Single population dynamics: differences between classical and individual-based modeling approaches

Динамика отдельной популяции: описание методами классического и индивидуально-ориентированного моделирования

Janusz Uchmański

Януш Ухманьски

Institute of Ecology and Bioethics, Faculty of Christian Philosophy, Cardinal Stefan Wyszyński University; e-mail: januch@cbe-pan.pl
Wóycickiego 1/3, 01-938 Warsaw, Poland.

ABSTRACT: Individual-based models describe interactions between individuals in the population. Dynamics of the number of individuals in the population is a byproduct of these interactions. Individual variability plays an important role in this type of models. Individual-based models can be formulated in the form in which individual variability is replaced by average values. In contrast, classical models describe dynamics of population density directly and assume that all processes in the populations can be expressed as functions of the density. Numerical results of classical and individual-based models can be similar but interpretations almost always are different.

РЕЗЮМЕ: Индивидуально-ориентированные модели описывают взаимодействия особей в популяции. В этих моделях динамика численности особей в популяции является побочным продуктом описываемых взаимодействий, а индивидуальная изменчивость играет важную роль. Индивидуально-ориентированные модели можно переформулировать так, чтобы индивидуальная изменчивость была заменена средними величинами. Напротив, классические модели непосредственно описывают динамику численности популяций и принимают, что все популяционные процессы могут быть функцией плотности. Численные результаты обоих типов моделей могут быть сходными, но их интерпретации почти всегда различны.

1. Introduction

Let us consider a population of a parthenogenetic insect species with a simple life cycle. Females lay eggs in autumn and die. Eggs overwinter in soil and hatch during the next spring. Larvae feed and grow. Adult females emerging in autumn do not feed, they only lay next generation of eggs and die.

Even such a simple case can be complicated by some hidden assumptions. The first question is the survival of eggs, larvae and adult females. Not all the eggs laid in autumn survive until spring. Similarly, not all the larvae survive during summer and not all the females, which emerge from the pupae, survive until egg laying. However, to make the following considerations simpler, let us assume that all individuals, starting their lives as eggs, survive until egg laying time in autumn. This assumption enables us to leave the problem of mortality aside and concentrate at this step on the reproduction process. We will come back, however, to the mortality problem in the end of this paper.

If we know the number of females in a given generation and the number of eggs laid by each female, it is easy to calculate the number of females in the next generation. Let us assume that in generation t there are N_t females. If i-th female lays r_i eggs, the number of the females in the next $[t+1]$ generation will be equal to the sum of r_i for all N_t females:

$$N_{t+1} = \sum r_i \qquad (1)$$

where summation is for $i = 1,\ldots, N_t$.

Two main questions arise now. The first one concerns the situation within the given generation. What are the values of r_i? Are they the same for each female or different? If they are different, can we predict their values? The second question extends for further generations. If we want to calculate the number of females in subsequent generations, we have questions about their r values: are they the same or different as for the females in the present generation? If different, can we predict these values?

This paper compares the possibilities of individual-based and classical modeling to describe dynamics of biological systems. It describes limitations and pitfalls of mathematical modeling in ecology. But it also stresses the usefulness of constant efforts to come with mathematical descriptions closer and closer to the reality. I

will consider an example of a simplest model of population growth — the exponential equation and its extensions.

2. Exponential equation — a short time scale

Let us assume that all females are equal. They lay the same number of eggs r. Now, the equation (1) can be transformed to the following form:

$$N_{t+1} = r N_t \qquad (2)$$

However, a more interesting situation is when not all the females are equal. Let us first assume that in generation t there are N_t^1 females, which lay r_1 eggs, and N_t^2 females, which lay r_2 eggs (of course $N_t^1 + N_t^2 = N_t$). The number of females in the generation $[t+1]$ will be described by the equation:

$$N_{t+1} = r_1 N_t^1 + r_2 N_t^2 \qquad (3)$$

We can calculate r_{ma} — the mean (arithmetic) number of eggs laid by a female from generation:

$$r_{ma} = (r_1 N_t^1 + r_2 N_t^2) / (N_t^1 + N_t^2) \qquad (4)$$

The number of females in generation $[t+1]$ can be calculated using the mean number of eggs laid by a female from generation t:

$$N_{t+1} = r_{ma} N_t \qquad (5)$$

Equation (5) is in the form similar to equation (2).

Nothing will change if we assume that all females are different and have different r_i values. Once again we can calculate N_{t+1} using equation (1) or equation (5) and the result of calculations will be the same. However, in this later case we should use the following equation for calculation the mean number of eggs laid by a female:

$$r_{ma} = (\textstyle\sum r_i) / N_t \qquad (6)$$

where summation is for $i = 1,\ldots, N_t$.

There are two approaches to calculate the number of females N_{t+1} in the next $[t+1]$ generation. The first one (represented by equation (1)), which can be named individual-based, uses individual numbers of eggs laid by each female from generation t. The second approach (represented by equation (5)), which can be named average-based, uses the average number of eggs laid by a female from generation t. Both approaches give the same result of N_{t+1} and need the same amount of information to perform calculations. To use equation (1), we need data on individual progeny production of each female. To use equation (5), we need average value of progeny production in a population. But to calculate the average number of eggs per female, we also need individual information about the number of eggs laid by each female. Thus, both approaches — individual-based and average-based — are entirely equivalent when applied at a short time scale (equal to one time step of a model, e.g.one year in our case population).

3. Exponential equation — a longer time scale

Let us now calculate the number of females for several consecutive years. We assume additionally, that

all females, both in the given and future generations, are the same and lay the same number of eggs equal to r. In generation t we have N_t females, the number of females n generations later will be:

$$N_{t+n} = r^n N_t \qquad (7)$$

When in each of n future generations the conditions are different, but it does not produce reproduction variability among females within the generations, we have the following equation:

$$N_{t+n} = (\textstyle\prod r_j) N_t \qquad (8)$$

where multiplication is for $j = 1,\ldots,n$ and r_j is number of eggs laid by a female from generation j.

Can we also apply in this case the approach which uses average value of the number of eggs per female? Let us calculate r_{mh} — a mean (harmonic) value of the number of eggs per female in consecutive generations:

$$r_{mh} = (\textstyle\prod r_j)^{1/n} \qquad (9)$$

where multiplication is for $j = 1,\ldots,n$. Then:

$$N_{t+n} = (r_{mh})^n N_t \qquad (10)$$

because

$$(r_{mh})^n = ((\textstyle\prod r_j)^{1/n})^n = \textstyle\prod r_j \qquad (11)$$

Equations (5) and (10) describe the number of females in consecutive generations using average values of eggs per female. Both equations are equivalent to the individual-based model (1). Equation (5) makes a prediction for one generation in advance, equation (10) for several subsequent generations. Prediction of equation (5) is exact. Equation (10) predicts exact number of females in the generation $[t+n]$, however predictions for generations different than $[t+n]$ can not be made with the help of harmonic average for generation $[t+n]$, but have to use harmonic averages calculated for proper generations.

Results presented in this section can be easily generalized to describe the case of a species with an individual variability in female reproduction in each generation. In this case, to estimate the number of females in the future generations, we have to calculate arithmetic averages of r values for each generation and harmonic average of these arithmetic averages.

Again both approaches — individual-based and average-based — are equivalent: they give the same results and they need the same information for model construction. There is only one difference between them but its nature is rather numerical. Individual-based models (1) use integers as input data and give results — numbers of females in the next generations — which are also integers. Average values calculated in average-based models (5) or (10) are in general real numbers and give numbers of females in the next generations in the form of real numbers. However, average-based approach does not use this fact. This effect can be easily deleted by using function *Trunc* in computer simulations.

4. Individual-based modeling

Individual-based approach is represented by equation (1). Let us replace r_i by r_i^t, to indicate that across

the different generations the number of eggs per female can be different. Then we have:

$$N_{t+1} = \sum r_i^t \tag{12}$$

where summation is for all females present in generation t: $i = 1,\ldots, N_t$.

The individual-based model does not consider the number of females as its main point. It concentrates rather on r_i^t values and tries to explain, are they the same or individually vary among females. The number of females is a byproduct which we obtain only after the proper description of ecological processes hidden behind r_i^t values — growth, assimilation of resources, competition between individuals, unequal partitioning of resources as the result of competition, and costs of life too, which also can be different for different individuals in the population.

Individual-based model (12) (and its average-based counterparts (5) and (10)) can describe not only an exponential increase of the number of females in the population, but also population oscillations, its stabilization, decrease or even the extinction of the population. The dynamics of the system depends on a distribution of r values among females in each generation.

5. Fitting to the empirical data

Let us assume that we have the numbers of females in several subsequent generations and expect that the data follow an exponential increase of a population. However, we have no data on the number of eggs laid by females. Fitting the parameters of mathematical model to the data is the most often action in such a case. We fit parameter r_f of the equation

$$N_{t+1} = r_f N_t \tag{13}$$

to the data we have.

Parameter r_f, which value has been obtained by fitting, is in general equal neither to r_{ma} nor to r_{mh}. It gives the best fit to the empirical data. The numbers of females in subsequent generations calculated by means of equation (13) with fitted value of parameter r_f, give on average the best approximation of experimental data.

6. Classical modeling

Classical approach to model the dynamics of ecological systems assumes that it is possible to describe the number of females in subsequent generations by the equation:

$$N_{t+1} = g N_t \tag{14}$$

Equation (13) is similar to equations (5) and (10) using a procedure of averaging values of parameters characterizing individuals. Parameter g is in some sense related to r_{ma} or to r_{mh}. Its values can also be obtained by fitting to empirical data. Individual-based models use specific submodels to calculate r values which are different for different patterns of interactions between individuals in the populations. Classical models use a different approach. They try to be as general as it is possible, to include all possible forms of interactions between the individuals and describe them by some general functions.

The main point of classical approach to population dynamics is an assumption concerning equation (14). It says that g is a function of N_t:

$$g = g(N_t) \tag{15}$$

For instance, assuming that

$$g(N_t) = 1 + (1 - N_t / K) \tag{16}$$

where K is a constant parameter, we obtain the well known logistic equation, used to describe populations in which the number of individuals becomes constant after a sufficiently long time.

Assumption (15) is of a statistical or empirical nature, because function (14) can be obtained by fitting to empirical data. It is also a very simplifying assumption, because it compresses complex ecological interactions (e.g., which influence female fecundity) in the form of a rather simple mathematical function. Only a very small set (two or three) of different functions is used to describe population dynamics of species with very different ecology.

This assumption has two important consequences. Applying equation (14) with assumption (15) shifts us from the world in which population number is expressed by natural numbers into another world in which it is expressed by real numbers. In that case N_t is no more the number of females, but becomes a density of females in the population. This automatically limits the application of the model to populations which are well mixed in space and without spatial effects. Sedentary species can not be described by the model (14).

Another consequence is very promising at a first glance. Classical models directly make use of the fact that population density is expressed by real numbers. Assuming that function (15) is a real function, we have an immediate access to all mathematic techniques which is used to analyze difference equations. In contrast, individual-based models are always simulation models. This is a strong limitation, which can partly be overcome by proper simulation procedures (see the ODD protocols in Grimm *et al.*, 2006, 2010, as an attempt to solve this problem). It is much more elegant to use methods available for analysis of difference equations when we are modeling with classical models. Analytical solutions are often possible in this case. However, the number of possible solutions is much more limited than in the case of individual-based models. These are asymptotic solutions, cycles or deterministic chaos. Individual-based models produce a wider range of possible solutions. Working with classical models we devote much more time to stability analysis. From the point of view of classical modeling, a stability of ecological systems seems to be much more important than it is in individual-based models. This emphasis on the stability of systems described by classical models is a direct result of mathematical methods used to model the dynamics of ecological systems. The importance of stability of ecological systems in nature is apparently much smaller (see Rhode 2005, for examples and arguments on so called nonequilibrium ecology).

Table 1. Data used to construct the individual-based model (number of females and *r* values). Harmonic average values of *r* (forth column) have been used to calculate results of the model (fifth column).
Таблица 1. Данные для построения индивидуально-ориентированной модели (число самок и значения *r*). Гармоническая средняя значений *r* (четвёртая колонка) использована для расчета результатов моделирования (последняя колонка).

Year	Number of females, empirical data	*r* values	Harmonic averages of *r* values	Number of females, calculated values
1	100	2.0	2.0	100
2	200	2.0	2.0	200
3	400	2.0	2.0	400
4	800	1.25	1.7783	800
5	1000	0.8	1.5157	1000
6	800	1.25	1.4422	800
7	900	1.0	1.3687	900
8	900	-	-	900

6. Conclusions

Let us illustrate the arguments above by an example. Fig. 1 presents the numberof females in consecutive years. Numerical values of the number of females are also presented in the second column of Tab. 1. Other parts of Tab. 1 include all elements necessary to construct an individual-based model or its average-based counterpart. Values of *r* are the same for all females in each generation, but *r* values in different generations are different. Tab. 1 also presents harmonic averages calculated starting from year 1 for the subsequent years 2, 3, 4, etc., and the number of females calculated according to equation (10).

Tab. 2 illustrates the construction of a classical model based on the same empirical data (number of females and *r* values). The second and the third column are the same as in the previous table. According to the rules of classical modeling, we have to present *r* values

Table 2. Data used to construct the classical model. The same number of females as in Tab. 1 was used. The function of the form (17) has been fitted to describe a relationship between *r* values and the number of females. Equation (18) was then used to calculate densities of females in the population in the consecutive years.
Таблица 2. Данные для построения классической модели, число самок такое же, как в Табл. 1. Функция (17) аппроксимирована для описания связи между значениями *r* и числом самок. Уравнение (18) было затем использовано для расчета плотности самок в популяции в ряду последовательных лет.

as related to the number of females and then fit a linear function to this relationship as it is most often done (see Fig. 2). The following function gives the best fit:

$$g(N_t) = 2.3022 - 0.0014\, N_t \qquad (17)$$

After substitution of equation (17) to equation (14) we obtain the classical model of the form:

$$N_t = (2.3022 - 0.0014\, N_t)\, N_t \qquad (18)$$

Forth column of Tab. 2 shows densities of females in the population calculated according to equation (18).

Fig. 3 compares the results of individual-based and classical models constructed for the same data set. The difference between the results of both models is not big if we compare their ability to predict numerical results. Individual-based model gives the exact data. The classical model only approximates the data, because it uses the function which is an approximation of a relationship between the population parameters. However, there exists a much deeper difference in the interpretation of these two models. Using function (18) we silently assume that independently of the population structure, at

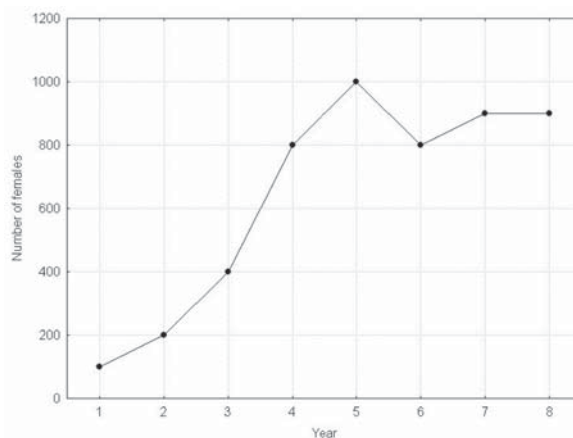| Year	Number of females, empirical data	*r* values	Density of females, calculated values
1	100	2.0	100.0
2	200	2.0	216.2
3	400	2.0	432.3
4	800	1.25	733.6
5	1000	0.8	935.5
6	800	1.25	928.5
7	900	1.0	930.6
8	900	-	930.0

Fig. 1. Number of females in the consecutive years, as based on data used for construction of individual-based and classical models.
Рис. 1. Число самок в ряду последовательных лет, на основе данных для построения индивидуально-ориентированной и классической моделей.
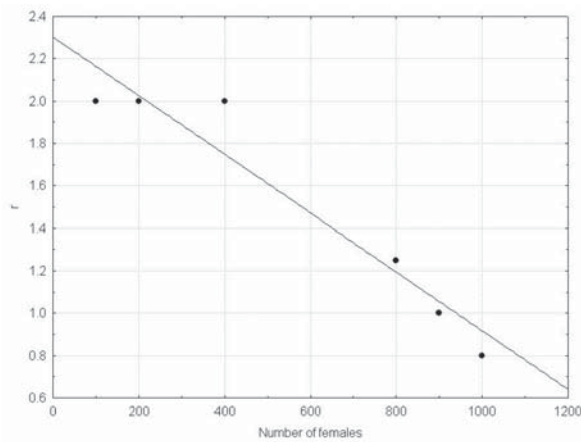
Fig. 2. Number of eggs per female in relation to the number of females in the population. It is assumed that female fertility does not vary within, but varies across the generations. Linear relationship in the form of (17) was fitted to the empirical data.

Рис. 2. Число яиц, отложенных одной самкой, в зависимости от числа самок в популяции. Принимается, что плодовитость самок неизменна в пределах генерации, но меняется между генерациями. Линейная зависимость в виде функции (17) была аппроксимирована к эмпирическим данным.
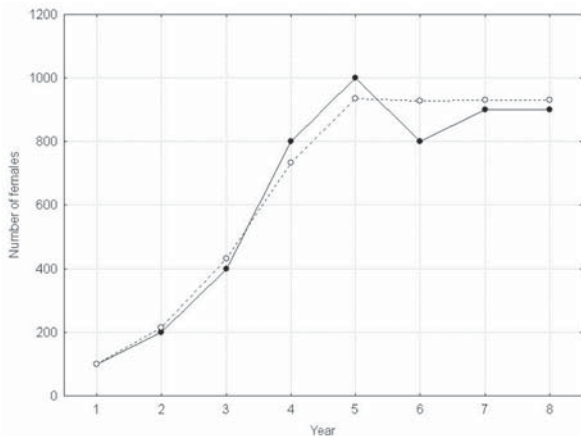

Fig. 3. Number of females in consecutive years: A comparison between results of an individual-based (solid circles) and classical (open circles) models.

Рис. 3. Число самок в ряду последовательных лет: сравнение между результатами индивидуально-ориентированной (черные символы) и классической (белые символы) моделей.

a given density all ecological processes in a population are fully and unambiguously explained by this density value. In contrast, during the construction of individual-based models we remember that for a given number of individuals in population there can be a variation in their demographic structure. For instance, the same number of individuals can have different weight distribution, different variability with respect to other im-portant individual features, or different proportion of females and males. Individual-based models always take such effects into account.

Working with the classical model we can always look for singular points and perform stability analysis. Solving the equation

$$1 = 2.3022 - 0.0014 \, N_t \qquad (19)$$

we find that $N_t = 930.1$ is the singular point. It is also the stable point, because the absolute value of the derivative of function (18) with respect to N is less then 1 for $N_t = 930.1$.

Methods of individual-based modeling do not allow to obtain results which are the product of mathematical technique used to construct and analyze the model. They do not make conclusions which have no roots in properties of the real system. After all, the data we have used in the example presented here do not indicate that there exists a stability of the population. Classical models directly use the fact that the population density is expressed by real numbers. This opens the opportunity to apply mathematical methods to perform stability analysis of the model equation. But as a consequence, it leads to the results which are the product of mathematical methods and may be far from the reality.

In the beginning of the paper we have assumed that there is no mortality in our case population. The introduction of mortality into the individual-based model does not change much the model and its interpretations. The mortality of the progeny decreases r values. The mortality of adult females changes only the distribution of r values among the females. Similar numerical changes will follow the introduction of mortality to the classical model. The analysis shows that relaxing the assumption about zero mortality does not change interpretations presented in this paper.

References

Grimm V., Berger U., Bastiansen F., Eliassen S., Ginot V., Giske J., Goss-Custard J., Grand T., Heinz S., Huse G., Huth A., Jepsen J.U., Jørgensen C., Mooij W.M., Müller B., Pe'er G., Piou C., Railsback S.F., Robbins A.M., Robbins M.M., Rossmanith E., Rüger N., Strand E., Souissi S., Stillman R.A., Vabø R., Visser U., DeAngelis D.L., 2006. A standard protocol for describing individual-based and agent-based models // Ecological Modelling. Vol.198. P.115–126.

Grimm V., Berger U., DeAngelis D.L., Polhill G., Giske J., Railsback S.F., 2010. The ODD protocol: a review and first update // Ecological Modelling. Vol.221. P.2760–2768.

Rhode K. 2005. Nonequilibrium ecology. Cambridge: Cambridge Univ. Press.