Article

Pine looper moth population dynamics in Netherlands: Prognosis with generalized logistic model

B.N. Nedorezova, L.V. Nedorezov

The Research Center for Interdisciplinary Environmental Cooperation (INENCO) of Russian Academy of Sciences, Kutuzova nab. 14, 191187 Saint-Petersburg, Russia

E-mail: l.v.nedorezov@gmail.com

Abstract

Current publication is devoted to analysis of well-known time series on the dynamics of pine looper moth (*Bupalus piniarius* L.) in national park De Hoge Veluwe (the Netherlands). For the approximation of every considering sample five various models with discrete time steps were used. Within the framework of every used model the influence of self-regulative mechanisms onto population size changing in time were took into account (and every used model had minimum number of unknown parameters). Estimations of model parameters were obtained at minimization of squared differences between theoretical (model) trajectories and empirical datasets (global fitting) for first ten values. Tails of samples (four or five points) were used for checking prognostic properties of models. Sets of deviations between theoretical and empirical trajectories were checked on Normality with zero average (Kolmogorov – Smirnov and Shapiro – Wilk tests), and were tested on absence/existence of serial correlation (Durbin – Watson criteria). Provided analysis showed that modified logistic model can only give sufficient approximation of empirical datasets. For parameters of this model confidence domains were determined, and for the situation when parameters belong to 90% confidence domain, forecasts of population size changing were constructed.

1 Introduction

Pine looper moth (*Bupalus piniarius* L.) is one of most dangerous forest pests (Schwerdtfeger, 1944, 1968; Isaev et al., 1984, 2001, 2009; Vorontsov, 1978, 1982; Nedorezov, Utyupin, 2011). Thus, all problems which have relation to modeling of this pest dynamics and with preparing of scientific-based forecasts are among the most actual IPM tasks (Klomp, 1966; Palnikova et al., 2002, 2005; Kendall et al., 2005; Nedorezov, 2010).

At present time it is possible to point out a big number of publications which are devoted to problems of analysis and description of pine looper moth population dynamics, and to collecting the respective datasets. On the other hand, modern mathematical models don't allow using this collected information for the description of population dynamics and for constructing forecasts. It can be taken into account within the framework of imitation models but its using leads to the necessity to have a lot of additional assumptions and hypotheses. These additional assumptions and hypotheses need in respective substantiations, bases and so on. The similar problem arises when non-imitative models with several variables are used for approximation of real datasets: incorrect assumptions can lead to qualitative change of the final result (Turchin et al., 2003; Nedorezov, 2007).

This is why the use of simple mathematical models which don't need in additional assumptions, special hypotheses, and take into account a small number of population regulative mechanisms, is very important element of initial stage of population dynamics analysis. If these models give sufficient description (approximation) of real datasets it gives certain possibilities for constructing scientific-based forecasts. Moreover, it gives a base for solution of the problem of population size optimal management.

As it was obtained before (Nedorezov, 2010), the well-known time series on fluctuations of pine looper moth in De Hoge Veluwe national park (the Netherlands) (Klomp, 1966) can be sufficiently described by the trajectories of generalized discrete logistic model. It means that observed fluctuations of pine looper moth can be effectively explained as a result of influence of intra-population self-regulative mechanisms only.

In current publication we use the same datasets (Klomp, 1966) but in other manner. Part of considering time series we use for estimation of model parameters, and the other part (tails of time series) we use for comparison of real trajectories with model forecasts (respectively, we use it for checking the prognostic properties of model). Taking into account that part of considering time series is used for model parameter's estimations the final results can differ from results pointed out in our previous publication (Nedorezov, 2010). For one of the tails of time series we construct the set of forecasts: for maximum, minimum, and average of population density which take into account that real values of model belong to the respective confidence domain.

2 Models

In modern literature it is possible to find a huge number of various mathematical models of population dynamics (see, for example, Kostitzin, 1937; Bazykin, 1985; Isaev et al., 1984, 2001, 2009; Brauer, Castillo-Chavez, 2001; Turchin, 2003; Nedorezov, 1986, 1997; Nedorezov, Utyupin, 2011; Kendall et al., 2005 and many others). In table 1 there are the simplest mathematical models of population dynamics which can be presented in the following form:

$$x_{k+1} = G(x_k, a, b), \tag{1}$$

where G is non-linear function, a, and b are non-negative parameters, and x_k is population size (or density) at time k (or moment of population size fixation). For obtaining model trajectories it is necessary to point out the initial value of population size x_0. Thus, all considering in current publication models contain three unknown parameters (i.e. parameters a and b which are presented in table 1 plus initial value of population size x_0 which must be determined by the real datasets too).

Table 1 Models used for approximation of datasets

Models[*]		Source	Name of the model (common or used in current publication)
1	$x_{k+1} = ax_k(1 + bx_k)^{-1}$	Kostitzin, 1937	Kostitzin model
2	$x_{k+1} = ax_k(b - x_k)$	Moran, 1950; Ricker, 1954	Discrete logistic model
3	$x_{k+1} = a(1 - e^{-bx_k})$	Skellam, 1951	Skellam model
4	$x_{k+1} = ax_k^{1-b}$	Morris, 1959; Varley, Gradwell, 1960, 1970	Morris – Varley – Gradwell model
5	$x_{k+1} = ax_k e^{-bx_k}$	Moran, 1950; Ricker, 1954	Moran – Ricker model

[*]Models have the same numbers in all tables

At the process of model parameter estimations all models were used in the form they are presented in table 1. Only discrete logistic model was used in other form. Within the framework of this model it was assumed that population size can intersect the threshold level b, but after that population size becomes equal to zero identically. Thus, we used the discrete logistic model in the following form:

$$x_{k+1} = \max\{0, ax_k(b - x_k)\}. \tag{2}$$

It can be interpreted in the following way. Intersection by the population size of the level b leads to local destruction of the whole ecosystem (Isaev et al., 1984, 2001, 2009). We have to note, that this modification of discrete logistic model (2) hasn't additional limits for the values of model parameters: originally we have to have the realization of inequality $ab \leq 4$ for obtaining non-negative trajectories for all values $x_0 \in [0, b]$ (table 1). At the same time without the inequality $ab \leq 4$ we have problems with prognostic properties of this model (2): if $ab > 4$ the origin becomes global stable state of equation (Nedorezov, 1986, 1997; Nedorezov, Utyupin, 2011).

3 Datasets

Analyzing time series on pine looper moth population dynamics (Fig. 1) can be free downloaded in Internet (NERC Centre for Population Biology, Imperial College (1999) The Global Population Dynamics Database, N 2727, N 2728 и N 2729). In the first case (time series N 2727, Fig. 1a) all values are presented in units "average number of eggs per squared meter"; in the second case (time series N 2728, Fig. 1b) values are presented in units "average of larva per squared meter"; in the third case (time series N 2729, Fig. 1c) values are presented in units "average of pupae per squared meter". In the first case the volume of sample is equal to 15 (first element of the sample was obtained in 1950). In the second case the volume of the sample is equal to 14: first element of this sample was also obtained in 1950 but the respective value for 1962 is absent. In the third case we have 14 elements in the sample: first value of this sample was obtained in 1951. On Fig. 1 time series are presented on the phase plane "population density x – birth rate y" where birth rate y equal to relation of values of population density of two nearest generations (Isaev et al., 1984, 2001, 2009):

$$y_k = \frac{x_{k+1}}{x_k}, \tag{3}$$

where k is the number of year.

All values were collected in the Netherlands, in the North-West part of the national park De Hoge Veluwe (total area of this park is equal to 20 ha) where Scottish pines are presented. Taking into account that all datasets were collected in one and the same place, in current situation we have strong correlated time series. And all these time series have relations to one and the same population process. In such a situation it is naturally to assume that if one of models allows us to obtain sufficient approximation of real dataset (with minimum of the sum of deviations squared between theoretical and empirical values among all considering models) then this concrete model must give us the best approximation for all other time series. As it was obtained before (Nedorezov, 2010), the generalized logistic model (2) gives the best approximation for time series on pine looper moth fluctuations (Fig. 1). All other models from the table 1 cannot give us a sufficient description of population dynamics. In current paper we use the first ten values of time series for determination model parameters, and in principle, it can lead to qualitative change of the situation with approximation of real datasets.
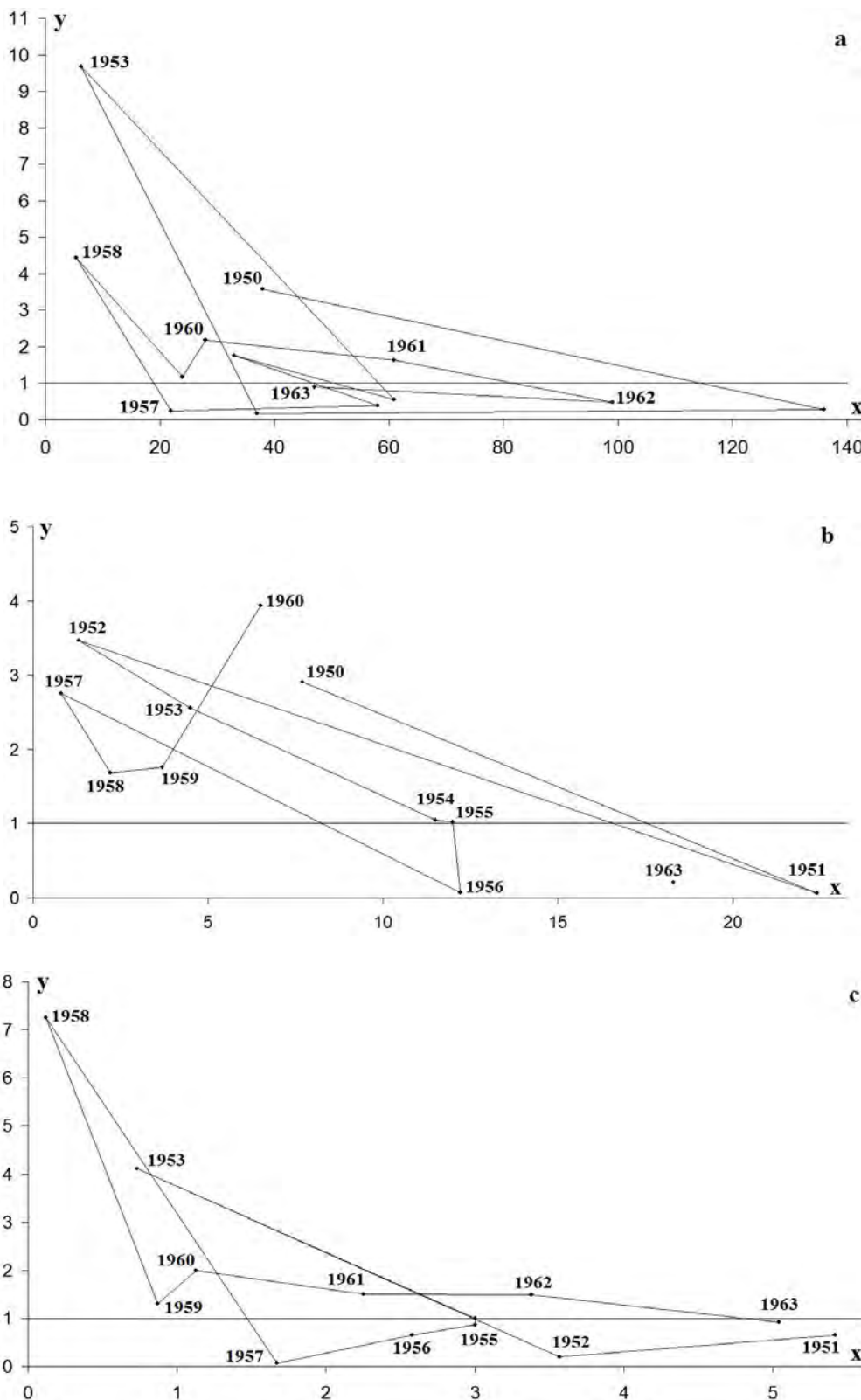
Fig. 1 Datasets on the pine looper moth fluctuations on the plane "population density – birth rate": a – dataset on population density (GPDD, N 2727) in "average number of eggs per squared meter"; b – dataset on population density (GPDD, N 2728) in "average number of larvae per squared meter"; c – dataset on population density (GPDD, N 2729) in "average number of pupae per squared meter". In first case the sample size is equal to 15 (first number was obtained in 1950), in the second case the sample size is equal to 14 (first number was obtained in 1950). In third case there are 14 numbers in the sample (first value of population density was obtained in 1951). x is population density, y is birth rate (3).

Visual analysis of pine looper moth population size changing (Fig. 1) shows that in some particular cases (see, for example, part of trajectory 1958-1962 for dataset N 2727, part of trajectory 1957-1960 for dataset N 2728, part of trajectory 1958-1963 for dataset N 2729; Fig. 1) we can observe the behavior of trajectory which corresponds to outbreak trajectories (Isaev et al. 1984, 2001, 2009). In particular, there is the decreasing of birth rate (3) on the phase of population size increasing. The absence of "ring movements" of the system in the domains with small or high population levels allows us to conclude that pine looper moth population fluctuations could be considered as an example of permanent outbreak. But for analysis of such difficult dynamic regimes we have to use more complicated mathematical models (which contains two or more dynamic variables).

Nevertheless, following the logic of population dynamics analysis described in details in our previous publications (Nedorezov, Lohr, Sadykova, 2008; Nedorezov, Sadykova, 2008, 2010; Nedorezov, Sadykov, Sadykova, 2010; Tonnang et al., 2009, 2010), we have to start with simplest mathematical models (table 1) before using more complicated models. We have to be sure that in considering situation we have a sufficient volume of a sample. If we can obtain a sufficient approximation of considering datasets with the help of simple mathematical models, it can be the reason for conclusion that time series are rather short (thus, we haven't sufficient volume of a sample), or for conclusion that hypothesis about difficult organized population dynamics must be rejected.

As it was obtained before (Nedorezov, 2010), this situation is realized for the Klomp's time series (Klomp, 1966). For whole samples generalized discrete logistic model allows us to obtain the sufficient approximation of real trajectories. Thus, the use of shorter time series (parts of initial samples) will not lead to the situation when all considering models (table 1) will not be applicable for the approximation.

4 Statistical Criterions

Let $\{\tilde{x}_t\}$ be the initial sample (results of observations of population size or density changing in time), $t = 0,1,...,N$, where $\tilde{x}_t$ is a value of pine looper moth density at time moment t. $N+1$ is sample size. Let also G be the non-negative non-linear function in the right-hand side of equation (1). Values of this function depend on population density and vector of unknown model parameters. The problem is: for existing sample $\{\tilde{x}_t\}$ we have to estimate parameters of model (1). For this reason the following criteria was used:

$$Q(x_0, \vec{\alpha}) = \sum_{t=0}^{M} (\tilde{x}_t - G^{(t)}(x_0, \vec{\alpha}))^2 \to \min_{\vec{\alpha}, x_0}, \qquad (4)$$

where $G^{(t)}(\cdot)$ are the iterations of function G, $G^{(0)}(x_0, \vec{\alpha}) = x_0$, $G^{(1)}(x_0, \vec{\alpha}) = x_1$ and so on; $\vec{\alpha}$ is vector of unknown model parameters, M is a number of used for estimation of model parameters values of the sample (it was assumed below that $M = 10$). In expression (4) the sum of squared deviations between theoretical (model) trajectory and empirical trajectory is minimized.

It is important to note that finding of minimum value of functional form (4) will allows us to rank all models. But it cannot allow giving the final decision about suitability or unsuitability of one or other model for the approximation of real datasets. We'll assume that model is suitable for fitting of empirical datasets if deviations between theoretical trajectory and empirical trajectory are the values of independent stochastic variables with symmetric (with respect to origin) distributions. It is obvious, if in the sequence of deviations we have any dependence, it means that model doesn't take into account the respective population process (or take it into account in incorrect way). Thus, this model we have to consider as unacceptable for fitting of empirical time series.

First of all, we have to be sure that arithmetic average of the set of deviations is equal to zero. More precisely, we have to check the hypothesis $H_0: Ee = 0$, where e is stochastic variable (deviation), Ee is expectation value, with alternative hypothesis $H_1: Ee \neq 0$. And we have to be sure that there are no reasons for rejecting of the hypothesis H_0. Additionally, distribution of deviations must be symmetric, uni-modal, and monotonic decreasing (in positive part of real axis) and monotonic increasing (in negative part of real axis) function. On the other words, methods for data collections must satisfy the following condition: we can observe deviations of any value in both sides with equal probabilities.

Checking of correspondence of the set of deviations to Normal distribution can be considered as sufficient condition for pointed out properties of the distribution of deviations. Moreover, if the set of deviations has this property (Normality) it gives a good base for checking of the independence of deviations: as it is well-known independence and noncorrelatedness are the same for Normal distributed stochastic variables. For checking of the Normality of deviations Kolmogorov – Smirnov criteria and Shapiro – Wilk criteria were used (Bolshev, Smirnov, 1983; Shapiro, Wilk, Chen, 1968). Additionally, as it was pointed out above, in the sequence of deviations we cannot have a serial correlation. For this reason Durbin – Watson criteria d was used (Draper, Smith, 1986, 1987). Critical values of this criteria are the following: $d_L = 0.7442$, $d_U = 1.16461$ (for sample size 10, 2.5% significance level, and one predictor variable). If amount of d which is determined by the formula:

$$d = \frac{\sum_{k=2}^{M} (e_k - e_{k-1})^2}{\sum_{k=1}^{M} e_k^2}, \qquad (5)$$

where e_k is the deviation (it is determined in formula (4)), satisfies the inequality $d < d_L$ or $d > 4 - d_L$ it means that hypothesis about the absence of serial correlation in the sequence of residuals must be rejected with 5% significance level. This hypothesis cannot be rejected if for statistics (5) the inequality $d_U < d < 4 - d_U$ is truthful. All other values of d (5) on the interval $[0,4]$ belong to the zone of ambiguity. For 1% significance level critical values of statistics (5) are the following: $d_L = 0.60452$, $d_U = 1.0007$.

If one of used statistical criterions gave the negative results then the assumption about suitability of model for fitting of empirical datasets was rejected. If all used statistical criterions gave positive results (i.e. we had no reasons to reject the hypothesis about the equivalence of average to zero, we couldn't reject the hypothesis about the Normality of the set of deviations etc.) then we can conclude that model gives the *sufficient approximation of* datasets. After calculations models which gave the sufficient approximation of datasets were compared between each other with criteria (4).

5 Results of Calculations

In Table 2 there are the estimations of model parameters (Table 1) for the considering time series (Fig. 1). For the first time series (Fig. 1a) Moran – Ricker model allowed obtaining the best approximation. At the same time the value of parameter a is in non-biological zone. The similar result for Moran – Ricker model is observed for the second time series (Table 2): in both situations the value of parameter a is much bigger than maximum of productivity of pine looper moth (Schwerdtfeger, 1944, 1968; Isaev et al., 1984, 2001, 2009; Vorontsov, 1978, 1982; Palnikova et al., 2002, 2005; Nedorezov, 2010). Application of Moran – Ricker model for thr fitting of third time series showed (Table 3) that there is the negative serial correlation: statistics (5)

$d = 0.47$ and it is less then critical level $d_L = 0.60452$. Thus, the hypothesis about the absence of serial correlation in the sequence of residuals must be rejected with 2% significance level (at two-side criteria).

Table 2 Estimations of model parameter's values and respective value of minimizing functional (for first ten values of time series)

Models	Estimations of model parameters			Functional
	x_0	a	b	Q_{min}
Results for the first sample (GPDD N 2727)				
1	72.92	0.87	$7.47 \cdot 10^{-20}$	10016.2
2	37.45	$4.57 \cdot 10^{-2}$	87.45	5403.4
3	72.92	873542022.06	$9.97 \cdot 10^{-10}$	10016.2
4	72.92	0.87	$5.14 \cdot 10^{-17}$	10016.2
5	34.49	5970.72	0.21	4003.4
Results for the second sample (GPDD N 2728)				
1	7.7	122484166054.27	15614128777.97	414.1
2	6.04	0.25	17.07	66.6
3	12.27	87332408.73	$1.03 \cdot 10^{-8}$	347.1
4	12.27	0.9	$2.4 \cdot 10^{-16}$	347.1
5	6.4	137665.77	1.65	105.8
Results for the third sample (GPDD N 2729)				
1	5.33	1.1	0.13	8.3
2	3.38	2.31	3.87	9.6
3	5.35	4.93	0.22	8.2
4	5.24	1.02	0.26	8.6
5	5.37	1.06	0.098	8.2

Table 3 Results of analyses of deviations between real datasets and theoretical trajectories

Models	Average ± S.E.	KS[1]	SW[2]	DW[3]
Results for the first sample (GPDD N 2727)				
1	0.1902±10.55	0.16139/p>0.2	0.93011/p=0.44902	1.505
2	-2.448±7.71	0.23105/p>0.2	0.93874/p=0.53909	1.601
3	0.1902±10.55	0.16139/p>0.2	0.93011/p=0.44902	1.505
4	0.1902±10.55	0.16139/p>0.2	0.93011/p=0.44902	1.505
5	-2.485±6.62	0.11945/p>0.2	0.98732/p=0.99236	1.997
Results for the first sample (GPDD N 2728)				
1	0.0±2.15	0.189/p>0.2	0.88739/p=0.15846	2.054
2	-0.273±0.86	0.20402/p>0.2	0.92769/p=0.42552	1.975
3	0.045±1.96	0.17802/p>0.2	0.94043/p=0.55779	2.477
4	0.05±1.96	0.17802/p>0.2	0.94043/p=0.55779	2.477
5	-0.954±1.06	0.14267/p>0.2	0.96208/p=0.80934	1.34
Results for the first sample (GPDD N 2729)				
1	-0.005±0.3	0.18203/p>0.2	0.92681/p=0.41728	0.471
2	-0.241±0.32	0.17798/p>0.2	0.93671/p=0.51706	1.986
3	-0.004±0.3	0.18239/p>0.2	0.92693/p=0.41835	0.47
4	-0.003±0.31	0.15946/p>0.2	0.93587/p=0.50802	0.496
5	-0.003±0.3	0.18275/p>0.2	0.92654/p=0.41472	0.47

[1]KS – Kolmogorov – Smirnov criteria; [2]SW – Shapiro – Wilk criteria; [3]DW – Durbin – Watson criteria

Let's note that for all models (excluding generalized discrete logistic model) the negative serial correlation is observed for third time series. Application of models for fitting of first and second time series gave positive results. There are no reasons for rejecting the hypotheses about the Normality of deviations (even at 10% significance level). There are no reasons also for rejecting the hypotheses about the absence of serial correlation in the sequence of residuals (also at 10% significance level and two-side criterion).

Thus, like in previous case (Nedorezov, 2010) when model from the Table 1 were compared at approximation of whole time series, we have the only model (generalized discrete logistic model) which gives us a sufficient approximation of real datasets. It is important to note that this model doesn't give the best approximation for some particular cases (Table 3).

6 Forecast

Prognostic properties of models can be compared with various methods. For example, it is possible to compare sums of deviations squared which were obtained for the tails of time series and model trajectories which were calculated with the best estimations of model parameters (Table 4). For the time series N 2727 the best approximation was obtained with the help of Moran – Ricker model (Table 2). At the same time the sum of deviations squared for the tail of time series is much bigger than the respective sums obtained for all other models (Table 4). The smallest sum for the tail of this time series was obtained with the help of generalized discrete logistic model, and this sum in ten times less than the respective sum obtained with the help of Moran – Ricker model.

Table 4 Total sums of squared deviations for theoretical and real datasets for "tails" of time series

	Time series		
Models	N 2727	N 2728	N 2729
1	11600.1	443.6	36.1
2	4047.6	556.2	85.5
3	11600.1	722.3	36.3
4	11600.1	722.3	33.3
5	86376.8	149951.3	36.5

For the dataset N 2728 the best approximation was obtained with the help of generalized discrete logistic model (Table 2). At the same time the best forecast was realized at using of Kostitzin model, which showed very bad approximation for first ten values of time series (Table 2). In third case N 2729 generalized discrete logistic model showed the most bad result in approximation (Table 4). For all other models we have better results. It is interesting to note that in two first situations the worst results were obtained with the help of Moran – Ricker model (Table 4); in the last case the prognosis obtained with Moran – Ricker model can be compared with results obtained with all other models.

Thus, we can conclude that model which gives the best approximation for considering datasets (among all models of any selected group) and satisfies to group of certain statistical criterions, may not give the best forecast of population dynamics. On the other words, the best approximation of real datasets by any mathematical model isn't the guarantee for obtaining the best forecast. Moreover, models which don't satisfy to some statistical criterions and, respectively, cannot be marked as suitable for fitting of empirical datasets, can give better prognosis with respect to models which are suitable for fitting datasets. It is necessary to point

out again: we talked above about the forecasts which can be obtained with the best values of model parameters only.

Let's consider shortly other possible variants of prognosis constructing for population size changing in time. We'll confine ourselves by the use of time series N 2727 and generalized discrete logistic model.

Taking into account that initial sample is the sequence of values of any stochastic variables, estimations of model parameters (which were obtained on the base of this sample) are the stochastic variables too. Thus, constructing of forecasts must take into account that real values of model parameters belong to any confidence domains. On Figures 2 and 3 there are the intersections of confidence domains (on Fig. 2 there is the intersection of confidence domain by the plane $x_0 = 37.45$; on Fig. 3: there is the intersection by the plane $a = 0.0456$) for the parameters of generalized discrete logistic model. Boundaries of confidence domains were determined with the help of formula (Draper, Smith, 1986, 1987):

$$Q(x_0, a, b) = Q(\hat{x}_0, \hat{a}, \hat{b})\left(1 + \frac{p}{N-p} F(p, N-p, \beta)\right), \qquad (6)$$

where $\hat{x}_0$, $\hat{a}$, and $\hat{b}$ are the optimal values of models parameters which give a global minimum for functional form Q (Table 2); p is number of unknown parameters of model ($p = 3$); $N - p$ is a volume of sample size minus number of unknown parameters; F is Fisher' distribution with p and $N - p$ degrees of freedom; β is significance level.
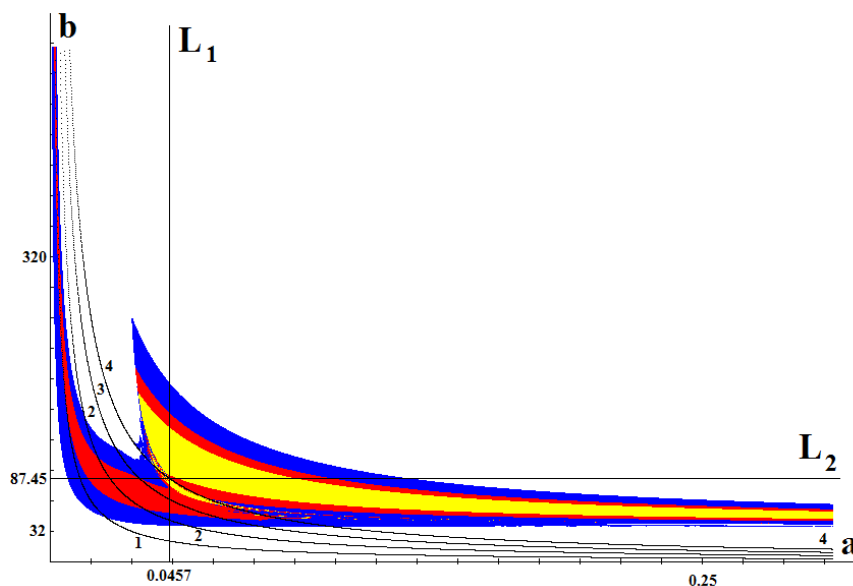

Fig. 2 Section of confidence domains by the plane $x_0 = 37.45$. 1, 2, 3, 4 are the bifurcation lines $ab = 1$, $ab = 2$, $ab = 3$ and $ab = 4$ respectively. Domain with yellow color corresponds to 90% confidence domain. Domain with yellow and red colors together corresponds to 95% confidence domain. Domain with yellow, red, and blue colors together corresponds to 99% confidence domain. Intersection of strait lines L_1 and L_2 gives the point of global minimum of the functional form Q.
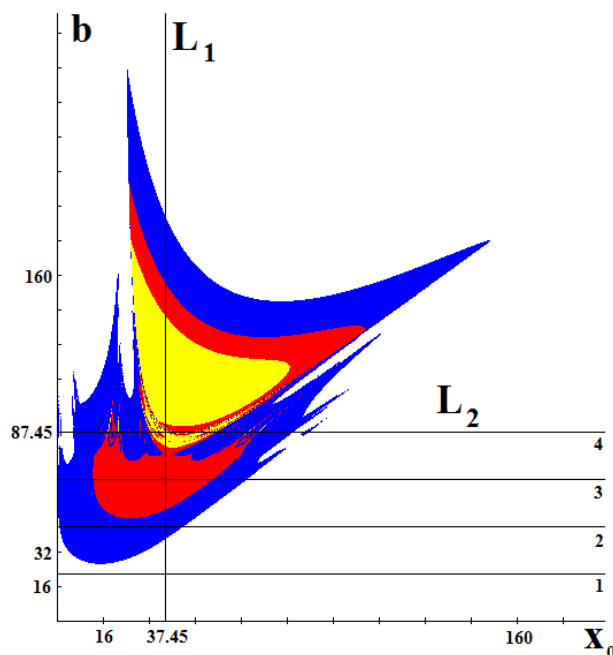
Fig. 3 Section of confidence domains by the plane $a = 0.0457$. 1, 2, 3, 4 are the bifurcation lines $ab = 1$, $ab = 2$, $ab = 3$ and $ab = 4$ respectively. Domain with yellow color corresponds to 90% confidence domain. Domain with yellow and red colors together corresponds to 95% confidence domain. Domain with yellow, red, and blue colors together corresponds to 99% confidence domain. Intersection of strait lines L_1 and L_2 gives the point of global minimum of the functional form Q.

Let Ω be the 90% confidence domain (Fig. 2 and 3), and let $\Delta = [0,150] \times [0,2] \times [0,210]$. Numerical calculations show that $\Omega \subset \Delta$. Modeling the stochastic points with rectangular distribution in Δ, and deleting all points belonging to $\Delta \setminus \Omega$ we obtain rectangular distributed points in Ω (Mikhilov, 1974; Ermakov, 1975). In initial sample all values of population density are positive, respectively, for constructing the forecast with generalized discrete logistic model we used trajectories with positive values only (for the first 15 values of trajectory – after that it could be equal to zero). It is important to note that we had about 98.7% of all trajectories with zero values at initial 15 steps (and these trajectories were deleted and didn't use for constructing of the prognosis). Additionally, effectiveness of algorithm (relation of number of points which were hit in Ω, to number of all modeled points) is about 0.028.

First of all, it is very interesting for practice to estimate the possible swing of population fluctuations. On Fig. 4 blue curves 2 and 3 correspond to maximum and minimum values of population density (these values were obtained for all trajectories with parameters from Ω). Note, that on Fig. 4 blue curve 3 is upper abscissa: values of this curve belong to the interval $[0.0495, 0.463]$. The second which is also of the practice interest, is the behavior of average of population density. But in this situation we have two qualitatively different possibilities: we can take into account or not the "probabilistic structure" of the confidence domain Ω (it is obvious, that various sub-domains of Ω have various probabilities – all confidence domains form one-parametric family with respect to significance level).
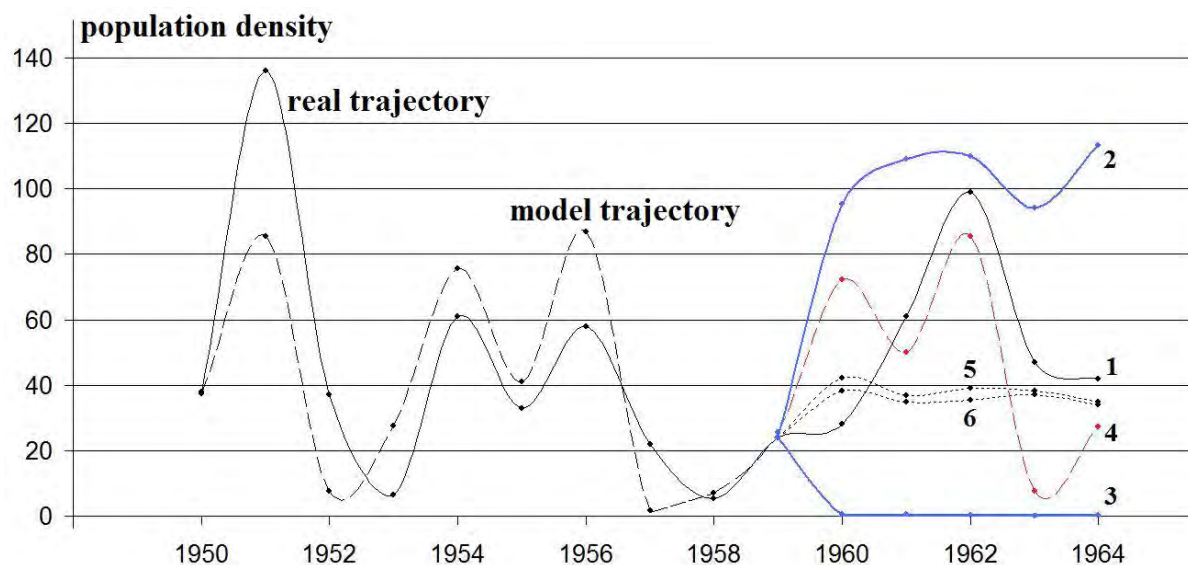
Fig. 4 Forecasts of the pine looper moth population dynamics with generalized discrete logistic model (for first time series, GPDD N 2727). Line 1 corresponds to real dataset; 2 and 3 (thick blue lines) correspond to maximum and minimum values of population densities respectively, line 4 (dashed-line red curve) is the trajectory of model, which was obtained with best model parameters, 5 and 6 (broken lines) correspond to mean values of population densities, which were obtained with an allowance for "probabilistic" structure of confidence domain and without it respectively.

If we don't take into account the "probabilistic structure" of confidence domain Ω, and assume that true values of population parameters can belong to various parts of Ω with equal probabilities, then we obtain the curve 6 (Fig. 4; average values were obtained for 7904 trajectories). Sum of squared deviations for this curve from empirical trajectory is equal to 4977.4. It isn't so good forecast like obtained with the help of model with best parameters. But it is much better than forecasts obtained with all other models (Table 4).

For taking into account the "probabilistic structure" of confidence domain Ω, we used the following algorithm:

1. Stochastic points $\gamma = (x_0, a, b)$ in Ω with rectangular distribution were obtained. For every obtained point γ the value of functional form $Q(\gamma)$ (4) was calculated.

2. On the interval $[Q(\hat{x}_0, \hat{a}, \hat{b}), Q^*]$ where $Q(\hat{x}_0, \hat{a}, \hat{b})$ is minimum value of functional form Q, and Q^* is the value of Q defined by the relation (6) with significance level $\beta = 0.1$, values of stochastic variable ξ with rectangular distribution were obtained.

3. If the inequality $\xi < Q^*$ was truthful we returned back to the point 1. If the inverse inequality, $\xi > Q^*$, was truthful point γ was used for calculating of the model trajectory. This algorithm (exclusion method by John von Neumann; Mikhilov, 1974; Ermakov, 1975) allows obtaining values of stochastic points with proportional to $1/Q$ distribution

Curve 5 (Fig. 4) was calculated with this algorithm (mean values were calculated for 1000 trajectories of the model). The sum of deviations squared for this curve 5 is equal to 4504.96. It is better than in previous case (when "probabilistic" structure wasn't take into account). But we have to note that forecast obtaining with model with best values of parameters is much better.

7 Conclusion

Analysis of time series on pine looper moth population densities fluctuations (Klomp, 1966) showed that in the Netherlands in national park De Hoge Veluwe where datasets were collected, this insect hasn't eruptive properties (which are observed in other locations; Isaev et al., 1984, 2001, 2009). For all analyzed time series (on eggs, larvae, and pupae densities changing in time) there exists one mathematical model only (generalized discrete logistic model) which gives sufficient approximation for empirical datasets. On the other words, all used statistical criterions didn't allow us to reject the respective hypotheses (about the equivalence of averages of deviations to zero, about the Normality of the sets of deviations, and about the absence of serial correlation in the sequences of residuals). These results of applications of statistical criterions mean that we have the respective base for conclusion that model gives us the sufficient approximation of real trajectories.

Existence of simple mathematical model with such a property, allows us to make the following conclusion: observed fluctuations of population densities can be explained as a result of influence of intra-population self-regulative mechanisms only.

Constructing of various forecasts with the help of simple mathematical models showed that not in all situations models (which satisfy to various statistical criterions and give sufficient approximation of datasets) allow us to obtain the best forecasts. Moreover, sometimes models (which don't satisfy to various statistical criterions and don't give sufficient approximation of datasets) allow obtaining the best prognoses. If it is observed within the group of models of the same level (which take into account the same sets of regulative mechanisms and the same sets of unknown parameters) it means that the process of finding of best model wasn't finished yet, and more complicated models must be applied for fitting of time series.

References

Bazykin AD. 1985. Mathematical biophysics of interacting populations. Nauka, Moscow, Russia (in Russian)

Bolshev LN, Smirnov NV. 1983. Tables of Mathematical Statistics. Nauka, Moscow, Russia (in Russian)

Brauer F, Castillo-Chavez C. 2001. Mathematical Models in Population Biology and Epidemiology. Springer-Verlag, New York, USA

Draper NR, Smith H. 1986. Applied Regression Analysis. V.1. Finance and Statistics, Moscow, Russia (in Russian)

Draper NR, Smith H. 1987. Applied Regression Analysis. V.2. Finance and Statistics, Moscow, Russia (in Russian)

Ermakov SM. 1975. Monta Carlo Method and Closely-related Questions. Nauka, Moscow, Russia (in Russian)

Isaev AS, Khlebopros RG, Nedorezov LV, et al. 1984. Forest Insect Population Dynamics. Nauka, Moscow, Russia (in Russian)

Isaev AS, Khlebopros RG, Nedorezov LV, et al. 2001. Population Dynamics of Forest Insects. Nauka, Moscow, Russia (in Russian)

Isaev AS, Khlebopros RG, Nedorezov LV, et al. 2009. Forest Insect Population Dynamics. KMK, Russia

Kendall BE, Ellner SP, McCauley E, et al. 2005. Population cycles in the pine looper moth: dynamical tests of mechanistic hypotheses. Ecological Monographs, 75(2): 259-276

Klomp H. 1966. The dynamics of a field population of the pine looper, *Bupalus piniarius* L. (Lep., Geom.). Advances in Ecological Research, 3: 207-305

Kostitzin VA. 1937. La Biologie Mathematique. A.Colin, Paris, France

Mikhilov GA. 1974. Some Questions of the Theory of Monte Carlo Methods. Nauka, Novosibirsk, Russia (in Russian)

Moran PAP. 1950. Some remarks on animal population dynamics. Biometrics, 6(3): 250-258

Morris RF. 1959. Single-factor analysis in population dynamics. Ecology, 40: 580-588

Nedorezov LV. 2007. Influence of food plant quality and parasitism on the cyclic fluctuations of larch bud moth. Euro-Asian Entomological Journal, 6(2): 229-244 (in Russian)

Nedorezov LV. 1986. Modeling of Forest Insect Outbreaks. Nauka, Novosibirsk, Russia (in Russian)

Nedorezov LV. 1997. Course of Lectures on Ecological Modeling. Siberian Chronograph, Novosibirsk (in Russian)

Nedorezov LV. 2010. Analysis of pine looper population dynamics with discrete time mathematical models. Mathematical Biology and Bioinformatics, 5(2): 114-123 (in Russian)

Nedorezov LV, Lohr BL, Sadykova DL. 2008. Assessing the importance of self-regulating mechanisms in diamondback moth population dynamics: Application of discrete mathematical models. Journal of Theoretical Biology, 254: 587–593

Nedorezov LV, Sadykov AM, Sadykova DL. 2010. Population dynamics of green oak leaf roller: applications of discrete-continuous models with non-monotonic density-dependent birth rates. Journal of General Biology, 71(1): 41-51 (in Russian)

Nedorezov LV, Sadykova DL. 2008. Green oak leaf roller moth dynamics: An application of discrete time mathematical models. Ecological Modelling, 212: 162-170

Nedorezov LV, Sadykova DL. 2010. Analysis of population time series using discrete dynamic models (on an example of green oak leaf roller). Lesovedenie, 2: 14-26 (in Russian)

Nedorezov LV, Utyupin YuV. 2011. Continuous-Discrete Models of Population Dynamics: An Analytical Overview. Ser. Ecology. Vip. 95. State Public Scientific-Technical Library, Novosibirsk, Russia (in Russian)

Palnikova EN, Sviderskaya IV, Soukhovolskii VG. 2002. Pine looper in Siberian forests: Ecology, population dynamics, and influence onto plants. Nauka, Novosibirsk, Russia (in Russian)

Palnikova EN, Soukhovolskii VG, Tarasova OV. 2005. Forest Entomology. Methods for Population Dynamics Analysis and Conditions of Individuals of Forest Insects. Siberian State Technological University, Krasnoyarsk, Russia (in Russian)

Ricker WE. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada, 11(5): 559-623

Schwerdtfeger F. 1944. Die Waldkrankheiten. Verlag Paul Parey, Berlin, Germany

Schwerdtfeger F. 1968. Okologie der Tiere. 2. Demokologie. Verl. Paul Parey, Berlin, Germany

Shapiro SS, Wilk MB, Chen HJ. 1968. A comparative study of various tests of normality. Journal of the American Statistical Association, 63: 1343-1372

Skellam JG. 1951. Random dispersal in theoretical populations. Biometrics, 38: 196-218

Tonnang H, Nedorezov LV, Ochanda H, et al. 2009. Assessing the impact of biological control of *Plutella xylostella* through the application of Lotka – Volterra model. Ecological Modelling, 220: 60-70

Tonnang H, Nedorezov LV, Owino J, et al. 2010. Host–parasitoid population density prediction using artificial neural networks: diamondback moth and its natural enemies. Agricultural and Forest Entomology, 12(3): 233-242

Turchin P. 2003. Complex Population Dynamics: A Theoretical/Empirical Synthesis. Princeton University Press, USA

Turchin P, Wood SN, Ellner SP, et al. 2003. Dynamical effects of plant quality and parasitism on population cycles of larch budmoth. Ecology, 84(5): 1207-1214

Varley GC, Gradwell GR. 1960. Key factors in population studies. Journal of Animal Ecology, 29: 399-401

Varley GC, Gradwell GR. 1970. Recent advances in insect population dynamics. Annual Review of Entomology, 15: 1-24

Vorontsov AI. 1978. Forest Pathology. Forest Industry, Moscow, Russia (in Russian)

Vorontsov AI. 1982. Forest Entomology. High School, Moscow, Russia (in Russian)