DETERMINISTIC MODELS AND IDENTIFICATION OF THEIR PARAMETERS

Václav Pink

ABSTRACT. This article deals with a possibility to identify parameters of a selected growth model of two populations coupled by a predator-prey interaction from a set of observed data. It starts with a brief description of the Gause-type model and of a property (interior equilibrium stability) important from a point of view of an application. Subsequently, data for four forms of the trophic function are simulated and then, a noise was added to the simulated data such that the coefficients of variation equal to 0.2, 0.3 and 0.4. For each data set, the parameters are estimated using a procedure implemented in the `R-language` package and the coordinates of equilibrium are computed. Then we can evaluate the effect of changing variation to the values of parameters and (un)stability of the equilibrium.

1. Introduction

The aim of this paper is to check whether parameters of a particular ordinary differential equations system that models some real process can be identified from a limited set of observed data.

The problem of parameters identification has appeared as long as real phenomena have been described by mathematical models, see, e.g., [11] and references therein, or [6] for special case of population models. A disadvantage of methods described in the literature consists in the fact that they require a set of observed data large and accurate enough. But such a data set is not available in many practical situations. An immediate motivation for the presented investigation comes out from a model of mites community in Moravian vineyards, [5]—population densities of a pest mite and of its natural enemy were checked 11-times during vegetation season and the collected data were fitted to a predator-prey model by an *ad hoc* method. The obtained results served as arguments to assertion that the mites community possesses an asymptotically

stable steady state and that a respective density of the pest lies under an injury level for grapes. This conclusion has been verified by several introductions of the predatory mite for grapevine protection. But a true consequent does not imply the correct assumptions or inference. Hence, the possibility of parameters identification from field data was reexamined in the paper [10] and the obtained results show that the identification of single parameters is not very reliable but the whole set of estimated parameters more or less correctly determines the location of the system equilibrium.

This paper searches answers to the questions:

1. How precise the observations should be to identified parameters be correct, i.e., how large relative deviation from theoretical values gives estimation of parameters with deviation small enough?

2. Do the parameters identified from data allow to establish stability of the system steady state?

The method consists in a simulation of data by adding some random deviation (a noise) to a numerical solution of a differential equation model with definite parameters and in a subsequent estimation of parameters from the generated data. The parameters are identified by recent methods provided by the software package `R-language`, cf. [12].

The subsequent section presents the predator-prey model forming a basis of the consideration and its elementary qualitative properties. The model is the standard Gause-type one introduced in the monograph [4] and dealt in all books on mathematical biology, e.g., in [2], [7]. The property important from the point of view of the application is just the asymptotic stability of interior equilibrium, not the existence or uniqueness of a limit cycle.

The simulation of data and identification of parameters are described in the third section. The obtained results are summarized in the fourth section. A brief discussion of them and an indication of an intended future research conclude the paper.

2. Model

Models of a predator-prey interaction form one of classical parts of mathematical biology, see, e.g., [2], [7]. We will use a model of prey population exhibiting intra-specific competition, i.e., a population in an environment with limited resources. In particular, we consider a population with logistic growth provided no predators are present. Predators are supposed to be specialized to the considered prey species, i.e., predatory population goes to exponential extinction if no prey is available. The predators destroy the prey population with a rate depending on the size (density) of the prey population and they convert the killed

---

[1]Gause, Georgyi Frantsevitch (December 27, 1910–May 4, 1986), Russian microbiologist.

Solution for (ax)/(ax+1), original parameters

Solution for ax/ax+1, original parameters

Solution for (ax^2)/(ax^2+1), original parameters

Solution for ax^2/ax^2+1, original parameters

a)

b)

FIGURE 1. Solution (left) and trajectory (right) of the system (1) with the parameters $r = 5$, $K = 300$, $d = 4$, $S = 15$, $\kappa = 0.7$, initial values $x(0) = 9$, $y(0) = 2$, and the trophic function $\varphi = \varphi_1$ given by the relation (2) with $a = 0.01$ and $k = 1$ (Fig. a) or $k = 2$ (Fig. b).

prey into their growth rate with some efficiency. That is, we use the Gause-type predator-prey ODE model in the form

$$x' = rx\left(1 - \frac{x}{K}\right) - S\varphi(x)y,$$
$$y' = -dy + \kappa S\varphi(x)y. \tag{1}$$

Here, $x = x(t)$ and $y = y(t)$ denote a time dependent size (or density) of the prey and predator populations, respectively, the parameters r, K, and d denote the prey growth rate, carrying capacity for the prey population, and the predator death rate, respectively. The real function φ is called trophic function; $\varphi(x)$

Solution for 1- e^(-ax), original parameters

Solution for 1-e^(-ax), original parameters

a)

Solution for 1- e^(-ax^2), original parameters

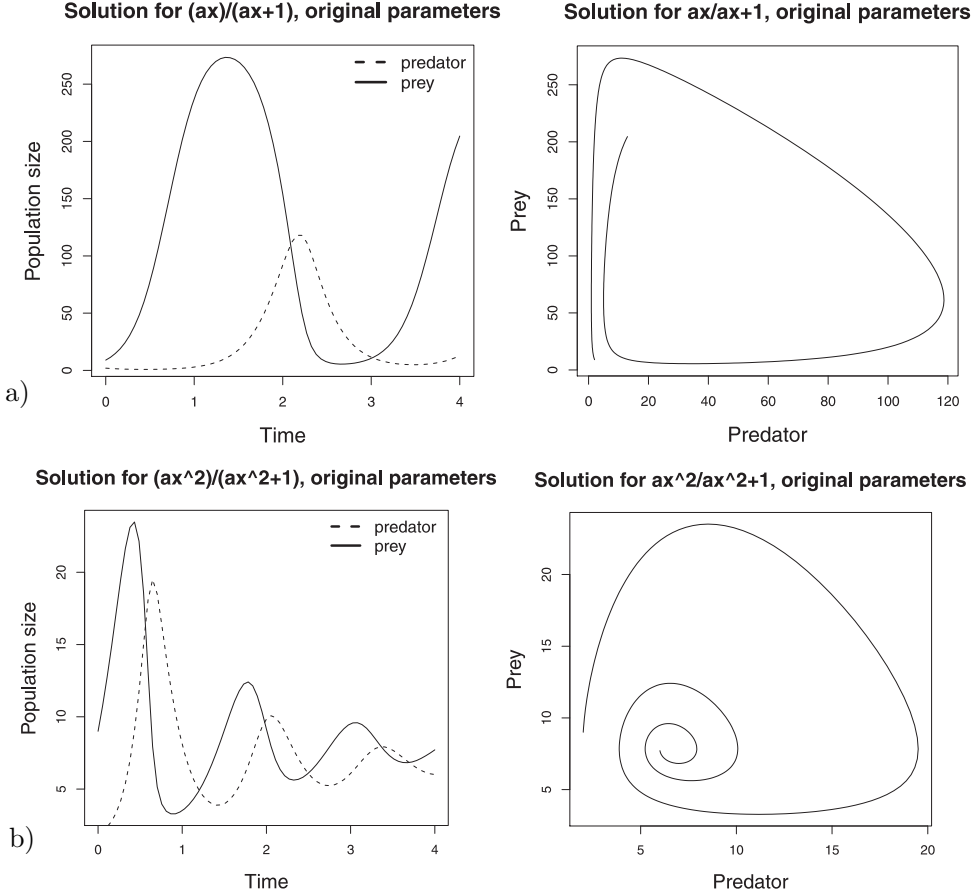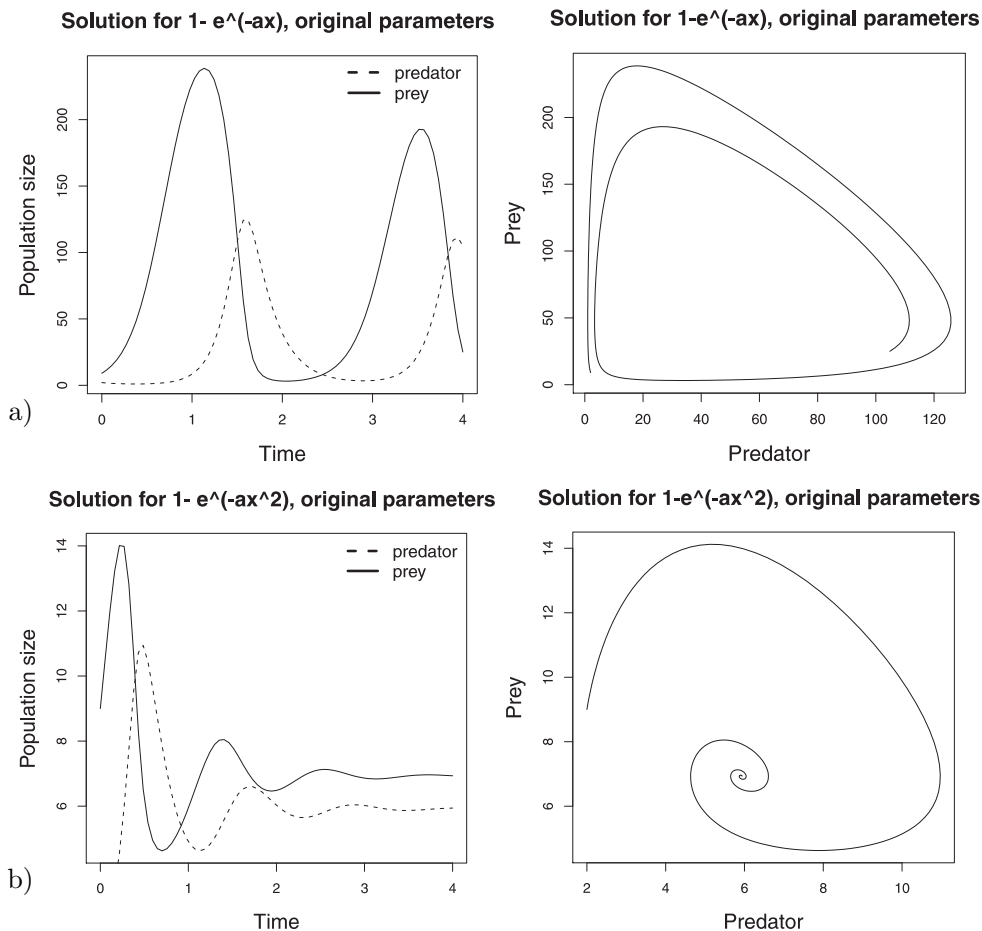**Solution for 1-e^(-ax^2), original parameters**

b)

FIGURE 2. Solution (left) and trajectory (right) of the system (1) with the parameters $r = 5$, $K = 300$, $d = 4$, $S = 15$, $\kappa = 0.7$, initial values $x(0) = 9$, $y(0) = 2$, and the trophic function $\varphi = \varphi_2$ given by the relation (2) with $a = 0.01$ and $k = 1$ (Fig. a) or $k = 2$ (Fig. b).

expresses a relative satiety of predator provided the size of prey population equals to x. We assume that the function φ is strictly increasing with the properties

$$\varphi(0) = 0 \quad \text{and} \quad \lim_{x \to \infty} \varphi(x) = 1,$$

i.e., if no prey is available the predators starve, if prey population is large, predators are totally satisfied. The parameter S represents the level of satiety of predator (the maximal possible size of prey population destroyed by predator population of unit size in a unit time), hence $S\varphi(x)$ represents a size of prey population destroyed by a predator population of unit size in unit time. Finally,

4

κ denotes the efficiency of conversion destroyed prey into the predator growth rate. All of the parameters are positive.

For brevity, let us introduce the notation

$$q(x) = r\left(1 - \frac{x}{K}\right).$$

The function q denotes the size (density) dependent growth rate of the prey population. We remind one property of the system (1) which is important from the point of view of the mentioned application: If $d < \kappa S$ then there exists an isolated steady state

$$(x^*, y^*) = \left(\varphi^{-1}\left(\frac{d}{\kappa S}\right), \frac{\kappa x^* q(x^*)}{d}\right).$$

Moreover, if $C_S^* := Sy^*\varphi'(x^*) + q(x^*) - x^*q'(x^*) > 0$, then (x^*, y^*) is locally asymptotically stable equilibrium.

We consider the trophic function φ in the form

$$\varphi(x) = \varphi_1(x) = \frac{ax^k}{ax^k + 1} \quad \text{or} \quad \varphi(x) = \varphi_2(x) = 1 - e^{-ax^k}, \tag{2}$$

where a is a positive parameter and $k \in \{1, 2\}$. The values $k = 1$ and $k = 2$ express the trophic function of the Holling type II and of the Holling type III, respectively. (The case $k = 1$ is used for invertebrates, the case $k = 2$ is used in environments where there are refuges for prey and/or predator has other food resources).

Numerical solutions of the system (1) with the trophic function defined by one of the relations (2) can be found by the procedure `lsoda` of the program `R-language`. The algorithm automatically selects either Adams method (order 1–12 for "non-stiff systems") or backward differentiation formulas (order 1–5 for "stiff systems"), for details see [8]. The results are plotted on Figures 1 and 2, the parameters were chosen such that the system possess asymptotically stable equilibrium.

3. Simulations

First, the set of equidistant nodal points $\mathbf{t} = (t_0 = 0, t_1 = 0.054, \ldots, t_{74} = 4)$ for independent variable (time) was chosen. The number 74 of nodal points is inspired by the mentioned application—it represents a maximal possible number of observations of mites community during a vegetation season of grapevine. For each of the nodal point, the corresponding values $x(t_i)$ and $y(t_i)$ were computed using numerical solution of the system (1) with one of the trophic functions (2); the parameters and the initial conditions were set to the values:

$r = 5$, $K = 300$, $S = 15$, $a = 0.01$, $\kappa = 0.7$, $d = 4$, $x(0) = x_0 = 9$, $y(0) = y_0 = 2$.

TABLE 1. Characteristics of estimated parameters of the system (1) with the trophic function (2), $\varphi = \varphi_1$, $k = 1$.

Coef. of Var.	Parameter	Min	Max	Mean	Median	Orig. Value
0.2	r	4.9054	5.5741	5.06388	5.01511	5
	K	283.7399	320.5527	304.9821	306.0766	300
	S	13.02903	16.3974	14.6521	14.6686	15
	a	0.009579	0.01241	0.01031	0.01024	0.01
	κ	0.6026	0.7770	0.7083	0.7088	0.7
	d	3.8685	4.2780	4.01945	4.01461	4
	x^*	58.6329	72.9569	63.5903	63.4767	61.5384
	y^*	39.4608	49.8418	44.1085	43.8443	42.8008
	C_S^*	4.5780	5.6015	5.2834	5.3214	5.4090
0.3	r	4.4935	5.8356	5.1448	5.1163	5
	K	273.5880	383.019	311.9426	306.7179	300
	S	12.4718	19.2276	14.7922	14.3179	15
	a	0.007819	0.0131	0.01025	0.0101	0.01
	κ	0.5226	0.8398	0.7114	0.7288	0.7
	d	3.3036	4.2815	4.04583	4.1057	4
	x^*	51.6688	79.6829	64.3225	64.0296	61.5384
	y^*	32.9512	52.4011	44.1766	44.9932	42.8008
	C_S^*	4.1302	6.06564	5.2486	5.3196	5.4090
0.4	r	4.4454	5.7311	5.1515	5.1355	5
	K	280.6219	368.2327	328.7662	326.8635	300
	S	12.6715	18.7131	14.9691	14.6178	15
	a	0.004752	0.01348	0.009779	0.009935	0.01
	κ	0.5176	0.8381	0.7116	0.7164	0.7
	d	3.4876	4.8488	4.0307	3.9476	4
	x^*	35.6058	84.7693	62.8715	62.7672	61.5384
	y^*	33.3950	51.3651	43.4612	44.1285	42.8008
	C_S^*	4.02395	7.2262	5.3936	5.3935	5.4090

The vector $\mathbf{t}$ represents observation times and the values $x(t_i)$ and $y(t_i)$ represent expected sizes of prey and predator populations, respectively. But observed sizes of population use to be affected by errors caused by inaccuracies in measurement, and/or by the fact that the real population sizes vary randomly around theoretical values, which are described by the deterministic model. We assume that the random variations are multiplicative and positive, since the population sizes need to be positive; the assumption is discussed in details in [6]. More precisely, we generate the values:

$$X_i = x(t_i) \cdot \varepsilon_i, \quad Y_i = y(t_i) \cdot \eta_i, \qquad i = 0, 1, \dots, 74,$$

TABLE 2. Characteristics of estimated parameters of the system (1) with the trophic function (2), $\varphi = \varphi_1$, $k = 2$.

Coef. of Var.	Parameter	Min	Max	Mean	Median	Orig. Value
0.2	r	4.7685	5.2005	5.01376	5.02558	5
	K	84.1849	999.2072	651.4418	813.8820	300
	S	11.5285	18.4706	15.1851	15.4108	15
	a	0.007918	0.01261	0.009776	0.009588	0.01
	κ	0.5713	0.9039	0.7048	0.6923	0.7
	d	3.6914	4.3556	4.02887	4.01409	4
	x^*	7.3344	8.4744	7.8251	7.8113	7.8446
	y^*	5.5374	8.5447	6.6661	6.5042	6.6846
	C_S^*	10.00529	11.1826	10.7963	10.8478	10.76711
0.3	r	4.7255	5.5051	4.9951	5.0009	5
	K	89.1605	999.9414	588.9833	549.5614	300
	S	12.7545	17.2549	15.3502	15.2993	15
	a	0.007366	0.01101	0.008946	0.008907	0.01
	κ	0.5948	0.8537	0.7056	0.7040	0.7
	d	3.6887	4.4385	4.005995	3.9922	4
	x^*	7.2991	8.2948	7.6983	7.6894	7.8446
	y^*	5.8480	7.7572	6.6016	6.5742	6.6845
	C_S^*	10.3520	11.2067	10.8692	10.9066	10.7671
0.4	r	4.4618	5.5870	5.03669	5.01698	5
	K	122.9133	999.7313	622.7539	763.145	300
	S	12.7716	19.5232	15.5764	15.6520	15
	a	0.006288	0.0115	0.008626	0.008826	0.01
	κ	0.5757	0.8564	0.7019	0.6896	0.7
	d	3.7573	4.1581	3.9866	3.9942	4
	x^*	7.0858	8.0877	7.6432	7.6878	7.8446
	y^*	5.3008	7.9398	6.5540	6.4349	6.6845
	C_S^*	10.5265	11.3413	10.9092	10.8880	10.7671

where ε_i, η_i are realizations of random variables with lognormal distribution and the mean value equal to 1. The variability of biological processes usually equals to 20 % (for details see, e.g., [1]) but it was higher in the mentioned mites community. Hence, the coefficient of variation of the random deviations ε, η from the set $\{0.2, 0.3, 0.4\}$ come into usage in simulations.

The generated data set X_i, Y_i, $i = 1, 2, \ldots, 74$ serve as input for estimation of parameters of the model (1). We have utilized the subroutine `fitOdeModel` that is contained in the library `simecol` of program `R-language`. The procedure computes parameter values that minimize the sum of squared differences between simulated data and data obtained by numerical solution of the differential

TABLE 3. Characteristics of estimated parameters of the system (1) with the trophic function (2), $\varphi = \varphi_2$, $k = 1$.

Coef. of Var.	Parameter	Min	Max	Mean	Median	Orig. Value
0.2	r	4.6606	5.6638	5.02882	4.9986	5
	K	272.9173	319.0244	299.3653	297.9745	300
	S	12.8386	16.3948	14.5279	14.3067	15
	a	0.00850	0.01225	0.01025	0.01019	0.01
	κ	0.6233	0.8199	0.7259	0.7403	0.7
	d	3.7207	4.2841	4.05623	4.06192	4
	x^*	43.5797	55.5036	48.9388	48.5068	47.9573
	y^*	32.1913	41.3325	36.5895	36.9921	35.2545
	C_S^*	6.02301	7.0246	6.6022	6.6354	6.6750
0.3	r	4.5358	5.7533	5.07464	5.07339	5
	K	275.2709	342.5684	315.8076	314.1207	300
	S	12.8822	18.08814	15.3884	15.3288	15
	a	0.007209	0.01413	0.009593	0.009596	0.01
	κ	0.5702	0.8724	0.6993	0.6909	0.7
	d	3.2270	4.7425	4.007732	3.9656	4
	x^*	37.9148	63.3217	47.4871	47.6209	47.9573
	y^*	29.2035	41.3320	34.7551	34.5605	35.2545
	C_S^*	5.5339	7.4522	6.7291	6.7110	6.6750
0.4	r	4.1351	5.8190	5.01904	5.1159	5
	K	281.3685	364.2503	322.4318	323.1469	300
	S	12.09450	18.8539	14.9866	14.6891	15
	a	0.006111	0.01138	0.009299	0.009957	0.01
	κ	0.5833	0.9466	0.7406	0.7150	0.7
	d	3.1852	4.8987	4.0490	4.02660	4
	x^*	32.4917	56.2687	46.5027	48.7507	47.9573
	y^*	27.8114	43.8549	35.7251	36.002010	35.2545
	C_S^*	6.05870	7.7723	6.8131	6.6329	6.6750

equations. It contains several optimization methods; we have chosen the Nelder-
-Mead algorithm for the presented computations. The procedure `fitOdeModel`
is described in the report [9].

4. Results

Twenty data sets were independently generated by the technique described in the previous section for each of the trophic functions φ_i, for each of the values $k = i$, $i = 1, 2$ and for each of the variation coefficient values from the set $\{0.2, 0.3, 0.4\}$. The parameters of the system (1) were estimated using

TABLE 4. Characteristics of estimated parameters of the system (1) with the trophic function (2), $\varphi = \varphi_2$, $k = 2$.

Coef. of Var.	Parameter	Min	Max	Mean	Median	Orig. Value
0.2	r	4.9085	5.3743	5.03162	4.9954	5
	K	59.9781	996.8408	536.0290	433.2503	300
	S	13.3435	15.7721	14.6935	14.6941	15
	a	0.008475	0.0120	0.009964	0.009860	0.01
	κ	0.6754	0.8238	0.7244	0.7190	0.7
	d	3.8631	4.5596	4.07960	4.04027	4
	x^*	6.4469	7.5550	6.9649	6.9645	6.9251
	y^*	5.7038	6.4327	6.03872	6.03350	5.9196
	C_S^*	11.9209	12.5377	12.3048	12.3332	12.38233
0.3	r	2.7641	5.8616	5.1312	5.1714	5
	K	64.4184	999.6789	745.4924	979.0528	300
	S	13.3865	25.5353	16.6916	16.3931	15
	a	0.004461	0.01131	0.008605	0.008602	0.01
	κ	0.5771	0.8282	0.6856	0.6866	0.7
	d	1.4987	4.5587	3.9844	4.08773	4
	x^*	3.6939	7.3532	6.6274	6.8422	6.9251
	y^*	4.2316	7.3433	5.6437	5.4995	5.9196
	C_S^*	11.7194	14.1465	12.5383	12.4699	12.3823
0.4	r	0.7209	8.8331	5.4147	5.2630	5
	K	219.3157	999.1424	804.2772	988.9057	300
	S	2.09622	24.6375	17.3078	16.3516	15
	a	0.005263	0.06712	0.01141	0.008943	0.01
	κ	0.3952	0.9636	0.6613	0.6752	0.7
	d	1.1449	5.0301	3.8762	3.9758	4
	x^*	4.5933	7.8663	6.5773	6.6754	6.9251
	y^*	3.7064	6.3016	5.1478	5.4117	5.9196
	C_S^*	11.8585	13.7483	12.6261	12.5825	12.3823

the procedure `fitOdeModel` and the coordinates x^*, y^* of the steady state together with the criterion of stability C_S^* were computed for each of the data sets. The values for the minimizing iterations were subjected to the restrictions:

$$0.1 < r < 15,\ 5 < K < 1000,\ 0 < S < 30,\ 0 < a < 10,\ 0 < \kappa < 1,\ 0 < d < 15.$$

Basic statistical characteristics of the estimates are summarized in the Tables 1–4.

The mean values and the medians of the parameters r, d, S, κ, a and of the coordinates x^*, y^* are sufficiently close to the true values. This optimistic observation is questioned by excessively out-lied extreme values of them. Moreover, the estimated carrying capacity K is unacceptable high for both of the trophic

functions with $k = 2$. Hence, an answer to the first question from the introduction should be that the relative deviations from expected values of populations sizes for a reliable identification of parameters need to be smaller than the biologically realistic ones. However, the parameters of the system (1) (with possible exception of the carrying capacity K) may be estimated from field data provided the community evolves according to the model.

Note also, that the range of estimated values increases with increasing coefficient of variation in majority of cases. But such an anticipated property is not true for all of the cases, for example, the range of estimated parameter d for the model with $\varphi = \varphi_1$, $k = 2$ for coefficient of variation equals to 0.4 is less than the one for coefficient of variation equals to 0.2.

The value of the criterion C_S^* is positive for all of the simulated data. Therefore, our investigation allows the positive answer to the second question formulated in the introduction.

5. Conclusions and directions for a future research

The provided work verifies that the `R-language` procedure `fitOdeModel` is an appropriate tool for identifying parameters of the ODE model of a natural process. In the particular case of the Gause-type predator-prey model (1), the prey intrinsic growth rate r, the predator death rate d, the satiety level S of predator, and the efficiency κ of processing destroyed prey by predator can be estimated from field data under realistic mild assumptions. A problem arises with identification of the prey carrying capacity K—it is not identified correctly for the model with exponential trophic function φ_2. This phenomenon may be caused by the fact that the solution of the ODE system tends quickly to the equilibrium value (cf. Figures 1 and 2) whose prey coordinate is far from the carrying capacity. Hence a comparison of parameter identification for systems with stable equilibrium point and with unstable steady state represents one line for a subsequent research.

The location and the stability of the interior equilibrium is identified correctly by the presented method. The question whether instability of steady state can be also recognized correctly remains open and it will be dealt in a future simulation study.

The presented results—identification of parameters from observed data—were obtained under the assumption that the underlying deterministic process of a natural phenomenon is known. But this need not be the actual case. Hence, investigation whether the proposed methods of a model identification are able to distinguish between two or more possible models of an observed ecological community poses a challenge for a future work.

REFERENCES

[1] BOLKER, M. B.: *Ecological Models and Data in R.* Princeton Univ. Press, Princeton, NJ, 2008.

[2] BRITTON, N. F.: *Essential Mathematical Biology.* Springer-Verlag, London, 2003.

[3] CARPENTER, S. R.—COTTINGHAM, K. L.—STOW, C. A.: *Fitting predator-prey models to time series with observation errors*, Ecology **75** (1994), 1254–1264.

[4] GAUSE, G. F.: *The Struggle for Existence.* Williams&Wilkins, Baltimore, 1934.

[5] HLUCHÝ, M.—POSPÍŠIL, Z.: *Use of the predatory mite Typhlodromus pyri Scheuten (Acari: Phytoseiidae) for biological protection of grape vines from phytophagous mites*, in: Proc. of the 8th Internat. Congress of Acarology, České Budějovice, CZ, 1990 (F. Dusbábek and V. Bukva, eds.), Modern Acarology, Vol. 2, Academia, Prague and SPB Academic Publishing BV, The Hague, 1991, pp. 655–660.

[6] MCCALLUM, H.: *Population Parameters. Estimation for Ecological Models.* Blackwell Science, Oxford, 2000.

[7] MURRAY, J. D.: *Mathematical Biology. I. An Introduction,* in: Interdiscip. Appl. Math., Vol. 17, Springer-Verlag, New York, NY, 2002.

[8] PETZOLD, L. R.: *Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations*, SIAM J. Sci. Stat. Comput. **4** (1983), 136–148.

[9] PETZOLDT, T.: *simecol-Howto: Tips, Tricks and Building Blocks.* Technische Universität Dresden, 2010,
`http://cran.uvigo.es/web/packages/simecol/vignettes/simecol-howtos.pdf`.

[10] POSPÍŠIL, Z.: *An identification of a predator-prey model. A simulation study.* Aplimat––J. Appl. Math. **1** (2008), 271–277.

[11] SCHITKOWSKI, K.: *Numerical Data Fitting in Dynamical Systems.* Kluwer Acad. Publ., Dordrecht, 2002.

[12] R DEVELOPMENT CORE TEAM: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2005,
`http://www.R-project.org`.

Department of Mathematics and Statistics
Faculty of Science
Masaryk University Brno
Kotlářská 2
CZ–611-37 Brno
CZECH REPUBLIC
E-mail: pink@mail.muni.cz