Parametric identification of the model with one predator and two prey species

M Shatalov[α,β], JC Greeff, SV Joubert and I Fedotov
Department of Mathematics and Statistics
Tshwane University of Technology Pretoria, South Africa

Abstract

In this paper a mathematical model describing the interaction of a lion population with that of the zebra and wildebeest populations is considered. The traditional method uses a model with known coefficients and a CAS numerical routine to determine a numerical solution that can be compared to historical data about the populations. The numerical values of the coefficients involved are usually "educated guesses" made by the team consisting of, for example, biologists, game rangers and experienced applied mathematicians. The coefficients are usually described in terms of quantities such as "carrying capacity", "birth rate" *et cetera*, and might mean little to the mathematician. In this paper an "inverse method" is considered, that is, a method easy enough for senior undergraduate and graduate mathematics majors to understand and apply as part of a "biomechanics" team in the field. This approach considers the model in question to have unknown coefficients. Using a CAS, numerical integration is applied using the historical data and then elementary statistical methods are used to determine the value of the coefficients.

1 Introduction

A deterministic mathematical model describing the interaction of one predator and two prey species is considered. Zebras and wildebeest, cooperating in nature, are considered as the prey and the lion, killing both zebras and wildebeests, is treated as the predator. The logistic mathematical model is formulated in terms of a non-linear dynamical system, described by three differential equations of first order, with each equation containing linear and quadratic terms. The model is linear with regard to unknown coefficients that must be derived from available statistical data. The method of parametric identification of the unknown coefficients is based on numerical time integration of both parts of the dynamical system, formulation of the goal function, subjected to minimization, which is the quadratic function of the parameters. Minimization of this function by means of the least squares method gives us a system of linear algebraic equations and solution of these equations gives us a set of unknown coefficients. The method is applicable to a broad spectrum of ecological systems described by a system of differential equations. To illustrate the application of this method, the problem of zebra, wildebeest and lion interacting is considered, and conclusions are formulated.

[α] mshatlov@csir.co.za
[β] Permanent address: Sensor Science and Technology of CSIR Manufacturing and Materials, P.O. Box 305, Pretoria 0001, South Africa

Explicit examples of how a model of mathematical biology is being developed, improved and tested are rare in scientific literature, probably because of many uncertainties, shortcomings in data, artificial assumptions and intuitive decisions in the development process. In the present paper a model with one predator and two prey based on the Lotka-Volterra model, in which logistic growth as well as mutualism are taken into consideration for the prey species, is discussed. This model fits the historical data from the Kruger National Park, which is located on the international border between the Republic of South Africa and Mozambique. The central grasslands of the Kruger National Park support huge herds of zebra (*Equus burchelli*) and wildebeest (*Connochaetes taurinus*), which are considered as prey and the lion (*Panthera leo*), which is their principal predator.

The general task of modelling is to design a model as simply as possible while still allowing specific problems to be addressed. The proposed method for identifying the unknown coefficients of the mathematical model is original and applicable to a broad class of dynamical systems, which are linear with respect to the unknown coefficients. The method consists of time integration of the system equations with variable upper time limit, by means of which the problem is converted to the solution of an over-determined system of linear algebraic coefficients. When applying the least squares method to this system, one can estimate the unknown coefficients and check the closeness of the resulting system of ordinary differential equation and its solution to statistical data. If the solution is sufficiently close to the statistical data, one can use the model for predictive purposes. If the closeness between the solution and statistical data is not satisfactory, the mathematical model should be modified to achieve an acceptable correspondence. Although the described model used in this exercise is simple and conservative, it nevertheless suggests directives for possible managerial actions.

2 The mathematical model with one predator and two prey species

The model describing interaction of one predator and two prey species was analyzed by Bazykin [1]. The logistic growth and mutualism effects for prey were outside the scope of this model. These effects were considered by May [5], and applied to a specific situation by Fay and Greeff ([2] to [4]). In the present paper a regular method of the parametric identification as applied to the model discussed by Fay and Greeff is considered.

The main assumptions are as follows:

- The model is of the Lotka-Volterra type and is described by the system of equations:

$$\begin{aligned}\frac{dx}{dt} &= x\cdot\left(a_{10}-a_{11}\cdot x-a_{12}\cdot y+a_{13}\cdot z\right)\\ \frac{dy}{dt} &= y\cdot\left(-a_{20}+a_{21}\cdot x+a_{22}\cdot y+a_{23}\cdot z\right)\\ \frac{dz}{dt} &= z\cdot\left(a_{30}+a_{31}\cdot x-a_{32}\cdot y-a_{33}\cdot z\right),\end{aligned} \tag{1}$$

where x, y and z represent the wildebeest, lion and zebra population densities respectively, and $a_{10}, a_{11},\ldots,a_{33}$ are zero or unknown positive coefficients.

- Independence of the coefficients in time, indicates the conservativeness of the model, that is, time independence of the variables.
- The model suggests a "linear response", which means that predators consume prey at a rate proportional to their number, and to the number of available prey. The effects of relative "fullness" or "saturation" of predators are not taken into consideration.
- Since the predator species is nomadic, the concepts of carrying capacity and intra-species competition is not applicable to this species.
- The reproduction function of the predator is assumed to be homogeneous in time, and the effects of seasonal calving for prey are not considered.

Logistic growth of the prey species is described by the terms $x\cdot\left(a_{10}-a_{11}\cdot x\right)$ and $z\cdot\left(a_{30}-a_{33}\cdot z\right)$, which in the absence of predators and mutualism gives the stable steady state values of the prey $x_0=\frac{a_{10}}{a_{11}}$ and $z_0=\frac{a_{30}}{a_{33}}$. Logistic growth of the prey in the absence of a predator but with mutualism taken into account, is described by the terms $x\cdot\left(a_{10}-a_{11}\cdot x+a_{13}\cdot z\right)$ and $z\cdot\left(a_{30}+a_{31}\cdot x-a_{33}\cdot z\right)$. The steady-state solution is now given by $x_0=\frac{a_{10}a_{33}+a_{30}a_{13}}{a_{11}a_{33}-a_{13}a_{31}}$ and $z_0=\frac{a_{30}a_{11}+a_{10}a_{31}}{a_{11}a_{33}-a_{13}a_{31}}$. By using standard eigenvalue methods, this solution is stable if $a_{11}a_{33}-a_{13}a_{31}>0$. Hence the steady-state solution of the System (1) is the root of the following system of linear algebraic equations:

$$\begin{aligned}a_{11}\cdot x+a_{12}\cdot y-a_{13}\cdot z &= a_{10}\\ a_{21}\cdot x+0\cdot y+a_{23}\cdot z &= a_{20}\\ -a_{31}\cdot x+a_{32}\cdot y+a_{33}\cdot z &= a_{30}.\end{aligned} \tag{2}$$

Note that $a_{22}=0$ since intra-species competition is not applicable to the predator species.

3 Solution of the problem

Te following problem will be solved: assume that the ecological process of the predator-prey interaction is described by equations (1) and that statistical data on population numbers is available on the limiting time interval $t\in\left[t_0=0,t_N=T\right]$, as given in table 1.

Table 1

$t_0=0$	t_1	t_2	$\cdots$	t_N
x_0	x_1	x_2	$\cdots$	x_N
y_0	y_1	y_2	$\cdots$	y_N
z_0	z_1	z_2	$\cdots$	z_N

The problem is to identify the coefficients $a_{10},a_{11},\ldots,a_{33}$ of System (1) from the experimental data in table 1. Here it is assumed that $N>>1$, that is, the data is rich enough to make this statistical inference.

Our approach to the problem is based on transformation of System (1) to an over-determined system of linear algebraic equations, using the experimental data. First the equations in (1) are integrated with regards to time to obtain:

$$
\begin{aligned}
a_{10}\cdot\int_0^t x(\tau)d\tau-a_{11}\cdot\int_0^t x^2(\tau)d\tau-a_{12}\cdot\int_0^t x(\tau)\cdot y(\tau)d\tau+a_{13}\cdot\int_0^t x(\tau)\cdot z(\tau)d\tau=x(t)-x(0)\\
-a_{20}\cdot\int_0^t y(\tau)d\tau+a_{21}\cdot\int_0^t x(\tau)\cdot y(\tau)d\tau+a_{23}\cdot\int_0^t y(\tau)\cdot z(\tau)d\tau=y(t)-y(0)\\
a_{30}\cdot\int_0^t z(\tau)d\tau+a_{31}\cdot\int_0^t x(\tau)\cdot z(\tau)d\tau-a_{32}\cdot\int_0^t y(\tau)\cdot z(\tau)d\tau-a_{33}\cdot\int_0^t z^2(\tau)d\tau=z(t)-z(0),
\end{aligned}
\tag{3}
$$

where $\tau\in[0,t]$. Now the problem of determination of coefficients $a_{10},a_{11},a_{12},a_{13}$ can be separated from the determination of coefficients a_{20},a_{21},a_{23} and $a_{30},a_{31},a_{32},a_{33}$. The algorithm of the coefficients' identification is obtained for the first set only, since the other coefficients can be obtained correspondingly. Coefficients $a_{10},a_{11},a_{12},a_{13}$ can be found from the first equation in System (3) and the statistical data in table 1. Table 1 is expanded as follows, using a Taylor expansion:

Table 2

$t_0 = 0$	t_1	t_2	$\cdots$	t_N
x_0	x_1	x_2	$\cdots$	x_N
y_0	y_1	y_2	$\cdots$	y_N
z_0	z_1	z_2	$\cdots$	z_N
${x_0}^2$	${x_1}^2$	${x_2}^2$	$\cdots$	${x_N}^2$
$x_0 \cdot y_0$	$x_1 \cdot y_1$	$x_2 \cdot y_2$	$\cdots$	$x_N \cdot y_N$
$x_0 \cdot z_0$	$x_1 \cdot z_1$	$x_2 \cdot z_2$	$\cdots$	$x_N \cdot z_N$
0	$\Delta x_1 = x_1 - x_0$	$\Delta x_2 = x_2 - x_0$	$\cdots$	$\Delta x_N = x_N - x_0$
0	$I_1^{(1)} = \int_0^{t_1} x(\tau)d\tau$	$I_1^{(2)} = \int_0^{t_2} x(\tau)d\tau$	$\cdots$	$I_1^{(N)} = \int_0^{t_N} x(\tau)d\tau$
0	$I_2^{(1)} = -\int_0^{t_1} x^2(\tau)d\tau$	$I_2^{(2)} = -\int_0^{t_2} x^2(\tau)d\tau$	$\cdots$	$I_2^{(N)} = -\int_0^{t_N} x^2(\tau)d\tau$
0	$I_3^{(1)} = -\int_0^{t_1} x(\tau)\cdot y(\tau)d\tau$	$I_3^{(2)} = -\int_0^{t_2} x(\tau)\cdot y(\tau)d\tau$	$\cdots$	$I_3^{(N)} = -\int_0^{t_N} x(\tau)\cdot y(\tau)d\tau$
0	$I_4^{(1)} = \int_0^{t_1} x(\tau)\cdot z(\tau)d\tau$	$I_4^{(2)} = \int_0^{t_2} x(\tau)\cdot z(\tau)d\tau$	$\cdots$	$I_4^{(N)} = \int_0^{t_2} x(\tau)\cdot z(\tau)d\tau$

All notation is introduced in Table 2. Corresponding integrals are calculated approximately, using one of the available quadrature formulae. For example, if the trapezoidal rule is used, then:

$$I_1^{(1)} = \int_0^{t_1} x(\tau)d\tau \approx \frac{t_1}{2}\cdot(x_0 + x_1);$$

$$I_3^{(2)} = -\int_0^{t_2} x(\tau)\cdot y(\tau)d\tau \approx -\left[\frac{t_1}{2}\cdot(x_0 y_0 + x_1 y_1) + \frac{t_2 - t_1}{2}\cdot(x_1 y_1 + x_2 y_2)\right]; \quad \textit{et cetera.}$$

Now the first equation of System (3) can be rewritten as an over-determined system of *N* ($N >> 4$) linear algebraic equations as follows:

$$\begin{gathered} a_{10}\cdot I_1^{(1)} + a_{11}\cdot I_2^{(1)} + a_{12}\cdot I_3^{(1)} + a_{13}\cdot I_4^{(1)} = \Delta x_1, \\ a_{10}\cdot I_1^{(2)} + a_{11}\cdot I_2^{(2)} + a_{12}\cdot I_3^{(2)} + a_{13}\cdot I_4^{(2)} = \Delta x_2, \\ \dots\dots\dots\dots\dots \\ a_{10}\cdot I_1^{(N)} + a_{11}\cdot I_2^{(N)} + a_{12}\cdot I_3^{(N)} + a_{13}\cdot I_4^{(N)} = \Delta x_N. \end{gathered} \tag{4}$$

Using the least squares method to solve this equation, the goal function is composed:

$$F(a_{10}, a_{11}, a_{12}, a_{13}) = \frac{1}{2}\sum_{k=1}^{N}\left[a_{10}\cdot I_1^{(k)} + a_{11}\cdot I_2^{(k)} + a_{12}\cdot I_3^{(k)} + a_{13}\cdot I_4^{(k)} - \Delta x_k\right]^2 \tag{5}$$

and then minimized. As a result the following system of linear equations is obtained:

$$\frac{\partial F}{\partial a_{10}} = a_{10} \cdot \sum_{k=1}^{N} I_1^{(k)2} + a_{11} \cdot \sum_{k=1}^{N} I_1^{(k)} \cdot I_2^{(k)} + a_{12} \cdot \sum_{k=1}^{N} I_1^{(k)} \cdot I_3^{(k)} + a_{13} \cdot \sum_{k=1}^{N} I_1^{(k)} \cdot I_4^{(k)} - \sum_{k=1}^{N} I_1^{(k)} \cdot \Delta x_k = 0\,,$$
$$\frac{\partial F}{\partial a_{11}} = a_{10} \cdot \sum_{k=1}^{N} I_1^{(k)} \cdot I_2^{(k)} + a_{11} \cdot \sum_{k=1}^{N} I_2^{(k)2} + a_{12} \cdot \sum_{k=1}^{N} I_2^{(k)} \cdot I_3^{(k)} + a_{13} \cdot \sum_{k=1}^{N} I_2^{(k)} \cdot I_4^{(k)} - \sum_{k=1}^{N} I_2^{(k)} \cdot \Delta x_k = 0\,,$$
$$\frac{\partial F}{\partial a_{12}} = a_{10} \cdot \sum_{k=1}^{N} I_1^{(k)} \cdot I_3^{(k)} + a_{11} \cdot \sum_{k=1}^{N} I_2^{(k)} \cdot I_3^{(k)} + a_{12} \cdot \sum_{k=1}^{N} I_3^{(k)2} + a_{13} \cdot \sum_{k=1}^{N} I_3^{(k)} \cdot I_4^{(k)} - \sum_{k=1}^{N} I_3^{(k)} \cdot \Delta x_k = 0\,,$$
$$\frac{\partial F}{\partial a_{13}} = a_{10} \cdot \sum_{k=1}^{N} I_1^{(k)} \cdot I_4^{(k)} + a_{11} \cdot \sum_{k=1}^{N} I_2^{(k)} \cdot I_4^{(k)} + a_{12} \cdot \sum_{k=1}^{N} I_3^{(k)} \cdot I_4^{(k)} + a_{13} \cdot \sum_{k=1}^{N} I_4^{(k)2} - \sum_{k=1}^{N} I_4^{(k)} \cdot \Delta x_k = 0 \tag{6}$$

This equation has a unique solution:

$$[a_{10}, a_{11}, a_{12}, a_{13}]^T = M^{-1} \cdot R \tag{7}$$

where matrices M and R are given by:

$$M = \begin{bmatrix} \sum_{k=1}^{N} I_1^{(k)2} & \sum_{k=1}^{N} I_1^{(k)} \cdot I_2^{(k)} & \sum_{k=1}^{N} I_1^{(k)} \cdot I_3^{(k)} & \sum_{k=1}^{N} I_1^{(k)} \cdot I_4^{(k)} \\ \sum_{k=1}^{N} I_1^{(k)} \cdot I_2^{(k)} & \sum_{k=1}^{N} I_2^{(k)2} & \sum_{k=1}^{N} I_2^{(k)} \cdot I_3^{(k)} & \sum_{k=1}^{N} I_2^{(k)} \cdot I_4^{(k)} \\ \sum_{k=1}^{N} I_1^{(k)} \cdot I_3^{(k)} & \sum_{k=1}^{N} I_2^{(k)} \cdot I_3^{(k)} & \sum_{k=1}^{N} I_3^{(k)2} & \sum_{k=1}^{N} I_3^{(k)} \cdot I_4^{(k)} \\ \sum_{k=1}^{N} I_1^{(k)} \cdot I_4^{(k)} & \sum_{k=1}^{N} I_2^{(k)} \cdot I_4^{(k)} & \sum_{k=1}^{N} I_3^{(k)} \cdot I_4^{(k)} & \sum_{k=1}^{N} I_4^{(k)2} \end{bmatrix}, \quad R = \begin{bmatrix} \sum_{k=1}^{N} I_1^{(k)} \cdot \Delta x_k \\ \sum_{k=1}^{N} I_2^{(k)} \cdot \Delta x_k \\ \sum_{k=1}^{N} I_3^{(k)} \cdot \Delta x_k \\ \sum_{k=1}^{N} I_4^{(k)} \cdot \Delta x_k \end{bmatrix} \tag{8}$$

Equation (7) solves the problem of the parametric identification of the first equation in System (1). Coefficients of the second and third equations of the system can be identified in a similar way.

4 Example

In Table 3 population census numbers for sixteen successive years are given, with wildebeest population numbers in the second row, lion numbers in the third row and zebra numbers in the fourth row. (Note that population numbers are measured in thousands).

Table 4

$XYZ^T =$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	12.01	10.15	7.67	6.01	5.46	5.66	6.31	7.22	8.18	8.94	9.15	8.59	7.5	6.55	6.11	6.17
1	0.31	0.54	0.68	0.58	0.4	0.27	0.2	0.17	0.18	0.22	0.32	0.45	0.52	0.49	0.4	0.31
2	10.02	9.63	7.82	6.33	5.8	6	6.67	7.63	8.7	9.61	9.98	9.49	8.37	7.34	6.82	6.82

All the unknown coefficients in the equations of System (1) are estimated according to the above-mentioned algorithm and the following results are obtained:

$$\begin{bmatrix} a_{10} \\ a_{11} \\ a_{12} \\ a_{13} \end{bmatrix} = \begin{bmatrix} 0.4257 \\ 0.0342 \\ 0.8397 \\ 0.0136 \end{bmatrix} \qquad \begin{bmatrix} a_{20} \\ a_{21} \\ a_{23} \end{bmatrix} = \begin{bmatrix} 1.5927 \\ 0.1139 \\ 0.0937 \end{bmatrix} \qquad \begin{bmatrix} a_{30} \\ a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} = \begin{bmatrix} 0.3563 \\ 0.0185 \\ 0.7824 \\ 0.0280 \end{bmatrix}.$$

These estimated coefficients are used to solve the system of three ordinary differential equations in (1) using the initial conditions given in the first column of table 3. Results of the proposed solutions are compared with the statistical data as shown in Figures 1 – 3.

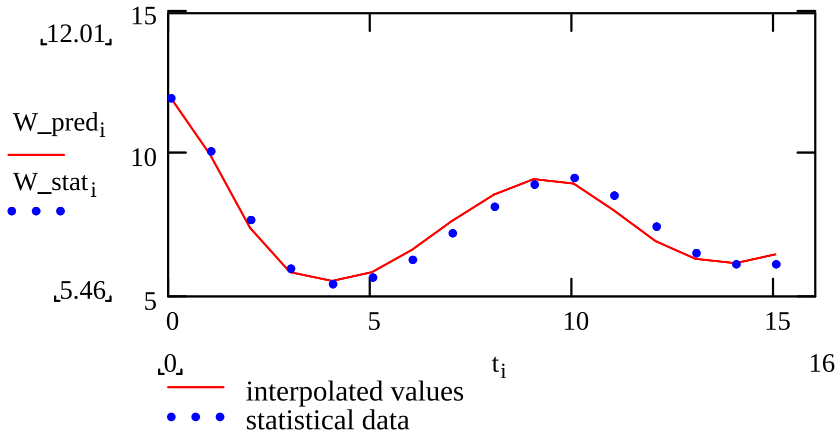

Figure 1: Comparison of interpolated population numbers of wildebeest according to System (1) and statistical data (both given in thousands)

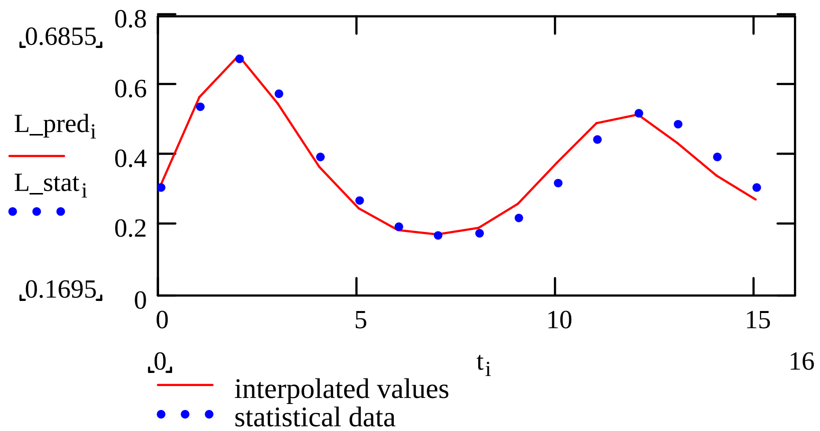

Figure 2: Comparison of interpolated population numbers of lion according to System (1) and statistical data (both given in thousands)

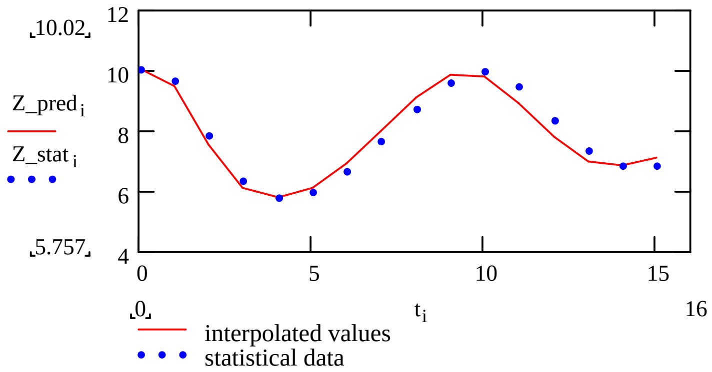

Figure 3: Comparison of interpolated population numbers of zebra according to System (1) and statistical data (both given in thousands)

Predicted population dynamics of wildebeest, lion and zebra for the next forty-five years are shown in Figures 4 -6. Predicted stable steady-state values of the populations over a period of 60 years are: $x_0 = 7.336$, $y_0 = 0.339$, $z_0 = 8.076$ (measured in thousands).

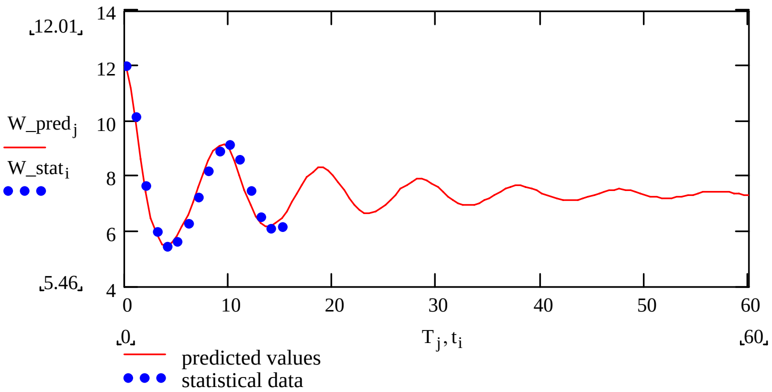

Figure 4: Predicted population dynamics of wildebeest over 60 years, according to System (1) shown with statistical data used in the algorithm

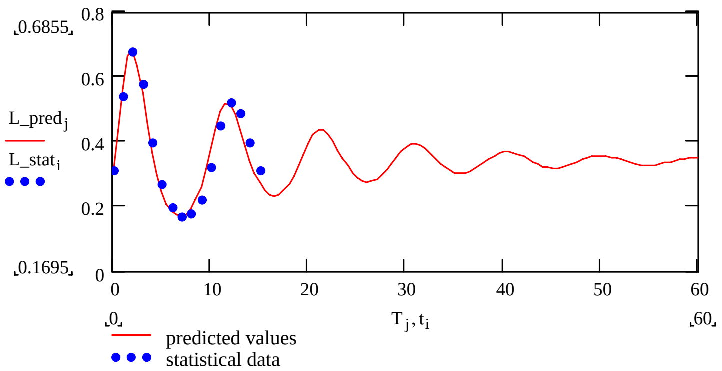

Figure 5: Predicted population dynamics of lion over 60 years, according to System (1) shown with statistical data used in the algorithm

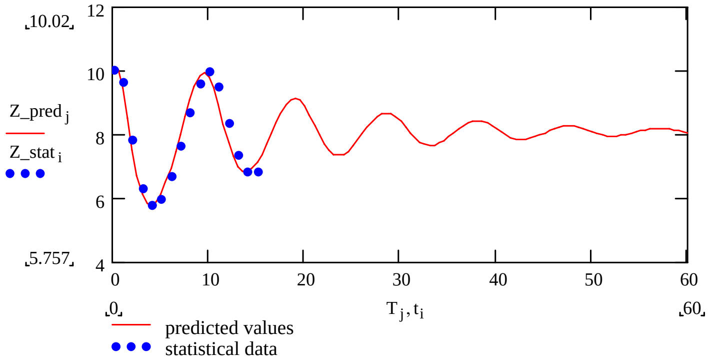

Figure 6: Predicted population dynamics of zebra over 60 years, according to System (1), shown with statistical data used in the algorithm

5 Discussion and conclusion

A mathematical model to describe the interaction between one predator and two prey species is formulated in this paper, taking logistic growth and mutualism effects of the prey species into consideration. The algorithm of parametric identification of this model is discussed and illustrated by example. It is shown that the algorithm gives good interpolation of the statistical data for the period of fifteen years of observation. Predictions on the dynamics of the predator-preys interaction are made for the period for the next forty-five years on the basis of the parametrically identified model. These predictions guarantee convergence of predator and prey numbers to the stable steady-state. Despite good correspondence of the interpolated and statistical data, the convergence of data to their steady-state is slow, yet the oscillatory behaviour of the populations corresponds to the field data. The accuracy of the mathematical model could further be improved by including the effect of the predator's "fullness" (see [3] & [4]), and the effect of seasonal calving of wildebeest and zebra.

References

[1] A. Bazykin (1998). Nonlinear Dynamics of Interacting Populations (Series in Neural Systems), *World Scientific Pub Co Inc*.

[2] T.H. Fay and J.C. Greeff (1999). Lion and Wildebeest: A Predator-Prey Model, *Mathematics and Computer Education*, 25, pp. 106-119.

[3] T.H. Fay and J.C. Greeff (2006). Lion, Wildebeest and Zebra: A Three Species Model, *Ecological Modelling*, 196, pp. 237 – 244.

[4] J.C. Greeff, H.T. Groeneveld & B.E. Eisenberg (2006). Testing the Model for one Predator and two Mutualistic Prey Species. *Ecological Modelling*, 196, pp. 245 – 255.

[5] R.M. May (1971). Stability and Multi-species Community Models. *Bulletin of Mathematical Biosciences*, 12, pp. 59-79.