Article

Application of nonlinear model of population dynamics with phase structure to analysis of pine looper moth time series

L.V. Nedorezov
Research Center for Interdisciplinary Environmental Cooperation RAS, Saint-Petersburg, Russia
E-mail: l.v.nedorezov@gmail.com

Received 25 October 2015; Accepted 30 November 2015; Published online 1 March 2016

Abstract
Current publication is devoted to application of ELP-model (egg – larvae – pupae) for analysis of pine looper moth (*Bupalus piniarius* L.) dynamics in Netherlands (Klomp, 1966) and determination of asymptotic stable dynamic regimes. Method for estimation of model parameters (without using of any minimizing functional forms) when several correlated time series must be taken into account is described. Parameters of ELP-model were estimated, and it allowed creating two following hypotheses about pine looper moth dynamics: it may correspond to strong 2-cycle or non-rigorous 3-cycle.

Keywords pine looper dynamics; mathematical model; estimation model parameters; asymptotic regimes.

Proceedings of the International Academy of Ecology and Environmental Sciences
ISSN 2220-8860
URL: http://www.iaees.org/publications/journals/piaees/online-version.asp
RSS: http://www.iaees.org/publications/journals/piaees/rss.xml
E-mail: piaees@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Estimation of ecological model parameters is one of the most important theoretical and practical problems (Isaev et al., 1984, 2001; Kendall et al., 1999, 2005; McCallum, 2000; Wood, 2001). In a case when empirical dataset contains values of population size changing in time only one can use simple mathematical models (like Moran – Ricker model, discrete logistic model etc.; Moran, 1950; Ricker, 1954; Pielou, 1977; Nedorezov, 2012) for fitting of considering time series. For these models various types of minimizing functions can be used when LSM (Least Square Method) is used for estimation of model parameters. After finding of (global) minimum of one or other minimizing function analysis of deviations (between theoretical/model and empirical values) allows obtaining a conclusion about suitability or non-suitability of model for approximation of datasets (Bard, 1974; Draper and Smith, 1998; McCallum, 2000).

Situation is changed totally when initial sample contains several correlated time series. For example, initial sample of pine looper moth (*Bupalus piniarius* L.) population dynamics in Netherlands (Klomp, 1966) contains three time series of densities of eggs, densities of larvae, and densities of pupae. These datasets were collected in one and the same location and for one and the same population. Thus, for correct description of population dynamics we have to have a model which contains respective variables (like LPA or ELP models;

Dennis et al., 2001; Desharnais et al., 2001; Nedorezov, 2014), and we have to use all these time series together for correct model parameter estimations.

In one-dimensional case there is a problem of finding of best minimizing function which corresponds to considering time series - it can be a sum of squared deviations, sum of absolute values of deviations, sum of deviations after logarithmic transformation of model and datasets etc. In the case with several time series problem of selection of minimizing function becomes stronger (Tonnang et al., 2009a, b): problem of combination of several minimizing functions must be added to set of previous problems.

In current publication method of model parameter estimation without necessity to minimize any function is considered (Nedorezov, 2015a; Nedorezova and Nedorezov, 2012). Method was applied for estimation of parameters of ELP-model (Eggs – Larvae – Pupae; Nedorezov, 2014). This model was applied for fitting of pine looper moth time series (Klomp, 1966). It allowed creating of hypothesis that pine looper dynamics is cyclic with rather short period in 2 or 3 years. This publication continues our investigations on analysis of pine looper moth dynamics with ELP-model and other mathematical models of population dynamics (Nedorezov, 2015a, b, c).

2 ELP-Model

Various species of forest insects have one-year generations, and during winter time individuals stay in pupae phase (for example, pine looper moth; Klomp, 1966; Isaev et al., 1984, 2001; Nedorezov, 1986; Kendall et al., 1999, 2005; Nedorezov and Utyupin, 2011 and others). Let P_k be a number of pupae at year k. Respectively, B_k is a number of butterflies, L_k is a number of larvae, and E_k is a number of eggs. Relation between B_{k+1} and P_k is determined by the following equation:

$$B_{k+1} = \mu_1 P_k . \qquad (1)$$

In (1) μ_1, $0 \le \mu_1 \le 1$, is quota of survived pupae during the winter period. Amount of this quota depends on weather conditions, but below it will be assumed to be constant. It also depends on food conditions for larvae: in model it can be described as dependence on L_k: $\mu_1 = \mu_1(L_k)$. Increasing of number of larvae leads to decreasing of food conditions for them and, respectively, to decrease of amount of μ_1. Asymptotically μ_1 goes to zero. Thus, following conditions are truthful for μ_1:

$$\mu_1(\infty) = 0 , \ \frac{d\mu_1}{dL_k} < 0 . \qquad (2)$$

The following function was used for fitting of time series by model trajectories:

$$\mu_1 = \frac{g_1}{1 + g_2 L_k^{g_3}} . \qquad (3)$$

In (3) all parameters g_j are non-negative constants, and $0 \le g_1 \le 1$.

Relation between variables E_{k+1} and B_{k+1} can be described with the following equation:

$$E_{k+1} = CB_{k+1}. \qquad (4)$$

In (4) C is productivity of butterflies. Below it will be assumed that productivity depends on L_k, $C = C(L_k)$. This function decreases with increase of number of larvae, and asymptotically it goes to zero:

$$C(\infty) = 0,\ \frac{dC}{dL_k} < 0. \qquad (5)$$

Simple function which satisfies conditions (5) can be presented in the form:

$$C = \frac{c_1}{1 + c_2 L_k^{c_3}}. \qquad (6)$$

In (6) all parameters c_j are non-negative constants. Taking into account that number of butterflies is real *invisible* variable (amount of this variable is rather difficult to determine in field conditions) it must be deleted from the model. Combining equations (1) and (4) we get

$$E_{k+1} = C\mu_1 P_k. \qquad (7)$$

In (7) functions in right-hand side satisfy to conditions (2) and (5), and in simple cases can be presented in forms (3) and (6).

Let μ_2 be a quota of eggs successfully transformed into larvae, $0 \le \mu_2 \le 1$. Below it will be assumed that $\mu_2 = const$. Thus, the following equation must be added to model:

$$L_{k+1} = \mu_2 E_{k+1}. \qquad (8)$$

The final equation is analog of Moran – Ricker model (Moran, 1950; Ricker, 1954):

$$P_{k+1} = L_{k+1} e^{-\alpha L_{k+1}}. \qquad (9)$$

In (9) parameter α corresponds to influence of self-regulative mechanisms on larvae's surviving, and expression $Exp\{-\alpha L_{k+1}\}$ is equal to quota of larvae successfully transformed into pupae. Combining equations (7), (8), and (9) we obtain ELP-model of insect population dynamics.

Model (7)-(9) has reach set of dynamic regimes (Nedorezov, 2014) including cycles of all lengths and chaotic trajectories. Model has eight non-negative parameters. Initial values of model variables are unknown parameters too, and must be determined at a process of parameter estimation. Note if parameters of model are known knowledge of initial value for E_1 is sufficient for calculation of model trajectory.

3 Used Time Series

Analyzing time series on pine looper moth population dynamics can be free downloaded in Internet (NERC Centre for Population Biology, Imperial College (1999) The Global Population Dynamics Database, N 2727, N 2728 and N 2729). In the first case (time series N 2727) all values are presented in units “average number of eggs per squared meter”; in the second case (time series N 2728) values are presented in units “average of larva per squared meter”; in the third case (time series N 2729) 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 a result of this gap in dataset first twelve values were used for statistical analyses only. In the third case sample size is equal to 14: first value was obtained in 1951.

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 and population, in current situation we have strong correlated time series. As it was obtained before (Nedorezov, 2010), the generalized logistic model (Moran, 1950; Ricker, 1954) gives the best approximation for time series on pine looper moth fluctuations for every separated time series. But some specific properties of this model didn’t allow determination of asymptotic stable dynamic regime of population.

4 Statistical Criterions

Before estimating of model parameters and/or constructing minimizing function it is necessary to describe a list of requirements to model. What does it mean that model corresponds to considering time series? In literature (Bard, 1974; Draper and Smith, 1998) one can find following basic requirements to set of deviations between model trajectory and empirical time series: model gives good fitting of time series if and only if deviations are values of independent stochastic variables with Normal distribution (with zero average), and there are no serial correlation in sequence of residuals.

Note, that normality of residuals is rather strong requirement for residuals. It can be changed onto requirement of symmetry of distribution with respect to origin and onto requirement to have monotonic branches of density function (monotonic decreasing branch in positive part, and monotonic increasing in negative part of straight line). Property of symmetry was checked with Kolmogorov – Smirnov test, Lehmann – Rosenblatt test, and Mann – Whitney test (Bolshev and Smirnov, 1983; Likesh and Laga, 1985; Hollander and Wolfe, 1973). Monotonic behavior of branches of density function was tested with Spearmen rank correlation coefficient. Testing of absence/existence of serial correlation in sequence of residuals was provided with Swed – Eisenhart test (Draper and Smith, 1998) and test on series of “jumps up” and “jumps down” (Likesh and Laga, 1985). Note, that all used tests pointed out above are nonparametric criterions.

We will say that Ω is a *feasible set* in a space of model parameters (with initial values for variable L) if for all element from this set hypotheses about absence of serial correlations, about symmetry of distributions, and monotonic behavior of branches of density functions cannot be rejected for every set of deviations (three sets) and for selected significance levels. In figure 1 there is a projection of feasible set Ω on plane (c_1, g_1) (5% significance level for all statistical criterions). As we can see in this Fig. 1 $\Omega \neq \varnothing$, and there are no black points in domain $c_1 g_1 < 1$; it means that observed time series do not correspond to regime of population extinction.

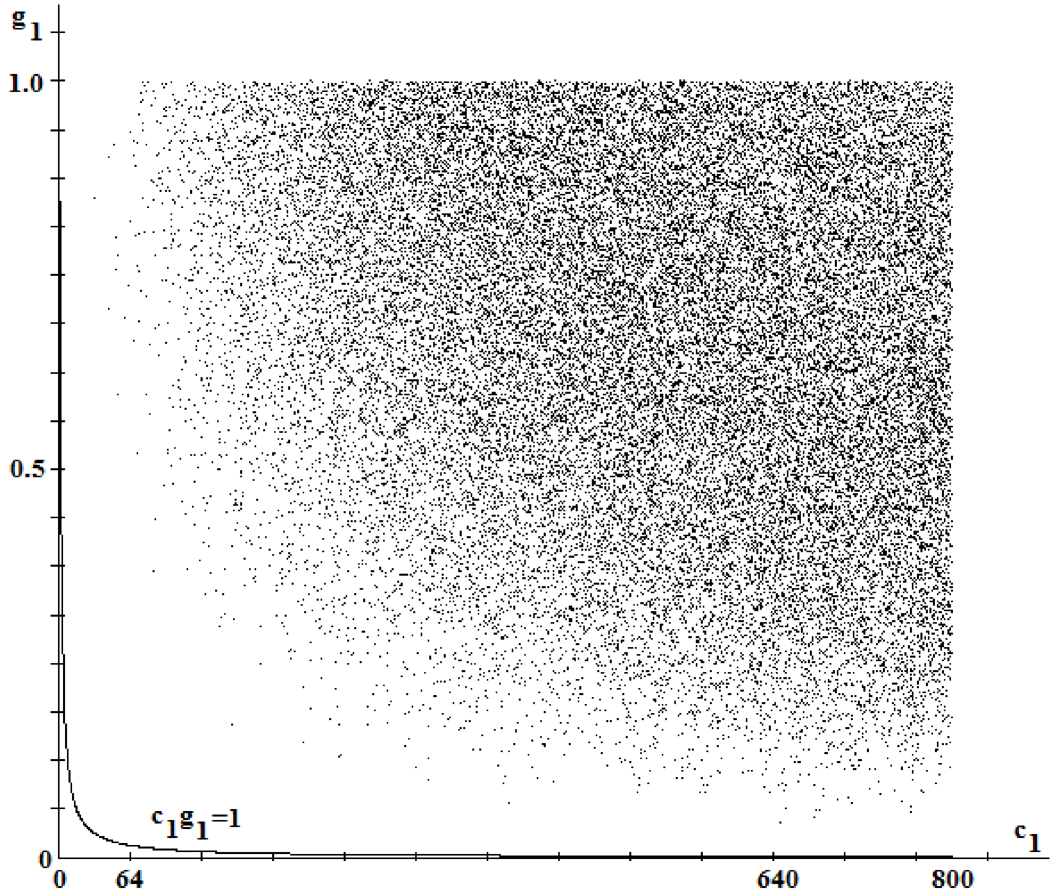


Fig. 1 Projection of feasible set Ω onto plane (c_1, g_1). $c_1 g_1 = 1$ is bifurcation curve which is a boundary of zone of population extinction (empty set).

Within the boundaries of feasible set one can use any minimizing function and find respective values of parameters. But it is not obligatory step in process of parameter estimations: statistical criterions can play roles of filter for points of feasible set. Let us assume that hypothesis about symmetry of distribution cannot be rejected with 5% significance level. If this hypothesis cannot be rejected with 10% significance level it means that we have stronger result. Strongest results can be obtained if we cannot reject hypothesis about symmetry with 95% or 99% significance level.

For Kolmogorov – Smirnov test significance level was changed on 95%, and for Lehmann – Rosenblatt test significance level was changed on 80%. All other criterions were used with previous significance level in 5%. It allowed finding two points in space of model parameters only. Below two dynamic regimes corresponding to determined points are considered separately. For third dynamic regime parameters were found at minimizing of loss-function which is equal to sum of squared deviations for all time series.

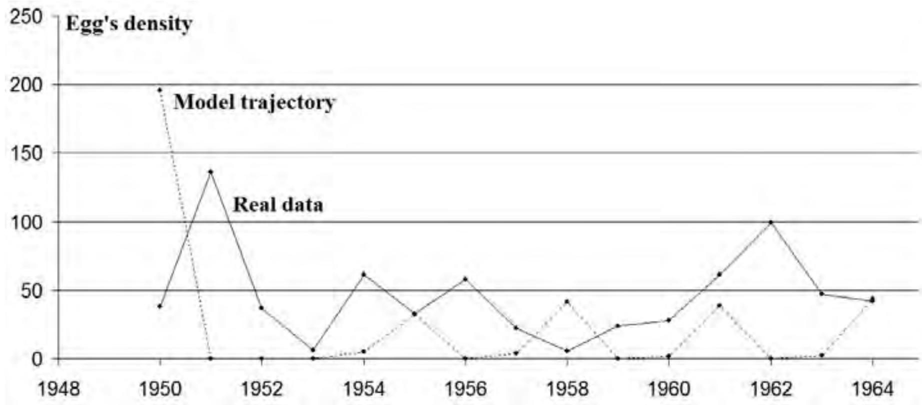
Remark. For determination of points of set Ω stochastic values with uniform distribution in a set were obtained in a following part of space of model parameters: $E_1 \in [0,200]$, $\mu_2 \in [0,1]$, $c_1 \in [0,800]$, $c_2 \in [0,10]$, $c_3 \in [0.5,2.5]$, $g_1 \in [0,1]$, $g_2 \in [0,6]$, $g_3 \in [0.1,4.1]$, $\alpha \in [0,2]$. In fig. 1 there are 10^6

points of set Ω. Probability p of event that stochastic point (with uniform distribution) belongs to Ω is equal to 0.0008182 approximately.

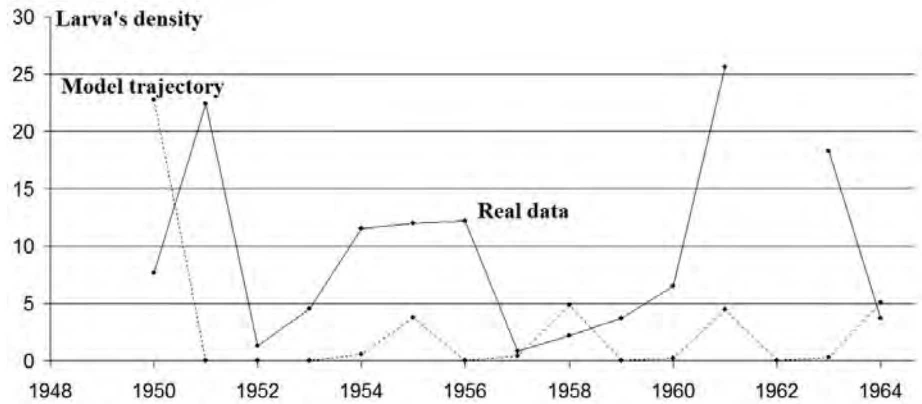
5 First Case

First point has following coordinates: $\mu_2 = 0.1162$, $c_1 = 350.34$, $c_2 = 5.31$, $c_3 = 1.41$, $g_1 = 0.92$, $g_2 = 4.0$, $g_3 = 3.27$, $\alpha = 0.021$. Initial value for eggs $E_1 = 195.76$. Results of fitting are presented in figure 2. Calculation of autocorrelation function for 20000 steps (after 10^6 free steps of process) shows that values of this function for lags $3k$, $k = 1,2,...$, are greater than 0.965. It means that asymptotic stable dynamic regime of pine looper moth is close to nonrigourous 3-cycle. Projection of asymptotic trajectory (for 20000 values) onto *plane density of eggs – density of larvae* is presented in Fig. 3. Similar types of pictures are observed for projections in other coordinate planes.

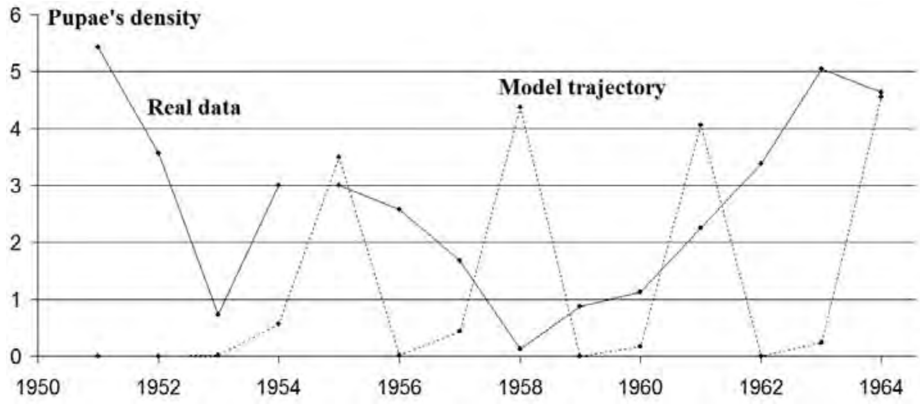
As one can see in Fig. 3, there are three non-intersected sets of points. First (smallest) set of points is close to origin. Second set of points belongs to domain $[0,5] \times [0,1]$. The last (biggest) set belongs to domain $[30,45] \times [3,6]$.



(a)



(b)



(c)

Fig. 2 Results of fitting of time series (first case). Broken lines are model trajectories. Solid lines correspond to empirical datasets. Cases: a – egg's density dynamics; b – larva's density dynamics; c – pupae's density dynamics.

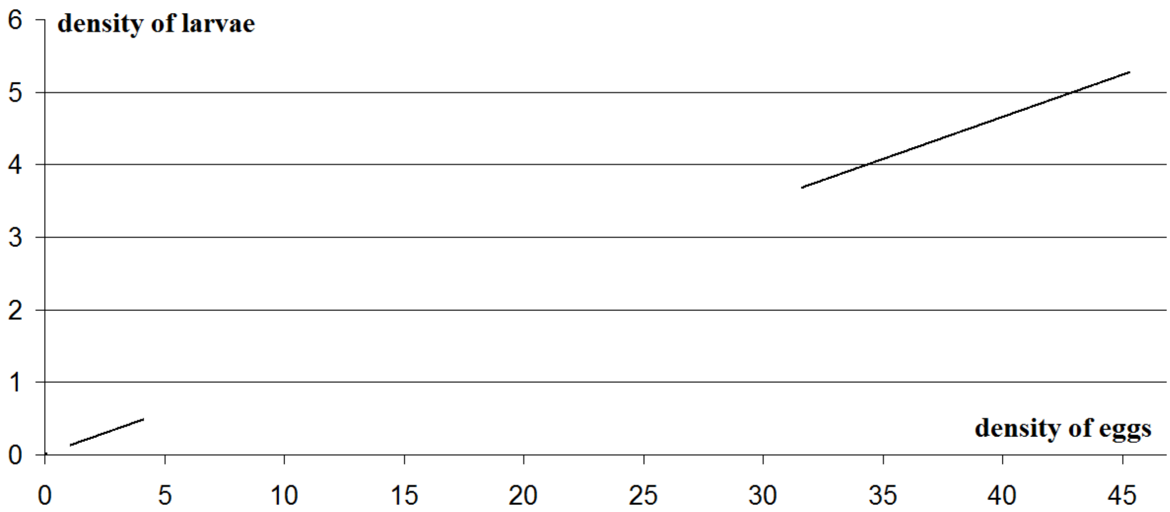


Fig. 3 Projection of asymptotic stable limit cycle on plane "density of eggs – density of larvae".

6 Second Case

Second point has following coordinates: $\mu_2 = 0.1285$, $c_1 = 213.96$, $c_2 = 0.91$, $c_3 = 1.08$, $g_1 = 0.506$, $g_2 = 0.446$, $g_3 = 1.983$, $\alpha = 0.105$. Initial value for eggs $E_1 = 182.26$. Results of fitting are presented in Fig. 4. Calculation of autocorrelation function for 20000 steps (after 10^6 free steps of process) shows that for these parameters asymptotic stable regime is strong 2-cycle of $ababab...$ type. Coordinates of cycle are following: $a = 35.49$ and $b = 5.35$ for eggs, $a = 4.56$ and $b = 0.69$ for larvae, $a = 2.83$ and $b = 0.64$ for pupae.

7 Third Case

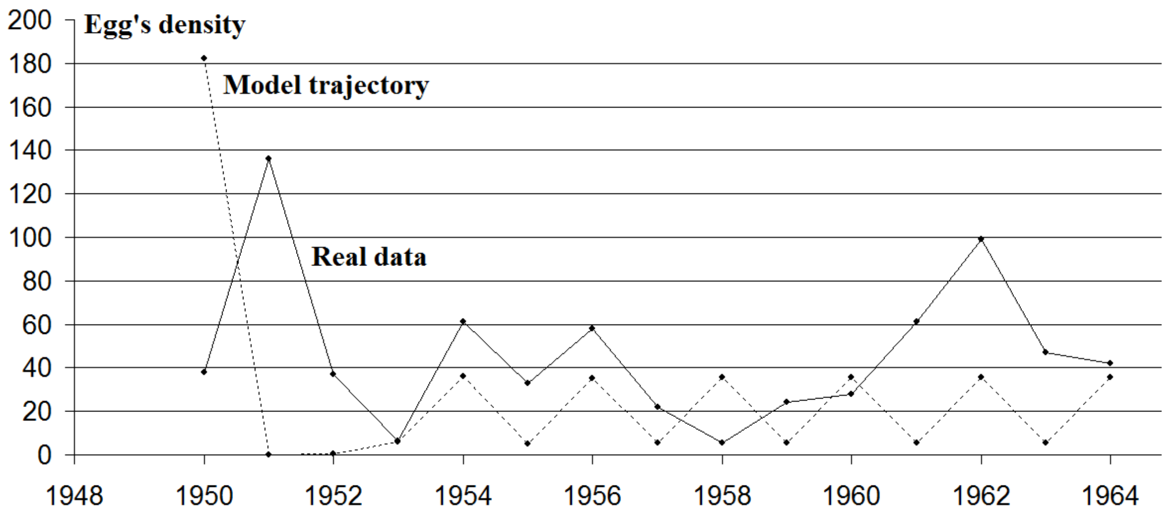
Denote as $\{E_k^*\}$, $\{L_k^*\}$, $\{P_k^*\}$ empirical samples corresponding to respective variables of ELP-model. Let β be a vector of model parameters, and $E_k = E_k(\beta)$, $L_k = L_k(\beta)$, $P_k = P_k(\beta)$ are the components of respective solution of ELP-model. For third case vector of parameters β^* was determined with minimization of the following loss-function:

$$Q(\beta) = \sum_{k=1}^{15}(E_k^* - E_k(\beta))^2 + \sum_{k=1}^{12}(L_k^* - L_k(\beta))^2 + \sum_{k=2}^{15}(P_k^* - P_k(\beta))^2 ,$$

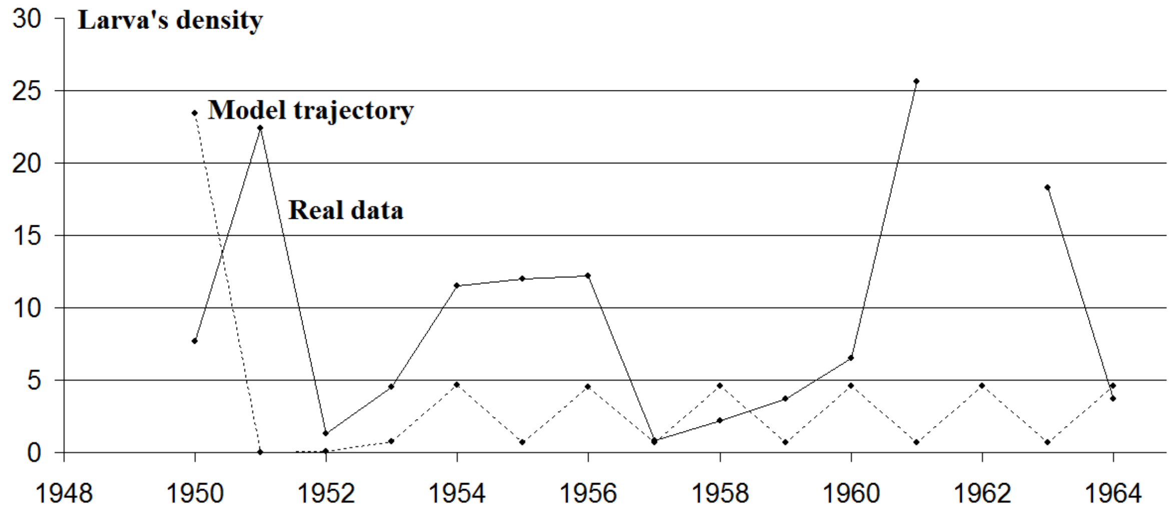
$$Q(\beta^*) = \min_{\Omega} Q(\beta) .$$

Minimum of loss-function Q was found for following values of parameters, $Q_{\min} = 20939.1$: $\mu_2 = 0.149$, $c_1 = 301.22$, $c_2 = 0.302$, $c_3 = 1.495$, $g_1 = 0.498$, $g_2 = 0.911$, $g_3 = 0.234$, $\alpha = 0.134$. Initial value for eggs $E_1 = 48.27$. For the first case $Q_{\min} = 68245.9$, for the second case $Q_{\min} = 54764.1$. Results of fitting are presented in Fig. 5.

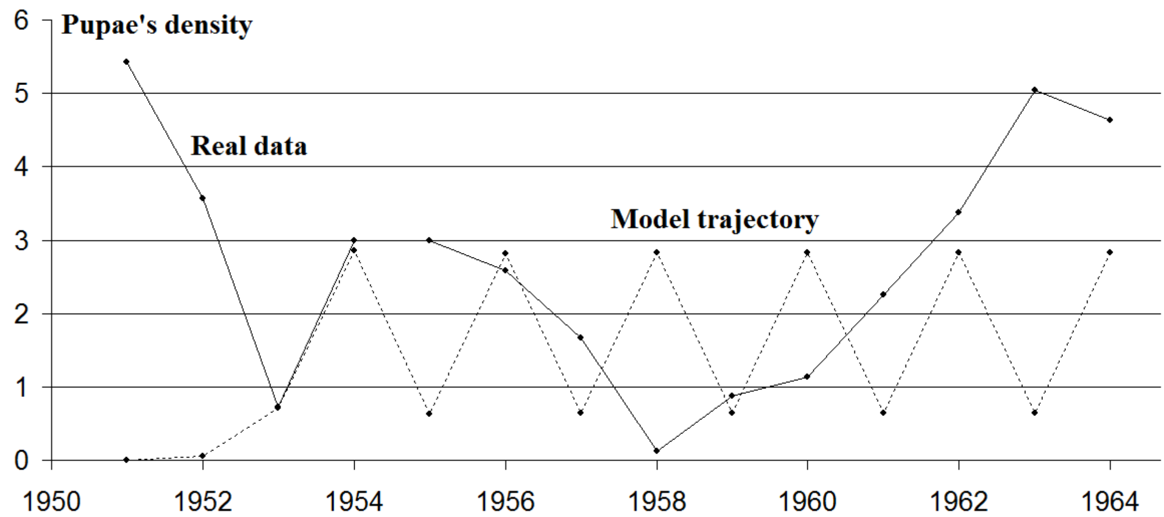
Calculation of autocorrelation function for 20000 steps (after 10^6 free steps of process) shows that for these parameters asymptotic stable regime is strong 2-cycle of *ababab...* type. Coordinates of cycle are following: $a = 47.06$ and $b = 25.71$ for eggs, $a = 7.01$ and $b = 3.83$ for larvae, $a = 2.73$ and $b = 2.29$ for pupae. Note that model trajectory presented in fig. 5 is close to stable limit 2-cycle.



(a)

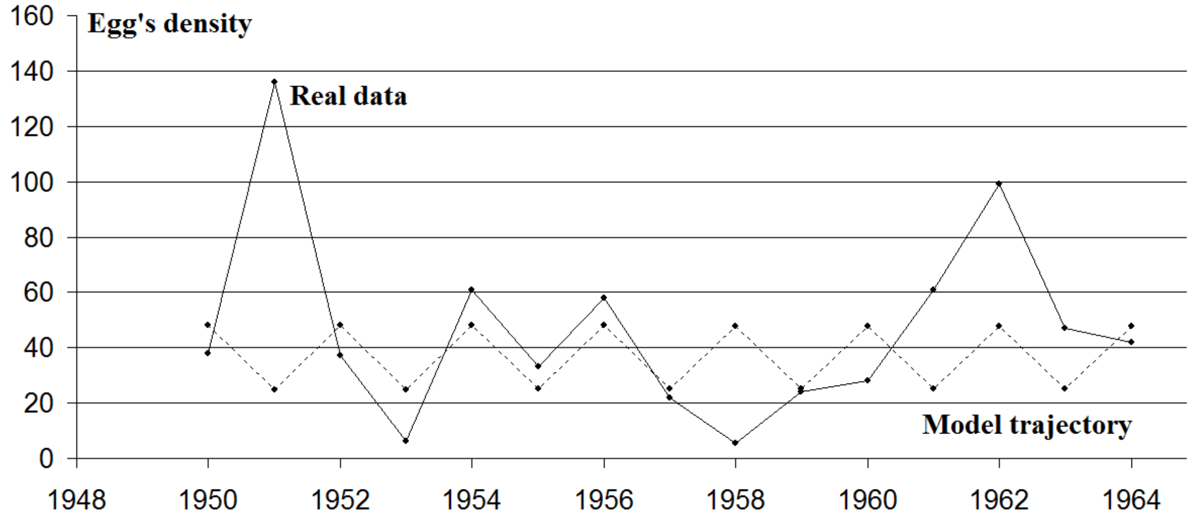


(b)

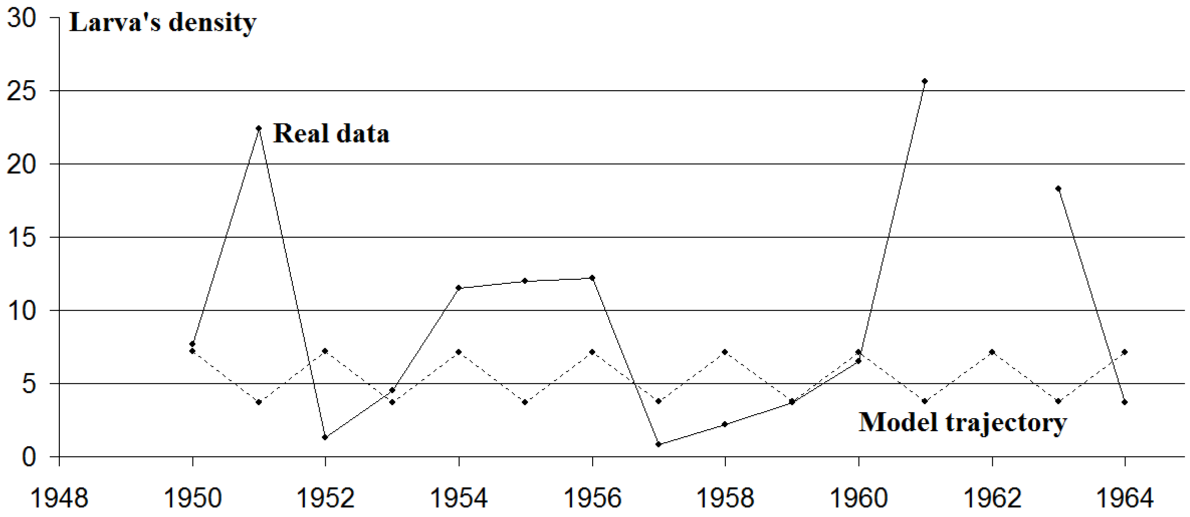


(c)

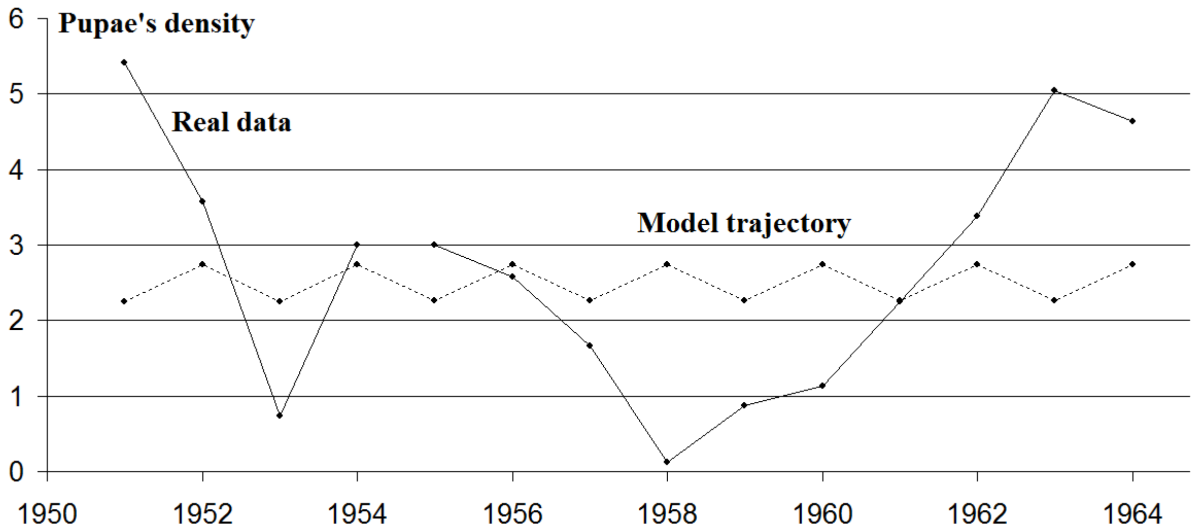
Fig. 4 Results of fitting of time series (second case). Broken lines are model trajectories. Solid lines correspond to empirical datasets. Cases: a – egg's density dynamics; b – larva's density dynamics; c – pupae's density dynamics.



(a)



(b)



(c)

Fig. 5 Results of fitting of time series (third case). Broken lines are model trajectories. Solid lines correspond to empirical datasets. Cases: a – egg's density dynamics; b – larva's density dynamics; c – pupae's density dynamics.

8 Conclusion

Presented method of model parameter estimations is based on set of (obvious) statistical criterions which must be satisfied for model when it is applied for fitting of empirical time series. It is based on assumption of symmetry of distribution of deviations between model trajectory and time series, on assumption of monotonic behavior of branches of density function, and on assumption of absence of serial correlation in sequence of residuals. These properties of sets of deviations were checked with following tests: Kolmogorov – Smirnov test, Lehmann – Rosenblatt test, Mann – Whitney test, Swed – Eisenhart test, and test on series of “jumps up” and “jumps down”. Spearmen rank correlation coefficient was used for checking of hypothesis about monotonic behavior of branches of density functions.

Set of points in space of model parameters was called as *feasible set* if deviations were satisfied to all pointed out statistical criterions (for fixed significance levels). For 5% significance level projection of fragment of feasible set onto plane “maximum of productivity of butterflies – maximum of quota of survived pupae during winter time” is presented in Fig. 1 for applied ELP-model to fitting of pine looper moth population dynamics in Netherlands.

It is known that changing of significance level can lead to increase or decrease of number of points in feasible set. In current publication it was assumed that symmetry of distribution is most important characteristics of set of deviations. Symmetry of distribution of deviations means that used method of data collection had not regular errors, and this property of distribution is truly important. For Kolmogorov – Smirnov test significance level was increased up to 95% (if we cannot reject Null hypothesis with this significance level it means that we *have to accept* hypothesis about symmetry of distribution), and for Lehmann – Rosenblatt test significance level was increased up to 80%. In a result of this operation of changing of significance levels number of points in feasible set was decreased to two points. Note that these points were found without constructing of any minimizing or maximizing functions which are normally used within the framework of least square method (LSM) and maximum likelihood method.

Inverse to LSM approach was used in third variant: sum of squared deviations was minimized within the limits of feasible set. In all considered cases we got that pine looper moth fluctuations correspond to regime of cyclic dynamics. In two cases it was strong 2-cycle, and in one case it was non-rigourous 3-cycle. It is also important to note that in all cases estimations of coefficients of transformation of pupae into larvae are very close to each other, estimations are bigger than 0.11. In two cases estimations of coefficients of surviving of individuals during the winter time, and coefficients of self-regulation (coefficient of transformation of larvae into pupae) are close to each other too. It can be considered as additional supporting background (first background is formed by the set of used statistical criterions) for hypotheses about dynamic regimes of pine looper moth.

References

Bard Y. 1974. Nonlinear Parameter Estimation. Academic Press, New York, USA

Bolshev LN, Smirnov NV. 1983. Tables of Mathematical Statistics. Nauka, Moscow, USSR

Dennis B, Desharnais RA, Cushing JM, et al. 2001. Estimating chaos and complex dynamics in an insect population. Ecol. Monographs, 71(2): 277-303

Desharnais RA, Costantino RF, Cushing JM, et al. 2001. Chaos and population control of insect outbreaks. Ecology Letters, 4: 229-235

Draper NR, Smith H. 1998. Applied Regression Analysis. John Wiley and Sons Inc, New York, USA

Hollander M, Wolfe DA, 1973. Nonparametric statistical methods. John Wiley and Sons, New York, USA

Isaev AS, Khlebopros RG, Nedorezov LV, et al. 1984. Forest Insect Population Dynamics. Nauka, Novosibirsk, USSR

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

Kendall BE, Briggs CJ, Murdoch WW et al., 1999. Why do populations cycle? A synthesis of statistical and mechanistic modeling approaches. Ecology, 80(6): 1789-1805

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

Likesh J, Laga J. 1985. Basic Tables of Mathematical Statistics. Finance and Statistics, Moscow, USSR

McCallum H, 2000. Population parameters estimation for ecological models. Brisbane: Blackwell Sciences Ltd.

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

Nedorezov LV. 1986. Modeling of Forest Insect Outbreaks. Nauka, Novosibirsk, USSR

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

Nedorezov LV. 2012. Chaos and Order in Population Dynamics: Modeling, Analysis, Forecast. LAP Lambert Academic Publishing, Saarbrucken, Germany

Nedorezov LV. 2014. Basic properties of ELP-model. Population Dynamics: Analysis, Modelling, Forecast, 3(4): 93-101

Nedorezov LV. 2015a. Application of generalized discrete logistic model for fitting of pine looper moth time series: Feasible sets and estimations of model parameters. Proceedings of the International Academy of Ecology and Environmental Sciences, 5(1): 38-48

Nedorezov LV. 2015b. Non-traditional approach to estimation of parameters of ecological models (on examples of larch bud moth and pine looper moth time series). Siberian Journal of Forest Science, 1: 72-84

Nedorezov LV. 2015c. Dynamics of a pine looper moth population in The Netherlands: estimation of parameters of a generalized discrete logistic model. EuroAsian Entomological Journal, 14(1): 93-99

Nedorezova BN, Nedorezov LV. 2012. Pine looper moth population dynamics in Netherlands: Prognosis with generalized logistic model. Proceedings of the International Academy of Ecology and Environmental Sciences, 2(2): 70-83

Nedorezov LV, Utyupin YuV. 2011. Continuous-Discrete Models of Population Dynamics: An Analytical Overview. State Public Scientific-Technical Library, Novosibirsk, Russia

Pielou EC. 1977. Mathematical Ecology. John Wiley and Sons, NY, USA

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

Tonnang H, Nedorezov LV, Ochanda H, et al. 2009a. 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. 2009b. Evaluation of discrete host – parasitoid models for diamondback moth and *Diadegma semiclausum* field time population density series. Ecological Modelling, 220: 1735-1744

Wood SN. 2001. Partially specified ecological models. Ecological Monograph, 71: 1-25