $0 \notin X_0$)

The sequence $F(k) : X \ni k \rightarrow E(\xi_t | \xi_{t-1} = k) \in \mathbb{R}^n$ can be extended as a continuous function in many ways. However, if the formula for the conditional expected value is also a real function $F(x)$, then the function $F(x)$ is most optimally extended. It occurs in all models, where the probabilities of certain events are estimated by regressions. All known regressions are chosen from the family of continuous real functions.

The definition of the predicted value has the following form:

Definition 2.6. *The predicted curve of a discrete-time stochastic process $(\xi_t)_{t \in \Upsilon}$ is formed by the function $F : X_0 \ni k \rightarrow E(\xi_{t+1} | \xi_t = k) \in \mathbb{R}^d$*

treated as a real function if possible. Then, the predicted curve is a sequence $F_{pr} : \Upsilon \ni t \rightarrow X$ such that:

$$\begin{aligned} F_{pr}(0) &= x_0 \\ F_{pr}(t+1) &= F(F_{pr}(t)) \end{aligned}$$

Both curves seem to be very similar, but it depends on the stochastic matrix. Predicted curves sometimes show non-fading fluctuations. Then, the process is realized with rapid fluctuations, but such fluctuations are not regular. The fluctuations of the expected curve always decreases to zero (due to the ergodic theorem) (Fig. 3).

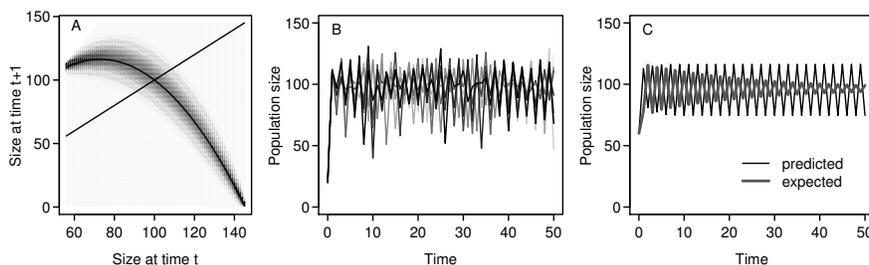


Figure 3: A. The stochastic matrix for the discrete-time birth-and-death process with probability of birth of an individual depending on the population size according to the formula $p_r(k) = 1.6 - 0.011k$, and probability of death of an individual depending on the population size according to the formula $p_s(k) = -0.6 + 0.011k$. B. Seven sample realizations of this process. C. Predicted and expected curves for this process.

The expected curve characterizes a probability distribution of states at a given time, and the predicted curve - a typical course of a stochastic process realizations.

2.3. Expected and predicted curves in discrete-time continuous-valued processes

Continuous-valued processes require some explanations. A stochastic kernel is defined in textbooks as a collection of densities of continuous distributions $p_x : X_0 \rightarrow [0, \infty)$. This is generalized in this paper. Definition 2.2 says that such a model is considered when the process tends not to change its state at many time steps. Then, a function $p_x(t)$ equals to the function $y \rightarrow p(x, y)(t, t+1)$ may not be a density because $\int_X p(x, y)(t, t+1) dy \leq 1$. The number $1 - \int_X p(x, y)(t, t+1) dy$ is a probability that the process doesn't change the state x at time $[t, t+1)$.

Let x be the state at time t . Let $x \in X_0$. Then:

$\int_{X_0} p(x, y)(t, t+1)dy$ is a probability that the next state at time $t+1$ is different than x and belongs to X_0 .

$\int_{X-X_0} p(x, y)(t, t+1)dy$ is a probability that the next state does not belong to X_0 . So, it is not equal to x . Because $X_0 \subset X$ then

$$\int_{X-X_0} p(x, y)(t, t+1)dy = \int_X p(x, y)(t, t+1)dy - \int_{X_0} p(x, y)(t, t+1)dy.$$

Ultimately $1 - \int_X p(x, y)(t, t+1)dy + \int_{X_0} p(x, y)(t, t+1)dy$ is a probability that the next state at time $t+1$ belongs to X_0 .

Since the expected value at time t for the process $(\xi_t)_{t \in \Upsilon}$ can be calculated only for the realizations with duration t or more, then the quadratic subkernel must be defined.

Definition 2.7. *A quadratic subkernel of the stochastic kernel $p : X_0 \times X \rightarrow \{f : \Upsilon \rightarrow [0, \infty]\}$ is the function $\check{p} : X_0 \times X_0 \rightarrow \{f : \Upsilon \rightarrow [0, 1]\}$ such that $\check{p}(x, y) = p(x, y) / (1 - \int_X p(x, y)(t, t+1)dy + \int_{X_0} p(x, y)dy)$.*

In analogy to the discrete-value process, we can calculate kernels for the transition from time t to time $t+2, t+3, \dots, t+s$ by defining the product of stochastic kernels. In these calculations, the transitions $x \rightarrow y \rightarrow z, x \rightarrow y \rightarrow y$ and $x \rightarrow x \rightarrow y$ for different x, y and z must be considered separately.

Definition 2.8. *The s -product of stochastic subkernels is the quadratic kernel of the functions $\check{p}^s(x, z)$. These functions are calculated as sums of integrals:*

For $s = 2$:

$$\check{p}^2(x, z)(t) = \int_{X_0} \check{p}(x, y)(t-1)\check{p}(y, z)(t)dy + \quad (2)$$

$$+ \check{p}(x, z)(t) \left(1 - \int_{X_0} \check{p}(z, y)(t+1)dy\right) + \check{p}(x, z)(t+1) \left(1 - \int_{X_0} \check{p}(x, y)(t)dy\right).$$

and for $s > 2$:

$$\check{p}^s(x, z)(t, t+s) = \int_{X_0} \check{p}^{s-1}(x, y)(t, t+s-1)\check{p}(y, z)(t+s-1)dy + \quad (3)$$

$$+ \check{p}^{s-1}(x, z)(t, t+s-1) \left(1 - \int_{X_0} \check{p}(z, y)(t+s-1)dy\right)$$

The expected value of this process at time t is equal to:

$$E\xi_1 = \int_X x_1 p(x_0, x_1)(0) dx_1 \quad (4)$$

$$E\xi_t = \int_{X_0} \int_X x_t \check{p}^{t-1}(x_0, x_{t-1})(0, t-1) p(x_{t-1}, x_t)(t-1) dx_t dx_{t-1}.$$

Now, we can define an expected curve for the discrete-time continuous-valued stochastic process.

Definition 2.9. *The expected curve for a discrete-time stochastic process $(\xi_t)_{t \in \Upsilon}$ is the function $F_{exp} : \Upsilon \ni t \rightarrow E\xi_t \in X$.*

The calculation of this curve is not simple, although it may require only a small value of time T and enable a good algorithmic calculation of the integrals. Estimation of their asymptotic values can be done using the ergodic theorem. However, the predicted curve is easy to define and very often easy to calculate.

Definition 2.10. *The predicted curve of the discrete-time stochastic process $(\xi_t)_{t \in \Upsilon}$ is a recursive sequence:*

$$F_{pr}(0) = x_0 \quad (5)$$

$$F_{pr}(t+1) = E(\xi_{t+1} | \xi_t = F_{pr}(t))$$

The kernel of many discrete-time continuous-valued stochastic processes is formed by densities of the normal distribution: $p(x, y)(t) = \frac{1}{\sigma(x)\sqrt{2\pi}} \exp(-\frac{(y-\mu(x))^2}{2\sigma(x)^2})$ where $\mu(x)$ and $\sigma(x)$ are given functions. This formula does not simplify the calculation of the expected curve, but the predicted curve has the formulas $F_{pr}(0) = x_0$ and $F_{pr}(t+1) = \mu(F_{pr}(t))$.

2.4. Expected and predicted curves in continuous-time discrete-valued processes

An intensity matrix consists of the instantaneous probability rates of changing a state x to y at time t . Let $p(x, y)(t_1, t_2)$ be the probability of a set of all realizations that have state x at time t_1 and state y at time t_2 . It is expressed by a very complex formula ([21]):

$$p(x, y)(t_1, t_2) = \sum_{r=0}^{\infty} \sum_{m_1 \neq x} \sum_{m_2 \neq m_1} \cdots \sum_{\substack{m_{r-1} \neq m_{r-2} \\ m_{r-1} \neq y}} H_{t_1, r}^{t_2}(x, m_1, m_2, \dots, m_{r-1}, y) \quad (6)$$

where:

$$H_{t_1,0}^{t_2}(x) = \exp\left(\int_{t_1}^{t_2} q_{x,x}(t)dt\right), \quad (7)$$

$$H_{t_1,1}^{t_2}(x, y) = \int_{t_1}^{t_2} q_{x,y}(t) \exp\left(\int_{t_1}^t q_{x,x}(s)ds + \int_t^{t_2} q_{y,y}(s)ds\right)dt, \quad (8)$$

$$H_{\tau_0,r}^{\tau_{r+1}}(m_0, m_1, \dots, m_r) = \int_{\tau_0}^{\tau_{r+1}} \int_{\tau_1}^{\tau_{r+1}} \dots \int_{\tau_{r-1}}^{\tau_{r+1}} \prod_{i=1}^r q_{m_{i-1},m_i}(\tau_i) \exp\left(\sum_{i=0}^r \int_{\tau_i}^{\tau_{i+1}} q_{m_i,m_i}(s)ds\right) d\tau_r \dots d\tau_2 d\tau_1 \quad (9)$$

In the last equation $\tau_0 = t_1$, $\tau_{r+1} = t_2$, $m_0 = x$ and $m_r = y$. This is done to reduce its notation. The expected value of the random variable ξ_t is equal to:

$$E\xi_t = \sum_{y \in X} \sum_{r=0}^{\infty} \sum_{m_1 \neq x} \sum_{m_2 \neq m_1} \dots \sum_{\substack{m_{r-1} \neq m_{r-2} \\ m_{r-1} \neq y}} y H_{0,r}^t(x, m_1, m_2, \dots, m_{r-1}, y) \quad (10)$$

The expected curve is the function $F_{exp} : [0, T) \ni t \rightarrow E\xi_t \in X$.

The calculation of this curve using presented formulas is not effective even by supercomputers. The number r above which the sum given by formula 9 is so small that it can be omitted, is usually very large. Presented formulas can be used proving some theorems about the expected curves. For calculations it is better to do a lot of simulations (10000 or more) and calculate means of population sizes at chosen times t (Fig.4).

The items in the intensity matrix are the derivatives at point 0^+ of the function $[0, T) \ni \Delta \rightarrow p(x, y)(t, t + \Delta) \in \mathbb{R}$, where $p(x, y)(t, t + \Delta)$ is the probability of changing state x to state y at time $[t, t + \Delta)$. These items are the rates of change of probabilities: that a process with state x at time t may change to state y at the next time. For each state $x \in X_0$ we can calculate:

$$\sum_{y \in X} y q(x, y)(t)$$

If its state at time t is equal to x , then this is the expected change in the state of the process after time t . We can plot such expected changes for different x on a chart (Fig.4).

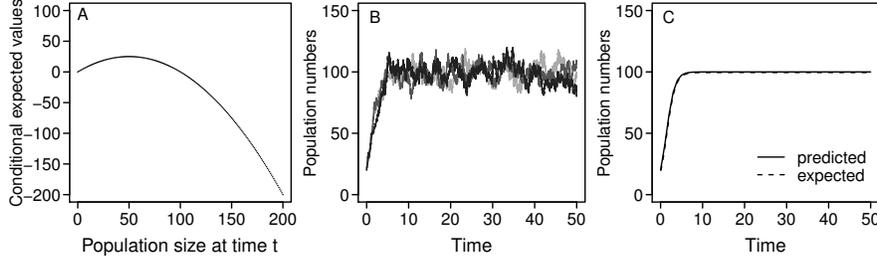


Figure 4: A. Conditional expected values of the rate of change in population size for a continuous-time birth-and-death stochastic process. The intensity matrix is given by the formula $q(n, k)(t) = 0$ if $|n - k| > 1$, $q(n, n + 1) = -0.005n^2 + n$, $q(n, n - 1)(t) = 0.005n^2$ and $q(n, n)(t) = -n$. The formula for the conditional expected values is equivalent to $F(n) = -0.01n^2 + n$. It can be extended to the real function $F(x) = -0.01x^2 + x$. This functions can be treated as a phase portrait of the differential equation. B. Three sample realizations of this process. C. The predicted and expected curves for this process. The mean population size for 100000 replicate simulations fluctuated between 99.46484 and 99.55412 at each time $t > 20$. The predicted curve is the solution of the differential equation $\frac{dF}{dt} = -0.01F^2 + F$ with the initial condition $F(0) = 20$. The solution of this equation is the function $F(t) = 100/(1 + 4exp(-t))$.

$X = \mathbb{N}^d$ is a discrete set. The sums $\sum_{y \in X} yq(x, y)(t)$ are defined only for the integer vectors x and y . However, the formula $x \rightarrow \sum_{y \in X} yq(x, y)(t)$ often has the form, which is also a function of real vectors. Then we can consider the differential equation:

$$\frac{dF}{dt} = \sum_{y \in X} yq(F(t), y)(t)$$

This is an idea of the predicted curve, which is a solution to the above equation.

Definition 2.11. For stochastic continuous-time discrete-valued stochastic processes with intensity matrix such that the function $X_0 \ni x \rightarrow \sum_{y \in X - \{x\}} yq(x, y)(t)$ is continuous and differentiable, the predicted curve is the solution of the differential equation such that:

$$F_{pr}(0) = x_0 \tag{11}$$

$$\frac{dF_{pr}}{dt} = \sum_{y \in X} yq(F_{pr}(t), y)(t)$$

For any two-dimensional intensity matrix, the expected and predicted curves are very similar (Fig.4). A greater difference is formed for four-dimensional matrices, as well as for specific parameters only.

2.5. The predicted curve and deterministic models

Although the definition of the predicted curve for Markov stochastic processes is new, these functions are already well known in the fields of applied mathematics, physics, theoretical biology, etc.

Examples of stochastic models and prediction curves are shown in Figs.1 and 2. Both figures show the same population model, in which the probability that some individual will give birth to one descendant in time $[t, t + 1)$ depends on the population size n and is equal to $p_r(n)$. At the same time, the probability of the death of an individual is equal to $p_s(n)$. The stochastic matrix for population dynamics (function $p(x,y)$) is shown in Fig.1. According to definition 2.6, the predicted curve for this process is a recursive sequence given by the formula:

$$F_{pr}(0) = n_0 \quad (12)$$

$$F_{pr}(t + 1) = F_{pr}(t) + F_{pr}(t)p_r(F_{pr}(t)) - F_{pr}(t)p_s(F_{pr}(t))$$

For the linear functions $p_r(n) = a_r n + b_r$ and $p_s(n) = a_s n + b_s$ this formula has the form:

$$\begin{aligned} F_{pr}(t + 1) &= F_{pr}(t) + F_{pr}(t)((a_r - a_s)F_{pr}(t) + (b_r - b_s)) = \quad (13) \\ &= F_{pr}(t + 1) = F_{pr}(t) + F_{pr}(t)(aF_{pr}(t) + b) \end{aligned}$$

where $a = a_r - a_s$ and $b = b_r - b_s$. Let $X_t = -\frac{a}{b+1}F_{pr}(t)$. Then:

$$X_{t+1} = X_t + X_t(-a\frac{b+1}{a}X_t + b) = X_t - X_t^2(b+1) + bX_t = (b+1)X_t(1 - X_t) \quad (14)$$

In theoretical biology, a recursive sequence of the form $X_{t+1} = \kappa X_t(1 - X_t)$ (for $\kappa > 1$ and $\kappa < 4$) is called a "discrete-time logistic population model". It was widely disseminated in biology by Robert McCredie May ([17]).

After Poisson, "birth-and-death" is the best known continuous-time Markovian stochastic process. Its intensity matrix is defined as $q(x, y)(t) = \mu(x)$ if $y = x + 1$, $q(x, y)(t) = \lambda(x)$ if $y = x - 1$ and $q(x, x)(t) = -\mu(x) - \lambda(x)$. For others y , $q(x, y)(t) = 0$.

The predicted curve satisfies the differential equation:

$$\frac{dF_{pr}}{dt} = \sum_{y \in X} yq(F_{pr}, y) = \quad (15)$$

$$(F_{pr} - 1)\lambda(F_{pr}) + (F_{pr} + 1)\mu(F_{pr}) - F_{pr}(\lambda(F_{pr}) + \mu(F_{pr})) = \mu(F_{pr}) - \lambda(F_{pr})$$

If the probability rate of birth of an individual depends on the population size according to the linear function $q_r(n) = a_r n + b_r$ and the probability rate of death of an individual depends on the population size according to the formula $q_s(n) = a_s n + b_s$, then the probability rate of change in population size from n to $n + 1$ is equal to $\mu(n) = n(a_r n + b_r)$ and the probability rate of change in population size from n to $n - 1$ is equal to $\lambda(n) = n(a_s n + b_s)$. Therefore:

$$\begin{aligned} \frac{dF_{pr}}{dt} &= F_{pr}(a_r F_{pr} + b_r) - F_{pr}(a_s F_{pr} + b_s) = \\ &= F_{pr}((a_r - a_s)F_{pr} + (b_r - b_s)) = F_{pr}(aF_{pr} + b) = rF_{pr}(1 - F_{pr}/K) \end{aligned} \quad (16)$$

where $r = b = b_r - b_s$ and $K = -\frac{b}{a} = -\frac{b_r - b_s}{a_r - a_s}$.

The last formula is known as the logistic model of the population. It was developed by a Belgian mathematician, Pierre Verhulst, in 1838 ([24]) and propagated by Raymond Pearl and Lowell Reed in 1920 ([19]).

Lotka-Volterra equations (also known as predator-prey models) are well known. They were described by Alfred J. Lotka in 1925 ([14]) as an alternative version of the autocatalytic chemical reaction model published in 1910 ([13]). 1927, Vito Volterra described the same set of equations ([25]), published it earlier in Italian.

Lotka-Volterra differential equations are the predicted values of the stochastic model for two population sizes. In this model, the number of predators decreases the probability of reproduction and increases the probability of death of a prey item, but the number of prey items increases the probability of reproduction and decreases the probability of death of a predator. This stochastic system can be described by a 2-dimensional Markov chain, which has a 4-dimensional intensity matrix of:

$$q((n_1, n_2), (k_1, k_2)) = \begin{cases} \mu_1(n_1, n_2) & \text{if } k_1 = n_1 + 1 \text{ and } k_2 = n_2 \\ \lambda_1(n_1, n_2) & \text{if } k_1 = n_1 - 1 \text{ and } k_2 = n_2 \\ \mu_2(n_1, n_2) & \text{if } k_1 = n_1 \text{ and } k_2 = n_2 + 1 \\ \lambda_2(n_1, n_2) & \text{if } k_1 = n_1 \text{ and } k_2 = n_2 - 1 \\ 0 & \text{if } |k_1 - n_1| > 1 \text{ or } |k_2 - n_2| > 1 \end{cases}$$

and $q((n_1, n_2), (n_1, n_2))(t) = -\mu_1(n_1, n_2) - \lambda_1(n_1, n_2) - \mu_2(n_1, n_2) - \lambda_2(n_1, n_2)$ for any n_1 and n_2 .

This process can be noted as: $(\xi_{1,t}, \xi_{2,t})_{t \in [0,t]}$ where $\xi_{1,t}$ denotes the number of predators and $\xi_{2,t}$ denotes the number of prey at time t . The conditional expected value of this process is equal to:

$$\begin{aligned} E(\xi_{1,t}, \xi_{2,t} | \xi_{1,t} = n_1, \xi_{2,t} = n_2) &= \\ &= (E(\xi_{1,t} | \xi_{1,t} = n_1, \xi_{2,t} = n_2), E(\xi_{2,t} | \xi_{1,t} = n_1, \xi_{2,t} = n_2)) \end{aligned} \quad (17)$$

and:

$$\begin{aligned} E(\xi_{1,t} | \xi_{1,t} = n_1, \xi_{2,t} = n_2) &= \\ &= (n_1 + 1)\mu_1(n_1, n_2) + (n_1 - 1)\lambda_1(n_1, n_2) - n_1\mu_1(n_1, n_2) - n_1\lambda_1(n_1, n_2) = \\ &= \mu_1(n_1, n_2) - \lambda_1(n_1, n_2) \end{aligned} \quad (18)$$

Similarly:

$$\begin{aligned} E(\xi_{2,t} | \xi_{1,t} = n_1, \xi_{2,t} = n_2) &= \\ &= (n_2 + 1)\mu_2(n_1, n_2) + (n_2 - 1)\lambda_2(n_1, n_2) - n_2\mu_2(n_1, n_2) - n_2\lambda_2(n_1, n_2) = \\ &= \mu_2(n_1, n_2) - \lambda_2(n_1, n_2) \end{aligned} \quad (19)$$

According to definition 2.8, the predicted curve for this process has the following formula:

$$\frac{dF_{pr}}{dt} = \sum_{(k_1, k_2) \in \mathbb{N}^2} (k_1, k_2)q(F_{pr}(t), (k_1, k_2))(t) \quad (20)$$

Thus, this is a function of the values in \mathbb{R}^2 . It can be written as $(F_{pr,1}, F_{pr,2})$ and these functions satisfy the system of equations:

$$\frac{dF_{pr,1}}{dt} = \sum_{(k_1, k_2) \in \mathbb{N}^2} k_1 q((F_{pr,1}(t), F_{pr,2}(t)), (k_1, k_2))(t) \quad (21)$$

$$\frac{dF_{pr,2}}{dt} = \sum_{(k_1, k_2) \in \mathbb{N}^2} k_2 q((F_{pr,1}(t), F_{pr,2}(t)), (k_1, k_2))(t)$$

According to equations: 18 and 19 the sums on the right-hand side can be replaced by their differences:

$$\frac{dF_{pr,1}}{dt} = \mu_1(F_{pr,1}, F_{pr,2}) - \lambda_1(F_{pr,1}, F_{pr,2}) \quad (22)$$

$$\frac{dF_{pr,2}}{dt} = \mu_1(F_{pr,1}, F_{pr,2}) - \lambda_1(F_{pr,1}, F_{pr,2})$$

Now, assuming that there is a linear dependency between the probability of birth-and-death of an individual and the population size of prey and predators. Then, probabilities of changes in population size are equal to:

$$\mu_1(n_1, n_2) = n_1(a_{r,1}n_2 + b_{r,1}) \text{ and } a_{r,1} > 0$$

$$\lambda_1(n_1, n_2) = n_1(a_{s,1}n_2 + b_{s,1}) \text{ and } a_{s,1} < 0$$

$$\mu_2(n_1, n_2) = n_2(c_{r,2}n_1 + d_{r,2}) \text{ and } c_{r,2} < 0$$

$$\lambda_2(n_1, n_2) = n_2(c_{s,2}n_1 + d_{s,2}) \text{ and } c_{s,2} > 0.$$

The numbers $b_{r,1}$, $b_{s,1}$, $d_{r,2}$, $d_{s,2}$ must have values such that all probability rates are not negative within a large interval of sizes of both populations. The system of differential equations 22 takes the following form:

$$\frac{dF_{pr,1}}{dt} = F_{pr,1}(F_{pr,2}(a_{r,1} - a_{s,1}) + (b_{r,1} - b_{s,1})) \quad (23)$$

$$\frac{dF_{pr,2}}{dt} = F_{pr,2}(F_{pr,1}(c_{r,2} - c_{s,2}) + (d_{r,2} - d_{s,2}))$$

It can be noted as:

$$\frac{dF_{pr,1}}{dt} = F_{pr,1}(-aF_{pr,2} + b) \quad (24)$$

$$\frac{dF_{pr,2}}{dt} = F_{pr,2}(cF_{pr,1} - d)$$

where $a = -(a_{r,1} - a_{s,1})$, $b = b_{r,1} - b_{s,1}$, $c = c_{r,2} - c_{s,2}$, $d = -(d_{r,2} - d_{s,2})$ are positive, real numbers. This is a classical form of Lotka-Volterra equations. This is the classical form of Lotka-Volterra equations. This model became very popular when D. A. MacLulich published an article on fluctuations in the numbers of hares and lynxes in Canada in 1937 ([16]). The graph of hare and lynx dynamics in this publication is the most frequently reprinted graph in science. It appears in many articles, manuals and textbooks around the world.

A known example of a Markov stochastic chain is the Wright-Fisher model of genetic drift [27]. Let each individual in the population have a gene pair of alleles A or B and produce many reproductive cells of one gene A or B. The cell fraction of allele A is equal $\frac{n_A}{n_A+n_B}$, where n_A is a number in the

genes of allele A for all individuals and n_B is many genes of allele B for all individuals. Each progeny is formed by the random connection of two reproductive cells. Genes cannot mutate. Generations do not overlap, so that genes from generation t cannot be combined with genes from the next generation $t + 1$. The initial generation (at $t = 0$) consists of N_{AA} individuals of genotype AA, N_{AB} individuals of genotype AB and N_{BB} individuals of genotype BB. What will be the proportion of individuals with the AA, AB and BB genotypes in each generation? What is the predicted value of this process?

The described model can be reduced to a pair of letters AA, AB and BB drawn multiple times, where the probability of drawing A is $p = \frac{2N_{AA} + N_{AB}}{2(N_{AA} + N_{AB} + N_{BB})}$, while the probability of drawing B is $(1 - p)$. So the probability of obtaining AA is equal to p^2 , the probability of obtaining AB is equal to $2p(1 - p)$, and the probability of obtaining BB is equal to $(1 - p)^2$.

The space of states of this model is a $(\mathbb{N} \cup \{0\})^3$. The stochastic process of the number of individuals of genotypes AA, AB and BB can be noted as $(\xi_{AA,t}, \xi_{AB,t}, \xi_{BB,t})$. The stochastic "matrix" of this process is equal to:

$$p((N_{AA}, N_{AB}, N_{BB}), (K_{AA}, K_{AB}, K_{BB}))(t, t + 1) = \quad (25)$$

$$\frac{K_{AA} + K_{AB} + K_{BB}}{K_{AA}!K_{AB}!K_{BB}!} p^{2K_{AA}} (2p(1 - p))^{K_{AB}} (1 - p)^{2K_{BB}}$$

where $p = \frac{2N_{AA} + N_{AB}}{2(N_{AA} + N_{AB} + N_{BB})}$.

The conditional expected value of this process is equal to:

$$E(\xi_{AA,t+1}, \xi_{AB,t+1}, \xi_{BB,t+1} | \xi_{AA,t} = N_{AA}, \xi_{AB,t} = N_{AB}, \xi_{BB,t} = N_{BB}) = \quad (26)$$

$$= (p^2 \xi_{t+1}, 2p(1 - p) \xi_{t+1}, (1 - p)^2 \xi_{t+1}) = \xi_{t+1} (p^2, 2p(1 - p), (1 - p)^2)$$

where $\xi_{t+1} = \xi_{AA,t+1} + \xi_{AB,t+1} + \xi_{BB,t+1}$ and $p = \frac{2N_{AA} + N_{AB}}{2(N_{AA} + N_{AB} + N_{BB})}$.

This is a real function. After several transformations of the formula in definition 2.6 the predicted curve of this process has a following form:

$$F_{pr}(t + 1) = N_{t+1} (p^2, 2p(1 - p), (1 - p)^2) \quad (27)$$

where N_t is population size at t and $p = \frac{2N_{AA,0} + N_{AB,0}}{2N_0}$.

The equation 26 or 27 is well known as "Hardy-Weinberg law".

In the original papers on genetic drift models, other stochastic processes were considered. It analyzed other variables: the number of genes of allele

A and the number of genes of allele B. It will be noted as $(\xi_{A,t}, \xi_{B,t})_{t \in \mathbb{N}}$. The space of states of this model is $(\mathbb{N} \cup \{0\})^2$. The probability of transition from the t -th generation (n_A, n_B) to the next generation (k_A, k_B) is equal to:

$$p((n_A, n_B), (k_A, k_B))(t, t+1) = \binom{k_A + k_B}{k_A} \left(\frac{n_A}{n_A + n_B} \right)^{k_A} \left(\frac{n_B}{n_A + n_B} \right)^{k_B} \quad (28)$$

To simulate this process, information about the population dynamics (function $t \rightarrow N_t$) is required. The assumption that this dynamics is constant is made very often. Genetic drift can then be observed in populations of different sizes (Fig. 5).

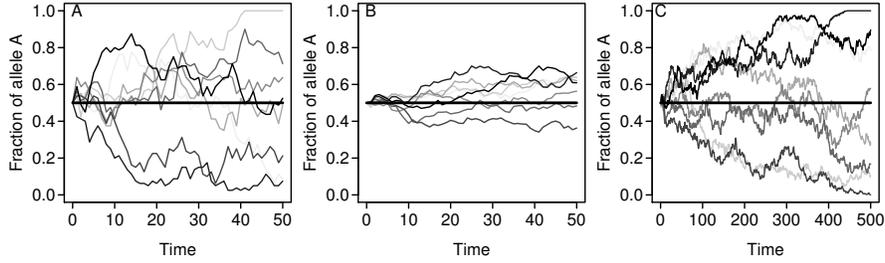


Figure 5: Results of eight replicate simulations of the Wright-Fisher genetic drift model in different population of sizes. A. $N_t = 40$, B and C. $N_t = 400$. C. The same results as in B, but presented by the time $T = 500$.

On each chart the line $P_{pr}(t) = 0.5$ is a predict fraction of allele A.

The conditional expected value of this process is equal to:

$$\begin{aligned} E(\xi_{A,t+1}, \xi_{B,t+1} | \xi_{A,t} = n_A, \xi_{B,t} = n_B) &= \\ &= \left(\kappa_{t+1} \frac{n_A}{n_A + n_B}, \kappa_{t+1} \frac{n_B}{n_A + n_B} \right) \end{aligned} \quad (29)$$

where $\kappa_{t+1} = \xi_{A,t+1} + \xi_{B,t+1}$

According to definition 2.6, the predicted curve of this process is equal to:

$$\begin{aligned} F_{pr}(0) &= n_{A,0} \\ F_{pr}(t+1) &= \left(\kappa_{t+1} \frac{n_{A,t}}{n_{A,t} + n_{B,t}}, \kappa_{t+1} \frac{n_{B,t}}{n_{A,t} + n_{B,t}} \right) = \\ &= \left(\kappa_{t+1} \frac{n_{A,t-1}}{n_{A,t-1} + n_{B,t-1}}, \kappa_{t+1} \frac{n_{B,t-1}}{n_{A,t-1} + n_{B,t-1}} \right) = \dots \end{aligned} \quad (30)$$

$$= \left(\kappa_{t+1} \frac{n_{A,0}}{n_{A,0} + n_{B,0}}, \kappa_{t+1} \frac{n_{B,0}}{n_{A,0} + n_{B,0}} \right)$$

The predicted fraction of allele A is constant and the same as in the initial population. In biology, this is a conclusion from the Hardy-Weinberg law.

Equation 26 was formed in 1908 by Godfrey H. Hardy and Wilhelm Weinberg ([9], [26]) as the limit of the fraction of alleles A for which the size of the population tends to infinity. There are other interpretations of the same equation as the predicted value: this is a central tendency characteristic of stochastic processes, which means that the mean fraction at each time is constant despite the drift of the fraction in different directions. No direction of the drift-model realizations is discriminated or stood out.

3. Conclusion

3.1. Similarity of deterministic models and predicted curves

Equivalence of the formulas of some deterministic models to the predicted values of stochastic processes with the same assumptions is not unusual. This applies to models that consider the size of the community. The rate of change of this size is a function of the unknown formula that depends on the size of the community. To establish the correct differential equation, it is necessary to divide the size of this formula by a value proportional to the community size and to consider the equation as the community size tends to infinity. The integer argument of the function is changed to a real one. This quotient is called "density", which confuses ecologists who calculate population density in a completely different way. Moreover, the non-linear functions of the community size are changed to a linear formula due to the possibility of developing all analytic functions as Taylor series.

If this model corresponds to a stochastic model in which the changes in the population are independent events with the same probability, then the stochastic matrix of its model is formed by binomial or multinomial distributions. When that the community increases to infinity, then the probabilities in this matrix will increasingly revolve around the conditional expected size. This is in accordance with Jacob Bernoulli's law of large numbers. In an infinite community, the stochastic process is realized with overlapping of the expected and predicted curves. However, they overlap with the solutions of the deterministic equations only if the functions of the probability or probability rate of the community size change are linear. For non-linear functions,

the differences between the differential model and the predicted curves can be quite large, as well as for specific parameters only.

The interpretation of the predicted curves may be quite different from the equations created by the deterministic models. This can be seen in the example of the Wright-Fisher genetic drift model and the Hardy-Weinberg law. In stochastic theory, they are complementary to each other. Nevertheless, some scientists believe that Hardy-Weinberg law and genetic drift model are mutually exclusive ([4]).

The similarity of some real processes to solutions of some deterministic models are treated as the proof that the world is deterministic. This article shows why the stochastic world is sometimes similar to the solutions of the differential equations. In other words, this article can be treated as a probabilistic explanation why the determinism of our world comes from.

3.2. Advantages of predicted curves

The predicted curve has a lot of advantages. It forms differential equations in which all parameters have a "biological" ("chemical", "physical", etc.) interpretation. They are values of measurement variables, or numbers, allowing the calculation of some relationships obtained by regression.

If the probabilities in the stochastic matrices are calculated based on experimental or field studies, then the time units for their calculation are known. Thus, probabilities are the mean fraction of events (transitions) over a specified time period (which is not large for a small time period, but increases if this time period increases). This time period is expressed in units of time (s, min, day, year, etc.), which can be displayed on the time axis of the predicted curve. Probability rates are derived quantities due to the time measured in the specified units. They are 1/s, 1/min, 1/day, 1/year, etc. For instance, the instantaneous probability rate of 0.24/day is equal to 0.01/h.

A stochastic model using such parameters has a time axis calculated in known units. The time axis of its implementation is expressed in the same units as the time axis of the predicted curve. This fact increases the application of predicted curves compared to deterministic models.

The predicted curve characterizes the central tendency of almost every stochastic process (excluding only processes without conditional expected values). It is often possible to derive its formula before specifying the values of the parameters. It allows one to perform mathematical analysis and to predict possible courses of stochastic process realizations. It can be used to

calibrate the parameters of stochastic models, because the predictions using them are always very good, even for very small communities.

The predicted curves as recursive sequences or differential equations can be characterized by stability or instability. It allows the use of these terms for stochastic models: it is stable if its predicted curve is stable. This is consistent with ecology, where the terms ‘stability’ and ‘instability’ are used to refer to ecosystems that are always stochastic. The definition of stability of stochastic systems allows a theoretical study between stability and durability of ecological systems.

4. methods

Statistical analysis of various existing Markov stochastic processes used in biology was performed.

References

- [1] P. Baldi; Y. Chauvin, T. Hunkapiller, M. A. McClure, *Hidden Markov models of biological primary sequence information* Proceedings of the National Academy of Sciences of the USA 91, (1994), pp 1059–1063.
- [2] V. Capasso, D. Bakstein, *An introduction to Continuous-Time Stochastic processes. Theory, Models, and Applications to Finance, Biology, and Medicine* Springer Science+Business Media New York, third eds. (2015).
- [3] M. Dehghania, M.H. Khoobanb, T. Niknam, *Fast fault detection and classification based on a combination of wavelet singular entropy theory and fuzzy logic in distribution lines in the presence of distributed generations* International Journal of Electrical Power & Energy Systems 78, (2016), pp 455–462.
- [4] W. Dembski, *The Design Inference. Eliminating Chance through Small Probabilities*. Cambridge University Press, (1998).
- [5] M. R. Fatmi, M. A. Habib. *Microsimulation of life-stage transitions and residential location transitions within a life-oriented integrated urban modeling system*. Computers, Environment and Urban Systems 69, (2018), pp 87–103.

- [6] G. C. Ghirardi, P. Pearle, A. Rimini, *Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles* Physical Review A 42(1) (1990) pp 78–89.
- [7] J. K. Grewal, M. Krzywinski, N. Altman, *Markov models—Markov chains* Nature Methods 16 (2019) pp 661–664.
- [8] R. W. Hamming, *The Art of Probability for Scientists and Engineers*. Addison-Wesley Publishing Company, (1991).
- [9] G. H. Hardy, *Mendelian Proportions in a Mixed Population*. Science XXVIII, (1908), pp 49–50.
- [10] C. G. Johnson, J. P. Goldman, W. J. Gullick, *Simulating complex intracellular processes using object-oriented computational modelling*. Progress in Biophysics & Molecular Biology 86, (2004), 379–406.
- [11] M. Kimmel, *Cellular population dynamics. I. Model construction and reformulation*. Mathematical Biosciences 48, (1980), pp 211–224.
- [12] J. T. Kool, *An object-oriented, individual-based approach for simulating the dynamics of genes in subdivided populations*. Ecological Informatics 4, (2009), pp 136–146.
- [13] A. J. Lotka, *Contribution to the Theory of Periodic Reaction*. Journal of Physical Chemistry A 14(3), (1910), pp 271–274.
- [14] A. J. Lotka, *Elements of Physical Biology*. Williams and Wilkins, (1925),
- [15] D. Lusseau, *Effects of Tour Boats on the Behavior of Bottlenose Dolphins: Using Markov Chains to Model Anthropogenic Impacts* Conservation Biology 17(6) (2003) pp 1785–1793.
- [16] D. A. Mac Lulich, *Fluctuations in the numbers of varying hare (Lepus americanus)*. University of Toronto studies., Biological series 43, (1937), pp 1–136.
- [17] R. M. May, *Simple mathematical models with very complicated dynamics*. Nature 261, (1976), pp 459–467.

- [18] M. Mazumdar, A. Kapoor, *Stochastic models for power generation system production costs* Electric Power Systems Research 35, (1995), pp 93–100.
- [19] R. Pearl, L. J. Read, *On the Rate of Growth of the Population of the United States Since 1790 and its Mathematical Representation*. Proceedings of the National Academy of Sciences of the United States of America 6(6), (1920), pp 275–288.
- [20] S. M. Ross, *Introduction to Probability Models. Eleventh Edition*. Academic Press is an Imprint of Elsevier, (2007).
- [21] M. Sokół, *A probability space of continuous-time discrete value stochastic process with Markov property*. Journal of advances in Mathematics 12, (2016) pp 5975–5990.
- [22] M. Sokół, *The relationship between intensity and stochastic matrices for continuous-time discrete value stochastic non-homogeneous processes with Markov property*. Journal of advances in Mathematics 12, (2017) pp 7244–7256.
- [23] M. Sokół, *Markov Stochastic Processes in Biology and Mathematics – the Same, and yet Different* . Journal of advances in Mathematics 14, (2018), pp 7540–7559.
- [24] P. F. Verhulst, *Notice sur la loi que la population suit dans son accroissement* . Mem. Acad. Lincei Roma. 2, (1838), pp = 31–113
- [25] V. Volterra, *Fluctuations in the abundance of a species considered mathematically*. Nature 119(12), (1927), pp 558–560.
- [26] W. Weinberg, *Über den Nachweis der Vererbung beim Menschen*. Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg 64, (1908), pp 368–382.
- [27] S. Wright, *Evolution in Mendelian Populations*. Genetics 16(2), (1931), pp 97–159.

5. Author contributions statement

This paper was written and prepared for publication by Miłoslawa Sokół: 100% of contribution.

6. Additional information

Funding Not applicable; **Conflicts of interest/Competing interests** Not applicable; **Availability of data and material** Not applicable; **Code availability** Not applicable; **Ethics approval** Not applicable; **Consent to participate** Not applicable; **Consent for publication** I consent for publication,

Figures

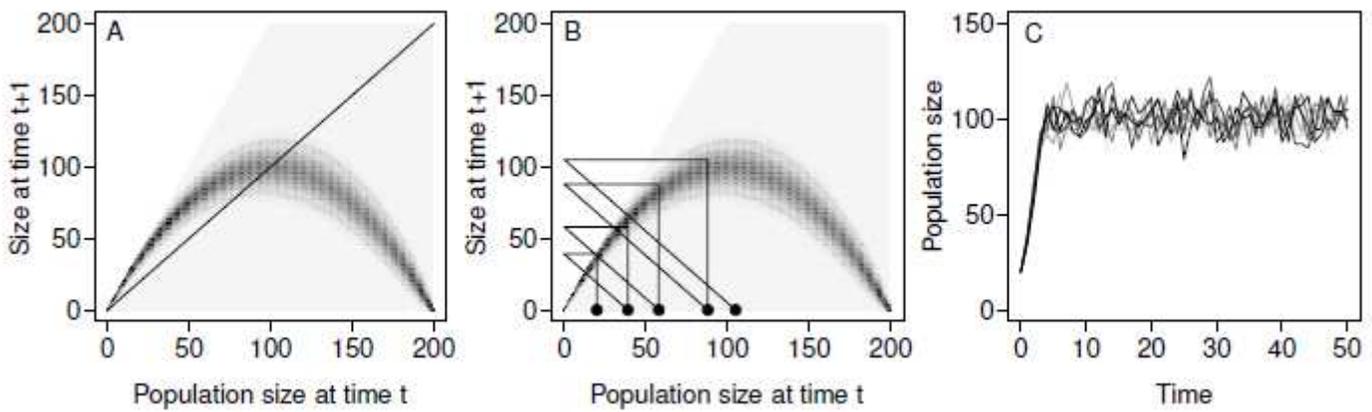


Figure 1

See the Manuscript Files section for the complete figure caption.

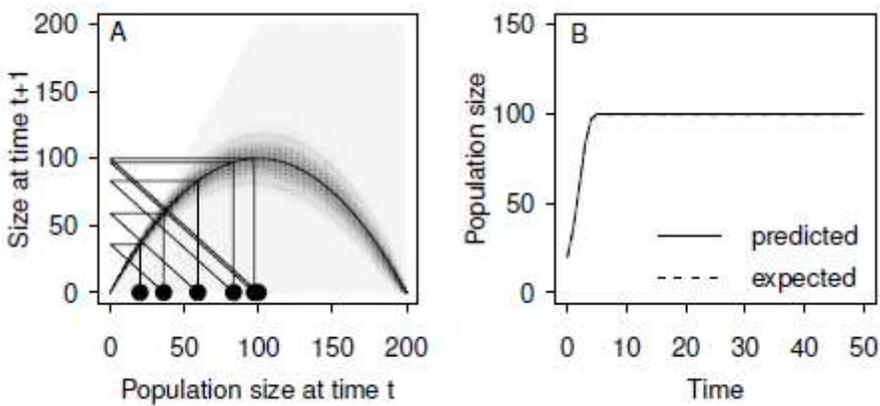


Figure 2

See the Manuscript Files section for the complete figure caption.

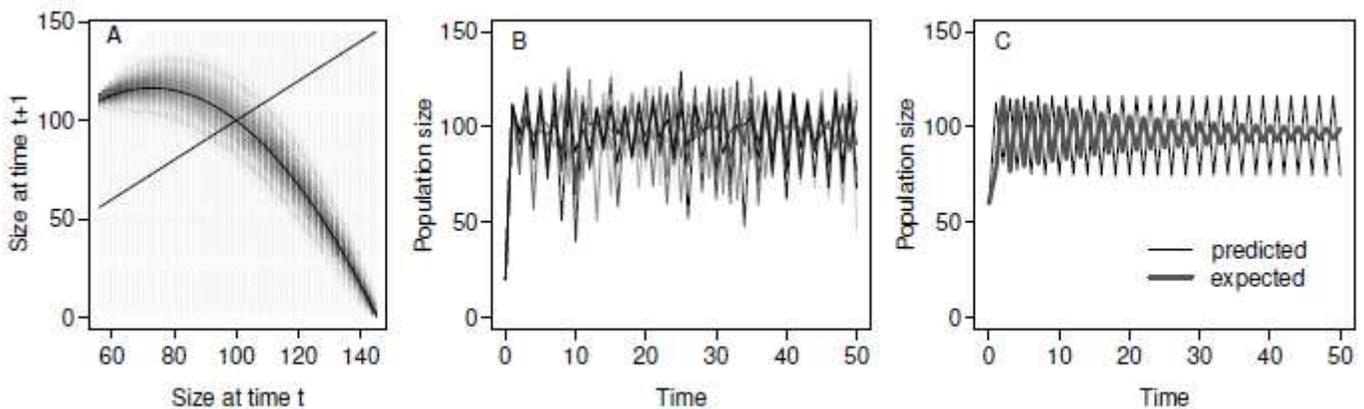


Figure 3

See the Manuscript Files section for the complete figure caption.

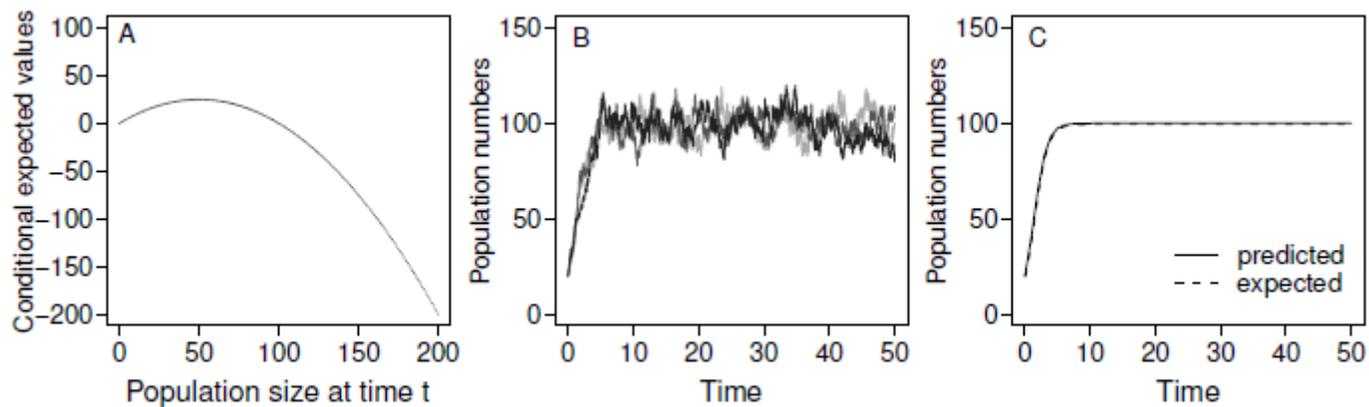


Figure 4

See the Manuscript Files section for the complete figure caption.

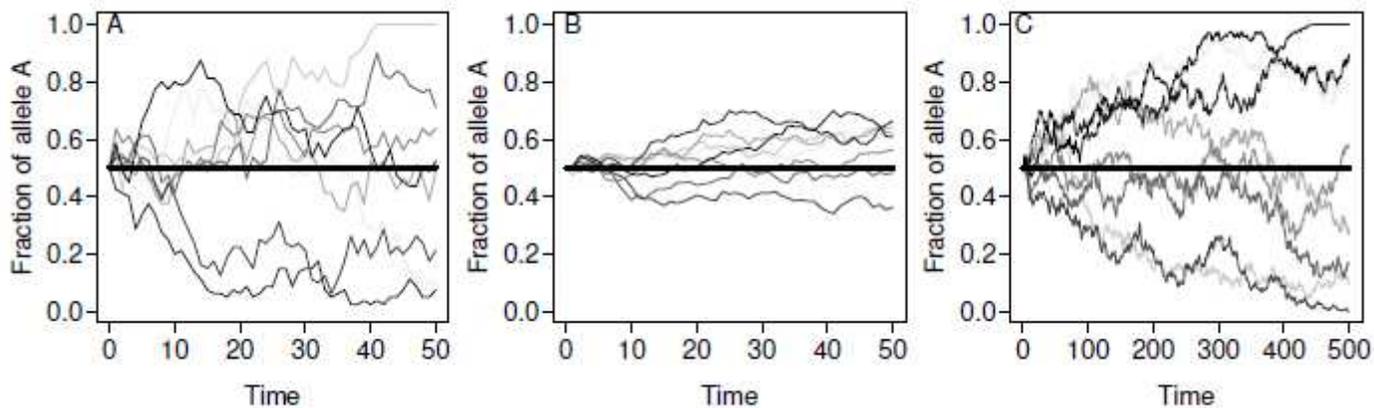


Figure 5

See the Manuscript Files section for the complete figure caption.